THE DEVELOPMENT OF EXECUTIVE FUNCTIONS AND INFORMATION PROCESSING SPEEDS IN TODDLERS BORN PRETERM

Kayleigh Lauren Day

University College London

A thesis submitted for the degree of

Doctor of Philosophy (PhD)

September 2017

Declaration

I, Kayleigh Lauren Day confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Supervisors

Dr Michelle de Haan **Professor Neil Marlow**

Abstract

Cognitive impairments are commonly reported in children born preterm, with particular difficulties in executive functions, information processing and attention. Yet, children that later present with mild to moderate impairments are typically missed in earlier standard developmental assessments. The current longitudinal investigation explores the development of executive function, information processing and attentional abilities in very preterm (<32 weeks of gestation) and term-born children at 3, 6, 12 and 30 months of age, corrected for prematurity. Performances on established paradigms assessing these cognitive abilities were also compared to the cognitive composite scores of the Bayley Scales of Infant and Toddler Development (third edition) at 12 months and 2 years of age.

Very preterm (n = 50) and term-born (n = 81) children were assessed in a multifaceted battery of behavioural, eye-tracking and event related potential tasks, to formulate a detailed understanding of developmental trajectories for executive functions, information processing and attention.

Overall, cohort performances were not differentiated within the first year and measures of attention were comparable for both groups over the two years. However, executive function and information processing differences were observed within the very preterm children during the second year assessments. These difficulties were independent of global cognitive performance, and variation on the executive function, information processing and attentional measures was poorly reflected in the Bayley-III cognitive scores at 2 years.

In conclusion, very preterm children display difficulties predominantly in executive function abilities by 2 years of age, independent of global cognitive scores. Longer-term follow-up of this cohort will highlight any links between these early deficits and later academic and social outcomes, and can aid in the development of tools for earlier identification of adverse cognitive outcomes in very preterm populations.

Acknowledgments

First and foremost, I would like to express my upmost gratitude to my two supervisors, Dr Michelle de Haan and Prof Neil Marlow, for their continuous support, patience and guidance over the past 4 years. I feel incredibly privileged to have had this opportunity to work with you both and I want to thank you for all of your motivation, encouragement and for the knowledge you have shared over the years.

This current investigation would not have been possible without the dedication and support of all the participating families and children involved. This has been an incredibly enjoyable project, and has been wonderful to see each little personality start to emerge over the two years. Thank you all!

To all of my wonderful friends and colleagues that have been part of the Preterm Development Project over the years, there are too many of you to name, but to each and every one, I thank you for the help and support you have given me personally and to the project as a whole. In addition, I would like to thank Deirdre Murray and her team in the University of Cork for allowing the use of the Babyscreen application, presented in this thesis, and particular thanks to Deirdre Twomey for her time helping with the analysis approach for the BabyScreen data.

To my Mum and Dad, in every decision I make, you are always there to support me. I cannot put into words how grateful I am, and I owe my every success to you. And to my sisters, Syanne and Aysha, thank you for your motivational phone calls and never failing to make me smile. I cannot thank you all enough for your endless support of my studies and for being there always.

Finally, to my Fiancé Tom, without your love, patience, understanding and ability to know the right things to say to pull me together when I needed it most, this whole process would not have been possible without you. Thank you, you are my rock.

I wish to express my gratitude to the funders of my PhD, ICH IMPACT studentships and Henrietta St George whose generous donation to Great Ormond Street Hospital Children's Charity went towards supporting this project. Without this financial support this research would not have been possible.

Table of Contents

Table of Tables

Table of Figures

Abbreviations

EF –Executive Function

IQ – Intellectual Quotient

Bayley-III – Bayley scales of Infant Development third edition

IVH = Intraventricular Haemorrhage

PVL = Periventricular Leukomalacia

ROP = Retinopathy of Prematurity

CLD = Chronic Lung Disease

BPD = Bronchopulmonary Dysplasia

NEC = Necrotizing Enterocolitis

IP – Information processing

ADHD – Attention deficit hyperactivity disorder

VDU – visual display unit

Chapter 1 Background

Survival following preterm birth is continually increasing due to advances in perinatal and neonatal care (WHO, 2012). Tracking the effects of preterm birth, in particular at very low gestations, is a focus of much research effort. As gestation at birth decreases a range of serious impairments affecting motor (cerebral palsy) and cognitive function become more frequent among survivors. Although these impairments may be decreasing in frequency (Moore, Hennessy, *et al.*, 2012), survivors of very premature birth (<32 weeks gestation) remain at risk of impairments in infancy and preschool ages, that later manifest as lower academic attainment compared to children born at full term (Kerr-Wilson *et al.*, 2012). The prevalence of global cognitive impairment, as measured by low Intellectual Quotient (IQ), increases with decreasing gestation (Mulder *et al.*, 2009; Kerr-Wilson *et al.*, 2012). However, academic difficulties observed within this population may be compounded by additional factors contributing to overall performances, with more prominent impairments in specific cognitive domains in addition to lower IQ (Aylward, 2002; Anderson, 2014).

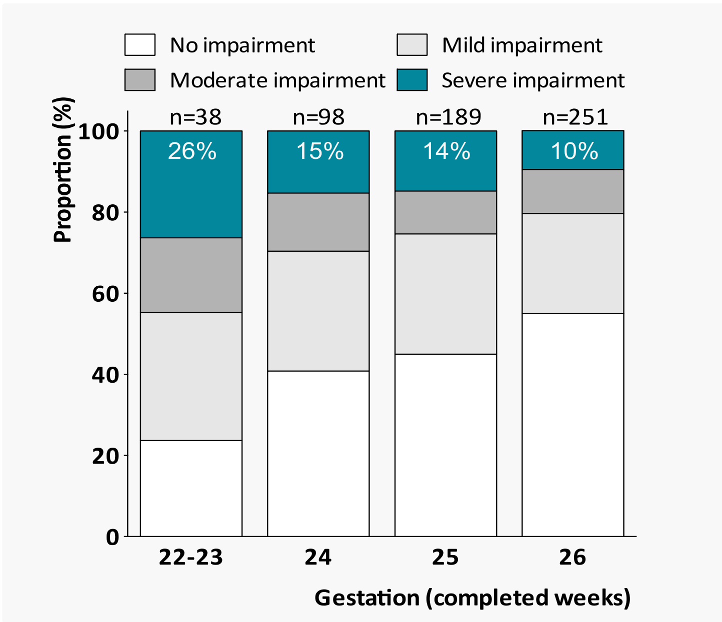
A particular focus within this population is executive function and information processing (Rose, Feldman and Jankowski, 2009, 2011; Mulder, Pitchford and Marlow, 2010, 2011b). Executive function or EF, is an umbrella term for a set of effortful cognitive processes responsible for directing focus and goal-setting behaviours (Diamond, 2006, 2013). The speed of information processing (or IP) refers to the speed at which the brain functions, both in terms of automatic responses and effortful mental processing in response to stimuli. Both areas have been previously reported as areas of difficulty within preterm populations (Johnson, 2007; Rose *et al.*, 2008; Mulder, Pitchford and Marlow, 2011b). Problems with EF become apparent and easily quantifiable at early school age. It is unclear whether EF dysfunction is not quantifiable until then or whether current developmental assessments, made usually at around 2 years of age, are too insensitive to detect subtle cognitive difficulties. The goal of current research is to improve identification methodologies, aiding interventions before the school years, and helping structure the development of those that require the support before it impacts their education. In order to achieve this, a better understanding is required regarding when the stated delays start to emerge. It is possible specific impairments in EF and IP are apparent from birth and, although likely to be subtle, could be detected if appropriate methods are identified. Equally these delays may emerge slowly across development, making early and accurate identification difficult.

This thesis seeks to evaluate the development of executive functions and information processing speeds through the first two and a half years of life in children born very preterm, with the focus of identifying emerging signs of dysfunction that are not detected by that of the current standard developmental assessments.

1.1 Preterm development

Infants born very preterm (<32 weeks of gestation) require significant and highly skilled medical interventions to survive. Advances in neonatology and obstetrics in the last few decades have led to a surge in preterm survival rates and at much younger gestational ages than ever before (Sun, Mohay and O'Callaghan, 2009). Thus developments in neonatal care are proving very successful. However, the risk of severe cognitive, sensory and motor impairments in these infants still remains high (see figure 1-1) (Moore, Hennessy, *et al.*, 2012). Impairments can range from severe mental and physical disabilities, including Cerebral Palsy (CP), Developmental Coordination Disorder (DCD) (Sun and Buys, 2012b), hydrocephalus, and neurosensory impairments such as blindness and deafness (Johnson, Wolke and Marlow, 2008), to mild cognitive impairment. Research to date is yet to identify biomarkers in infancy that could predict the level of impairment likely to be observed in a child at a later stage of development. Whether these biomarkers exist, remains to be seen.

Overall impairment:



Motor grades:

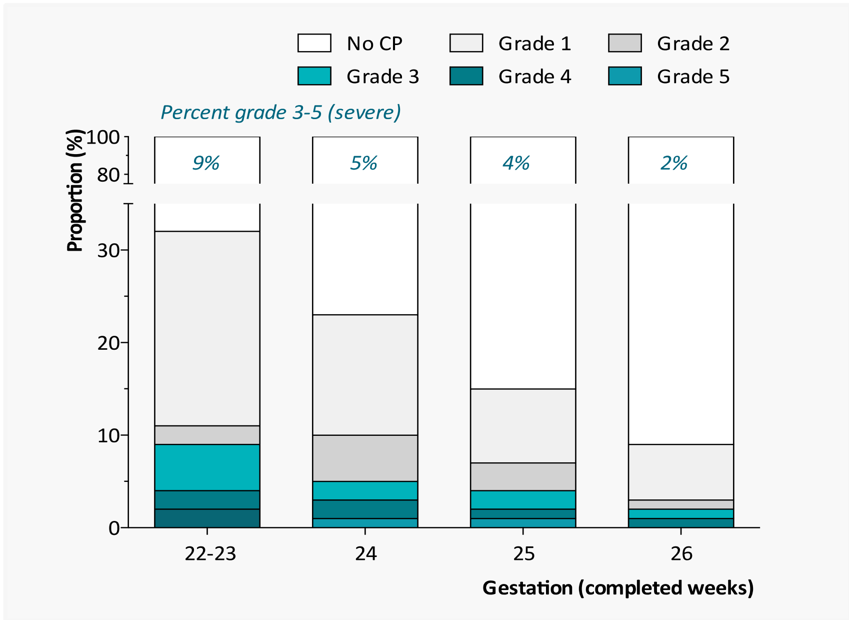


Figure 1-1. Both images taken from Moore *et al.*, (2012). Impairment rates in EPICure2 cohort, including infants born before 27 weeks in 2006. To the left: overall level of impairment including motor, sensory, communication and developmental domains, categorised by national consensus recommendations at the time of publication. To the right: functional outcomes according to the Gross Motor Function Classification System.

During the final stages of fetal development, the brain is differentiating at its highest rate, making this an extremely sensitive phase of development (Huppi, 2010), see figure 1-2. Even subtle disturbances of the intrauterine environment, especially if they occur in the mid second and early third trimesters, can lead to serious impairments later in life (Johnson, Wolke and Marlow, 2008; Ishii and Hashimoto-Torii, 2015). The brain outcome for very preterm children is related to a complex amalgam of destructive and developmental influences. Both focal injuries and changes to regional developmental trajectories are influenced by both external and internal factors. Examples of external factors include perinatal infection and perfusion fluctuations that lead to haemorrhagic and ischaemic injuries. Internal factors are disturbed by the switch from intrauterine to independent life, causing alterations to the developing organisation of the brain in terms of differentiation, migration, myelination and synaptogenesis (Volpe, 2009). In a number of cases, it is likely that severe motor and cognitive outcomes are related to the topography of acquired brain injury (Krägeloh-Mann, 2004), however, biomarkers for specific cognitive deficits are yet to be determined.

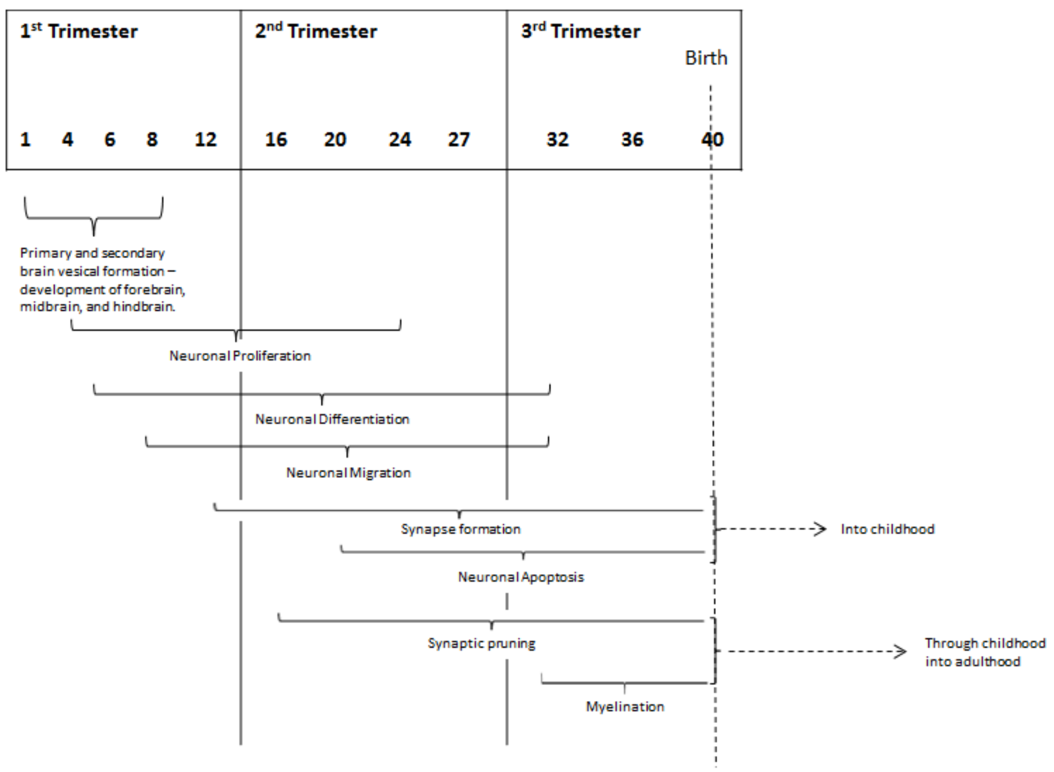


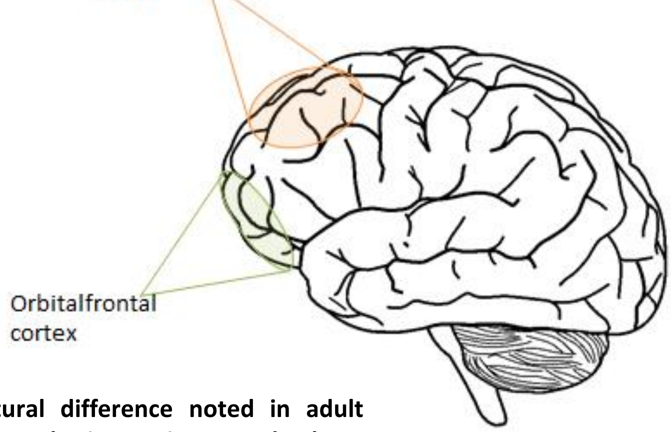
Figure 1-2. Neuronal growth during fetal development. Adapted from (Allen and Kelly, 2015)

Premature infants do not always present with focal neurological injury, yet can still later display cognitive, behavioural and/or neuropsychological deficits (Bhutta *et al.*, 2002). These deficits are termed by some as 'hidden disabilities' due to the absence of neurological injury, and can thereby be difficult to predict early on (Johnson, Wolke and Marlow, 2008). These high prevalence/low severity deficits are most often reported when children reach school age and can significantly impact early learning and influence later school achievements, both academically and in a social context. This puts those that present with these cognitive and neuropsychological deficits at a disadvantage to their term-born peers in many areas (Johnson, Wolke and Marlow, 2008).

Figure 1-3 illustrates and summarises the brain regions associated to executive functions (EF), the cognitive abilities most closely linked to academic attainment, and details neuroanatomical differences reported in the preterm literature.

Brain regions associated with EF:

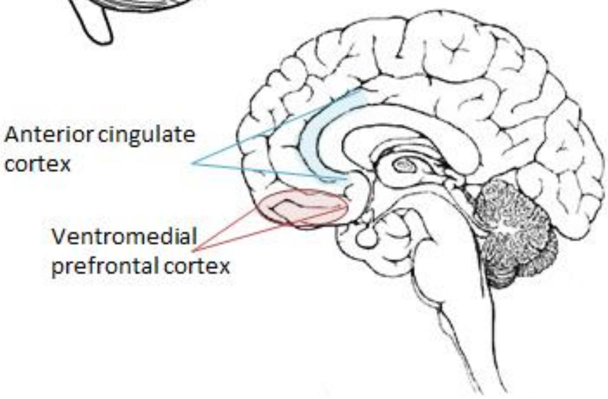
- Dorsolateral prefrontal cortex: the predominant region associated to EF. Used to maintain information in mind, set shifting, planning and problem solving (Alvarez and Emory, 2006; Diamond, 2013)
- Orbitalfrontal/ventralmedial cortices: responsible for socially acceptable behaviour and linked to emotion. Therefore linked to inhibition but often debated as self-regulation oppose to EF (Alvarez and Emory, 2006; Diamond, 2013).



- Anterior cingulate cortex: activation has been linked to selective attention and although not typically considered part of EF, this ability and therefore this region has been considered to be essential for EF task completion (Peterson *et al.*, 1999; Alvarez and Emory, 2006).

Structural difference noted in adult preterm brain regions and those associated with EF tasks:

- Reduced white matter volume in posterior corpus callosum, thalamus and fornix in combination with reduced grey matter volume in temporal gyri have been reported to account for 21% of performance differences in EF in an ex-preterm population (Nosarti *et al.*, 2014)



- Reduced activity has been reported in the dorsolateral prefrontal cortex in response to a working memory task in ex-preterm adults that suffered severe neonatal brain injury (Kalpakidou *et al.*, 2014).

Figure 1-3. The predominant brain regions involved in EF in typically developing populations and a brief summary of difference reported in brain region functionality in preterm populations associated with EF performance.

1.1.1 Influence of neonatal factors on cognitive function

Neonatal complications following premature birth are extensive. Accounting for the numerous complications and possible influences on later development is almost impossible due to the varying degrees of illness during the perinatal period. However, although complex, it is likely a relationship exists between adverse clinical events and cognitive function later in development (Linsell *et al.*, 2015).

Before the age of 5 years, global cognitive performance is commonly measured by the cognitive composite score on the current third edition of Bayley Scales of Infant and Toddler Development (Bayley-III) or previously, the Mental Development Index from the Bayley-II. After 5 years of age, IQ is typically used to define global cognitive performance. Consistently, correlations are observed between poorer cognitive outcomes with decreasing gestational age (Bhutta *et al.*, 2002; Johnson, 2007; Orton *et al.*, 2015; Johnson and Marlow, 2016). In infancy, male sex, lower birth weight, non-white ethnicity and parental education (or social economic status/SES) all have been shown to be strong predictors of poorer global cognitive outcomes. However, apart from parental education, these prognostic effects appear to diminish in later childhood (Linsell *et al.*, 2015). The effect of SES on cognitive outcome continues to be reported in the adult literature, above other biological factors (Tideman, 2000; Hack, 2006). Although the health of ex-preterm adults can be effected by the complications associated to early life experiences, there is no evidence to suggest the perinatal complications consistently influence outcomes beyond that of gestation age at birth and known brain injuries (Marlow, 2004).

Reports of specific neonatal risk factors associated with deficits in EF are mixed. A meta-analysis by Linsell *et al.,* (2015) identified 7 studies investigating prognostic factors for EF deficits in preterm cohorts born after 1990, with a wide variety of tests used to assess EF. The inconsistencies of measures used, combined with the small number of studies identified, rendered it difficult to meaningfully combine the results to ascertain the predominant factors associate with EF impairments in early childhood (Linsell *et al.*, 2015). More definitive correlations have been reported in the MR-based studies (Howard, K., Anderson, P. J., & Taylor, 2008; Vollmer *et al.*, 2017). Diffuse white matter abnormalities identified within the neonatal period correlate with overall IQ performance and specific EF and attentional deficits within preterm populations (Vollmer *et al.*, 2017). White matter tracts such as the fronto-striatal pathway, connecting the frontal lobe to the basal ganglia, and fronto-occipital pathway, connecting the frontal lobe and visual cortex, have been identified as two of many long association fibre tracks at risk within this population. These alterations in white matter tract integrity appear to be present even in the absence of major focal brain injury (Li *et al.*, 2015; Vollmer *et al.*, 2017). Conventional radiological images do not identify such problems, and computational MRI techniques are not used in clinical practice to date (Duffau, 2014). These alterations are not immediately apparent yet are likely to have functional implications (Vollmer *et al.*, 2017).

Although difficult to translate to human observations, animal studies have observed functional outcomes to be predominantly affected if an injury to the brain, such as hypoxia, occurs during neuronal migration (Kolb *et al.*, 2013). Very preterm delivery occurs during the developmental stage of neuronal migration (figure 1-2). Poor oxygenation of the brain during the preterm birth is a common occurrence and has been associated to specific deficits within the EF sub-skill, working memory (Taylor *et al.*, 2004). These sub-skills and associated deficits will be discussed in greater detail in subsequent sections. In a more recent investigation into Apnoea of Prematurity (AOP), intermittent hypoxia used to model AOP, had a neuroprotective effect in rodents with brain lesions and in response to behavioural stressors (Bouslama *et al.*, 2015). This highlights the inconsistencies within the literature and the need for further research.

In the current thesis, very severe brain injuries were excluded. Neonatal complications will be summarised, however, they will not be taking into consideration within the main study analyses relating to cognitive performance given the mixed reports within the literature. The factors considered to be most influential of later outcome in the early years and therefore will be considered within all analyses are SES as defined by the Index of Multiple Deprivation quintile (a nationally available score related to postcode (NPEU, 2013)) and male sex.

1.1.2 Cognitive impairments and later academic achievements in children born preterm

Global cognitive impairment in the early years, as defined by low scores on cognitive scales in measures such as the Bayley-III is a common finding following preterm birth (for review see (Anderson and Doyle, 2008)). This impacts learning in early school years and into middle childhood; see Figure 1-4.

Ex-preterm children can present with a wide range of developmental problems including; speech and language delay; impaired attentional abilities; working memory impairments; reduced processing speeds and later specific difficulties in academic areas such as reading and mathematics (Johnson, Hennessy, *et al.*, 2009; Woodward *et al.*, 2009). Difficulties can range from mild through to severe and are not consistent across preterm cohorts; a proportion of children do not present with any notable difficulties. The problem that faces clinical professionals is the sensitivity and predictive validity of the developmental assessments used to identify these difficulties. The Bayley-III (Bayley, 2006) is the current standard assessment tool in the UK to determine the achievement of developmental milestones in the early years. Significant cognitive impairments are detected by scores 2 standard deviations (SD) below the mean on measures such as the Bayley-III and are often found to be predictive of later learning and cognitive difficulties. However, those with milder impairments are often missed in the initial developmental assessments (Aylward, 2002; Hack *et al.*, 2005). A discernible pattern to predict those that will and those that will not present with cognitive impairments later in life is yet to be uncovered.

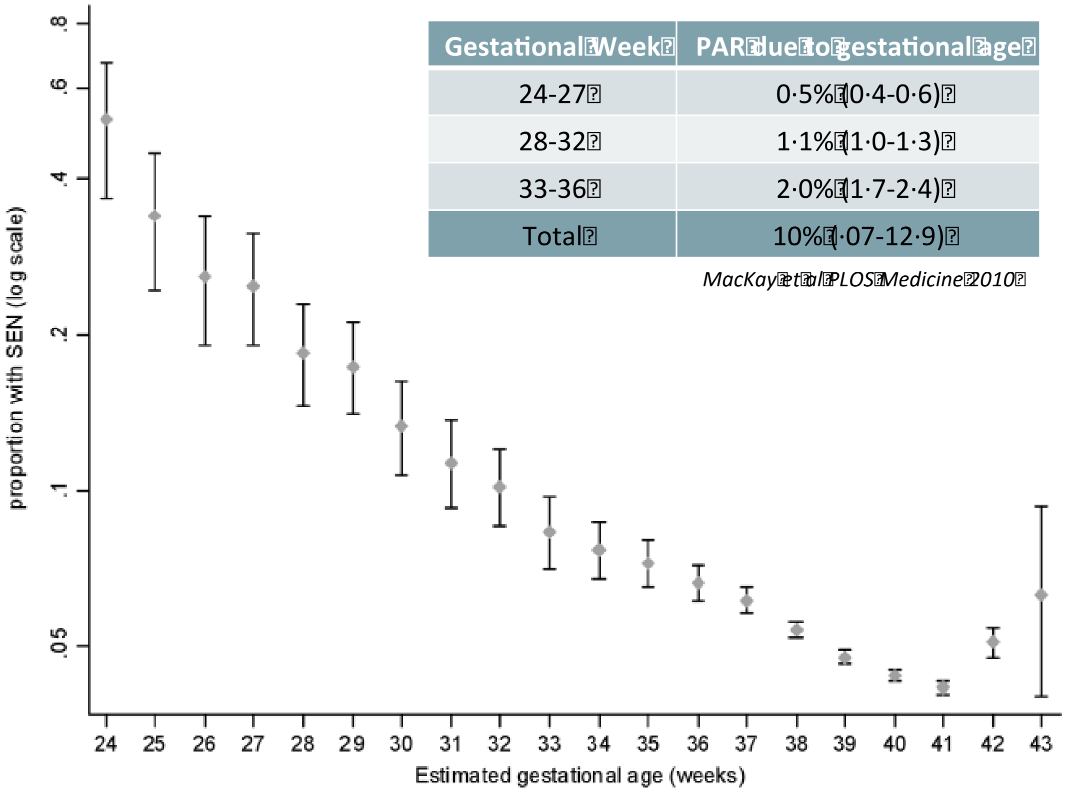


Figure 1-4. Continuum of special needs by gestational age at birth by Mackay (2010)

Academic attainment in middle childhood, specifically in reading and maths, appear to be areas of greatest difficulty compared to other academic abilities. A significant proportion of ex-preterm children require additional educational support, Figure 1-4 (Bull and Johnston, 1997; Bowen, Gibson and Hand, 2002; Johnson, Hennessy, *et al.*, 2009; Odd, Evans and Emond, 2013). Low IQ scores influence academic achievement in middle childhood for ex-preterm populations with these difficulties persisting into adulthood (Nosarti *et al.*, 2007; Breeman *et al.*, 2015; Burnett *et al.*, 2015; Eryigit-Madzwamuse and Wolke, 2015; Johnson and Marlow, 2016). However, low IQ does not appear to account for all difficulties observed. Those with less severe impairments, although scoring in the low average range for IQ, appear to have problems relating to specific functions, namely: visual-perceptual/motor abilities, attention, reading, writing, spelling and mathematical skills (Aylward, 2002). The neuropsychological functions reported here utilise a set of cognitive abilities termed Executive Functions (EF). EF is commonly considered an umbrella term encompassing multiple sub skills, essential for problem solving and cognitive regulation, assessments of which, have been found to account for the variability in academic attainment in healthy control cohorts (Blair and Razza, 2007; Bull, Espy and Wiebe, 2008; Mulder, Pitchford and Marlow, 2010). Poor performance in EF assessments in preterm populations compared to term born peers following adjustment for IQ are repeatedly found (Espy *et al.*, 2002; Bohm, Smedler and Forssberg, 2004; Aarnoudse-Moens, Smidts, *et al.*, 2009; Mulder *et al.*, 2009; Sun and Buys, 2012a). EF sub-domains typically include inhibition, working memory, and cognitive flexibility. However, some authors consider these processes to linked with an underlying mechanism and therefore view EF as a unitary concept (Miyake *et al.*, 2000).

Mathematical abilities appear to be a particular area of difficulty in preterm children, independent of the lower global cognitive scores (Johnson, Hennessy, *et al.*, 2009). In line with this, studies show preterm children are more likely to fail age-appropriate maths questions than their term born peers, after adjustment for general cognitive ability (Simms *et al.*, 2013). These children are reported to find more complex, simultaneous mathematical problems more difficult to those that are sequential in structure, implying their difficulties are associated with specific cognitive functions involved in complex mathematical processing. Working memory, perceptual and attentional problems, and visuo-spatial inabilities are all speculated to contribute to these difficulties (Marlow, 2004; Johnson, Hennessy, *et al.*, 2009).

This is in contrast to the reading and spelling difficulties in which delays are accounted for by general cognitive ability (Johnson et al., 2009). The processes of learning to read, write and solve mathematical problems require the acquisition of simple skills, such as letter and number recognition, before complex skills, such as reading and addition, can be achieved. It is speculated that crystallized knowledge, information that is gleamed from past experiences, forms the basis of these simple skills. However in order to acquire this knowledge, it has been argued a biological basis is involved in learning. EF abilities, specifically working memory performance, influence the acquisition of new information and dictate a child's capacity to learn. Mathematical abilities appear to be particularly dependent upon working memory performance. A predictive relationship has been frequently reported between working memory abilities and early mathematical scores in typically developing cohorts (Gathercole *et al.*, 2004). Given such reports, it is perhaps unsurprising to observe difficulties within the working memory domain in preterm populations correlating to performance in mathematical assessments (Mulder, Pitchford and Marlow, 2010).

The nature of mathematical difficulties observed in this population is understood to be different to that of development dyscalculia (Simms *et al.*, 2015). Developmental dyscalculia is a condition whereby an individual has specific numerical difficulties related to the approximate number system, a system that handles quantity information, both in terms of representation and manipulation, in order to achieve mathematical problems (De Smedt *et al.*, 2013; Simms *et al.*, 2015). Within preterm populations, more complex calculations appear to be the most prominent difficulty and it is thought deficits in working memory may explain this, not the way the brain characterises numerical presentations. However, it is not clear why mathematics is an area of particular difficulty within this population, and literacy skills are spared (Isaacs *et al.*, 2001). Current research suggests a possible explanation for these difficulties lies in the reduced grey matter volume reported within the intraparietal sulcus (IPS) in preterm populations with numerical difficulties (Isaacs *et al.*, 2001); a region commonly associated to number processing and calculation ability (Klein *et al.*, 2014). This research correlates with other studies reporting grey matter reductions within this region of the brain and associated EF difficulties, see Figure 1-3 (Kalpakidou *et al.*, 2014; Nosarti *et al.*, 2014).

The most consistent domains found to account for academic difficulties, and the strongest predictors of academic attainment, are speed of processing and working memory abilities (Rose and Feldman, 1996; Rose, Feldman and Jankowski, 2002; Mulder, Pitchford and Marlow, 2010). Commonly these two domains are considered to be linked: processing speed may mediate EF performance (Rose and Feldman, 1996; Bull and Johnston, 1997; Rose, Feldman and Jankowski, 2002, 2009, Mulder, Pitchford and Marlow, 2010, 2011b). The study by Mulder *et al.* (2010), found a global deficit in processing speed the likely explanation of some of the difficulties in academic attainment observed within their study (Mulder, Pitchford and Marlow, 2010). However, this group also reported working memory abilities were independently predictive of later academic achievements.

Bull and Johnston (1997) proposed that mathematical difficulties are likely to be explained by information processing speed. Two theories were put forward by this group; firstly, reduced speed of processing could reflect the way the brain processes information as a whole and therefore the speed of completion of mathematical problems is thereby reduced. Alternatively, recall of basic information within crystallised knowledge and failure to automate simple mathematical operations could also explain performance difficulties within these children (Bull and Johnston, 1997).

The inclusion of processing speed, specifically verbal processing, and working memory in a predictive model of academic outcome, the same amount of variance was accounted for as full scale IQ (Mulder, Pitchford and Marlow, 2010). Due to the relationship between working memory and processing speeds in determining academic attainment, it would be valuable to study longitudinally from the early years to investigate how early these functions predict later capabilities.

When exploring EF abilities in the early years, it is valuable to consider when the construct is evaluable in terms of first emergence, the differentiation and the trajectory of the construct through childhood. The literature is divided on the topic of EF and its developmental trajectory, and in order to make predictions about later functioning based on specific aspects of this construct, it is important to consider the likelihood of targeting such domains when they are in their infancy. Section 1.2 explores the multiple theories surrounding EF.

1.1.3 Current clinical practice

Due to the recognised risk associated with preterm birth, several strategies are in place to evaluate outcomes in very preterm children (Kallioinen *et al.*, 2017; NICE, 2017). Before the release of the updated NICE guidelines in August 2017, standard practice in the UK following preterm birth of before 32 weeks gestation required infants to be followed up at 4 time points after discharge from hospital; at 3, 6, 12 and 24 months of age. The most recent guidelines include an additional follow-up within the fourth year (NICE, 2017). The most frequently used standardised assessment in preterm follow-up clinics in the first two years after discharge is the Bayley-scales of Infant and Toddler Development (current version, edition III) (Bayley, 2006). These follow-ups are recommended to all clinics, however, variations in practical and economic allowances across difference hospital trusts may mean assessments such as these are not always performed. Tracking the development of these children in this crucial and the nature of these assessments should work in theory, but this does not always translate into clinical practice. Multiple hour-long assessments are expensive and require well-structured comprehensive follow-up services. The difficulties in implementing these practices suggest a need to address the ease of follow-up in this population, potentially calling for faster and more efficient methods of assessment.

In addition to these clinical implications, many studies have reviewed the current edition of the Bayley following concerns regarding the sensitivity and predictive validity of the tool (Milne, McDonald and Comino, 2012; Aylward, 2013; Spittle *et al.*, 2013; Johnson, Moore and Marlow, 2014; Mansson and Stjernqvist, 2014; Spencer-Smith *et al.*, 2015; Anderson and Burnett, 2017). The overwhelming conclusion of the majority of these studies finds the current edition not sensitive enough to detect children with mild cognitive impairments. Although not originally designed to predict IQ, the Bayley-III is commonly reported to under-identify those likely to later present with cognitive impairments, and exhibits particular difficulty in recognising those that fall within the mild impairment range that are likely to require additional assistance before starting school (Anderson and Burnett, 2017). Prior to the Bayley-III, the Bayley-II was the most frequently used infant developmental assessment and yielded much professional confidence (Johnson and Marlow, 2006). Upon an update to the test, the re-standardisation procedure used different reference population strategies, including seeding the population with poorly performing groups. Final scores were found to be 7 points higher than Bayley-II according to the publisher (Bayley, 2006; Moore, Johnson, *et al.*, 2012). This lower sensitivity at lower scores and a non-linear relationship with the Bayley-III brings challenges when targeting the detection of mild impairments as in preterm cohorts (Moore, Johnson, *et al.*, 2012). The proposition from a number of studies to overcome this insensitivity in the first instances is to use 1 standard deviation below the norm, or to formulate a local control reference for the purposes of research (Marlow, 2013; Spencer-Smith *et al.*, 2015).

Assessments at 2 years of age also show poor predictive validity of cognitive performance at 5 (Potharst *et al.*, 2012). Development of better measures is therefore essential before successful interventions can be put in place that work to improve later outcomes. This emphasizes the importance of obtaining a greater understanding of the domains that underpin the cognitive deficits later observed, and the subsequent need to identify the development profile and trajectory of specific domains. This will allow the necessary predictions to be made regarding which infants may benefit from intervention services. More broadly, early interventions have shown to have advantageous effects on the developmental profile within infancy and leading on to improved cognitive performance at preschool age, including in preterm populations (Spittle *et al.*, 2007; Hadders-Algra, 2011). However, evidence is limited that interventions significantly improve longer term cognitive performances within ex-preterm children, supporting the requirement of more extensive research (Spittle *et al.*, 2007).

A key factor to consider in the assessment of premature cohorts is the adjustment for age at birth. Accounting for prematurity in developmental assessments has been a long running discussion within the literature. The disagreements stem from two viewpoints: the biological opinion and the environmental opinion. The biological perspective states development takes a set time from conception. Originally proposed by Gesell and Amatruda (1947) (for review see Wilson and Cradock, 2004), they stated development was dictated by time itself and was not influenced by external factors. As such, a preterm child would lag behind a term born child on a developmental basis, at least initially, due to the underdeveloped central nervous system. Once the nervous system was fully developed the opinion within the literature employed the concept that 'catch up' growth would occur and the preterm child would meet the developmental stage of a term born within the first few years after birth. On the contrary, the environmental perspective suggests that development is primarily driven by external influences and the exposure to the outside world, with factors such as medical care and parental stimulation, advances and aids development (Wilson and Cradock, 2004). In support of this is the research into the benefits of higher social economic status effects on developmental achievement (Tideman, 2000; Hack, 2006). In any regard, the set clinical practice following preterm birth is to utilise correction for gestational age as proportionately using a biological basis, the assessment at chronological age puts preterm children at a disadvantage, and this disadvantage increases with lower gestational age (Wilson-Ching *et al.*, 2014). Therefore in the current investigation adjusted ages will be utilised.

1.2 Development of Executive function

Executive function, executive or cognitive control (Diamond, 2006), as discussed, are terms used to encompass a number of cognitive abilities. Multiple attempts have been made to clearly define the term 'Executive Function' (EF), and as described by Bohm *et al.* (2004), these various different explanations clearly demonstrates the complexity of the system. The general consensus is EFs include our ability to plan, inhibit behaviours, shift between tasks, use and understand verbal and non-verbal communication, sustain and manipulate information within our working memory, and in some instances, our attentional abilities. These are often condensed into three main areas: cognitive flexibility, working memory, and inhibition (Diamond, 2013). The purpose of these domains are to work together to in order to achieve a personal or social goal (Bohm, Smedler and Forssberg, 2004; Mulder *et al.*, 2009). This three part model is supported by studies in pre-adolescent children and adults. However, Miyake and colleagues have recently proposed a new framework that includes an additional common EF factor, considered to account for the similarities that overlap the three factor model (Miyake and Friedman, 2012).

There are several theories regarding the development of EF. Some preschool children are described as having executive dysfunction due to the lack of control over these abilities within the early years. Exactly when and how these different domains come into play is still a matter of debate (Isquith *et al.*, 2005). Multiple models have been put forward to explain the emergence of these abilities, three of most the predominant models are briefly discussed and summarised in Figure 1-5.

Miyake *et al.*, (2000), built a model around the focus of the inter-relatability and the independent nature of the three main EFs; cognitive flexibility, working memory and inhibition. 'The unity and diversity framework' derived from their research in 2000, concluded that there is both unity and diversity within the EF domains, with it possible to measure each as separate entities, but all correlate in terms of overall function. When investigating traditional paradigms in relation to their model, Miyake and colleagues results suggested that each EF domain contributed differently to the performance on each task, reaffirming the consideration of each as distinct functions with underlying commonality (Miyake *et al.*, 2000). Following additional research, the group went on to further develop this model, delving deeper into the unity that connects the three sub-skills. This led to the group proposing an additional factor within the model, termed common EF, and refers to the cognitive underpinnings consistent across the three domains, whilst simultaneously considering what makes each domain unique. When using this model including the unity across measures, flexibility and updating (working memory), all appeared independent factors and explained differences in inhibition (Figure 1-5). The authors drew three further conclusions when considering variability in EF. Firstly that genetic variations may contribute to the overall functioning of the sub-domains; secondly that individual variability may relate to 'clinically and societally important behaviours' and by extension therefore effecting the performance on EF tasks; and lastly that variability also presents with developmental stability (Miyake and Friedman, 2012). The final point is fundamental in relation to the current thesis. Although their work focused around the composition of EF in adult populations, the changes in relationship across the different domains remained consistent over longitudinal measures. Many authors subscribe to the view that childhood EF is a unitary concept (Wiebe, Espy and Charak, 2008; Hughes *et al.*, 2010) as developmental studies into the three factor view have produced mixed results (Lehto *et al.*, 2003; Huizinga, Dolan and van der Molen, 2006; Howard, Okely and Ellis, 2015).

Reports of differentiation in performance across EF sub-skills have been observed in the early year studies (Lehto *et al.*, 2003; Huizinga, Dolan and van der Molen, 2006). However, the interpretation of this differentiation and the extent of each sub-skills contribution to the overall construct is not consistent (Howard, Okely and Ellis, 2015). The predominant finding, and the most consistent interpretation independent of EF tasks administered, suggests the unified concept best fits EF performance in the early years and differentiates in later childhood (Tsujimoto, Kuwajima and Sawaguchi, 2007; Wiebe, Espy and Charak, 2008; Hughes *et al.*, 2010; Brydges *et al.*, 2014).

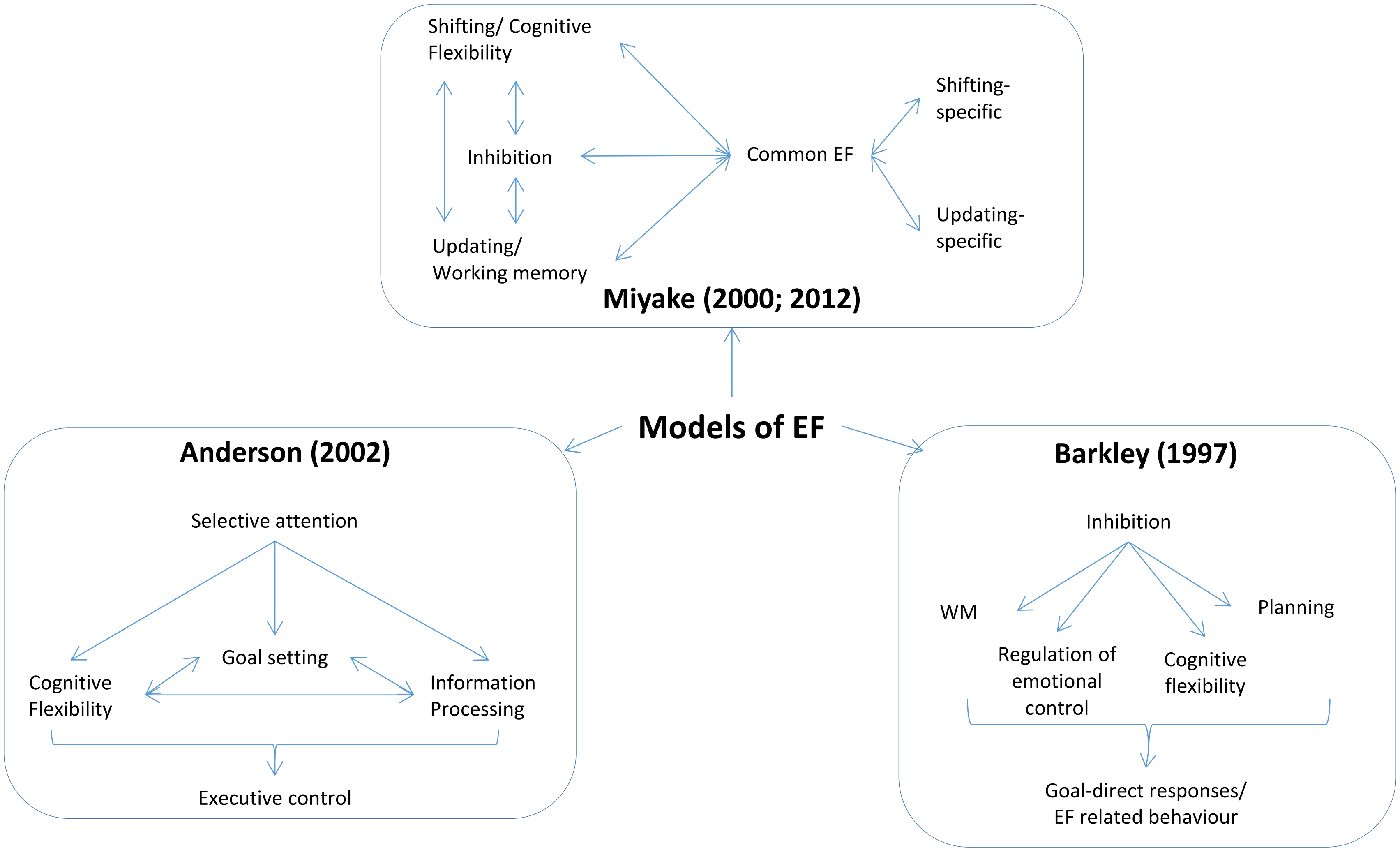
Aspects of the Miyake model resonate with the EF model put forward by Anderson in 2002. The proposition in this model is EF performance is conditional on selective attentional processes. This hierarchical view of EF suggests a gradual development of the other EF sub-skills. In this model Anderson considered the following composition as EF: cognitive flexibility, goal setting and information processing. Within this model all sub-skills develop independently and combine over time, finally presenting as overall executive control or EF (Anderson, 2002). Anderson concluded that the different executive domains come ‘on-line’ at different points in development: attentional control appears between 9 and 12 months, cognitive flexibility and goal setting behaviours at 3 to 4 years, and information processing speed only shows measurable improvements in later childhood, although it is measureable in infancy (Anderson, 2002). Not typically considered a sub-skill to EF, the inclusion of information processing in this construct was justified due to strong correlation between EF abilities and speed of response defining performances. By including the neural integrity of the circuits involved, speed of performance is accounted for within the model (Anderson, 2002). Both models proposed by Anderson and Miyake consider the sub-skills of EF to be inter-related. Although Anderson proposes a hierarchical emergence of these abilities, the continuity of an underlying relationship between the EF sub-domains is clear in the two theories (Miyake *et al.*, 2000; Anderson, 2002).

An alternative widely accepted model of EF was offered by Barkley (Barkley, 1997, 2001). Barkley was the first to propose a hierarchical, evolutionary based model for the development of EF throughout the early years, with a focus on the development of inhibitory abilities in the first instance (Barkley, 1997). Barkley states that inhibitory abilities are an essential part of EFs as they predominantly determine self-directed behaviours; without inhibition, self-directed decisions and goals are not possible as prepotent responses interfere with the control. It is proposed that the development starts with simple motor inhibition, proceeding to aspects of working memory, internal thoughts, flexible thinking and planning. Subsequently developmental advancements during childhood creates goal-directed behaviour, impulse control becomes more refined and complex attentional skills develop (Barkley, 1997, 2001).

Opinions continue to differ on the true likely representation of EF structure within the developmental pathways. As better understandings of how the EF sub-domains emerge, improvements in the detection of early signs of dysfunction would be anticipated (Isquith *et al.*, 2005). Garon *et al.* (2008) reviewed the unitary and differentiated models in regards to the developmental literature, concluding that before 3 years of age, basic EF abilities are emerging, but once beyond this age, crucial integration and coordination of the cognitive processes occurs, advancing EF performance. If considering this in the context of the preterm literature, whilst mindful of findings suggestive of domain specific difficulties within this population later in life, it remains to be seen whether it is possible to detect early EF difficulties emerging before 3 years, irrespective of the structure of EF abilities in these early years.

The review in the following sections will provide further insight into what is known about emergence of EF in both typically developing and preterm populations, before revisiting and reviewing the three models in relation to the preterm literature.

Figure 1-5. Schematic representation of three EF models commonly referred to within developmental literature.



1.2.1 Cognitive flexibility

Cognitive flexibility, or the shifting of mental states (Miyake *et al.*, 2000) is the ability to focus, shift and combine information from various sources, moderate and adapt behaviour following mistakes, and formulate plans for action (Anderson, 2002; Diamond, 2012). The ability to plan is vital for achieving specific goals and first appears around 7-8 months, becoming more advanced with age. The development of this skill is seen to advance over childhood and improves greatly between 18 to 27 months. At this age, children can recognise a goal and construct a sequence of actions in order to achieve it (Sun and Buys, 2012a). Plans that need conceptual reasoning appear too difficult before the age of 4 and only advance considerably between the ages of 7 to 11 years (Anderson, 2002).

The use of the other EF domains is often considered necessary in order to achieve shifting in mental state. Inhibition is required to ignore a previous perspective, and working memory is required in order to actively process the new view (Diamond, 2013). The dependence of flexible thinking on the other two EF sub-skills creates problems when attempting to identify pure cognitive flexibility, and often, if difficulties are observed in one domain, another is likely impaired (Nosarti *et al.*, 2007; Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, 2009).

Immaturity of this domain is thought to be reflected in perseveration responses to cognitive flexibility tasks. Perseveration in this context, refers to the failure to modulate a response following the presentation of new information, instead the previously acquired response is repeated. Tasks that typically assess this domain challenge a child's ability to use information provided by creating an established response set before introducing a rule change that challenges that response. The more complex the rule change, the harder it is for children to incorporate the information to modify their behaviour. This skill continues to develop and advance throughout middle childhood and into adolescence (Anderson, 2002).

Cognitive flexibility in later childhood

In general, performance deficits on cognitive flexibility tasks appears to be a consistent finding in preterm populations from childhood and into adulthood (Nosarti *et al.*, 2007; Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, 2009; Rose, Feldman and Jankowski, 2011; Eryigit Madzwamuse *et al.*, 2015). However, cognitive flexibility deficits are often reported alongside other EF domain difficulties. The meta-analysis of neurobehavioral outcomes in children born preterm by Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, (2009) found a decrement in cognitive flexibility performance across 12 studies with a moderate effect size. This, however, was not a specific deficit, as working memory and verbal fluency performances were also weaker in the preterm group compared to the controls. Nosarti *et al*., (2007) investigated EF performance in ex-preterm adults, with the predominant deficits observed in cognitive flexibility, inhibition and visual motor speed. When looking between these age groups, 11 year old ex-preterm children have been reported to present with deficits in all 3 EF domains (Rose, Feldman and Jankowski, 2011).

Although not directly comparable, discrepancies across these studies could suggest developmental changes in EF within ex-preterm cohorts, as domain specific differences impact performance differently with age. Detailed longitudinal exploration into cognitive performance would be required to explore this further as each study above utilised different paradigms and had differing inclusion criteria. What is transparent from previous research, however, is that ex-preterm populations demonstrate cognitive flexibility difficulties from childhood extending into adult life.

Emergence of cognitive flexibility during infancy and early childhood

When difficulties in cognitive flexibility emerge is not clear. The focus of many research studies is in the determination of which domain displays the greatest area of difficulty in ex-preterm children. Although poorer performances are detected in cognitive flexibility in older childhood, such difficulties have not been a consistent finding in toddlers aged 2 to 3, although research in this age range is minimal. Espy *et al*., (2002), did not find differences in term and preterm performances on a spatial reversal task, thought to assess shifting abilities. In contrast, a multi-facet study by Pozzetti *et al.*, (2014) identified cognitive flexibility as the only EF domain in a multi-factor analysis to show significant differences between groups. Both studies utilised reversal paradigms whereby the child was instructed to find a reward repeatedly in a specific location and once a criterion of correct responses was reached, the location was switched. Rule reversal paradigms are often considered an assessment of cognitive flexibility. However, in order to achieve a correct response, inhibition of the previous rule is fundamental, presenting a problem in interpretation. The multi-factorial approach by Pozzetti *et al.*, (2014) attempted to account for the interrelated nature of such tasks, creating composite scores from different tasks for each EF domain. Tasks such a Multi-Location Multi-Step (MLMS) paradigm, although predominantly is thought to target cognitive flexibility, likely recruits all EF sub-skills. This task therefore provided multiple variables to add to the Exploratory Factor Analysis conducted by Pozzetti *et al*. (2014). Although in principle this exploration could provide insight into the variability of EF domain performances in a preterm population, there are many limitations. Fundamentally, although the authors provided rationale in regards to the primary outcome variables that contributed to each EF factor composite score, as this was a new approach, there was no evidence to support the choice of variables selected (Pozzetti *et al.*, 2014).

These findings stress the difficulties in separating the performances of the different EF domains. Although attempts have been made to distinguish between EF sub-skills, even in the adult literature there it is a challenge in particular with cognitive flexibility due to the nature of the sub-skill. The adult literature is fairly consistent, as EF difficulties appear to be a typical finding within preterm populations, irrespective of whether flexibility is a specific issue. In early childhood, although some findings suggest cognitive flexibility may be the preliminary domain displaying difficulty in ex-preterm toddlers, this is far from consistent and research is sparse.

1.2.2 Working memory

Working memory is the ability to retain information in mind whilst executing a given task, or manipulating information in order to achieve a desired goal. Working memory helps us to plan and enables us to hold alternative views in mind, connecting it to cognitive flexibility. This ability also facilitates giving directions and linking current, future and past events (Diamond, 2006). Working memory has been reliably observed to be typically detectable between 8 and 12 months, however a review of the literature by Reznick et al., 2004, suggests working memory abilities can be detected in those as young as 6 months. This ability develops and matures with age, increasing in capacity and ability to handle more complex information (Sun and Buys, 2012a).

Working memory in later childhood

Within the preterm literature, difficulties with working memory, although not the exclusive problem, have been widely reported in school age children (Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, 2009; Hutchinson *et al.*, 2013). As illustrated in the previous section, results are often mixed as to whether the deficit resides independently within the working memory domain and/or is coupled with other EF subskills (Bohm, Smedler and Forssberg, 2004). Working memory has been reported as a specific deficit in preterm cohorts in some instances. A study by Mulder *et* al., (2010) explored the impact of EF deficits on academic outcomes in middle childhood. Working memory appeared to predict attainment outcomes independent of verbal processing speed, a measure that mediated the relationship between the other EF domains and academic outcomes. Research in ex-preterm adolescent and adult cohorts suggest difficulties in working memory may improve with age (Rushe *et al.*, 2001; Saavalainen *et al.*, 2007), however this is not a consistent result (Breeman *et al.*, 2015).

These independent and coupled deficits within the EF domains have been linked with poorer mathematical abilities in typically developing cohorts (Bull and Scerif, 2001). As discussed in section 1.1.2, difficulties in mathematics are frequently reported within preterm populations and many studies have found associations between this academic area and working memory deficits (Bull, Espy and Wiebe, 2008; Mulder, Pitchford and Marlow, 2010; Simms *et al.*, 2013).

Emergence of working memory during infancy and early childhood

Few studies have explored possible early signs of working memory deficits in ex-preterm children. Traditional assessments of working memory in infancy use delayed-response type tasks, in which an object of interest is hidden and following a set period of delay, the infant is required to attempt to locate it (Diamond and Doar, 1989; Reznick *et al.*, 2004). A study by Sun *et al.* (2009), found preterm infants as young as 8 months showed poorer performance on an A-not-B tasks (A*B* task) (Piaget, 1954) compared to term born peers, after adjusting for global cognitive performance (Sun, Mohay and O'Callaghan, 2009). However, the A*B* task is highly dependent upon inhibition of prepotent responses and requires cognitive flexibility to modulate responses, thus interpretation of performance cannot be assigned to one domain (Espy *et al.*, 1999).

Preterm born children aged 2 to 3 years have been reported to display poorer performances in tasks such as the Delay Alteration Task, designed to assess spatial working memory abilities (Espy *et al.*, 2002). Behavioural differences to spatial working memory tasks have also been reported at 3 to 4 years using delayed location recall tasks (Vicari *et al.*, 2004; Baron *et al.*, 2010). In contrast, Pozzetti *et* al., (2014) did not observe specific working memory deficits when using a 'spin the pots' paradigm. The discrepancy between studies could be explained by the different tasks used to explore these abilities. The delay alteration/delay location, require toddlers to wait before the retrieval of a reward from one of 2 or 3 set locations, the time delay before retrieval gradually increasing as a measure of performance (Vicari *et al.*, 2004; Baron *et al.*, 2010). The spin the pots paradigm require toddlers to locate a number of different rewards from different locations that were spun following the start of the trial, no time delay is imposed (Pozzetti *et al.*, 2014). It could be speculated that different aspects of working memory are being targeted in the two paradigms; the delayed alteration imposing a greater demand on spatial working memory, whereas the spin the pots task demands greater memory for colour and shape of reward location. Both require working memory of visual information, however, the specific mechanisms underlying these differences require further exploration.

The proportion of studies to observe differences within this EF domain is certainly suggestive of a working memory deficit in preterm populations. Further clarification on when these differences first emerge, and if it is possible to reliably identify specific working memory abilities in the first years of life, still remains to be seen.

1.2.3 Inhibition

Inhibition, is the ability to ignore irrelevant stimuli to the task at hand and to control behaviour (or thoughts) by preventing a response and/or replacing it with another (Diamond, 2006). It is thought to first appear between 7 and 12 months of age (Sun and Buys, 2012a). The signs of significant improvement within EF tasks that require working memory and inhibition occur between the ages of 3 to 5 years (Rennie, Bull and Diamond, 2004). Inhibition continues to mature throughout childhood, when exercising discipline and controlling emotions (Diamond, 2012).

Deficits in inhibition can lead to uncontrolled impulsive behaviours and are associated with disorders such as ADHD, the disorder that formed the basis of Barkley's EF model, which considers inhibition as the fundamental domain of executive control (Barkley, 1997; Diamond, 2012; Sun and Buys, 2012a).

Inhibition in later childhood

ADHD is often a diagnosis made in children born preterm (Bohm, Smedler and Forssberg, 2004), although there is speculation whether this population has a 'purer' form of attentional disorder, specifically inattention, as hyperactive behaviour is not always reported (Johnson, 2007). Nevertheless, the impulsivity observed in those with ADHD is often apparent within ex-preterm cohorts from childhood through to adolescence (Bohm, Smedler and Forssberg, 2004). Impulse control is commonly defined within the development literature as the inhibition of actions and control of emotions (Hammond, Potenza and Mayes, 2011). Bohm *et al.,* (2004) found impulse control deficits in a cohort of preterm 5 year olds after adjusting for IQ, which were associated with later academic difficulties (Bohm, Smedler and Forssberg, 2004). This observation is reflected in a later study using a rapid visual processing task in ex-preterm 11 year olds, where the preterm participants displayed a greater rate of false alarms. These deficits do not appear to dissipate, as studies of ex-preterm adults present with marked deficits in tasks requiring response inhibition. These fundamental difficulties are speculated to stem from deficits in inhibitory control and mental flexibility (Nosarti *et al.*, 2007).

Emergence of inhibition during infancy and early childhood

Few studies have investigated inhibitory control in the preschool years (Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, 2009). Tasks such as the Bear Dragon paradigm (Kochanska, Murray and Harlan, 2000), where the child has to follow the instructions of one puppet and ignore those given by the other, or the Gift Delay Open task (Carlson, 2005), where the child has to wait to open a present until told, are both classic tasks of inhibition. In children born preterm, gestational age has been shown to be a significant predictor of performance on these tasks (Duvall *et al.*, 2015). However, paradigms that include conflicting instructions, such as the bear/dragon, have a high working memory load (Carlson, Mandell and Williams, 2004). Delayed response tasks, such as the Gift Delay Open task (Carlson, 2005), have a lower working memory load, but limited sensitivity, because they are typically scored as pass or fail according to an arbitrary predefined time limit. Delayed response tasks are commonly used to assess inhibition from the ages of 3 onwards (Carlson, Mandell and Williams, 2004); before this age, children lack the ability to comprehend task instructions. Other paradigms, such as the Snack Delay (Kochanska *et al.*, 1996), have been used on children under 3 years, but in populations where language is delayed, as in some cases of children born premature, this can confound study interpretations as language abilities are influential on task performance (Cuskelly, Einam and Jobling, 2001).

Within the first year, the traditional Piagetian A-not-B paradigm, considered later in this thesis, has previously been considered to represent inhibitory abilities. As noted however, working memory is also fundamental to this tasks success (Diamond and Goldman-Rakic, 1989). Difficulties in attempts to parcel out the different EF domains in the preschool period are in part constrained by the limited assessment structure available at these younger ages due to restricted behavioural repertoires (Wiebe, Espy and Charak, 2008). Many tasks require infants either to make repetitive responses, respond to rule changes, and/or retrieve hidden objects. These factors make targeting one ability without the incorporation of others very difficult. Although a combination of tasks that are thought to predominantly target different domains can be used in an attempt to highlight specific differences, the interpretation and validity of findings is challenging (Pozzetti *et al.*, 2014).

1.2.4 EF overview

The literature is clearly undecided regarding which domain has the greatest influence on later cognitive performance. There is an emerging consensus that a specific working memory deficit is likely to be the primary cause of the difficulties observed in this population in older children and into adulthood. When exploring the extent of problems over a range of functions in preterm cohorts, many studies report deficits within the other EF domains, of which are not fully explained by IQ scores (Botting *et al.*, 1998; Bohm, Smedler and Forssberg, 2004; Neil Marlow *et al.*, 2007). The possibility of domain specific differences in later childhood/early adulthood are in agreement with the literature exploring the structure of executive control, suggesting differentiation of EF occurs later in development (Miyake *et al.*, 2000). As the current review suggests, in infancy and later childhood, assessing one of these domains without the influence of another is challenging (Beauchamp *et al.*, 2008).

There are mixed reports regarding EF abilities in ex-preterm children aged 2 to 3 years, possibly due to attempts to address when specific EF domains emerge (Espy *et al.*, 2002; Pozzetti *et al.*, 2014). Although tasks can be argued to target a predominant EF domain, the other sub domains are typically involved in task performances. It is therefore not possible to categorically conclude where specific deficits originate with the use of these tasks. Ex-preterm toddlers EF abilities require further investigation, perhaps with a focus of exploring EF as a unitary concept, in order to clarify overall stability of the construct and not in an attempt to parcel out the different domains.

Very few studies track cognitive performance longitudinally to explore the developmental trajectory of EF (Anderson, 2014). To the authors knowledge, the only study investigating cognitive performance at multiple time points across the first two years after birth in a preterm cohort was the study conducted by Lobo and Galloway (2013), who investigated the stability of learning in this early period of life. The focus of this study was related to the trajectory of EF, but rather fixated on the infants' ability to learn within the early months, as a means of identifying learning difficulties later in life (Lobo and Galloway, 2013).

Irrespective of differences in opinion regarding the structure of EF and which abilities make up each sub-domain, it is clear that ex-preterm children, and later adults, have difficulties in these cognitive abilities. In later life, there is evidence to suggest that specific domain differences may account for these cognitive deficits, but which is primarily responsible, is still a matter of debate. When difficulties in EF are first detectable remains unclear. Typically, studies have approached this question with attempts to determine when specific domain differences emerge in preterm populations. However, the evidence for the emergence of EF domains within typical cohorts also divides opinions. In studies of early development, it is apparent that classic paradigms do not allow for clear differentiation of the EF sub-domains. With this taken into account, it is highly plausible that EF is a unitary construct that differentiates in later life. This is in accordance with the unity and diversity model presented by Miyake *et al.,*(2000) and is supported by previous reports in typically developing cohorts as the best fit for the development of EF (Wiebe, Espy and Charak, 2008). Thus in the current investigation we evaluated a series of classic EF tasks in very preterm and term infants, with the aim of identifying the early signs of differential EF performance in the context of a longitudinal study.

1.3 Attention

First described by Posner and colleagues, attention is commonly conceived as three networks; orienting or selective attention; alerting, arousal or sustained attention; and executive control or simply executive attention (Posner and Petersen, 1990; van de Weijer-Bergsma, Wijnroks and Jongmans, 2008; Mulder *et al.*, 2009). In this model, the orienting attentional network describes the spatial positioning of attention to surrounding stimuli, and is thought to be fully developed by 6 months of age (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008). The alerting or the arousal network dictates focus in order to maintain continuous information processing abilities, and upon unanticipated stimulation, creates a state of arousal (Amso and Scerif, 2015). Finally, the executive network refers to self-directed attentional behaviours and is largely connected to the wider range of EFs (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008; Diamond, 2013). This multidimensional construct brings together these three systems in order to achieve higher-order processing as well as coordinating and responding to sensory and motor stimulation (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008; Scerif, 2010).

The developmental trajectory of these networks is again complex because they are inter-related. In infancy, the first of the networks to be observed is that of arousal and orienting of visual attention. Evaluation of this network is often driven by the response to novel stimuli and a large proportion of investigations focus on the infant's attraction to faces (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008; Scerif, 2010). This focus on highly salient stimuli can lead to difficulties in disengagement of attention, termed 'sticky attention'. This 'blank stare' is not necessarily a measure of visual processing, as discussed in detail within the visual habituation literature (Stechler and Latz, 1966). Visual habituation is a technique developed as a means of trying to assess early information processing in infants. Initially when presented with a novel stimulus, an infant will maintain visual focus. It is assumed that the time spent observing the stimulus reflects internal processing. With repeated exposure to a stimulus, the looking time decreases, termed a habituation response. Current theories suggest the faster the speed of habituation the more efficient the visual processing of the infant. Look durations increase with age, and are thought to reflect improved disengagement of attention and improved efficiency of information processing (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008; Kavsek and Bornstein, 2010).

In infancy, attentional networks have been found to be highly correlated with processing speeds and early EF abilities (Garon, Bryson and Smith, 2008). Although the precise mechanisms are not yet clearly defined, paradigms such as the 'Visual Search task' (Scerif *et al.*, 2004) and the Gap Overlap task (Atkinson *et al.*, 1992; Hood and Atkinson, 1993) are thought to tap into this relationship. The Gap overlap paradigm requires disengagement of visual attention from a central stimulus and shifting gaze to the periphery (Atkinson *et al.*, 1992; Hood and Atkinson, 1993). It is proposed that those who disengage at a faster rate are then able to reengage with an alternative image and process scenes quicker than those with difficulties (Rose, Feldman and Jankowski, 2002). This is a paradigm that will be used in the current study as it may be used in infancy through to the preschool years, thereby providing longitudinal observation of the development of disengagement measures within a preterm population.

Alerting, arousal or sustained attention is observed in infancy and early childhood. Although a 3 network model is often reported within the attentional literature, the overlap between the development trajectories of sustained and selective attention has led some to postulate whether a two arm model is more appropriate; selective and sustained attention as one network, executive attention as the other (Steele *et al.*, 2012). In any regard, the capacity of the sustained attentional network allows for active information processing due to a sustained state of arousal. This is often observed during play sessions, when infants or toddlers display a prolonged interest in a specific object (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008). In laboratory settings, the Continuous Performance Task has been used to assess sustained attention, whereby the child presented with a stream of continuous stimuli is asked to respond to a specific infrequent target (Akshoomoff, 2002).

Over the toddler years, playing behaviours develop becoming self-directed and planned; this is considered a function of the executive attentional network. This network has strong associations to the EF domains: cognitive flexibility, working memory and inhibitory abilities (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008). Executive attention is often associated with activity in the Dorsolateral Prefrontal Cortex, a region also highly linked to EF abilities (see Figure 1-3) (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008). The 'flanker task' has previously been considered a measure of executive attention, where a child or adult, is asked to select a responses that correspond to a visual target stimulus, and ignore the distractor stimuli, presented either side (Rueda *et al.*, 2004). The interrelated nature of these attentional networks and EF means that obtaining a 'pure' measure of the any attentional networks is not easily achieved and is often assessed as part of EF tasks. For example, the Flanker task is predominantly regarded as an assessment of inhibition, although attentional abilities are certainly likely to impact performance (Steele *et al.*, 2012).

The dual network model proposed by Steele *et al*. (2012) suggests the selective and sustained attentional networks, although they may be more closely related in childhood, differentiate during development. The developmental trajectory of these networks may show a similar structure to the emergence of EF domains (Steele et al 2012; Wiebe et al., 2011).

In ex-preterm infants, paradigms that assess visual orienting within the first 6 months have observed difficulties with gaze shifting behaviour. Errors in gaze shift tasks include specific looks away from central fixation stimuli and more general looks away from task equipment. Butcher *et al*. (2002) explored the developmental trajectory of shift patterns in term and preterm born cohorts. Over development, term born infants were observed to make more errors compared to preterm infants. The authors interpreted this increase in looks away from task stimuli as the emergence of disengagement behaviour, the ability to inhibit attention to salient stimuli, and required for high level cognitive processing. This more mature looking behaviour was not observed in the ex-preterm infants (Butcher *et al.*, 2002). However, these differences in gaze shift patterns are not consistent across the literature. Rose *et al*., (2001) reported comparable gaze shifting behaviour across the first year after birth in a longitudinal cohort of term and preterm infants. Gaze patterns advanced with age at a similar rate within both cohorts, with shorter looks to targets and faster shift rates at the later time point (Rose, Feldman and Jankowski, 2001). Shift rates, although indicative of attentional processes, are also utilised in the investigation of information processing speeds, another area postulated to be affected by preterm birth (Rose, Feldman and Jankowski, 2002; Mulder, Pitchford and Marlow, 2011a) and explored in section 1.4.

Preterm children and adults are more frequently assigned diagnoses of ADHD (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008), alongside a range of psychological disorders, in particular the inattentive subtype (Johnson, 2007; Lawrence *et al.*, 2009; Jaekel, Wolke and Bartmann, 2013). Among those who do not meet diagnostic criteria there is an excess of sub-clinical symptoms (Johnson, 2007). The propensity for inattention may explain some of the learning difficulties found among preterm populations and the effect may be independent of general cognitive performance (Jaekel, Wolke and Bartmann, 2013). In one study a battery of EF tasks, including attentional measures, explained variance in cognitive and behavioural scores between a very preterm and term born populations, but the majority was explained by working memory and visual processing speeds (Mulder, Pitchford and Marlow, 2011b). Behavioural inattention observed in the classroom may be related to what we understand to be the neuropsychological attentional networks (discussed below), but to date limited evidence is available to connected the two (Steele *et al.*, 2012). On the contrary, behavioural inattention has been shown to correlate to a greater extent with working memory, processing speed and inhibitory deficits (Espy *et al.*, 2002; van de Weijer-Bergsma, Wijnroks and Jongmans, 2008; Scerif, 2010). Thus behavioural inattention may be considered a reflection of EF deficits, as opposed to specific difficulties in attentional networks.

Investigations into the functionality of attentional networks in ex-preterm adults report lower performances compared to controls; however, as discussed previously, tasks utilised in older populations are often highly confounded by other cognitive domains. Nosarti *et al.* (2007) utilised the Test of Attention Performance to investigate attentional abilities in ex-preterm adults, but by evaluating performance based on response times, there will be confounds by slower processing speeds typically observed within this population (Nosarti *et al.*, 2007).

1.4 Information processing

'Information processing speed' is the term given to the speed of communication between different brain regions in order to complete specific cognitive goals. The speed in which information is transferred is fundamental to the success of basic cognitive tasks and is considered an essential cognitive resource (Turken *et al.*, 2008). It has been long since established that speed of information processing (IP) is positively correlated with IQ scores in typical adult population studies, and it is speculated that a link between speed of processing and working memory drives this correlation (Jensen, 1993). The same relationship has been seen in preterm cohorts. Possible associations have been observed between working memory deficits and reduced processing speeds in preterm children (Mulder, Pitchford and Marlow, 2010).

Reduced speeds in information processing is commonly reported in preterm cohorts both in terms of sensory information (Rose, 1983; Ramon-Casas *et al.*, 2013) and higher level cognitive information reflected in reduced cognitive performance (Rose, Feldman and Jankowski, 2009). Studies of preterm cohorts consistently report strong associations between processing speed and individual variability in EF performance and overall academic achievements (Mulder, Pitchford and Marlow, 2010, 2011b; Rose, Feldman and Jankowski, 2012). Up to 60% of the differences in global cognitive performance scores between term and preterm participants may be accounted for by variations in processing speed alone (Rose and Feldman, 1996). In addition to the association with IQ, processing speed is linked to the performance variations in multiple independent domains across many studies, from working memory (Rose, Feldman and Jankowski, 2002), to language deficits (Ortiz-Mantilla *et al.*, 2008) and attention (Mulder, Pitchford and Marlow, 2011a).

The traditional view of early cognitive investigations is that they showed poor predictive validity for later cognitive performance (Fagan and Singer, 1983). Tests in infancy are speculated not to tap into the same processes that are apparent in childhood and beyond (Rose and Feldman, 1990; Colombo, 1993). In this vein, improvements in this area have been the focus of infancy research, with processing speed measures indicating a level of continuity and stability in the prediction of later cognitive abilities (Rose, Feldman and Jankowski, 2009).

Preterm infants have been observed to perform poorly in cognitive tasks that additionally assess speed of information processing. In 2002, Rose, Feldman and Janowski*,* utilised habituation, gaze shifting and recognition paradigms with poorer performances observed in the preterm cohorts. In contrast, and in an earlier study Rose *et al., (*2001) found no group differences between term and preterm infants in an attentional gaze shift paradigm in the first year after birth. The authors speculate that discontinuity between their findings reflect different processing skills in the paradigms used. In the second study in 2002, a familiarisation paradigm may have created a greater cognitive load than the first in 2001, a simple visual expectation paradigm; thereby the greater complexity provided a more detailed investigation of the preterm infants cognitive abilities (Rose, Feldman and Jankowski, 2001, 2002).

Deficits in processing speed within the preterm population are observed in middle childhood and into adulthood. The link between processing speed and working memory seen in typically developing populations, is highly associated with overall academic attainment in preterm populations (Rose and Feldman, 1996; Fry and Hale, 2000; Mulder, Pitchford and Marlow, 2010). In the study by Mulder *et al.,* (2010), verbal processing speeds accounted for the variations in attentional abilities, inhibitory performance, semantic fluency and shifting abilities. Working memory was independently predictive of academic attainment. This finding echoed a previous study by Rose *et al.,* (1996). In contrast, Bull and Johnston (Bull and

Johnston, 1997), reported that the typical arithmetic difficulties observed in preterm children were best predicted by motor processing speed, and independent of memory. Although there are still variations amongst study results, there remains a strong suggestion that processing speed is a good predictor of later academic performance.

In terms of continuity, there is a lack of investigations within the preterm literature exploring processing speeds from infancy into the toddler years. Rose *et al.*, (2009) reported toddlers born preterm to display a persistent deficit in a range of cognitive abilities from the first year and into the second. Information processing speeds in combination with recognition memory, recall and attention accounted for the variation in general cognitive ability with this cohort (Rose, Feldman and Jankowski, 2009). To date, there is limited additional evidence investigating such trajectories.

Information processing speed may also be investigated using neural processing techniques. Event Related Potentials (ERPs) provide more accurate temporal measures of specific processing and may be used to assess preterm-term differences. Processing difficulties related to slow cognitive performance and social inabilities are correlated to the 'lower-order' processing speeds of sensory information, particularly that of visual and auditory stimuli (Fellman *et al.*, 2004; Mikkola *et al.*, 2007; Sokhadze *et al.*, 2017). If preterm children do not process sensory information at the same rate as typically developing individuals, this will impact the performance of cognitive tasks. In particular, infancy research is largely dependent on looking times and the speed of gaze-shifting to evaluate early cognitive abilities (Butcher, Kalverboer and Geuze, 2000). Exploring neural correlates of sensory systems may inform on the evaluation of mechanisms underpinning cognitive performance.

ERPs are averaged voltage deflections produced by the brain in response to any sensory modality (Woodman, 2010). In this brief review, ERPs to visual and auditory stimuli are summarised. Upon detection of a stimulus, an initial change in polarity is observed in the ERP waveform; in the case of visual stimuli, the C1, reflects the location of the stimulus within the visual field (Clark, Fan and Hillyard, 1994). The N1 and/or P1 are typically the first components observed in an ERP waveform in response to auditory stimuli. These are postulated to reflect the physical attributes of the stimulus oppose to cognitive evaluations and indicate the detection of the stimulus in the primary visual or auditory cortices (Herrmann and Knight, 2001). Following this, the P300 or P3 is produced. The P3 is a positive deflection, peaking at approximately 300 milliseconds post stimulus onset. This component reflects the attention to the stimulus and is typically a larger response when an infrequent stimulus is detected (Herrmann and Knight, 2001). This technique is therefore tailored to explore differences in information processing speeds in early attentional networks to any sensory modalities between different cohorts.

Within this thesis, the focus is on neural correlates associated to auditory processing. Northam *et al.* (2012) used MRI and diffusion tractography techniques to investigate the integrity of the interhemispheric pathways associate with language in a group of extremely preterm infants. They found a significant reduction in volume in the posterior Corpus Callosum that accounted for 57% of the variance in language abilities within the preterm cohort (Northam *et al.*, 2012). This in part echoed the findings of Nosarti *et* al. (2004) who had previously reported correlations between the posterior corpus callosum volume and verbal fluency and IQ scores, but only in preterm males. These findings suggest that transfer of auditory information related to speech and language across the hemispheres is compromised, and potentially slower, in children born preterm. To the authors' knowledge, there are no current studies exploring the attentional response of the brain to speech sounds in a cohort of preterm toddlers. A more detailed review of auditory processing in relation to language can be found in Appendix 4, but will not be considered as the primary focus in this thesis.

1.5 Conclusion

In contrast to the large body of research into cognitive processes in school age ex-preterm children, there is a paucity of research that seeks to evaluate the early trajectory of emerging cognitive processes that will underpin later performance and allow for the early identification of children at clinical risk of later learning difficulties. Such evidence is needed to formulate targeted interventions to ameliorate the high prevalence of special needs and cognitive deficits in very preterm children at school age. There is genuine uncertainty about the early emergence of executive functions and information processing speed differences in infancy.

In this thesis, studies of the early emergence of EF skills up to 30 months of age will be presented whilst exploring the influence of processing speed and attentional differences in a group of very preterm infants (VP henceforth) with relatively uncomplicated neonatal courses, compared to a group of term children (term henceforth).

1.6 Research Aims and Objectives

The overall aim was to use targeted neuropsychological assessments in combination with ERP techniques and eye-tracking technology to identify at which point within the first two years after birth differences begin to emerge in EF abilities between VP and term controls, and to what extent information processing (IP henceforth) abilities impact global cognitive performance. The specific study objectives were:

1) To explore the differences in EF, attention and IP speeds at 3 time points within the first year and at 30 month of age between the term and VP cohorts. Each EF task incorporates the different EF subskills and performances will be adjusted for global cognitive score (defined by the cognitive composite score of the Bayley-III). This will seek to identify the emergence of EF, attention and IP difficulties not accounted for by global cognitive performance.
2) The first year of assessments will be used as predictors of cognitive score of the Bayley-III at both:
    i) 12 months
    ii) 2 years
3) Finally, the EF tasks at 30 months will be used to investigate the variation of the Bayley-III at 2 years.

The final two objectives will explore to what extent the performances in EF tasks in the first and second year predict the variation on the Bayley-III cognitive scores and will examine the effectiveness of the cognitive scale at detecting any variation in EF abilities in a VP and term cohort.

1.7 Hypotheses

This thesis is broken down into 5 chapters: Global measures of cognitive, language and motor development; Executive Functions; Attention; Information Processing speeds; and Prediction of global cognitive performance at 12 months and 2 years from earlier measures of EF, IP and Attention. Accordingly, the following hypotheses where made for each aspect of the study:

1.7.1 Global measures of cognitive, language and motor development

- VP infants will score lower in the Bayley-III cognitive composite scores compared to term infants, at both 12 months and 2 years of age.
- The cognitive composite score at 12 months will predict the cognitive composite score at 2 years of age.
- The cognitive composite score at 30 months will account for some of the variation in other neuropsychological measures throughout the other chapters but will not explain all differences seen between the two cohorts.

1.7.2 Executive Functions

- VP infants will show a reduced ability in EF tasks compared to term controls.
- Any differences observed within the first year will also be observed in the second year performance in corresponding tasks.
- Any differences seen in the two cohorts will not be completely accounted for by the global cognitive score.

1.7.3 Attention

- The information processing speeds obtained from the attentional task will show slower processing and more immature looking behaviour in VP children compared to term controls at each time point that it is measured (6, 12 and 30m).
- Overall global cognitive differences at 12 and 30 months account for a proportion of the variation in performance within the two groups.

1.7.4 Information processing speed

- VP children will display slower response times across all IP measures.
- Any differences observed will still be present after adjusting for global cognitive performance.

1.7.5 Prediction of global cognitive performance at 12 months and 2 years from earlier measures of EF, IP and Attention

- Poor correlations will be observed between the variation in the Bayley-III cognitive scores at 12 months and 2 years and the EF, IP and attention performances from the first year
- Likewise, the proportion of variation accounted for by EF, IP and attentional performances in the Bayley-III cognitive score at 2 years will be low.

Chapter 2 Methods

2.1 Preterm Development Project

The Preterm Development Project: Growing up after preterm birth (PDP) was established in 2011, by Professor Neil Marlow and Dr Michelle de Haan, with the assistance of Dr Lara Platten, and Dr Charlotte Sanderson-Brindle. This prospective cohort study was primarily put together to further our understanding of the early brain and social-cognitive development of children born very preterm. The long term aim of the study is to use the information collected to help improve early identification methods of those at risk for later social, cognitive and academic difficulties, and to develop targeted interventions in order to reduce levels of developmental delay seen within this population, thereby reducing the social cost these delays have on the education system.

The main objectives of the PDP are outlined in the UCH PDP protocol, please see Appendix 1. The structure of the study is presented in Figure 2-1.

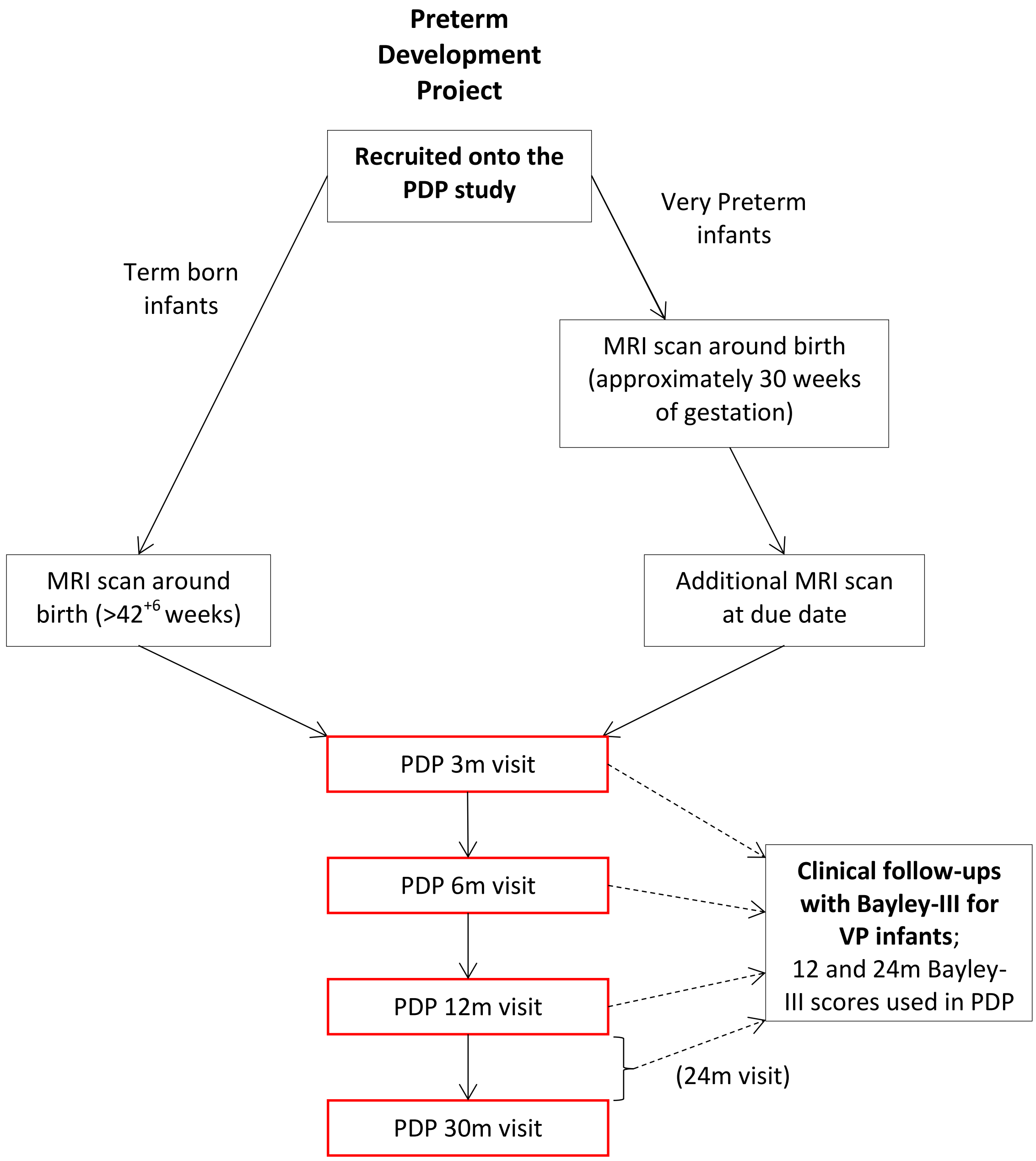


Figure 2-1. The Preterm Development Project (PDP) study structure. The red boxes indicate the stages of the study where the data reviewed in the thesis were collected.

On discharge from the hospital, the term and VP infants received the normal clinical outpatient follow ups as standard, with additional visits to University College Hospital to visit the UCH PDP Baby lab within the Clinical Research Facility in the Elizabeth Garrett Anderson wing. The infants were asked to attend 3 visits within their first year at 3, 6 and 12 months of age. Following this first year of assessment, consent was sought to allow subsequent assessments at a rate no greater than 1 per annum. The outcome for this PhD project is based on the data obtain at the follow-up assessment at 30 months of age, however utilises the data collected at the previous time point to address the study aims. The VP infants were corrected for gestation by using their Expected Date of Delivery (EDD) for the purpose of these assessments, the reasons for which were discussed in Chapter 1.

This chapter will detail the recruitment process of the infants into the study, the inclusion and exclusion criteria, the ethical approval for the study, a brief description of the tasks administered and reviewed within this thesis, and finishing with the statistical plan for all analyses conducted in later results chapters. The detailed methodologies for each task administered will be in the relevant chapters.

2.1.1 Participant recruitment

The PDP study, originally funded by SPARKS, aimed to recruit fifty preterm infants born at <32 weeks of gestation and fifty term born controls as a comparison group. Recruitment for the study was initiated before the start of this PhD project. Support from this studentship, allowed for additional recruitment of term born controls and then aided in the completion of recruitment of the VP infants. Eighty one term born and fifty VP born children were included in the data analysed within this thesis; of these, forty-four term born participants and thirty-nine VP infants were recruit during the PhD period. The author completed a minimum of one follow up assessment for sixty term born and forty seven VP born participants during the 3 years of data collection. The attrition rate for the study is show in Figure 2-2.

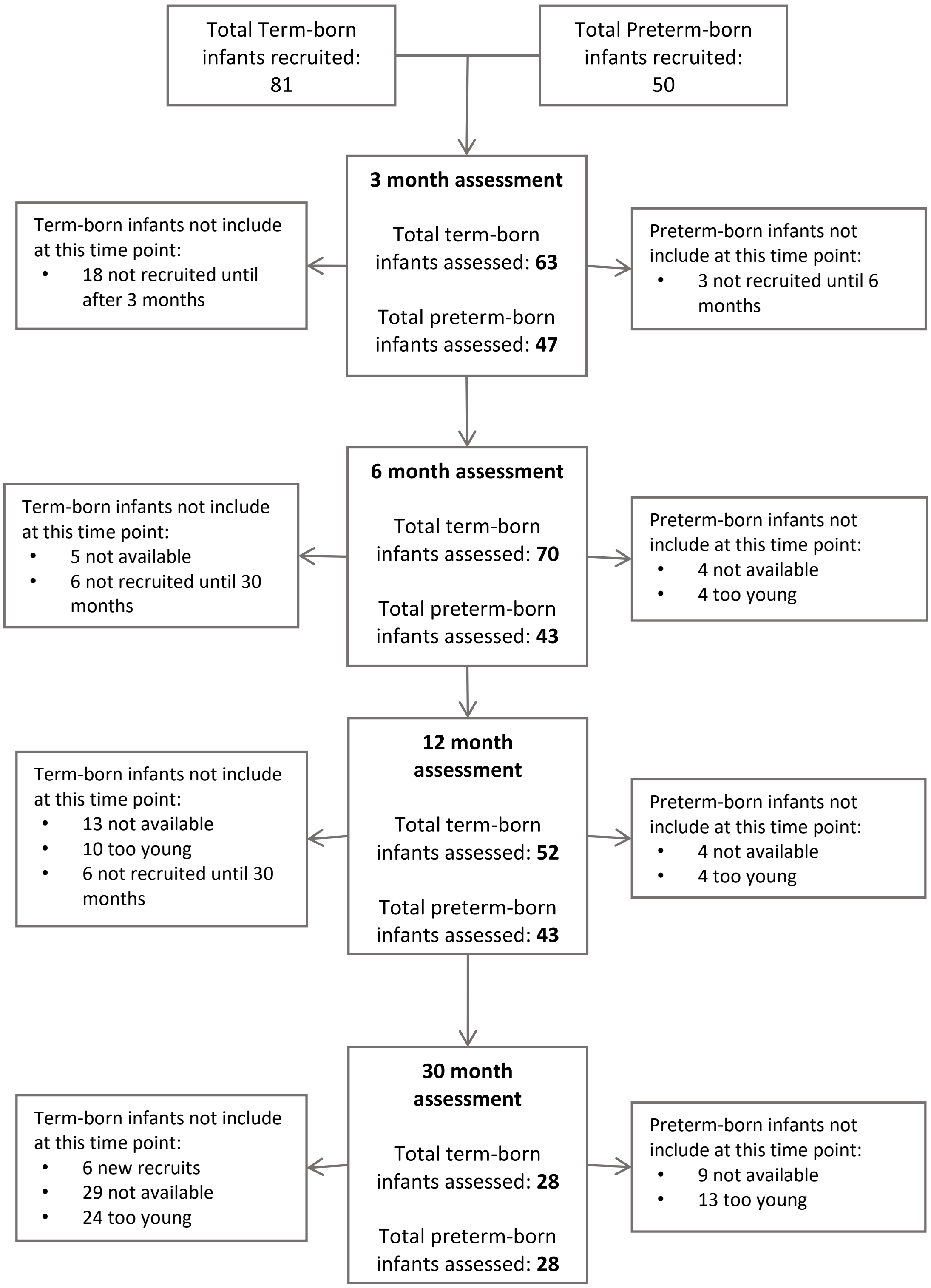


Figure 2-2. The attrition rate of the participants included in this thesis from the PDP study.

2.1.1.1 Preterm Inclusion criteria

Infants born at <32 completed weeks of gestation were recruited from the Neonatal Unit at University College Hospital London. In the first instance, VP parents were approached at the end of the first week after birth and permission was sought to include them in the study. Exclusion criteria were low likelihood of survival and severe congenital abnormality. An information leaflet was given to parents at the time of consent, so to provide information on the study follow-ups, separate from routine medical follow-ups. No alteration to clinical care was necessary as part of the project other than the two imaging procedures around birth. Copies of consent forms and study documents were included in the infants' medical files, and copies were given to the parents for future reference. Parent contact details were passed on to the study team following consent from the participating parents in order to organise the follow-up appointments in the UCH Babylab.

2.1.1.2 Term inclusion criteria

Term born children included in the control group were recruited from antenatal classes and postnatal wards at University College Hospital London. Inclusion criteria for the term group were: gestation between 37-42 weeks, birthweight between 10^{th} and 90^{th} percentile for gestation, no perinatal complications and Apgar score at 5min >7. A leaflet detailing the study and providing contact details was distributed via local parent and infant groups and their venues. Participants were also recruited via email notices.

2.1.1.3 Attrition rate

Due to the nature of the study, a number of children originally recruited did not complete all assessment phases. As highlighted in Figure 2-2, detailing the attrition rate of infants through the study, a proportion of infants did not fully withdraw from the study within the first year of assessment, but were unable to attend specific assessments either due to illness or other family circumstances. The timeline of the study and the start of this PhD project were such that a number of the term born infants first recruited onto the study were missed at the 30 month time point. Due to this, a small cross-section of term children were recruited at this age to ensure a balanced sample of term and VP children at the 30 month time point.

2.1.2 Ethical Approval

The study was approved by the NW London Research Ethics Committee 2 (Reference 10/H0720/80) and is registered with the Research and Development Department of UCLH.

The study incurred a number of amendments to include new researchers, additional tasks and new assessment batteries; all changing the documentation and requiring review. For the purpose of this PhD project, Appendix 1 and Appendix 2 detail the amended protocol and approval letter for assessment administered up to and including the 30 month follow up.

	Term		*Preterm*					
	Total term n = 81		Total Preterm (n = 50);		Very Preterm (n = 17)*;* (27-31+6weeks)		Extremely Preterm (n = 33)*;* (<27weeks)	
	n=81		n=50		*n=17*		*n=31*	
Infant sex (M:F)	40:41		31:19		12:5		19:14	
Maternal education (<GCSE:>GCES)	96% (3:77)		87% (6:41)		86.7% (2:13)		87.5% (4:28)	
	Median	***IQR***	***Median***	***IQR***	***Median***	***IQR***	***Median***	***IQR***
Gestation (Weeks + days)	40^{+2}	$39^{+2} – 41^{+2}$	26^{+0}	$25^{+0} – 28^{+0}$	28^{+4}	$28^{+0} – 29^{+3}$	25^{+2}	$24^{+6} – 26^{+0}$
Birth Weight (g)	3380	3175 – 3850	767	670 – 922	956	736 – 1130	730	657 – 785
SDS	-.05	-.49 – .47	-.41	-.89 – .05	-1.37	-2.12 – -.50	-1.85	-.52 – .13
IMD quintile (SES)	4	2 – 4	3	2 – 4	3	2 – 4	3	2 – 4

Table 2-1. Total population demographics for infants included within this thesis. Male sex and maternal education reported separately as ratio data; remaining characteristics reported as Median and IQR. Maternal education categorised by those with qualifications greater than GSCE and those below; the IMD quintiles is the Index of Multiple Deprivation for UK postcodes categorised into 5 groups with 1 = least deprived and 5 = most deprived (NPEU, 2013).

Ethnicity	*Term*		*Preterm*	
	Mother (n=73)	Father (n=73)	Mother (n=45)	Father (n=44)
White British	29	42	14	16
White Irish	6	1	1	0
White and Asian	1	1	0	0
White and Black Caribbean	1	0	0	0
White and Black African	1	0	0	1
Any other white background	26	21	6	7
Chinese	3	2	0	0
Indian	3	2	8	7
Black African	1	1	4	4
Black Caribbean	0	0	1	1
Bangladeshi	0	0	1	1
Pakistani	0	0	3	4
Arab	0	0	1	1
Any other Asian background	0	0	3	1
Any Other Mixed/Multiple Ethnic Background	0	3	1	0
Any other ethnic group	1	0	2	1

Table 2-2. Maternal and paternal ethnicity of all infants included within this thesis.

2.1.3 Primary language of cohort

At the time of recruitment, the study was explained in English and consent was given and taken in English. Families whose mother tongue was not English were not excluded from the study, but needed a level of understanding to consent to the study in English. The researcher or physician taking consent utilised their professional judgment to assess whether the level of understand was present. If this level of understanding was not there, consent was not taken and the child was not entered onto the study.

Throughout the study communication with the infant or toddler was in English. This was particularly important during the 30 month follow-up given the nature of the assessments. It was therefore at this point that the predominant language spoken within the home was recorded. The proportion of English spoken within the home is detailed below in table 2-3.

	Only English		Bilingual or greater		
			Predominant language		***Percentage of time English spoken in home***
	n	%	English (n; %)	Other (n; %)	Median (range)
Total (n=49)	28	57.14	39 (92.86)	3 (7.14)	100 (50-100)
Term (n=26)	14	53.85	21 (87.5)	3 (12.5)	100 (50-100)
Preterm (n=23)	14	60.87	18 (100.0)	0	100 (90-100)

Table 2-3. Proportion of English spoken within the cohort at 30 month follow-up, detailing the predominant language spoken within the home and the percentage of time spoken in English.

If the mother tongue of the toddler was not English, and the toddler displayed signs of misinterpretation of task instruction, the parent was instructed to give an exact translation to the mother tongue during specific tasks.

2.1.4 Medical factors

As discussed in section 1.1, numerous neonatal complications can impact the brain development and later function of infants born preterm. Below is a summary of key neonatal characteristics of the VP infants within the current cohort.

	Total cases; n=50		*Very Preterm; n=17 (27-31+6weeks)*		*Extremely Preterm; n=33 (<27weeks)*	
	n	%	n	%	n	%
Cases of IVH/PVL: I-II	25	50	6	35.28	19	57.58
IVH+ Intraventricular dilation	17	34	4	23.52	13	39.39
Intraparenchymal lesion/PVL	8	16	2	11.76	6	18.18
ROP:	29	58	3	17.65	26	78.79
Stage 1	8	16	0	0	8	24.24
Stage 2	17	34	2	11.76	15	45.45
Stage 3	4	8	1	5.88	3	9.09
CLD/BPD:	42	87.5	10	62.5	32	75
Mild	11	22.92	3	18.75	8	25
Moderate/ Severe	31	64.58	7	43.75	24	75
NEC	7	14	2	11.76	5	15.15
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Days in ITU	33.19	22.19	18.88	20.24	40.34	19.75
Total days in hospital	107.26	37.52	89.87	49.71	115.41	27.52

Table 2-4. Neonatal characteristic of the infants born very preterm included within this study, total, and subdivided into very preterm 27-31+6weeks; and Extremely preterm <27weeks gestation.

2.2 Assessment Methodology and summary of study paradigms

The experimental paradigms included in the PDP study aimed to assess the development of EF, attention, and IP speed differences in a cohort of term and VP infants and will be categorised accordingly; the results of each category will be reported in separate chapters.

As detailed in the literature, targeting one cognitive domain within EF without incurring the use of others is very difficult (Mulder *et al.*, 2009); therefore the predominant cognitive domain will be discussed for each task, however overall, the paradigms have been considered measures of more general EF functioning. In each chapter, the paradigms will be reported in order of age of assessment. Below, table 2-5 briefly introduces the tasks utilised within this thesis and the chapters where they are reviewed. Table 2-6 details the additional questionnaire measures utilised in order to acquire relevant information use in the thesis. Within each chapter, a background of their standing in current literature will be reviewed, including how they fair in terms of known preterm research. Full methodologies including apparatus and procedures will be explained within the methods section of the relevant chapters.

Chapter	Assessment and age performed	Description
Chapter 3*:* *Global measures*	***Bayley Scales of Infant and Toddlers Development 3rd edition*** *– cognitive, motor and language scales (Bayley-III);* *12 and 24/30 months*	*A well-established global assessment scale with normative data available for population comparisons. Gold standard clinical assessment used to determine development milestones according to age in 3 main areas: cognition, language and motor skills.*
Chapter 4: *Executive Functions*	***Delayed Response Task (DRT)****;* *6 months*	*The infant was sat on a parent or guardians lap in front of a large black screen with two windows. The windows were occluded by a blind. Upon raising the blind, a stimulus was presented at one of the two windows, making a noise to capture the infants' attention. The blind was then lowered and a 5 second delay administered. The blind was then raised and the direction of the first eye movement from the infant*

		was recorded. The procedure was repeated with the stimulus presented randomly to each window. The DRT administered in this study included both social and non-social stimuli, which were summarised to provide performance measures.
	A not B paradigm*;* *12 months*	*The infant observed the hiding of an object in one of two locations and was then asked to retrieve it without any delay imposed. After two correct retrievals in the same location, the hiding location of the object was changed to the other location in plain sight, and the child was again asked to retrieve the object. If the child searched correctly, the procedure was repeated, accept with a delay of 5 seconds implemented before being allowed to search for the object in all instances. This procedure was continued until the child incorrectly searched for the object following a switch. This error has been termed as an 'AB error'* (Diamond, 1985).
	Dimensional Change Card sort task; *30 months*	*The DCCS task comprised of two sorting boxes and a selection of sorting cards. The cards varied on two dimensions, with the sorting box displaying the same, yet inverse dimensions. The task required the child to sort according to each of the dimensions in turn. If both dimensions were correctly sorted, the cards were changed and the next level was administered; the complexity of the different dimensions increased with each level.*
Chapter 5: *Attention*	***GAP Task****;* *6, 12 and 30 months*	*Used eye-tracking technology to assess visual reaction speed and attentional processes. The task challenged the child's ability to disengage from a central stimulus when presented with a peripheral target. The task*

		was designed to highlight any deficits in attention by looking at the time taken to disengage (e.g. an inability to disengage may be a sign of 'sticky attention', typically seen in children with Autism Spectrum Disorder).
Chapter 6: *Information Processing-*	***Conjugate Mobile Reinforcement paradigm****;*	*A behavioural paradigm designed by Rovee & Rovee (1969) targeting the operant learning response. A reinforcement paradigm, where the infant learnt specific movements displaced a mobile suspended*
Behavioural measures	*3 months*	*above them, rewarding them with sounds and visual movement. The infants were considered to have learnt during the paradigm if a pre-determined criterion is reached.*
	Babyscreen App;	*A touchscreen based assessment designed by Twomey et al* (Twomey *et al.*, In Press). *A newly developed*
	30 months	*application based on classic EF tasks but in a touchscreen environment enabling the investigation of speed of processing in relation to EF abilities. Tasks on which the application was designed include the A not B, and Dimensional card sort task (DCCS).*
	Multi-Location Multi-step task (MLMS)*;*	*An extension to the A not B paradigm. An object was hidden in one location for one or more trials but retrieval of the object required the completion of a multiple step process. After correctly locating the*
	30 months	*object in 3 consecutive trials, the hiding location was switched to an alternative location. Perseverative errors and time to completion were summarised to provide performance measures.*
Neural measures	***Auditory ERP paradigm***;	*Designed to assess speed of auditory information processing with particular interest in the speed of information transfer across the corpus callosum.*

	30 months	*Infants watched an unspecific visual presentation on a screen whilst listening to simple auditory syllables presented to left and right ear independently (monaurally). The auditory N1 and P3 amplitudes and latencies will be compared between cohorts.*

Table 2-5. Name and brief description of assessments evaluated within this thesis

Assessment age	**Questionnaire**	**Description**
All time points	*Demographics (see Appendix 3)*	*General information about the family; address, siblings, medical history, educational background, employment history, language dominance.*
30 months	*The Oxford Communicative Development Inventory (OCDI)*	*A UK adaptation of the MacArthur-Bates CDI. A parental report of the receptive and expressive language produced by the child at the time of the assessment. The OCDI is used with the analyses in the subsequent chapters if the language scale of the Bayley-III is not completed.*

Table 2-6. Summary of questionnaires considered within this thesis.

2.3 Statistical analyses

Excluding the Bayley-III and the BabyScreen tasks, which were coded online, all other tasks were video recorded with data acquired off line. 20% of the data from each task were double coded for reliability, which was set at a minimum of 80% to be considered accurate. Due to the variation in participant attendance during the study, and the nature of neuropsychological assessments in infants and young children, full datasets for each child were not possible. This thereby led to variations in participant numbers for each task. Given this, population demographics will be given at the beginning of each results section. Data were analysed following a set procedure detailed below.

Normality of each key variable within a data set was explored using histograms and the Shapiro Wilk test of normality. If data were not normally distributed a transformation was attempted to achieve normality. If normality could not be reached, or if a transformation was not appropriate, for example in cases where the data was based on standardised population scores, non-parametric tests were carried out on the raw data. For normally distributed (parametric) data, the mean and standard deviations will be reported; for non-normally distributed data (non-parametric) medians and ranges will be reported.

Following data exploration, variables were assessed for equality of variance using an analysis of variance test and then compared using the appropriate statistical test for differences in the term and VP groups. Significance was set $p < .05$; and a trend identified as a p-value between .05 and .1. In cases where the variables had repeated measures for each participant, a repeated measured analysis of variance (ANOVA) was carried out, where main effects and interaction effects were explored. Post-hoc analyses in the form of stepwise multiple comparisons using Bonferroni adjustments were run where appropriate. The statistical tests used will be reported in each results section.

In section 1.1.1, upon review of the literature, the neonatal factors previously determined to be most influential of outcome in the early years and therefore were considered within all analyses were gestational age, social economic status as determined by the Index of Multiple Deprivation quintile score (IMD quintile; calculated from the family postcode (NPEU, 2013)) and male sex. For each task, one variable was selected a priori on a theoretical basis to best reflect the task performance and will be stated in each results section. A regression model was applied to investigate the relationship between this outcome variable whilst adjusting for study group (term/VP), male sex and IMD quintile.

For the tasks conducted at 30 months of age, an additional regression model was fitted to each task data adjusting for the cognitive composite score on the Bayley-III at 2 years. This investigated whether any differences seen in the task outcomes at this age were explained by overall cognitive scores or demonstrated domain-specific variations.

For the Gap data in Chapter 5, the longitudinal observations were investigated using a multi-level mixed effects model. These were generated with the main dependent variable cluster over the two or three age points, depending on the model.

Ultimately, a longitudinal investigation of the data will be conducted within Chapter 7 to explore the predictive validity of the EF, IP and attentional measures in relation to the Bayley-III cognitive scores. Sequential regression models will be produced to examine the contribution of each task score to the overall proportion of variance accounted for in the Bayley-III results. Z-scores will be produced for continuous variables using the term-born population mean and standard deviations. These will then be entered into the models alongside the pre-determined demographic confound variables: Male sex, Index for Multiple Deprivation quintile score as the measure of Social Economic Status, and study group. The baseline group in for all regression models unless otherwise stated, will be term-born females with an IMD quintile of 1.

Chapter 3 Global Measures of Cognitive, Language and Motor Development

The Bayley Scales of Infant and Toddler Development – Third edition (Bayley-III), is an internationally recognised developmental assessment, standardised across a representative sample. Developed and published originally by Nancy Bayley in 1969 (Bayley, 1969), the scale has been repeatedly used as a gold standard for assessing developmental delay in at risk populations, going through multiple updates to keep up with current research. The most recent updated was to the Bayley-III in 2006. Produced by Pearson Education Ltd., the measure assesses 5 key developmental domains: cognition, language, motor skills, social-emotional skills and adaptive behaviour; in children as young as 1 month up to 42 months of age.

As discussed in section 1.1.2, global cognitive abilities are an indication of later IQ, and assessments such as the Bayley-III are designed to highlight children whose developmental progression is out of the normal range. Since its development in 1969 (Bayley, 1969), the Bayley-III has been consistently utilised in clinics and developmental research. Particularly within individuals born preterm, there is an imperative need to track their neuropsychological development within the first few years after birth, to ensure the infants hit their developmental milestones; the Bayley-III has been and currently is the gold standard measure used. The Bayley-III, however, has previously undergone scrutiny due to its lack of sensitivity to those showing mild cognitive delay and its inability to detect subtle domain specific performance differences within the first two years after birth (Johnson, Moore and Marlow, 2014). This research alone highlights the need for improvements in early identification measures for those at risk of mild delays.

The Bayley-III, and its predecessor, the Bayley-II, have been consistently used in preterm literature to further explore the cognitive impairments seen later in development with mixed results (Lobo and Galloway, 2013; Bode *et al.*, 2014; Spencer-Smith *et al.*, 2015). Although clinical practice consists of the use of the Bayley-III at 3, 6 and 12 months of age following preterm birth, research suggests that the measure has poor predictive validity before 24months of age (Lobo and Galloway, 2013). Once at 24 months, studies have more frequently found correlations to later IQ (Bode *et al.*, 2014). A consistent finding however, is the overestimation of developmental abilities (Anderson and Burnett, 2017). Although studies report a strong correlation between those identified as delayed in the Bayley-III at 2 years and those that later present with cognitive delays (Spencer-Smith *et al.*, 2015); a considerable proportion of children that later present with delays are not identified by the Bayley-III measures at 2 years, with this consistent across all Bayley subscales (Spittle *et al.*, 2013; Spencer-Smith *et al.*, 2015).

Discussed in section 1.1.1, the effect of SES (social economic status) on cognitive outcome has been continually reported in the preterm literature (Tideman, 2000; Hack, 2006; Moore, Hennessy, *et al.*, 2012). Poorer SES has been found to be associated to moderate cognitive difficulties, defined as <85 on the cognitive scales (Hack, 2006; Beaino *et al.*, 2011). Those from poorer family backgrounds are less likely to show improvements in cognitive score from childhood into later school years (Hack, 2006). It is also possible that the difficulties reported in preterm populations at schools are not representative of the levels of impairments detected in infancy due to the SES related biases to follow-ups; the more disadvantaged families appear to require greater persuasion to attend follow-up assessments (Moore, Hennessy, *et al.*, 2012).

A second factor consistently associated to cognitive outcome is sex of the infant. Male survivors of premature birth are typically reported to present with poorer cognitive abilities than ex-preterm females (Moore, Hennessy, *et al.*, 2012; Månsson, Fellman and Stjernqvist, 2015). Boys have frequently been reported to be at greater risk for brain injury and respiratory problems compared to girls and appear to be at a greater risk for sensory, motor and communicative problems (Elsmen, Pupp and Hellstrom-Westas, 2004; Peacock *et al.*, 2012; Månsson, Fellman and Stjernqvist, 2015). Both factors are therefore accounted for when exploring cognitive outcome within the current investigation.

The Bayley-III was conducted at two time points within the current study, 12 months as 24/30months; with the VP toddlers assessed during clinical follow-up at 2 years, the following section will elaborate, and the term toddlers during the PDP 30 month assessment. Given the evidence from the literature, it will not be possible to conclude which of the assessments holds the strongest predictive validity within the current cohort as further follow-ups would be required. However, it will be possible to explore the continuity of the measure longitudinally. In subsequent chapters, performances in EF, attention and IP will be adjusted for cognitive score performances, and inversely, the proportion of variation in the Bayley-III cognitive scores accounted for by the EF, attention and IP measures will be explored.

3.1 Methodology of the Bayley-III

Infants born very preterm are as standard in the UK, followed up at 4 time points after leaving hospital as part of their routine care: at 3, 6, 12 and 24 months of age. These assessments overlapped in part with the PDP assessment timeframe. Due to the nature of hospital appointments, in practice, the age at which the infants were seen varied. It was therefore not advisable for the PDP to repeat the Bayley-III assessment due to the possibility of practice effects and scores not reflecting true abilities. Permission was therefore sort from the parents of the infants to access medical records and the relevant Bayley-III scores were obtained. The results from the clinical 24 month Bayley-III cognitive scale were used to adjust for global cognitive performance at 30 months of age in the VP cohort and the term infants were assessed on the Bayley-III during the PDP assessments at 12 and 30 months, performed by the researchers involved in the PDP study.

The PDP researchers were taught the administration procedure of the Bayley-III by Mrs B Hutchon, the paediatric occupational therapist and National Trainer for Bayley assessments, who is responsible for the follow-up clinics within the North Central London Network, including UCH. Consistency between assessments the term and VP cohorts was strived for by following the same administration practices as those adopted in clinic.

The administration of the Bayley-III took place in the Clinical Research Facility at University College Hospital and typically lasted between 30-90 minutes depending on the age and the developmental stage of the child. The cognitive scale was typically the first scale to be performed in order to ensure scores reflected the full functional capacity of the child. The challenge of infant and toddler research is obtaining the necessary data, but in the most optimal conditions before the child tires. The cognitive scale was selected as the primary outcome measure for the assessment in this instance and therefore completion of this scale was essential.

The cognitive scale was selected and justified as the predominant outcome measure in review of the numerous investigations that previously observed a link between global cognitive performance and EF measures (Potharst *et al.*, 2012; Lobo and Galloway, 2013; Bode *et al.*, 2014; Spencer-Smith *et al.*, 2015). Although the scale has not been found to account for all variation in performance at later ages, it does appear to predict those with severe developmental impairment (Anderson and Burnett, 2017). This current investigation set out to explore whether any initial indications of EF differentiation were apparent at 2 and a half years of age in a cohort of term and VP toddlers. It was therefore essential to adjust EF performances by a measure of global cognitive performance to see if any differences in EF still remained. In a couple of instances in the current investigation, the impact of language comprehension on task performances was questioned; the language scales were therefore additionally explored in greater detail. Although all scales of the Bayley are reported below, the cognitive and language scores were predominantly taken forward though the subsequent chapter analyses.

The results of the Bayley-III comprise of 5 raw scores from each of the scales. In the following datasets, the raw scores were converted into scaled scores which took into consideration the infants' age at assessment. The scaled scores were then converted into composite (or standardised) scores; this normalised the scores around a mean of 100 and standard deviation of 15. The composite scores are calculated based on a normative sample of typically developing children, and are used to determine the developmental stage of a child's performance during the assessment according to typically developing peers. The measure has been performed on thousands of children in different countries so to provide a standardised representative dataset for each country. The UK sample included 221 children aged between 12 and 41 months and is said to take into consideration 'geographic region, gender, age, ethnicity and parental education' (Bayley, 2006). The use of composite scores reduced the reliance of carrying out the assessment within a narrow age-range. As discussed, this enabled the clinical 24 month follow-up Bayley-III scores of the VP children to be utilised in the 30 month PDP visit instead of repeating the assessment again during the PDP visit.

A topic that frequents the literature surrounding the Bayley-III is the insensitivity of the tool to mild cognitive impairments. The results of the study by Johnson, Moore and Marlow (2014), reported a low sensitivity of the Bayley-III at 24 months of age, with a number of children in the mild neurodevelopmental disability range speculated to have been missed. Given this, the cut-off score of <85 was used to identify those at possible risk of delay within this thesis (Aylward, 2013; Anderson and Burnett, 2017).

3.2 Results

3.2.1 Bayley-III scores at 12 months of age

Fifty-three full term and 32 VP infants completed the cognitive scale of the Bayley-III at 12 months (see Table 3-1 for population demographics). The mean cognitive composite score for the VP infants was 98.28 compared to 107.83 for the term infants.

		Term (n=53)	VP(n=32)
Gestational age	*Median (range); weeks^{+d}*	40^{+2} (37^{+1} – 42^{+0})	26^{+2} (24^{+0} – 29^{+4})
Male sex		28 (52.83)	23 (71.88)
IMD Quintile	1	7	5
	2	9	9
	3	7	13
	4	17	11
	5	13	6

Table 3-1. Population demographics for the cognitive scale on the Bayley-III at 12 months.

	Total cases; n=32		*Very Preterm; n=11 (<32>27 weeks)*		*Extremely Preterm; n=21 (<27weeks)*	
	n	%	n	%	n	%
Cases of IVH/PVL: I-II	14	43.75	3	27.27	11	52.38
IVH+ Intraventricular dilation	7	21.88	1	9.09	6	28.57
Intraparenchymal lesion/PVL	7	21.88	2	18.18	5	23.81
ROP:	18	56.25	2	18.18	16	76.19
Stage 1	6	18.75	0	0	6	28.57
Stage 2	8	25.00	1	9.09	7	33.33
Stage 3	4	12.50	1	9.09	3	14.29
CLD/BPD:	28	87.50	7	63.64	100	100
Mild	8	25.00	3	27.27	5	23.81
Moderate/ Severe	20	62.50	4	36.36	16	76.19

Table 3-2. Neonatal characteristics of the VP infants that completed the cognitive scale on the Bayley-III at 12 months.

At 12 months of age, VP infants cognitive composite scores were on average 9.55 points (±.73 (95%CI) lower compared to term infants (t(83) = 3.90, p < .001), with an effect size of .8 (±.06 (95%CI). The motor composite scores were on average 12.53 points (± .87 (95%CI) lower (t(78) = 4.63, p < .001) and consistent over the two subscales, fine (1.73: .16 (95%CI); (t(78) = 3.50, p < .001)) and gross motor (2.11: .62

(95%CI); (t(78) = 3.79, p < .001)). Language had similar scores (2.11: .62 (95%CI); (t(74) = .60, p = .55)).

12m Bayley-III scale		*Term (n=53)*	*Preterm (n=32)*	*Mean difference (95%CI)*	*p*
Cognitive Composite	N	53	32		
	Mean (SD)	107.83 (11.87)	98.28 (9.12)	9.55 (± .73)	***
Cognitive z-score	Mean (SD)	0 (1.00)	-.80 (.77)	0.8 (± .06)$^{\Psi}$	
Language Composite	N	44	32		
	Mean (SD)	100.55 (10.86)	98.44 (17.73)	2.11 (± 1.12)	
Receptive Language Scaled score	N	44	32		
	Mean (SD)	9.96 (2.64)	9.34 (2.66)	.62 (± .20)	
Expressive Language Scaled score	N	44	32		
	Mean (SD)	10.27 (2.18)	10.09 (3.80)	.18 (± .24)	
Motor Composite	N	46	32		
	Mean (SD)	101.87 (12.24)	89.34 (10.99)	12.53 (± .87)	***
Fine Motor Scaled score	N	46	32		
	Mean (SD)	11.07 (2.53)	9.34 (1.81)	1.73 (± .16)	***
Gross Motor Scaled score	N	46	32		
	Mean (SD)	9.47 (2.83)	7.09 (2.67)	2.38 (± .20)	***

Table 3-3. 12 month Bayley-III composite scores; *p<.05; **p.01; ***p<.001;

$^{\Psi}$ Effect size of primary measure, cognitive composite score.

Although there was a significant difference between the two study groups, only one VP infant scored within the clinically significant range with a score of 75 (Table 3-4). Three term infants and four VP infants scored <85 on the language scale.

12m Bayley-III scale		*Term (n=53)*	*Preterm (n=32)*
Cognitive Composite	N<85	0	1
	Mean score (SD)	-	75
Language Composite	N	3	4
	Mean score (SD)	80 (5.2)	73.5 (6.81)

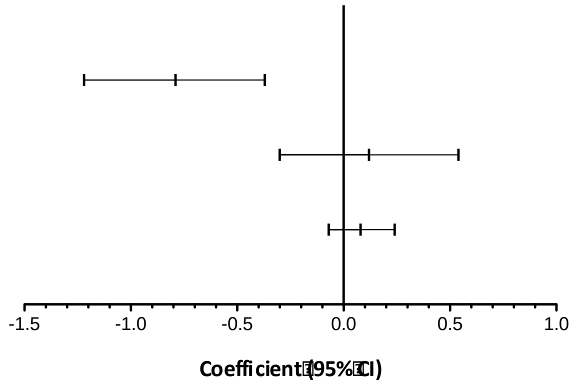
Table 3-4. Count of infants to score within the clinical range (<85)

Due to the term infants scoring on average 7 points higher than the expected standardised norms, z-scores utilising the term mean and standard deviation of the composite scores were produced in order to account for the variation between the two study cohorts.

Table 3-5 investigates the effect of male sex and IMD quintile on the 12 month cognitive z-score of the Bayley-III. At 12 months, study group was the only variable to have a predictive effect on the 12 month cognitive score outcome.

Table 3-5. Linear regression model with outcome as 12 month cognitive z-scores (F(3, 81) = 5.45, p = .001). The base group was set as the term born females with an IMD quintile of 1.

Predictor	Term (n=53)	Preterm (n=32)			
	Median (range)	Median (range)	Coef	95%CI	P
Study Group	-	-	-.79	-1.22 – -.37	.000
Male Sex	28 (52.83)	23 (71.88)	.12	-.30 – .54	.56
IMD Quintile	4 (1-5)	3 (1-5)	.08	-.07 – .24	.30
Cog score (const.)			-.34	-.99 – .31	.31



Overall model fit R^2 = 0.17

3.2.2 Bayley-III scores at 2 years

Twenty-four full term and 26 VP infants completed the cognitive scale of the Bayley-III at 2-2.5 years (see Table 3-6 for population demographics). The mean cognitive composite score of the VP infants' was 100 compared to the term infants' score of 105.

		Term (n=24)	***VP (n=26)***
Gestational age	*Median (range); weeks^{+d}*	40^{+2} (37^{+0} – 42^{+2})	26^{+2} (23^{+4} – 29^{+4})
Male sex		14 (50%)	21 (72.41%)
IMD Quintile	1	5	5
	2	4	3
	3	4	10
	4	9	7
	5	6	4

Table 3-6. Population demographics for the cognitive scale on the Bayley-III at 2-2.5 years.

	Total cases; n=26		*Very Preterm; n=8 (<32>27 weeks)*		*Extremely Preterm; n=18 (<27weeks)*	
	N	%	N	%	N	%
Cases of IVH/PVL: I-II	12	46.15	2	25.00	10	66.66
IVH+ Intraventricular dilation	7	26.92	0	0	7	38.89
Intraparenchymal lesion/PVL	5	19.23	2	22.22	3	16.67
ROP:	13	50.00	1	12.50	14	70.00
Stage 1	5	19.23	0	0	6	30.00
Stage 2	5	19.23	0	0	6	30.00
Stage 3	3	11.54	1	12.50	2	10.00
CLD/BPD:	25	96.15	7	87.50	18	100.00
Mild	9	34.62	3	37.50	6	33.33
Moderate/ Severe	16	61.54	4	50.00	12	66.67

Table 3-7. Neonatal statistics for the VP children that completed the cognitive scale on the Bayley-III at 2 years.

2-2.5 year Bayley-III scale		Term (n=24)	Preterm (n=26)	Mean difference (95%CI)	p
Cognitive Composite	N	24	26		
	Mean (SD)	107.71 (12.42)	101.92 (12.81)	-5.79 (-12.98 – 1.40)	
Cognitive z-score	Mean (SD)	0 (1)	-.47 (1.03)	-.47 (-1.05 – .11) Ψ	
Language Composite	N	14	26		
	Mean (SD)	119.43 (10.41)	97.81 (18.23)	-21.62 (-12.45– -30.79)	***
Receptive Language Scaled score	N	14	26		
	Mean (SD)	13 (1.75)	9.65 (2.80)	-3.35 (-4.81 – -1.88)	***
Expressive Language Scaled score	N	14	26		
	Mean (SD)	13.57 (2.68)	9.5 (3.82)	-4.07 (-6.18 – -1.97)	***
Motor Composite	N	9	26		
	Mean (SD)	117.44 (15.91)	94.58 (11.86)	-22.87 (-35.58 – -10.15)	***
Fine Motor Scaled score	N	9	26		
	Mean (IQR)	12.33 (1.66)	9.92 (1.92)	-2.41 (-3.83 – -.99)	**
Gross Motor Scaled score	N	9	26		
	Mean (IQR)	13.44 (4.18)	7.88 (1.93)	-5.56 (-8.82 – -2.3)	***

Table 3-8. Bayley-III scores at 2 year time point; *p<.05; **p<.01; ***p<.001

Ψ Effect size of primary measure, cognitive composite score.

In contrast to the scores at 12 months, at 2-2.5 years the cognitive composite scores of the VP children showed a lesser deficit compared to term children (difference in means: 5.0; -12.98 – 1.40 (95% CI); (z(48) = 1.84, p = .07)).

VP infants scored significantly lower in language (difference in means 21.62 points; 12.45-30.79 (95% CI); (t(38) = 4.77, p < .001)) which was consistent over the two subscales, receptive (difference in means 3.35; 4.81-1.88 (95% CI); (t(38) = 4.64, p < .001)) and expressive language (difference in means 4.07: 1.97-6.18 (95% CI); (t(38) = 3.93, p < .001)); and motor composite scores (difference in means 22.87 points; 10.15-35.58 (95% CI); (z(33) = 3.39, p < .001)) and subscales, fine (difference in means 2.41: .99-3.83 (95%CI); (z(33) = 2.93, p < .003)) and gross motor (difference in means 5.56: 2.3-8.82 (95%CI); (z(33) = 3.41, p < .001)).

Both the cognitive and motor composite score were compared with Mann Whitney U tests due to the marginal skew in both data sets (Figure 3-1).

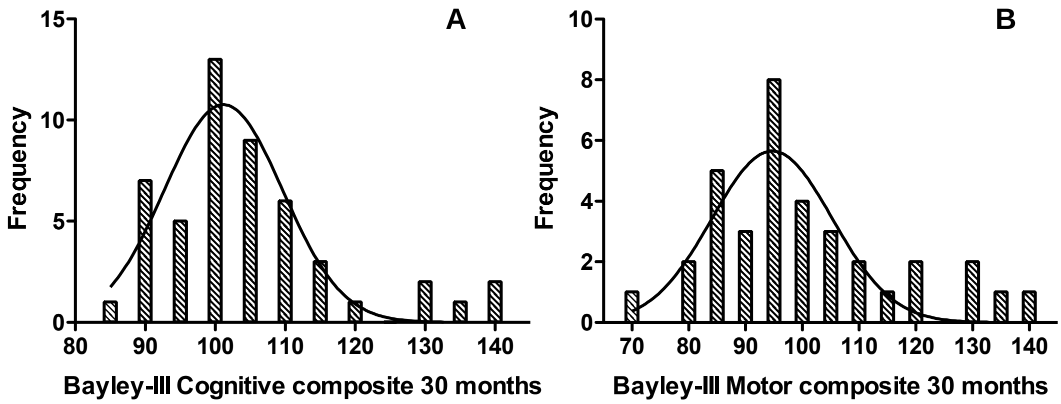


Figure 3-1. Frequency distribution of the cognitive (A) and motor (B) composite scores collapsed across study groups.

No term or VP born children scored within the clinically significant range for the cognitive scales at the 2 year assessment, when taking <85 as the cut off (Table 3-9). Nine VP children displayed language scores within the clinical range, with one infant scoring within the clinical range at for both language scores.

2 year Bayley-III scale		*Term (n=24)*	*Preterm (n=26)*
Cognitive Composite	N<85	0	0
	Mean score (SD)	-	-
Language Composite	N	0	9
	Mean score (SD)	-	78.33 (7.05)

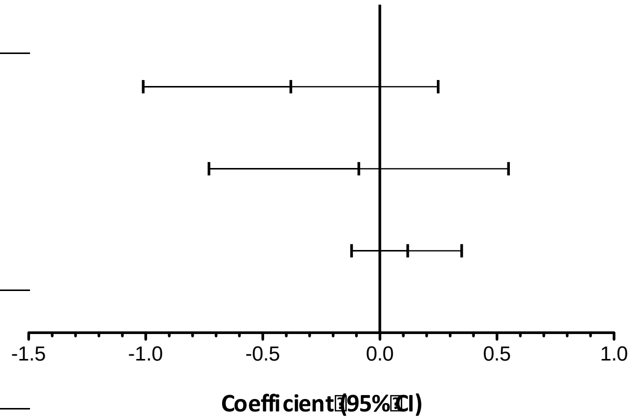
Table 3-9. Count of children within the clinically significant range of <85

Again, the term infants scored on average 5 points higher than the expected standardised norms, therefore the z-scores utilising the term mean and standard deviation of the composite scores were produced in order to account for the variation between the two study cohorts.

Table 3-10 explores the effect of study group, male sex and IMD quintile on the 2 year cognitive z-score of the Bayley-III. At 30 months, study group did not have any predictive effect on the cognitive score as the outcome.

Table 3-10. Linear regression model with outcome as 30 month cognitive z-scores (F(3, 46) = 1.21, p = .32). The base group was set as the term born females with an IMD quintile of 1.

Predictor	Term (n=24)	Preterm (n=26)	Coef	95%CI	P
	Median (range)	Median (range)			
Study Group	-	-	-.38	-1.01 – .25	.23
Male Sex	14 (50)	21 (72.41)	-.09	-.73 – .55	.78
IMD Quintile	4 (1-5)	3 (1-5)	.12	-.12 – .35	.32
Cog score (const.)			-.36	-1.30 – .57	.44



Overall model fit R^2 = 0.07

3.2.3 Longitudinal Bayley-III cognitive scores

Eighteen term and 22 VP infants completed the cognitive scale of the Bayley-III at both time points months (see table 3-11 for population demographics). The mean cognitive composite score for the VP infants increased by 2.5 points (± 1.51 (95%CI)) compared to smaller .27 point rise (± 2.08 (95%CI)) for the term born infants. The effect size within the term-born infants was .03 and within the VP .27.

		Term (n=18)	***VP(n=22)***
Gestational age	*Median (range); weeks^{+d}*	40^{+3} (38^{+5} – 42^{+0})	26^{+3} (24^{+0} – 29^{+4})
Male sex		28 (52.83%)	23 (71.88%)
IMD Quintile	1	2	5
	2	4	3
	3	3	5
	4	4	7
	5	5	2
Cognitive composite 12m	Mean (SD)	105.56 (13.92)	99.09 (9.08)
Cognitive composite 2 years	Mean (SD)	105.83 (11.28)	101.59 (13.31)
Mean Difference	(95%CI)	.27 (± 2.08)	2.5 (± 1.51)

Table 3-11. Longitudinal population demographics for the cognitive scale on the Bayley-III at 12 months and 2-2.5 years

Table 3-12 and figure 3-2 shows the correlation coefficients between the cognitive scores collected at 12 months and 2 year time points.

Bayley-III scale	*Cognitive Composite*	
	N	Correlation Coefficient (p)
Total Cohort	40	.31 (.05)
Term	18	.31 (.22)
Preterm	22	.27 (.23)

Table 3-12. Correlation coefficients between the 12 month and 2 year cognitive Bayley-III scores.

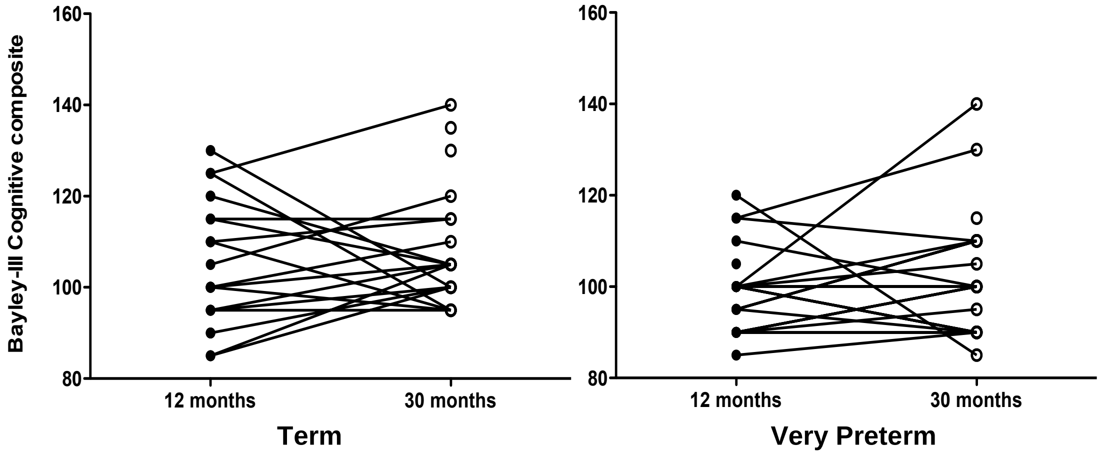
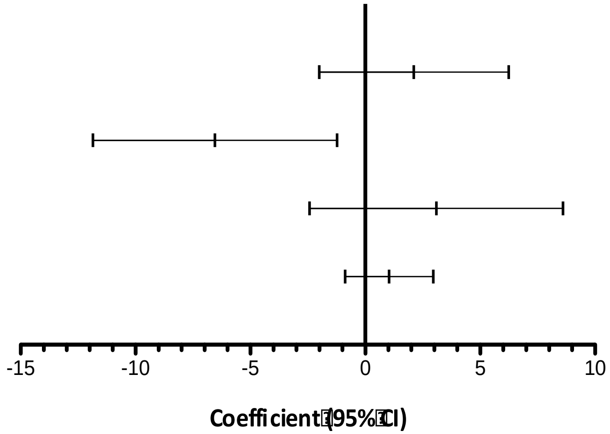


Figure 3-2. 12 month and 2 year Bayley-III correlation data for Term and Very Preterm children

Table 3-13 displays the outcome of a random intercept linear mixed-effect model with cognitive composite scores at both time points as the dependent variable, participant ID as the random effects identifier and age of testing nested within the model. Study group was a significant predictor of the overall cognitive score, with the preterm regression coefficient of -6.55. The age of the assessment was not a significant predictor of the outcome, nor was male sex or IMD quintile

Table 3-13. Mixed effects regression model with outcome as cognitive scores over the 12 month and 2 year Bayley-III assessments. Only children with both time points were included within the model. The predictor 'Age' was the longitudinal index within the model. The base group was set as the 12 month cognitive score for the term born females with an IMD quintile of 1.

Predictor	Term (n=18)	Preterm (n=22)	Coef.	95%CI	P
	Median (range)	Median (range)			
Age	-	-	2.11	-2.01 – 6.24	.31
Study Group	40^{+3} (38^{+5}-42^{+0})	26^{+3} (24^{+0}-29^{+4})	-6.55	-11.86 - -1.23	.02
Male sex	9 (50%)	18 (81.82%)	3.09	-2.42 – 8.60	.27
IMD Quintile	4 (1-5)	3 (1-5)	1.04	-.88 – 2.96	.29
Cog score (const.)	-	-	100.73	92.93 – 108.63	.000



Wald chi^2 = 8.94, p = .06

3.2.4 Z-score analysis

Consistent with previous research, the term infants in the current cohort scored higher on the cognitive scale of the Bayley-III than typically expected when utilising the normalised means provided by the tool. Due to this higher mean within the term group, z-scores were calculated for both cognitive composite scores at 12 months and 2 years and will be used for all global cognitive performance adjustments within subsequent analysis chapters. By formulating a z-score based on the term born infants performances, a clearer interpretation can be made as to how the VP infants perform in relation to the controls on each of the EF, IP and attention tasks once global cognitive performance is adjusted for. Table 3-14 collates the scores at both assessment ages.

		Term	***VP***
12m Cognitive scores (n=T:53; VP:32)	**Male**	28 (52.83%)	23 (71.88%)
	IMD Quintile (Median; IQR)	4 (2)	3 (2)
	Composite (mean; sd)	107.83 (11.87)	98.28 (9.12)
	z-score (mean; sd)	0 (1.00)	-.80 (.77)
24/30m Cognitive scores (n=T:24; VP:26)	**Male**	14 (50%)	21 (72.41%)
	IMD Quintile (median; IQR)	4 (2)	3(2)
	Composite (mean; sd)	107.71 (12.42)	101.92 (12.81)
	z-score (mean; sd)	0 (1.00)	-.47 (1.03)

Table 3-14. Cognitive composite and z-scores from the Bayley-III assessment at both 12 and 24/30 month time points.

3.3 Discussion

The primary goal of the current thesis is to further our understanding of the VP children's cognitive abilities through the first two years of life, with a particular interest in determining any specific EF, IP and attentional difficulties. Thus, it is of high importance to have a global measure of cognitive performance to understand if any difficulties observed in later specific tasks are in line with general ability, or impaired above the level expected.

The results of the Bayley-III cognitive scale within the current investigation found the term infants scored on average 7 points higher than the VP infants at 12 months of age, and 5 points higher at 2 years. Although the term infants displayed a level of consistency in the scores at each age point, when looking within infants that had two measures there was a no longitudinal correlation observed.

Within the preterm literature, a prominent question is how early delays or impairments can be accurately detected in order to develop successful, targeted interventions (Spittle *et al.*, 2007). Previously, Bayley-III cognitive scores at 2 years have been reported to correlate with later cognitive performance abilities (Bode *et al.*, 2014); limited evidence is available for the predictive validity at 12 months (Lobo and Galloway, 2013). The VP infants displayed a marginal increase in cognitive score at 2 years, however, when exploring those with two time longitudinal measures, using pairwise correlation, figure 3-2 displays the weak correlation between the scores across the two years. This draws into question the reliability of the scores as a measure of cognitive performance over the two years. Although it could be argued that the absence of a relationship between the 12 and 30 month scores reflects a discontinuity between the measures; the 12 month Bayley-III scores cannot be discounted without additional follow-ups. The increase in VP scores at the two year assessment but poor correlation to the score at 12 months could suggest that some of the VP infants are displaying an improvement in cognitive ability, but others a decrease. This would not necessarily infer the cognitive score at 12 months as incorrect, but it would make it a poor predictor at 2 years. As seen in previous studies, for example the investigation by Lobo and Galloway (2013), the 24 month assessment is a better reflection of later performance in preterm infants; it could therefore be postulated that the performance differences here are showing more stability by the age of 2 years. Which of the scores, either at 12 months or 2 years, is the more accurate reflection of later ability within the current cohort will require further follow-up beyond the current investigation. However, the validity of this measure can be explored in relation to the two cohort's performances on EF, IP and attention specific tasks. This will provide an insight into how well the measures are evaluating the abilities in relation to these cognitive skills.

There has been much discussion within the literature regarding the score used to define developmental delay within the Bayley-III (Aylward, 2013; Johnson, Moore and Marlow, 2014; Anderson and Burnett, 2017). Previously, the Bayley-II had a combined cognitive and language score called the Mental Development Index (MDI) and was highly regarded for many years (Johnson and Marlow, 2006; Johnson, Moore and Marlow, 2014). Upon the introduction of the Bayley-III, where the cognitive and language assessments were separated into independent scales, although the Bayley-III scores correlated with the previous MDI scores, the Bayley-III appeared to be producing scores approximately 7-10 points higher, making scores of 107 the norm, not appropriate for a standardised measure (Aylward, 2013; Johnson, Moore and Marlow, 2014).

Johnson, Moore and Marlow explored the differences in scores in a study in 2014 with a cohort of extremely preterm infants. The conclusion reached was a cut off of <85 on either the cognitive or language scales is more representative of moderate to severe neurodevelopmental delay than the previously used score of <75. This higher cut off was more in line with those that had previously scored within the neurodevelopment impairment range of <70 on the MDI (Johnson, Moore and Marlow, 2014).

Other methods of dealing with this discrepancy include using Developmental Quotient (DQ), generated by dividing the developmental age by the chronological age and multiplying by 100 (Milne, McDonald and Comino, 2012). This theoretically provides an estimated rate of development relative to a standardised sample. However, when investigating a group of children born preterm, particularly beyond the age of 2 years of age, there is often disagreement on whether to continue to adjust for the child's corrected age or use their chronological age. The use of this measure can therefore be hard to justify (Rickards *et al.*, 1989; Sugita *et al.*, 1990; de Jong *et al.*, 2015). This method also assumes that the standard deviations of the scores are comparable for all ages, however, that is not always the case and therefore the DQ could be considered less precise (Anderson and Burnett, 2017).

Within the current cohort, no toddler scored within the mild to moderate clinical range for the cognitive score at 24 months (cognitive score <85). It could be postulated that the current results illustrates the poor sensitivity and predictive validity of the Bayley-III commonly reported in the literature, and mild cognitive impairments are not reflected in the scores (Anderson and Burnett, 2017). When comparing to previous studies, the current results of the Bayley-III appear unusual from a cohort of children born before 32 weeks of gestation. Previously, in a cohort of toddlers born at <30 weeks, 11% of the infants tested scored within the mild to moderate impairment range at 24 months (<85) (Spencer-Smith *et al.*, 2015). These results echoed the study by Bode *et al.*, (2014) who reported 18% of VP infants to score within mild to moderate range (GA <30 weeks). An investigation into the cognitive performance of extremely premature infants (birth at < 27 weeks gestation) at 2 to 3 years of age reported 10.2% to score <85 on the cognitive scale (Johnson, Moore and Marlow, 2014). These results are not supported in the current cohort, where 69% of the infants were born at <27 weeks gestation.

These high Bayley-III scores could be a reflection of lower neonatal risks within the cohort. However, in the investigation by Spencer-Smith *et al.*, (2015), 9% infants were reported to have had a IVH grade 3 or over compared to 19% of the current cohort, and 31% of the infants reportedly suffered with BPD compared to 61% of the current cohort. Although these characteristics have not been consistently related to later outcome, this is likely due to the difficulty in categorising the severity of illness within premature cohorts; previous reports have found associations between these conditions and later cognitive performances (Luu *et al.*, 2009; de Mello, Rodrigues Reis and da Silva, 2017). As the prevalence levels of these neonatal conditions do not differentiate the cohorts discussed from the current cohort, the high Bayley-III scores of the current population could therefore indicate the infants recruited onto the PDP are displaying a more typical trajectory. However, given what is known about preterm development and given the mean gestational age of the cohort (approximately 26 weeks gestation), evidence would suggest this is not likely to be the case.

Although the sensitivity and specificity of the Bayley-III scores are still being explored within the literature, there is some evidence that suggests performance at 2 years may have a level of predictive validity for cognitive functioning later in life (Bode *et al.*, 2014; Spencer-Smith *et al.*, 2015; Breeman *et al.*, 2016; Linsell *et al.*, 2017, In press). Bode and colleagues found a correlation coefficient of .81 for the cognitive scale and .78 for the language scale of the Bayley-III when comparing scores to those collected in the WPPSI-III at 4 years of age in a group of preterm children (Bode *et al.*, 2014). Spencer-Smith and colleagues also investigated the predictive nature of the Bayley-III at 2 years on later functioning using the DAS-II in a group of very preterm children, reporting a low sensitivity of the measure at detecting those within the mild to moderate range (Spencer-Smith *et al.*, 2015). In contrast, an investigation into the relationship between cognitive function in childhood through to adulthood in a cohort of very preterm or low birth weight infants, by Breeman *et al.*, (2015), reported a level of consistency within cognitive scores measured at 20 months of age to IQ scores reported in adulthood, even when excluding those within the severe range. This clearly illustrates the discrepancies within the literature regarding the use of the Bayley-III.

From this exploration into the Bayley-III scores at 12 and 2 years, it can be concluded that there is an overall performance difference between the term and VP cohorts on global cognitive function. Although, none of the toddlers score within the clinically significant range at the 2 year time point for the cognitive scale, this does not rule out subtle impairments within the VP group, and compared to the term born infants there is a significant difference between the cohorts over the two time points. Due to this difference and given the discrepancies with the Bayley-III normative data within the literature, it is likely that the term born infants within this study are a better reference point for global cognitive performance. In order to ensure performance across subsequent task analyses is comparable, the z-scores calculated based on the term performance at each age point will be used to adjust for global cognitive performance. For the tasks analysed in the first year, the 12 month cognitive z-scores will be used, and for those conducted at 30 months the 2 year z-scores will be used.

Chapter 4 Executive Functions

EFs are considered top-down processes that influence the more routine or fundamental cognitive skills that have been automated over time following learning and repeated practice (Burgess, 1997). The progression of automatic processing to higher order functioning, allows for more control and process-driven behaviours that facilitate our ability to plan, troubleshoot and handle novel situations (Gilbert and Burgess, 2008). Deficits in these abilities can lead to impulsivity and distractible behaviours (Hughes, 2002) as well as difficulties in conceptual reasoning and later academic achievements (Aylward, 2005). Children born preterm consistently show poorer performance on EF tasks (Howard, K., Anderson, P. J., & Taylor, 2008; Mulder *et al.*, 2009; Mulder, Pitchford and Marlow, 2011b; Rose, Feldman and Jankowski, 2011, 2012; Aarnoudse-Moens *et al.*, 2013).

Quantifying these cognitive processes has proved a challenge for many research groups due to the interrelated nature of these skills, in addition to the associated social and emotional influences (Burgess, 1997; Gilbert and Burgess, 2008). The understanding of the three sub domains of EF: inhibition, working memory, and cognitive flexibility have driven the development of EF assessments to predominantly focusing on one EF domain over the others. For example, inhibitory based tasks commonly require the individual to overcome a strong stimulus-associated response; working memory tasks require the holding of information in mind over a period of delay; lastly, cognitive flexibility tasks require switching between rules or conditions or between two or more stimulus-based responses (Gilbert and Burgess, 2008). However, as discussed, assessing one domain incurs the use of the others because of the common factors between them, described by the unity and diversity model (Miyake *et al.*, 2000; Miyake and Friedman, 2012).

When studying the development of EF, research suggests that these skills improve with age (Beveridge, Jarrold and Pettit, 2002), with a substantial amount of EF research focusing on preschool and early school aged children. Age-related changes in EF abilities are often reflected in task complexity. For example, younger children are more likely to score poorly on tasks with complex rules than older children who have the abilities to comprehend, remember and execute the task instructions (Hughes, 2002). In older children, this allows task parameters to be altered in order to assess the contribution of different EF domains on cognitive performance, providing a within subject understanding of EF maturity, such as working memory load capacity (Hughes, 2002). However, in very young children, this is not necessarily reflective of EF ability, but rather of language and/or motor abilities that could hinder the completion of the task. Therefore emphasis has been placed on the importance of age appropriate assessments for targeting EF in infancy and early childhood (Best and Miller, 2010), and can constrain assessments to specific methodological structures.

Studies into the developmental trajectory of EF and the emergence of the sub-domains have produced conflicting results, potentially due to the restrictive nature of infancy capabilities when attempting to target the different EF domains. A general consensus within the literature proposes all domains emerge and show signs of development during preschool years, with working memory and flexible thinking showing continued development into adolescence and beyond (Best and Miller, 2010; Roebers, 2017), yielding support for the unity and diversity model (Miyake *et al.*, 2000). These developmental changes in EF ability are thought to reflect, in part, the adaptations of the frontal lobe during development. The prefrontal cortex is considered to predominantly govern EF abilities (Sun and Buys, 2012a).

The growth of the frontal lobe is protracted in human development. Prefrontal adaptations including synaptogenesis and myelination occur late in the pre- and perinatal period. It has been postulated that this area may be vulnerable to disruptions, such as hypoxic events, during preterm birth (Espy *et al.*, 2002). Although the causality of EF deficits observed in ex-preterm populations is unknown, there are repeated reports of smaller regional volumes, such as the frontal lobes, basal ganglia and cerebellum, as well as disturbances in subcortical white matter, correlated with poorer EF outcomes in ex-preterm populations (Nosarti *et al.*, 2008; Sun and Buys, 2012a; Taylor and Clark, 2016).

Impairments in EF are typically reported in later childhood for ex-preterm populations, the size of the deficit being proportional to gestational age at birth (Aarnoudse-Moens *et al.*, 2011). Whether the deficits observed are specific to one EF domain or a more general disability, in the literature is undecided, as detailed in section 1.2. In any regard, two meta-analyses investigating ex-preterm performances on EF related tasks reported a consistent deficit in EFs across the preterm literature (Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, 2009; Mulder *et al.*, 2009). In contrast to the widely investigated difficulties observed in EF in later life, there is a limited understanding in the literature of when these deficits first emerge in preterm populations.

In infancy, a number of studies have reported poorer performances in EF-related tasks in preterm cohorts (Sun, Mohay and O'Callaghan, 2009; Lobo and Galloway, 2013). For example, preterm infants have been reported to display poorer working memory abilities in response to the A-not-B paradigm at 7 to 8 months of age (Sun, Mohay and O'Callaghan, 2009), and have displayed poorer learning to the conjugate mobile reinforcement paradigm at 3 months of age (Lobo and Galloway, 2013). However, to the authors' knowledge, no studies to date have examined the trajectory of these abilities over the first year into the toddler years. It is unclear how deficits observed in the first year fit into the older cognitive profile of children born preterm.

Pre-school and older children have been extensively studied in the very preterm population (Espy *et al.*, 2002; Vicari *et al.*, 2004), but very few studies have investigated EF performance in toddlers (Ross *et al.*, 1996; Pozzetti *et al.*, 2014). In those that have, the results of EF tasks are mixed. For example in a study by Ross *et al.*, (1996), significant differences between 28 month old term and preterm toddlers were reported on a hidden object task (a working memory assessment) and a reverse response set paradigm (a cognitive flexibility assessment). In contrast, Pozzetti *et al.* (2014) only reported differences between preterm and term born toddlers on cognitive flexibility measures. Differences between these investigations, from population demographics, to task procedures, could explain these findings. The absence of investigations into this age range leaves a large gap in our understanding on the emergence of deficits in preterm populations.

The current investigation aims to provide additional evidence by exploring EF abilities within a longitudinal cohort.

The following chapter uses established investigations to understand the relationship between emerging EF over the first 2 years after birth in a very preterm population and more conventional measures of developmental outcome, the cognitive scale of the Bayley-III. It is hypothesised that differences will be observed in EF performances across both the first and second year assessments. Due to the nature of the development of EF, currently there are no established tasks available that assess EF abilities in the both first and second years. The necessity of increasing task complexity with age in order to challenge EF performances unfortunately hampers direct comparisons between the tasks from the first to the second year. Nevertheless, a detailed observation of EF abilities over the first two years, in a population in whom difficulties are predicted could provide a clearer theoretical understanding of how EF develops. With this greater understanding, better detection of early difficulties could be established. Subsequently, custom targeted inventions then could be developed to optimise the developmental trajectories of children showing signs of early delay.

The EF tasks will be reported in age order to acquire an understanding of how differences in EF performance evolve with age.

4.1 Methodologies and procedures

4.1.1 Delayed Response Task (DRT)

The Delay Response Task (or DRT) was administered at the 6 month time point, with the version used similar to that used by Schwartz and Reznick, (1999) and Noland *et al.* (2010), using ocular movements to determine correct responses (developed originally by Gilmore and Johnson (1995)). The current task design was developed by Natasha Mooney, who has an abstract on her work with this task published in DMCN.

The premise of this task in the first instance is to assess early working memory capabilities by presenting the infant with a stimulus in a specific location and assessing their ability to remember this location after a delay. The DRT format has been a long established assessment of frontal lobe function both in human and non-human primates (Fuster, 1973). Improved performance in such task were noted from 6 to 12 month infants by Diamond and Doar, in line with prefrontal cortex development (1989). A delayed response task was selected due to its established success in the literature for determining early cognitive functioning in infants (Garon, Bryson and Smith, 2008), and the well-known link between the frontal lobe and EF performance (Diamond and Doar, 1989). At 6 months of age, infants have been shown to be able to retain information in mind for a period of a few seconds, and this ability increases with age. The success of other EF tasks, for example the A-not-B paradigm, are typically confounded by immature motor and planning abilities at this age (Thelen, Corbetta and Spencer, 1996). Therefore ocular movements were utilised for this paradigm to provide the best reflection of EF performance.

In preterm infants, there is limited research utilising Delayed Response paradigms at 6 months of age. The majority of investigations occur in the second half of the first year, utilising paradigms such as the A-not-B (Sun, Mohay and O'Callaghan, 2009; Sun and Buys, 2012a).

4.1.1.1 DRT Apparatus and methodology

The DRT apparatus consisted of a 90 x 60cm screen securely fastened to a narrow table 72cm high. The screen contained 2 windows of the same dimension (15cm x 21cm) cut 9cm in from each side of the screen and 7cm from the top. On the back of the screen, a roller blind was secured at the top with a pull that would lower and raise the curtain during the task. The infant was sat on the parent or guardians lap on a chair in front of the screen. The chair was adjusted so that the infant was positioned in the centre of the screen and approximately 150cm way. During the task, a camera was secured to the top centre of the screen so that the infant was in full view and their eye gaze fully visible. The experimenter sat behind the table. A black cloth was laid over the table, beneath the screen, to obscure the experimenter sat behind.

During the task, 2 stimuli were presented to the infant from the windows in the screen. In this version of the task, a non-social stimulus, or a social stimulus was presented. The non-social stimulus was one of two rattles, either a round rattle comprised of 2 pink bowls filled with rice, and decorated with pompoms, or the typical shape of an infant rattle with a stem and oval end in blue, yellow and red colours, but was not a common toy that the child was likely to have come into contact with. Neither rattle displayed any social reference and when shaken in the windows of the screen during the task, was held in a way that did not display any part of the experimenters' body. The social stimulus was the experimenter themselves, presenting their face in the windows of the screen. Upon presenting their face in the window, the experimenter would say 'Hello *infants name*, hello', in an enthusiastic manner to capture the infants' attention.

The DRT comprised of 3 test phases; a pre-test phase and a social and a non-social condition. The infant was first shown the pre-test phase where a paired comparison screening took place. The comparison displayed both the non-social stimulus and the experimenters face simultaneously, one in each window. Simultaneous comparison presentation was repeated twice, alternating the stimuli between windows in the two trials. The order of the side of social stimulus presentation was counterbalanced between infants. Preceding the pre-test phase, either the social or non-social condition was initiated, the order of which was counterbalanced between infants. Both conditions started with the lowering of the roller blind. Once at the bottom the experimenter waited for 3 seconds before lifting the blind again, being careful not to expose any of their body to the infant through the windows of the screen. Once the blind was lifted, the experimenter waited for 5 seconds before either presenting the non-social stimulus or social stimulus in one of the windows. This 5 second delay was the response window for the infant to direct its gaze to the previous stimulus presentation window, where they should be expecting the stimulus to re-appear (this was not possible on the first trial as no trial had preceded this). The 'call' of each condition lasted approximately 5 seconds each, in order to orient the infants' attention to the window of presentation. The non-social stimulus was rattled for 5 seconds, and the experimenter called the infant using the phrase above whilst at the window. The roller blind was then lowered and the next trial would begin. The infants gaze was recorded in the response window after the completion of the first trial, when the curtain was re-opened before the presentation of the second stimulus condition.

The infant's initial gaze direction after the opening of the curtain on each trial was recorded as correct or incorrect according to the side of the preceding stimulus presentation (Reznick *et al.*, 2004). Due to this, the first trial performed could not be recorded as a test trial for the first condition as there were no preceding trials. A total of 10 responses for each condition was required from the infant, therefore to achieve this, 10 trials were conducted for the first condition, and 11 for the second. The first trial of the second condition was still and assessment of working memory for the previous condition as the response window precedes the next new stimulus presentation. For clarity, please see figure 4-1.

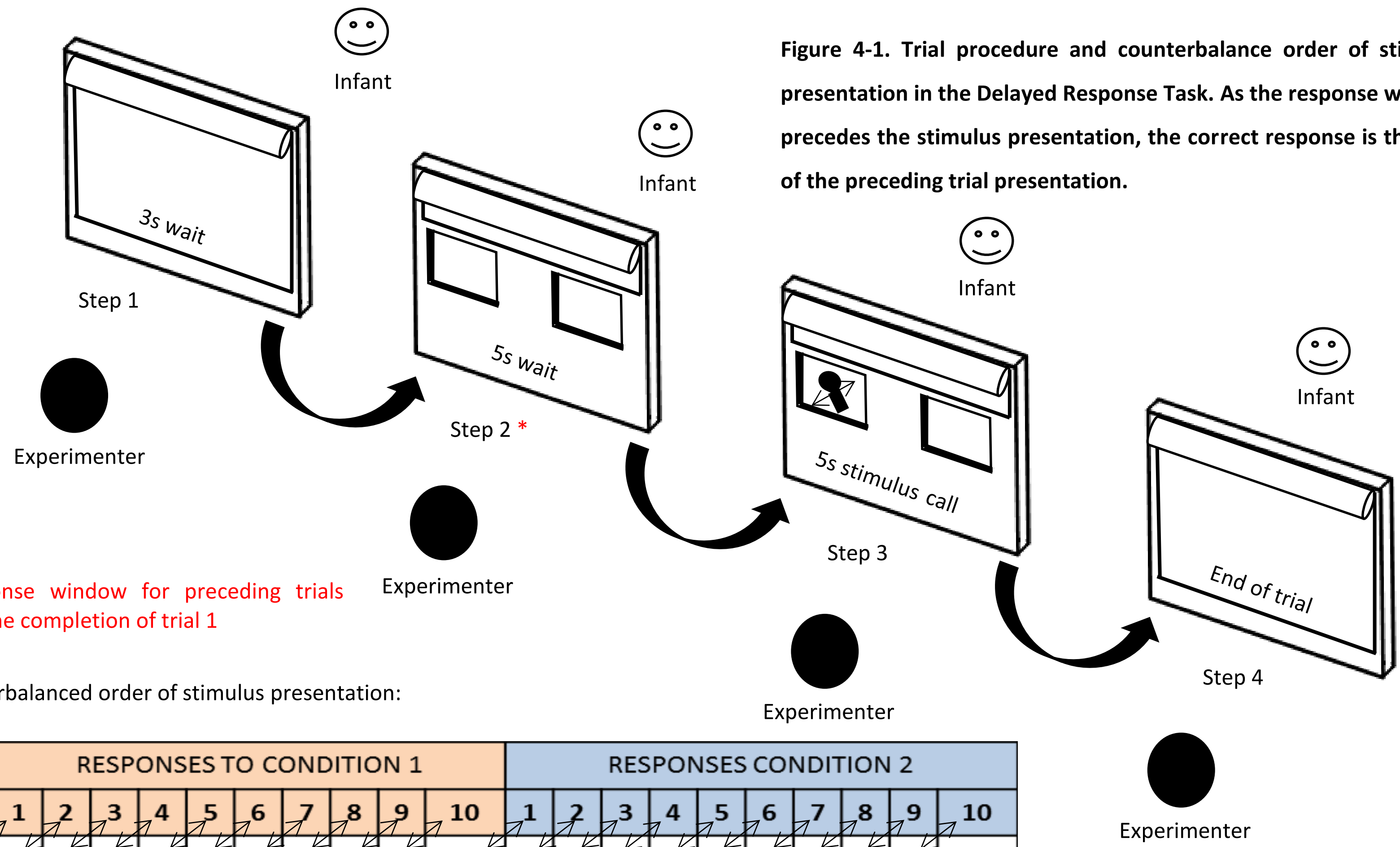
The sound of the roller blind was clearly audible within the videos and was used to determine the start and finish of each trial. Upon hearing the roller blind lift, the direction of the infant's first look was marked as the response to the previous trial.

If the infant looked towards the location of the previous stimulus, a correct response was awarded. If the infant looked towards the incorrect direction, an incorrect response was recorded. If the infant failed to look at the apparatus in the response window, this was recorded as not looking.

The total number of correct trials for each condition was recorded then divided by the total number of trials for that block. This gave a 'proportion of correct' for each condition. Although the aim was to obtain 10 trials for each condition, this was not always possible due to the temperament of the infant. The results were collapsed across conditions, but a minimum of 8 completed trials per condition were required for infants to be included in the analysis, to ensure there was no social bias within the results.

Normal distribution of the 'proportion correct' variable meant a t-test was carried out to compare the study group performances (term vs VP). The number of trials where the infants were not looking (total not-looking), were not normally distributed and therefore were compared using Mann-Whitney U-test. The variable selected a priori to best reflect the performance on this task was the 'proportion correct'. A regression model was produced with this main performance variable as the outcome measure, with the specified predictors stated in chapter 2 section 2.3 additionally included.

Figure 4-1. Trial procedure and counterbalance order of stimulus presentation in the Delayed Response Task. As the response window precedes the stimulus presentation, the correct response is the side of the preceding trial presentation.



*Response window for preceding trials after the completion of trial 1

Counterbalanced order of stimulus presentation:

	RESPONSES TO CONDITION 1										RESPONSES CONDITION 2									
0	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
L	L	R	R	L	R	L	L	R	R	L	L	R	R	L	R	L	L	R	R	L

Stimulus change

4.1.2 A not B Paradigm (A*B* task)

At 12 months of age, EF was measured using a task with a similar format to the delayed response task, the Piagetian A not B paradigm (A*B*) (Piaget, 1954; Diamond, 1985). The A*B* paradigm is a long-standing assessment of EF, particularly working memory in children within their first year and has long been considered to reflect developmental milestones (Rose, 1983).

The administration used the standard method. The task required infant to attend to a stimulus whilst it was hidden in location 1 in the first instance. This primary step of the task recruits attentional control in order to later correctly reach for the object in location 1 (Reynolds and Romano, 2016). Following two correct retrievals, the object hiding location was switched and the infant had to inhibit the established prepotent motor response that gave rise to the reward of the toy in the first two trials. If the infant correctly identified the toy following the switch, the procedure was repeated and a delay was imposed before the infant was allowed to search on subsequent trials. Working memory networks were therefore challenged following the short delays (Schwartz and Reznick, 1999; Espy *et al.*, 2002; Reynolds and Romano, 2016). When the infant did not correctly identify the toy following the switched hiding location, the task was terminated and the infant was termed to have shown the preservative 'A-not-B' error.

The 'A-not-B' error (A*B* error) has been found to emerge between 7 and 8 months (Wellman, Cross and Bartsch, 1987). From this age onwards, performance has been shown to display marked improvements with age, with the length of delay tolerated increasing over the first year (Diamond, 1990; Thelen, Corbetta and Spencer, 1996; Garon, Bryson and Smith, 2008). This task therefore is seen a good reflection of EF capabilities at 12 months.

Previous studies with preterm infants have found mixed results on this task (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008), highlighting inconsistencies in this population. Significantly higher A*B* errors have been reported in preterm infants compared to term born peers at 8 (Sun and Buys, 2011), and 10 months of age (Ross *et al.*, 1992) when corrected for prematurity. In contrast, in 6 to 14 month old infants longer delays have been tolerate in preterm infants over terms born controls before the A*B* error was observed (Matthews, Ellis and Nelson, 1996). The latter study however, utilised a non-reaching version of the A*B* paradigm, allowing for the younger ages to be assessed and included preterm infants considered to be low-risk as their mean prematurity was -31.9 days. This therefore could explain the discrepancies with other population reports. In any regard, performance differences are not conclusive, and the literature calls for further investigations for additional clarity.

4.1.2.1 A-not-B Apparatus and methodology

Figure 4-2 illustrates the apparatus used for this paradigm. The infant sat on a parent/carer's lap on one side of an elongated table with the experimenter on the other. The testing table contained two wells (10cm x 10cm) of a depth of 8cm. Within these wells, the experiment hid a toy for the infant to find. In a pre-test period, the infant was given the toy to play with for a period of time before commencing the task, so that there was a desire to locate the toy. If one toy did not create any level of enjoyment, the toy was changed as it was important the child displayed some interest in the object being hidden. Once the pre-test period was complete, the task initiated by hiding the toy in the left or right well. The toy was placed into a well when full attention of the child was on the toy. As the toy was placed into the well, the experimenter stated 'I am hiding the toy in here'. The experiment then simultaneously covered both wells with the 2 orange cloths illustrated in Figure 4-2. The experimenter then asked the infant 'where is it?' in combination with a hand gesture where they opened both hands and raised their shoulders. If the infant correctly identified the well containing the toy, the child was allowed to play with it for a short period. If the toy was not correctly identified, the experimenter initiated the second trial. Each trial followed the same format.

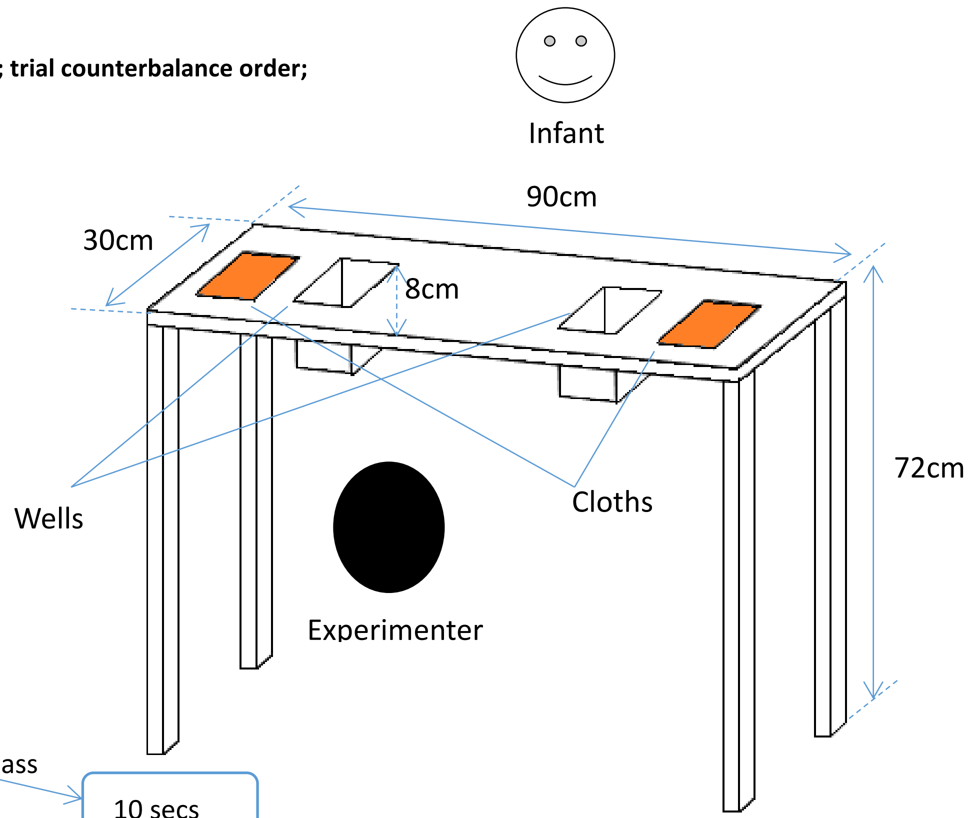
The initial well the toy was hidden in, was counterbalanced between infants to avoid any bias (please see Table 4-1). The criterion set for the paradigm, asked the infant to correctly identify the toy in the same location on 2 consecutive trials before the location of the toy was switch to the alternative well. Upon correct identification after a switch, a delay of 5 seconds was introduced, where the experimenter paused after placing the toy in the well and covering the cloths. The 5 seconds was counted out loud whilst maintaining the infant's eye gaze. The infant was then asked to locate the toy again. If the infant again located the toy on 2 consecutive trials and on the switch, the delay was increased by another 5 seconds. This continued until the child made an error on the switch trial. If the infant did not correctly identify the toy on the switch, this was considered the 'AB' effect and the task was terminated.

The infants' responses were video recorded and scored offline. The A*B* error was coded according to the delay that the infant reached. For example, if the infant reached the 10 second delay after previously correctly selecting the location of the toy at 5 seconds, and then proceeded to meet the criterion and find the toy correctly on two consecutive trials with a 10 second delay before not correctly finding the toy on the switch, the infant was given the score of 10 for the A*B* error. All infants included in the analyses were required to meet the criteria for the switch at each level. Although an infant may pass the first level, on occasions, the task was not completed, and therefore the level at which the A*B* error occurred was not achieved.

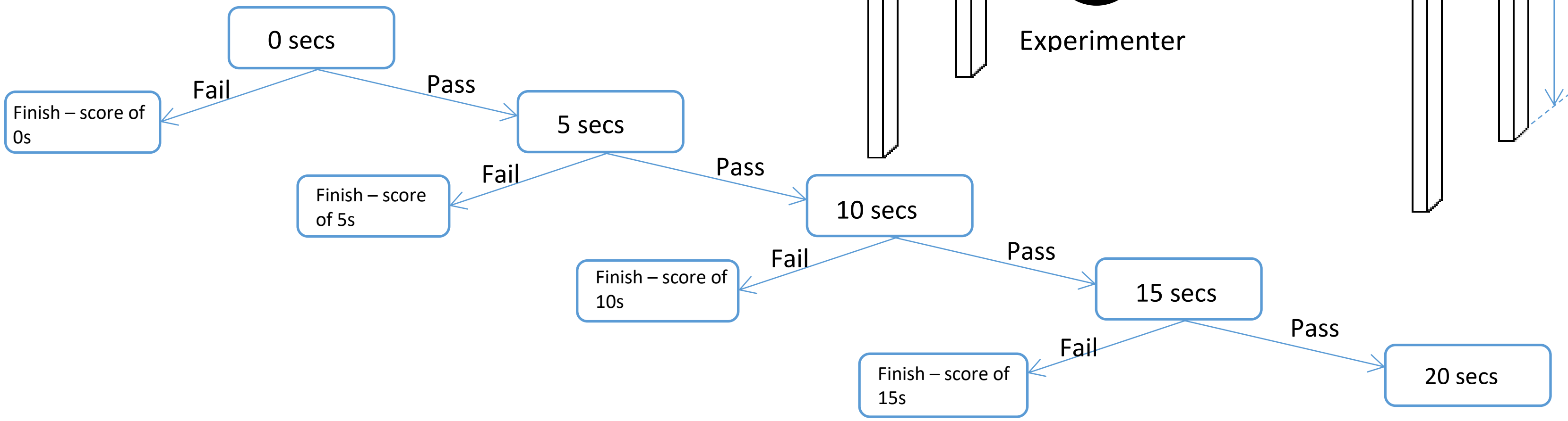
The variable selected a priori to best reflect the performance on this task was the A*B* error. The ordinal nature of this variable dictated the use of a Mann-Whitney U-test to explore any group differences, although the same statistical procedure stated in section 2.3 was followed to explore the data. The regression model fitted was an ordinal logistic regression with the A*B* error as the dependent variable and predictors included were consistent with the procedure stated in section 2.3.

Figure 4-2. Illustration of A not B apparatus including table dimensions; trial counterbalance order; and paradigm procedure table

Counterbalance order of trials	**0 sec**	**5 secs**	**10 secs**	**20 secs**
Counterbalance order 1	RRL	LLR	RRL	LLR
Counterbalance order 2	LLR	RRL	LLR	RRL



<u>Paradigm trial procedure:-</u>



4.1.3 Dimensional Change Card Sort Task

The 'Dimensional Change Card Sort Task' or DCCS (Zelazo, Frye and Rapus, 1996; Zelazo, 2006) was the EF task administered at 30 months of age with a predominant focus on cognitive flexibility. The standard procedure detailed by Zelazo was the protocol utilised for the task (Zelazo, 2006), with the stimuli produce by the Carlson lab, (Carlson, 2013).

The DCCS task comprised of two sorting boxes and a selection of sorting cards. The cards varied on two dimensions, with the sorting box displaying the same, yet inverse dimensions. The task required the child to sort according to each of the dimensions in turn. If both dimensions were correctly sorted, the cards were changed and the next level was administered. The complexity of the different dimensions increased with each level.

It has been observed that children of 30 months of age find it difficult to integrate two aspects of a picture that are not part of the same object, or, separate the colour of an object from its shape. Diamond and Kirkham (Diamond, Carlson and Beck, 2005) suggest this is not an inability to recognise the two features, rather seeing the same picture from two different perspectives and integrating this information is too challenging. When given pictures with only one discernible feature, 3 year olds can sort with ease. It is only when the second dimension is added that confusion arises.

The rules of the paradigm consistently switch, requiring the child to adapt their behaviour according to the rule changes. The more complicate the instruction or rule, the greater the working memory load. Success at this tasks is theorised to require all of the EF subdomains (Diamond, Carlson and Beck, 2005; Garon, Bryson and Smith, 2008). Although, fundamentally, the child needs to inhibit the secondary feature of the picture to sucessfully pass the trial, without the flexibility to adapt their behaviour or the ability to hold the rule in mind, failure is likely to occur (Diamond, Carlson and Beck, 2005).

Unlike three year olds who have been extensively investigated with the DCCS paradigm, research into EF performance is limited in the toddler age range (Garon, Bryson and Smith, 2008; Pozzetti *et al.*, 2014). The use of this task, although well established in older children, was experimental in the current investigation in regards to the toddler's capacity to understand what was required from them. Two additional measures of EF, the BabyScreen App and Multi-Location Multi-step paradigm were additionally included in the 30 month assessment battery in support for the use of the DCCS at this age. However, both had predominant processing speed measures and will be considered in the information processing chapter (Chapter 6).

4.1.3.1 DCCS Apparatus and Methodology

The apparatus consisted of two sorting boxes, with a selection of sorting and target cards. The apparatus was set up so that the two sorting boxes were placed between the experimenter and the child, within reaching distance of both. Each sorting box, with the dimensions of 9cm X 12cm X 16cm, had a target card placed on the front and back of the box. All sorting cards displayed the target image. This image differed in dimension and complexity with advancing conditions. The sorting cards were white on the back and laminated with the approximate dimensions of 8cm X 13cm.

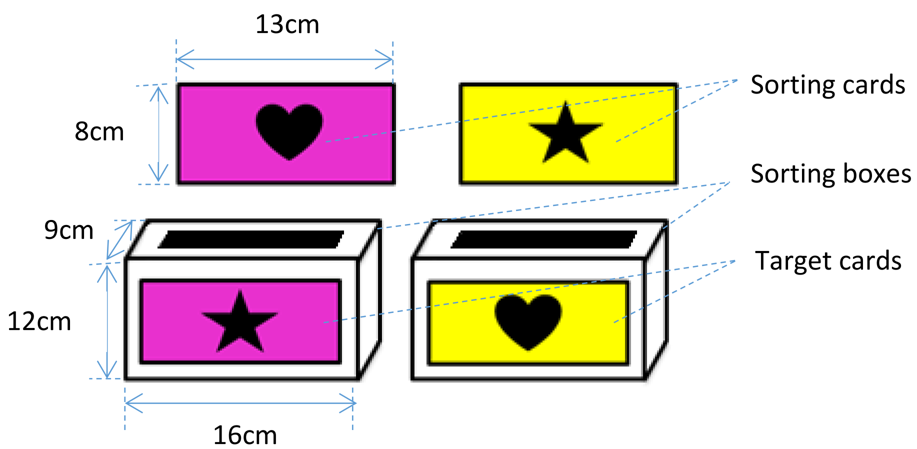


Figure 4-3. Dimensions and set up of the Dimensional Change Card Sort task sorting boxes, cards and target cards. The example shows level 3 of the DCCS task where the child is asked to sort first by colour, then by shape.

At 30 months, the task begun with the simplest condition. Condition 1 saw a target card of an elephant and fish placed on the two separate sorting boxes. The experimenter read from the standardised script written by Carlson *et al.,* (2013) in combination with gestures: 'we have these two boxes here, this box has a fish on it [gestured to the left box], this box has an elephant on it [gestured to the right box]. This is the fish game. In the fish game, all the fish go in the fish box, because that is where they belong. See here is a fish [held up fish demo sorting card], fish go here [placed the fish card into the fish sorting box]'. The child was then asked 'which box do the fish go in?'. They were given 2 opportunities to correctly identify the appropriate sorting box. Irrespective of the response on the second check, the trial continued.

Each condition comprised of 10 trials split into 2 sections, a and b. The sorting rule changed between part a and b, with the target cards remaining the same for both parts. For example, condition 1, the child was asked to sort the fish cards for part a and then was verbally instructed that the rule had changed, and they now needed to sort the elephants cards into the other sorting box. 4 or more cards needed to be correctly sorted in order to move on to the next section. If the child passed the 2 sections, the target cards were removed and the new condition was introduced.

The coding of the task required the recording of the total correct trials achieved and the highest level passed. All subjects had the same basal level, 1a due to their age and understanding. The main outcome measure for this task was the highest level completed. As noted, each condition comprised of parts a and b. If the child successfully passed part a, but failed to sort ≥4 cards on part b, the highest level passed was part a. For the purposes of the data analysis, each level was numbered incrementally, e.g. 1a and 1b was coded, level 1 and 2. An additional measure that was investigated within this paradigm was the total number of correctly sorted cards (total number correct trials).

Two sets of analyses were completed with this task. Primarily all participants to successfully complete the task were analysed, followed by a second analysis which took language abilities into consideration. Participants are excluded on the basis of the following criteria: score lower that 1SD below the standardised mean on the language composite score in the Bayley-III (<85) at 2 years, a proportional comprehension score <1SD below the cohort mean on the OCDI (<.76), or if no language score was available. A number of the term born children (11/25) did not complete the language scale of the Bayley-III at the 30 month assessment, typically due to fatigue during the assessment. Although there were no language concerns from the assessment team, the Oxford Communicative Development Inventory (OCDI) was completed by all parents. The prospective cohorts will be described in detail in section 4.2.3.

The variable selected a priori to best reflect the performance on this task was the highest level completed. Due to the ordinal nature this variable, although the same statistical procedure stated in section 2.3 was followed to explore the data, a Mann-Whitney U-test was the most appropriate for an outcome variable of this nature. The total number of trials completed was a secondary focus within the results and displayed a marginally positively skewed, with a Shapiro-wilk test result of 0.04. The data was transformed, but normality was not reached, therefore a Mann-Whitney U-test was performed.

4.1.4 Longitudinal exploration of EF analyses

The final set of results reported in the current chapter looks at the relationship between the EF task performances in the first year to the EF measure at 30 months. Due to the ordinal nature of the DCCS, an ordinal logistic regression model was fitted to the data including the previously defined demographic variables: study group, male sex, and IMD quintile, and included the primary outcome variables for the DRT and A*B* paradigm: total proportion correct and time to A*B* error respectively. This model was then repeated with the additional inclusion of the 30 month cognitive z-scores of the Bayley-III.

4.2 Results

The results of the three EF tasks are detailed below in age of administration order. For each task, a summary of the demographic information is initially provided as each task had a different sub population of the overall longitudinal cohort. Within the demographic information the Bayley-III cognitive scores provided are from the assessment nearest in age to the experimental task; for the DRT at 6 months and the A*B* task at 12 months, the 12 month Bayley-III score is provided; for the DCCS, the 30 month cognitive scores are provided. Unfortunately due to a number of factors, including infant temperament, incomplete datasets and missed appointments, not all participants completed all assessments at all ages. This explains the n number differences for the Bayley-III scores and the total number of children to complete each task as not all completed both.

4.2.1 DRT at 6 months age

During the 6 month assessment phase, 57 term born and 33 VP infants completed the DRT (Table 4-2) assessed on the ability to remember the location of the previous stimulus. During the paradigm, neither study group displayed a response above chance when ocular movements were used to identify the infant's response (Table 4-3; Figure 4-4). This was not confounded by the number of infants not looking towards the apparatus as there was no difference observed between groups in the number of trials where the infant was not looking (Table 4-3).

		Term (n=57)	VP (n=33)
Gestational age	*Median (range); weeks^{+d}*	40^{+3} (37^{+1} – 42^{+1})	26^{+2} (23^{+6} – 31^{+4})
Male sex		28 (49%)	22 (67%)
IMD Quintile	1	4	5
	2	10	8
	3	8	12
	4	21	9
	5	16	4
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=44; VP=25)	110.23 (11.26)	99.6 (8.89)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=44; VP=25)	.2 (.95)	-.69 (.75)

Table 4-2. Demographic details of infants included in Delayed Response Task analysis.

No effect of group was seen in the total proportion of correct trials (t(88) = -.14, p = .89; see Figure 4-4); nor was there an effect of group on proportion of not-looking trials (z(88) = 1.41, p = .158). No correlation was observed between the total number correct and the 12 month cognitive z-scores (r = -.16, p = .17).

Variable	*Term, n=57*	*Preterm, n=33*
Mean total proportion correct (SD)	.46 (.12)	.47 (.13)
Mean number of trials not-looking (IQR):	1.72 (2.34)	0.88 (1.16)

Table 4-3. Mean proportion of correct trials in Delayed Response Task collapsed across the social and non-social conditions (total correct looks/total trials completed). Mean number of 'trials not looking' during DRT where the infant failed to look towards the equipment during the response window.

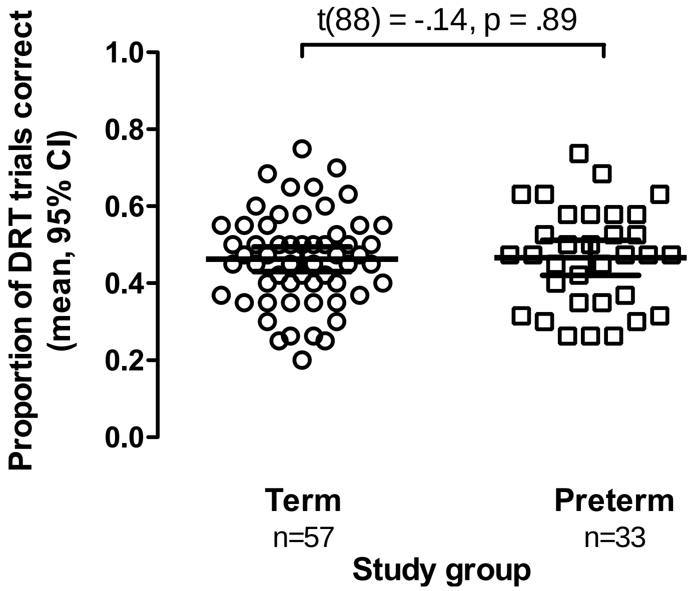
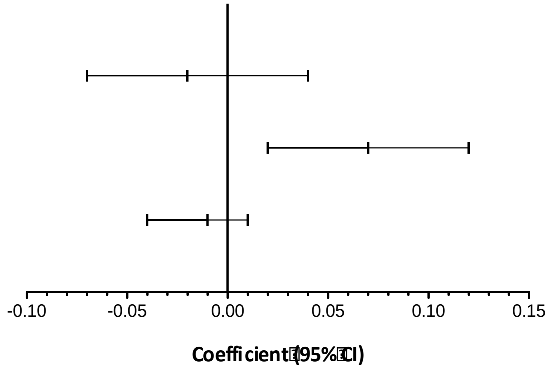


Figure 4-4. Total proportion of trials correct in Delayed Response Task by study group. A t-test was used to compare the proportions in each group that correctly identified the location of the previously observed stimulus.

Table 4-4, presents the results of the linear regression model for the total proportion correct as the primary outcome measure. Following regression, no differences remained between the VP and term groups but better test performance was independently associated with male sex. The model only accounted for 10% of the variance in the outcome.

Table 4-4. Linear regression model looking at the proportion of correct looks in DRT ($F(3, 85) = 3.19$, $p = .03$). SES for one term-born infant missing, therefore this infants data was excluded from the model. The constant for the model was coded as term-born females with an IMD quintile of 1.

Predictor	Term (n=56)	Preterm (n=33)	Coef.	95%CI	P
	Median (range)	Median (range)			
Study Group	40^{+2} (37^{+1} – 42^{+1})	26^{+5} (23^{+6} – 31^{+4})	-.02	-.07 – .04	.50
Male sex	27 (48.2%)	22 (66.7%)	.07	.02 – .12	.01
IMD Quintile	4 (1-5)	3 (1-5)	-.01	-.04 – .01	.20
Const	-	-	.46	.39-.54	.000



Overall model fit R^2 = 0.10

4.2.2 A*B* task at 12m age

During the 12 month assessment phase, 43 term born children and 36 VP children completed the A*B* task. Performance was defined by the number of seconds delay before the A*B* error occurred, also termed 'time to A*B* error'. Both study groups tolerated the same length of delay.

		Term (n=43)	***VP (n=36)***
Gestational age	*Median (range); weeks^{+d}*	40^{+1} (37^{+1} – 42^{+0})	26^{+2} (23^{+4} – 31^{+4})
Male sex		21 (48.84%)	24 (66.67%)
IMD Quintile	1	6	5
	2	7	7
	3	5	10
	4	15	9
	5	10	5
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=43; VP=29)	107.79 (12.64)	98.79 (9.03)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=43; VP=29)	0 (1)	-.76 (.76)

Table 4-5. Demographic details of infants included in A-not-B paradigm analysis.

Seven of the 43 term infants and 7 of the 36 VP infants did not to reach the task criterion of 2 correct retrievals of the toy at the start of the paradigm (z = .48; p = .63); the infants that did not pass this initial criterion therefore did not continue through the rest of the paradigm (Table 4-6).

Having passed the initial criterion, 6 term born and 3 VP infants did not complete the task and were excluded from the analyses as the paradigm was terminated early, leaving 30 term and 26 VP infants who displayed an A*B* error (Table 4-6). Overall, no effect of group was seen in the number of seconds to AB error after excluding those that did not meet the task criterion (z(68) = 1.41, p = .158; see Figure 4-5). A positive correlation was observed with the 12m cognitive z-scores (r =

.27, p = .05), which when further investigated, was driven by the term group (T: r = .41, p = .02; PT: r = -.03, p = .91).

Time to AB error	*Term (n=43)*	*Preterm (n=36)*
0s criterion not achieved	7/43	7/36
0s	12/36	12/29
Not reached criterion	0	0
5s	11/20	7/16
Not reached criterion	4	1
10s	6/8	6/8
Not reached criterion	1	1
15s	0/2	1/1
Not reached criterion	1	1
20s	1/2	0
Not reached criterion	0	0
>20 seconds but AB error not achieved.	1	0

Table 4-6. Proportion of each study group to display the A*B* errors at each time delay; including the proportion of infants to not achieve the criterion in the first instance.

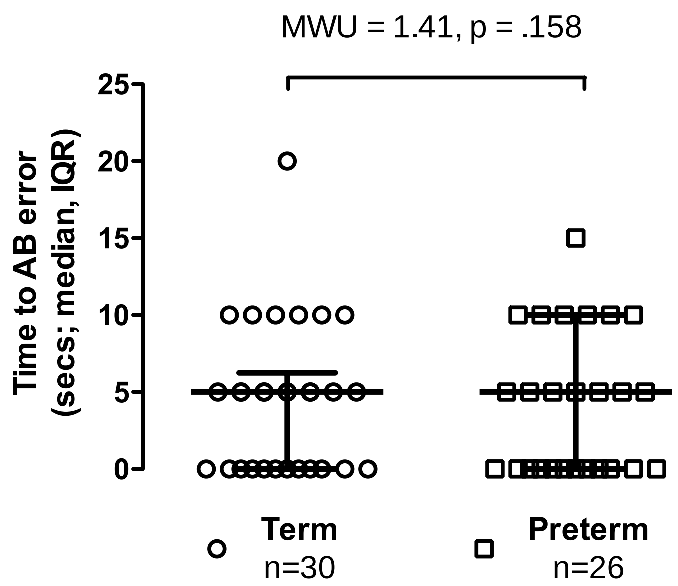


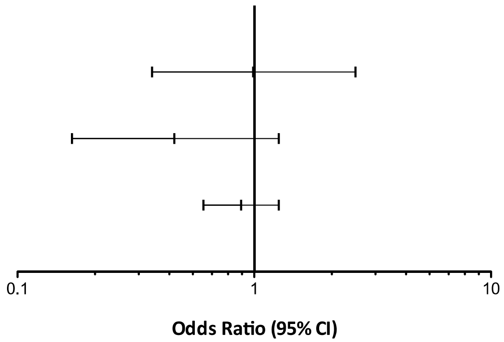
Figure 4-5. Time to A-not-B error in seconds delay across study groups.

Table 4-7 presents the results of an ordinal logistic regression model with time of A*B* error as the primary outcome measure reported as an odds ratio. Given the positive correlation in the term group to the Bayley-III cognitive z-score at 12 months, an additional model was fitted to investigate this relationship. However the original result was unaffected by this adjustment and the second model has not been reported.

The regression model in Table 4-7 shows no predictive effects of group, male sex or IMD quintile on the performance in the AB paradigm at 12 months of age. The model was not considered a good fit of the data as it only accounted for 2% of the variance.

Table 4-7. Ordinal logistic regression modelling the A*B* error as the main outcome of the A*B* paradigm; reporting odds ratio with 95% confidence intervals (LR chi2 =2.48, p =.48). The baseline for the model was coded as term-born females with an IMD quintile of 1.

Predictor	Term (n=30)	Preterm (n=26)	OR	95%CI	P
	Median (range)	Median (range)			
Study Group	39^{+6} (37^{+1} – 42^{+0})	26^{+4} (23^{+4} – 31^{+4})	.99	.37 – 2.68	.99
Male sex	13 (43.33%)	16 (61.54%)	.46	.17 – 1.27	.13
IMD Quintile	4 (1-5)	3 (1-5)	.88	.61 – 1.27	.49



Overall model fit pseudo R^2 = 0.02

4.2.3 DCCS at 30 months age

The Dimensional Change Card Sort task saw 25 term-born and 25 VP toddlers complete the paradigm at 30 months of age. The term born children completed a greater number of levels during the DCCS paradigm compared to VP toddlers and the term toddlers successfully passed more trials than the VP toddlers. Language was not seen to affect the performance of either group.

		Term (n=25)	*VP (n=25)*
Gestational age	*Median (range); weeks^{+d}*	40^{+2} (37^{+0} – 42^{+1})	26^{+2} (23^{+4} – 31^{+4})
Male sex		11 (44)	18 (72)
IMD Quintile	1	3	5
	2	4	1
	3	4	9
	4	9	6
	5	5	4
Bayley-III cognitive composite score at 30m	Mean (SD) (n:T=23; VP=23)	108.26 (12.40)	101.09 (10.55)
Bayley-III Cognitive z-score at 2 years	Mean (SD) (n:T=23; VP=23)	0 (1)	-.53 (.85)
Bayley-III language composite score at 2 years	Mean (SD) (n:T=14; VP=23)	119.43 (10.41)	97.95 (18.07)
Proportional comprehension score on OCDI	Mean (SD) (n:T=17; VP=22)	.93 (.17)	.93 (.17)
Number of toddlers with no language score:		4	1

Table 4-8. Demographic details of all infants to complete the DCCS paradigm.

The first set of analyses included all participants; the second removed infants with possible indications of language delays.

In the full cohort, a significant difference was observed between the term and VP toddlers in ‘highest level completed’, with the term-born children scoring higher as a cohort compared to the VP children (z(48) = 3.27, p = .001; see Figure 4-6). Within each level a variation was observed in the total number of trials correct; the median number of trials correct for term children was 17 and for VP children was 13.

Highest level achieved	*Term (n=25)*	*Preterm (n=25)*
0 levels passed	0	1
1 (level 1A: fishes)	0	1
2 (level 1B: elephants)	3	11
3 (level 2A)	12	10
4 (level 2B)	4	0
5 (level 3A)	6	1
6 (level 3B)	0	0
7 (level 4A)	0	1
8 (level 4B)	0	0
	Mann Whitney U: 3.27; p = .001	
Median total number of trials completed (range)	17 (11-28)	13 (1-34)
	Mann Whitney U: 2.77; p = .006	

Table 4-9. Highest level passed and total number of correct trials of all participants to complete the DCCS paradigm.

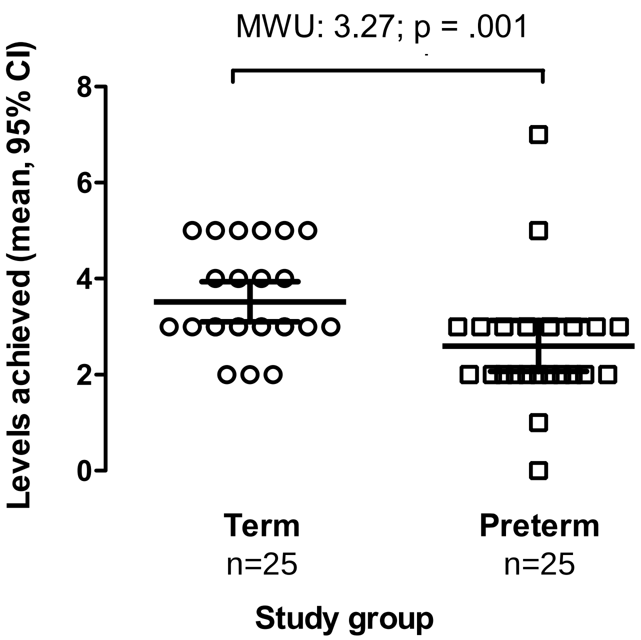


Figure 4-6. Highest level achieved on the Dimensional Change Card Sort Task by study group. MWU: Mann Whitney U Statistic

		Term (n=19)	***VP (n=17)***
Gestational age	*Median (range); weeks^{+d}*	40^{+2} (38^{+4} – 42^{+1})	26^{+3} (23^{+4} – 31^{+4})
Male sex		8 (42.11%)	13 (76.47%)
IMD Quintile	1	3	5
	2	2	1
	3	3	4
	4	6	4
	5	5	3
Bayley-III cognitive composite score at 2 years	Mean (SD) (n:T=17; VP=16)	111.18 (13.05)	104.38 (9.81)
Bayley-III Cognitive z-score at 30m	Mean (SD) (n:T=17; VP=16)	.28 (1.05)	-.27 (.79)
Bayley-III language composite score at 2 years	Mean (SD) (n:T=14; VP=16)	119.43 (10.41)	107.13 (12.71)
Proportional comprehension score on OCDI	Mean (SD) (n:T=23; VP=23)	.98 (.02)	.96 (.07)

Table 4-10. Demographic details of infants included in secondary DCCS paradigm analysis excluding infants with low language scores.

The second set of analyses excluded any children with a low or absent language score (as categorised in section 4.1.3.1). Six children within the term group and 8 children within the VP group were identified, all of which were excluded from the subsequent analyses; the remaining cohort summarised in Table 4-10.

Highest level achieved	*Term (n=25)*	*Preterm (n=25)*
0 levels passed	0	0
1 (level 1A: fishes)	0	1
2 (level 1B: elephants)	1	5
3 (level 2A)	8	9
4 (level 2B)	4	0
5 (level 3A)	6	1
6 (level 3B)	0	0
7 (level 4A)	0	1
8 (level 4B)	0	0
	Mann Whitney U: 2.67, p = .008	
Median total number of trials completed (range)	20 (11-28)	15 (7-34)
	Mann Whitney U: 2.41, p = .02	

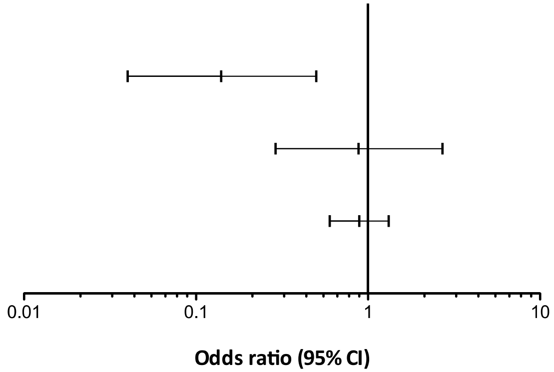
Table 4-11. Highest level passed and total number of correct trials to during the DCCS paradigm of participants within the sub-cohort; excluding those with low language scores.

Term infants continued to perform to a higher level on the DCCS than that of the VP infants after excluding those with low language score (z(36) = 2.67, p = .008; Table 4-11). The difference between the total number of trials still remained, with term children successfully passing a median of 20 trials and VP children a median of 15, in those with good language scores (z(36) = 2.41, p = .02; Table 4-11).

In an ordinal logistic regression against highest level completed, VP infants performed less well after allowing for the effect of male sex and IMD quintile (Table 4-12). A further analysis including the 2 year Bayley cognitive z-scores, revealed that both DCCS level and Bayley-III cognitive z-score had independent effects in the model; a better performance on the DCCS was associated with higher cognitive scores on the Bayley-III, but the effect size was small, .13 (Table 4-13). Both models included the full cohort to complete the DCCS paradigm as low language abilities did not appear to significantly impact the differences seen in study group performances.

Table 4-12. Ordinal logistic regression (LR chi2 = 12.42, p = .006), with highest level achieved defined as the primary outcome for the DCCS; reporting odds ratio with 95% confidence intervals. The baseline for the model was coded as term-born females with an IMD quintile of 1.

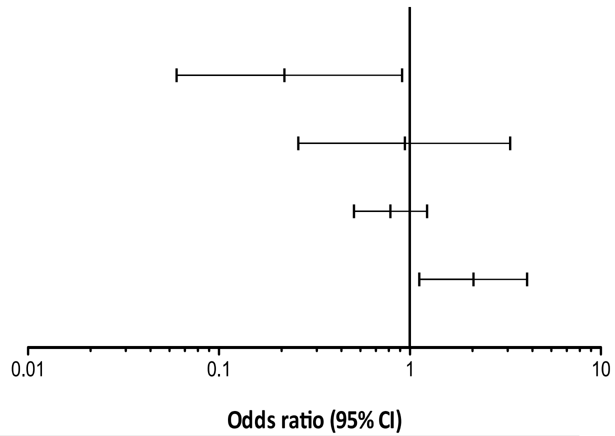
Predictor	Term (n=25) Median (range)	Preterm (n=25) Median (range)	OR	95%CI	P
Study group	40 (41^{+3} – 38^{+4})	26+3 (28^{+3} – 24^{+3})	.14	.04 – .50	.002
Male sex	11 (44%)	18 (72%)	.88	.29 – 2.71	.82
IMD quintile	4(2)	3(1)	.89	.60 – 1.32	.57



Overall model fit was Pseudo R^2 = 0.09

Table 4-13. Ordinal logistic regression (LR chi2 = 16.15, p = .003) with highest level achieved during DCCS as the outcome and including the additional adjustment of the 2 years z-score; reporting odds ratio with 95% confidence intervals. The baseline for the model was coded as term-born females with an IMD quintile of 1.

Predictor	Term (n=23)	Preterm (n=23)	OR	95%CI	P
	Median (range)	Median (range)			
Study group	40^{+1} (41^{+4} – 38^{+3})	26^{+2} (28^{+0} – 24^{+3})	.22	.06 – .91	.04
Male sex	11 (44%)	18 (72%)	.94	.26 –3.35	.92
IMD quintile	4(2)	3(1)	.79	.51 – 1.23	.30
2 year Cog Z-score	.04 (1.00)	-.53 (.85)	2.15	1.12 – 4.11	.02



Overall model fit Pseudo R^2 = 0.13

4.2.4 Longitudinal exploration of EF performance over first and second year assessment

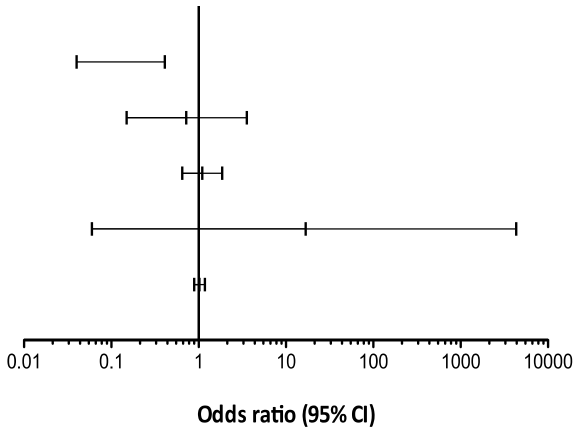
Exploration of the longitudinal relationship of the data collected across the 3 EF tasks included the responses from 14 term children and 17 VP children. With performance on the DCCS at 30 months set as the dependent outcome variable, there was no predictive effect of DRT or A*B* task performance after allowing for the effect of study group, male sex and IMD quintile (Table 4-15). VP infants continue to perform less well on the DCCS however, when adjusting for the additional EF measures, the model accounts for a greater proportion of the variation within the DCCS outcome. A further analysis with the addition of the 2 year Bayley-III cognitive score again saw an independent effect of study group within the model, unaffected by the additional EF predictors. This model accounts for even greater variance in the DCCS outcome at 24% (Table 4-16).

		Term (n=14)	***VP (n=17)***
Gestational age	*Median (range); weeks^{+d}*	40^{+1} (38^{+4} – 42^{+0})	26^{+3} (24^{+0} – 31^{+4})
Male sex		6 (42.86%)	13 (76.47%)
IMD Quintile	1	2	4
	2	4	1
	3	2	4
	4	4	5
	5	2	3
Bayley-III cognitive composite score at 2 years	Mean (SD) (n:T=12; VP=16)	107.5 (12.15)	102.5 (11.11)
Bayley-III Cognitive z-score at 2 years	Mean (SD) (n:T=12; VP=16)	-.02 (.98)	-.42 (.89)

Table 4-14. Demographic details of all infants to complete all EF paradigms and included in longitudinal analysis.

Table 4-15. Ordinal logistic regression (LR chi2 = 14.64, p = .01) with highest level achieved during DCCS as the dependent variable and including the additional adjustment of the total proportion correct in DRT and time to A*B* error in A*B* task; reporting odds ratio with 95% confidence intervals. The baseline for the model was coded as term-born females with an IMD quintile of 1.

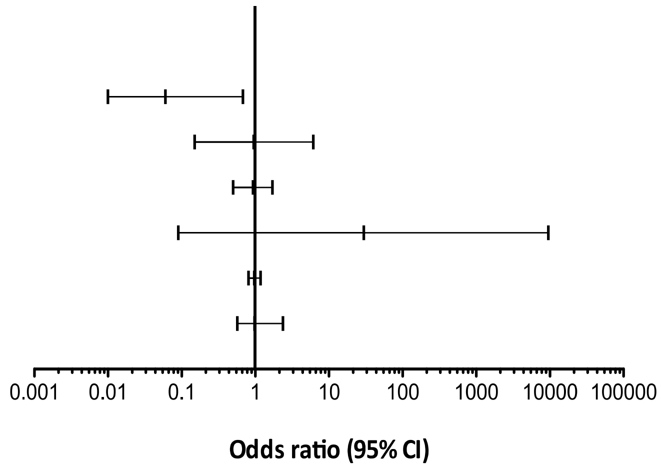
Predictor	Term (n=14)	Preterm (n=17)	OR	95%CI	P
	Median (range)	Median (range)			
Study group	40^{+1} (38^{+4} – 42^{+0})	26^{+3} (24^{+0} – 31^{+4})	.04	.00 –.41	.007
Male sex	6 (42.86%)	13 (76.47%)	.72	.15 – 3.56	.69
IMD quintile	3 (1-5)	3 (1-5)	1.10	.65 – 1.86	.71
DRT	-	-	16.78	.06 –4340.95	.32
A*B*	-	-	1.02	.89 –1.18	.74



Overall model fit Pseudo R^2 = 0.19

Table 4-16. Ordinal logistic regression (LR chi2 = 15.69, p = .02) with highest level achieved during DCCS as the dependent variable, and including the additional adjustment of the total proportion correct in DRT, time to A*B* error in A*B* task, and 2 years z-score; reporting odds ratio with 95% confidence intervals. The baseline for the model was coded as term-born females with an IMD quintile of 1.

Predictor	Term (n=12) Median (range)	Preterm (n=16) Median (range)	OR	95%CI	P
Study group	40^{+4} (38^{+5} – 42^{+0})	26^{+3} (24^{+0} –29^{+4})	.06	.01 –.68	.02
Male sex	5 (41.67%)	13 (81.25%)	.95	.15 – 6.15	.96
IMD Quintile	3 (1-5)	3 (1-5)	.93	.50 – 1.71	.82
DRT	-	-	29.73	.09 – 9542.32	.25
A*B*	-	-	.97	.81 – 1.18	.79
2 year Cog Z-score	-	-	2.37	.98 –.57	.06



Overall model fit Pseudo R^2 = 0.24

4.3 Discussion

The development of EF abilities across the first two years after birth has not been extensively studied in ex-preterm populations. The aim of the study was therefore to explore the advancement of EF abilities across 3 time points up to the age of 30 months in a cohort of very preterm children. Established EF paradigms were utilised to ensure the abilities assessed depicted classically defined EFs as predetermined by a general consensus in the developmental literature. The use of such paradigms allows for comparisons to be made to previous observations both from typically developing populations and cross-sectional reports of ex-preterm children to fully evaluate the current findings. The paradigms selected were the delayed-response task (Schwartz and Reznick, 1999; Noland *et al.*, 2010), the Piagetian A-not-B task (Piaget, 1954) and the Dimensional Change Card Sort task (Zelazo, 2006).

With the use of the DRT and A-not-B paradigms, we were unable to demonstrate differences in performance between the term and VP infants at 6 or 12 months of age. By 30 months, we have demonstrated differences in achievement using the DCCS paradigm. These differences were only partially explained by performance in the Bayley-III cognitive scale. From the measures utilised in the current investigation, there were no predictive relationships observed within the measures from the first year to that of the DCCS outcome at 30 months. Although there is rationale for a relationship between the abilities used within these paradigms, none are direct replicates of tasks from previous time points therefore the absence of relationship is not unexpected.

Both the DRT and A*B* tasks are delayed-response-type paradigms, where the infant has to wait to respond, assessing what is thought to reflect the emergence of working memory abilities, coupled with inhibitory control. Before 8 months of age, it is thought that 'inhibitory control' refers to inhibition of reflex behaviours rather than prepotent responses, which emerge with the maturation of the motor cortex (Diamond, 1991). It is the utilisation of this latter inhibitory behaviour that is crucial for success in the A*B* paradigm. By including both paradigms, both the reflex and motor-related inhibitory behaviours were assessed in relation to working memory. In this instance, there was no effect of prematurity seen in either set of results.

The DRT paradigm required the infants to remember the location of either a social or non-social stimulus presented to them 3 seconds prior. This paradigm was unsuccessful at measuring working memory within this cohort, as neither group appeared to be performing above chance for the main outcome variable 'proportion of trials correct'. This task has been reported to measure working memory capacity at 6 months of age (Reznick *et al.*, 2004). However, Diamond has speculated that the paradigm may be more effective at later ages (7-8 months) and different tasks may be required to access working memory at earlier ages, if it is possible to do so (Diamond, 1990).

When considering the model of human brain development, it is possible 6 months is too early to detect working memory performance. The frontal cortex has a more extensive developmental period than that of the smaller structures in the brain, in order to encompass the experiences during the postnatal period, thereby shaping and structuring connections during development (Johnson, 2001). Around 6-8 months the cortex appears to go through large structural and metabolic changes, with an increase in glucose metabolism and reduction in synaptic density (Chugani, Phelps and Mazziotta, 1987; Paterson *et al.*, 2006; Sun and Buys, 2011). This coincides with the timescale of EF emergence (Pelphrey and Reznick, 2003), at which 6 months marks the start of these changes. It is possible it is not until 7 to 8 months that these abilities can be reliably investigated. This is supported by the numerous studies that appear to more consistently report successful working memory assessment at 7-8 months (Diamond, 1990; Schwartz and Reznick, 1999).

Evidence from both animal and human infant studies suggest DR tasks are processed by the prefrontal cortex (Diamond and Doar, 1989). Adverse experiences within first year of life have been observed to correlate with abnormal brain development, particularly in the frontal lobe due to its protracted developmental period over the first year (Espy *et al.*, 2002). The multitude of perinatal risk factors associated with preterm birth have been speculated to cause altered connections within the frontal lobe, potentially resulting in cognitive impairments later in life (Diamond, 2006; Sun and Buys, 2012a). However, the evidence to support the disruption that is responsible for later difficulties is limited, as classified neurological injury does not occur in all preterm cases, yet cognitive impairments are still observed later in life (Vollmer *et al.*, 2017). If to view the structural changes that are occurring in the brain within the first year from a neuroconstructivist approach, it could be argued that an abnormal developmental trajectory is responsible for the delays observed later in life, therefore at 6 months structural changes are still occurring and any mild abnormalities are yet to functionally present themselves (Oliver *et al.*, 2000). On the contrary, if substantial metabolic and structural changes occur in the brain between the ages of 6-8 months as discussed above (Chugani, Phelps and Mazziotta, 1987; Paterson *et al.*, 2006; Sun and Buys, 2011), one might expect functional differences be start to emerging by 12 months of age. Differences may be present but were not detectable within the current investigation.

An alternative explanation for the null result in the DRT is that the delay imposed during the current study may have been too challenging for infants of this age. Previous studies have used delays of 1-2 seconds (Reznick *et al.*, 2004). Although the investigation was on information processing speeds, Rose *et al*. (2002) explored preterm infant performance on a familiarisation task, with preterm infants significantly slower than term born infants on this task. This was in contrast to a previous study, where the infants were assessed on a visual expectation paradigm and no study group differences were observed (Rose, Feldman and Jankowski, 2001). The authors postulated that this was due to the type of information being measured. The first study assessed the simple detection of stimulus and motor response; the second required the encoding of the stimulus to compare it to the next and was therefore considered a higher order cognitive task for an infant within the first year. The absence of group difference in the DRT task could be viewed similarly. It could be speculated that the motivation necessary for the updating of working memory systems with location of the stimulus may not be present within this paradigm, and eye-gaze to the stimulus location may not be stimulating any active working memory processes. In combination with the greater delay, this could explain the absence of responses to this paradigm. Further exploration of the DRT at this age with shorter time delays would be required to advance our understandings of working memory abilities at 6 months of age.

The AB task appeared successful: 80% of the VP infants and 83% of term infants displayed some level of working memory and inhibitory abilities at 12 months of age. As discussed, the AB paradigm may reflect early working memory capabilities and inhibitory abilities to prepotent responses. In older cohorts, working memory has been identified as a specific area of difficulties in preterm cohorts (Mulder, Pitchford and Marlow, 2010). In the current investigation, performance was similar in both study groups. This absence of performance differences in a working memory paradigm at 12 months is not a first for a cohort of preterm infants (Wilcox, Nadel and Rosser, 1996). As summarised by Jongbloed-Pereboom *et al*., (2012) results on A*B* paradigms and related delayed response tasks have produced very mixed findings across the preterm literature. A study by Matthews *et al*., (1996), for example, observed age-corrected preterm infants to outperformed term born infants on an A*B* paradigm when exploring the maturation of performance from 7 to 15 months of age; on the contrary Ross *et al*., (1992) found preterm infants to be significantly less successful of the A*B* paradigm at 10 months of age. These differences across studies are likely to reflect the variety amongst cohort demographics and differences in assessment parameters, making comparisons across study findings challenging.

It is often thought that the predominant domain assessed in this paradigm is that of working memory over inhibitory behaviours due to the reduced performance with increased time delay. However, obtaining the toy following the reaching behaviour reinforces the associated motor response and therefore inhibitory processes are almost certainly required to ensure this is not utilised after the hiding location is changed (Sun and Buys, 2012a). A study by Sun and Buys (2011) found preterm infants to display a poorer performance on the A*B* task compared to term born infants (Sun and Buys, 2011). An important characteristic of the current cohort is the absence of any scores in the clinical range of the Bayley-III at 12 months of age, with only one infant in the mild range (<85). This could be considered a high performance on the Bayley-III for a preterm cohort, suggesting the infants are performing at an appropriate level at 12 months. However, as discussed previously, given the range of perinatal complications observed in this cohort, which have been indicative of poorer outcome in other studies (Aylward, 2002; Taylor and Clark, 2016), it could be argued lower cognitive scores might be have been expected. Given the similar performance on the A*B* task however, these results could be a good representation of the cohorts' performance.

On the contrary, differences are observed both in the DCCS and in the 2 year Bayley-III cognitive scale when comparing term and VP performances. Speculation could therefore be made regarding the first year EF and Bayley-III measures. It is plausible the first year measures are not sensitive enough to later developmental abilities, potentially due to the restrictive nature of the assessments, or due to the developmental trajectory these abilities take within the second year of life. These results could be suggestive of domain differentiation beginning to occur and signs of delay only starting to emerge as the functions develop.

At 30 months, term born toddlers completed the DCCS task to a higher level compared to the VP cohort and successfully completed more trials throughout the assessment. This result was still apparent after adjusting for the cognitive z-score in the Bayley-III at 2 years. As discussed in Chapter 3, at this age, the cognitive scale on the Bayley-III has been shown to reflect cognitive performance later in life (Bode *et al.*, 2014; Spencer-Smith *et al.*, 2015). However, as the difference between term and preterm performance on the DCCS remained following the correction for global score, it could be postulated that the Bayley-III is not sensitive enough in its assessment to detect this variation in EF performance.

Very few studies have investigated the influence of preterm birth on EF at 2.5 years of age, and there are very few data available on the performance of a preterm cohort on the DCCS task. A recent study by Duvall *et al.,* (2015) explored the performance of EFs in relation to medical characteristics, using a modified version of the DCCS in infants born preterm at 3 to 4 years of age. They found a greater percentage of the typically developing infants passed their version of the DCCS but no effect of study group as a predictor of performance. This group also performed additional tasks targeting purer inhibitory skills, such as the bear-dragon paradigm (Carlson, 2005). It was in the additional tasks where an effect of study group was observed. The authors argued the absence of gestational age effect on the DCCS results was due to this task assessing more complex brain networks. An effect of gestation observed in a purer measure of inhibition, indicated gestational age was not likely to be sensitive enough to the specific deficits in the DCCS task (Duvall *et al.*, 2015). These data may be viewed as supporting the Barkley hypothesis of inhibition as a fundamental problem underpinning EF dysfunction (Barkley, 1997). Although it is not possible to identify whether one EF domain is fundamentally affecting the outcome on the DCCS, it is clear that study group is having a significant impact in the performance on this task at 30m of age in our current observations.

One criticism might be that infants do not have the language comprehension for the DCCS paradigm at 30 months. Although the designers of the DCCS materials (Diamond, Carlson and Beck, 2005) advise a starting age of 2 years is possible, if there are developmental delays, particularly with concerns in language development as commonly reported in the preterm population (Ramon-Casas *et al.*, 2013; Vohr, 2014), performance could be influenced by these factors over and above the EF task demands. However, the difference observed between the two cohorts remained after excluding those with low language scores at 2 years of age, with term toddlers achieving a higher level during the paradigm and completing a greater number of trials successfully compared to the VP toddlers. It is clear that the overall success at this task was not language dependent within this cohort.

Overall, the EF measures obtained during the first year of the PDP study were not indicative of EF performance at 30 months of age. Poorer EF performances in the VP group are first evident during the second year, with these differences in performance across the two study groups only partially explained by the Bayley-III cognitive scores. These results could be reflective of the developmental trajectory of the EF sub-domains, and differentiation may begin around 2 years of age. Difficulties within the preterm literature in identifying specific domain differences in the first year could be due to undifferentiated structure of EF. These results are consistent with previous research that suggests the problems seen in this population are not easily detectable until later in childhood.

Chapter 5 Attention

Orienting our attention is a crucial part of how we start to learn and take in information from our surroundings. This system emerges and undergoes significant changes within the first year of life (Elsabbagh *et al.*, 2009). Understanding how this develops in the typical population has been crucial to our comprehension of the adult attentional systems and its association to other higher cognitive processes (Amso and Scerif, 2015). Success in achieving goal directed behaviour in later life relies on disengagement from distraction and to be able to flexibly switch focus, making attention orientation vital to more complex EFs (Sun and Buys, 2012a). Some adopt the opinion that attention is the predominant skill for executive control as detailed in the model by Anderson (2002).

Clinical inattention is amongst the most commonly reported deficit following premature birth. Some ex-preterm children are diagnosed with attention deficit hyperactivity disorder (ADHD), however, it is typical for ex-preterm children to not display the impulsive behaviour and hyperactive movement and talking elements that manifest within this disorder (Johnson, 2007). This has led to speculation that this population may display a purer form of attentional deficit (Wolke, 1998). Nevertheless, the difficulties in attentional abilities have been frequently related to the later academic difficulties reported in this population (Jaekel, 2013). This gives rise to the need to further explore the emergence of these difficulties to primarily obtain a more comprehensive view on the developmental trajectory of attentional abilities, with the subsequent drive to investigate whether interventions would be beneficial to long term academic attainment.

The clinical definition of 'attention deficit' and deficits in what research classifies as attentional networks may not completely overlap. Clinical attentional difficulties related to behavioural manifestations have been associated to working memory difficulties or problems with inhibition that lead to the poor attentional focus (Bohm, Smedler and Forssberg, 2004). However in research terms, attention refers to the network of cognitive process that appear to emerge throughout childhood and when fully developed are classified into sustained, selective and executive attentional networks (Steele *et al.*, 2012). Although there are different definitions for clinical inattention and attentional network difficulties, it is likely both are interrelated. For example, attentional networks have been associated to EF task performances and thereby by could be considered to be related to behavioural responses (Kane and Engle, 2003; Cowan, 2011). Given this relationship between the attentional networks and EF in the current context, it is important to understand whether the poorer performance observed in the EF chapter are related to the cognitive attentional networks.

Difficulties have been reported across the different attentional networks in children born preterm and lower global cognitive abilities do not account for the differences observe in all cases (N Marlow *et al.*, 2007; Mulder *et al.*, 2009). A meta-analysis by Mulder *et* al., (2009) summarised the selective and sustained attentional performances in preterm cohorts over 2 years. Selective attention, where specific information is selected from the environment whilst ignoring other distractions, is often investigated with this use of visual attention paradigms. In the preterm literature these abilities appear to be less well developed and associated to gestational age at birth in preterm populations (Mulder *et al.*, 2009). The ability to maintain focus on a task at hand, or sustained attention, is often assessed by continuous performance tasks in older children, and again, the reports suggest a poorer performance correlated with the age of gestation at birth (Mulder *et al.*, 2009). In infants, the predominant network is considered to be that of selective or orienting attention (van de Weijer-Bergsma, Wijnroks and Jongmans, 2008).

Before the age of 3 months, infants often focus on stimuli in the centre of their visual field and struggle to disengage. It is not until half way through the first year that shifting attention is observed (Butcher, Kalverboer and Geuze, 2000; Colombo, 2001). The development of the ability to disengage and the speed of disengagement has been a large focus within the literature. There is speculation that the inability to disengage could be indicative of later attentional deficits (Elsabbagh *et al.*, 2009; Wass, Porayska-Pomsta and Johnson, 2011; Hitzert *et al.*, 2014; Green *et al.*, 2015).

This is an area typically reported within preterm populations later in life (Taylor, Klein and Hack, 2000; Bhutta *et al.*, 2002; Aarnoudse-Moens, Smidts, *et al.*, 2009). If difficulties can be detected within the first year of life, this could help predict later impairments.

Orienting to stimuli can be observed in new-born infants, however, it is not until around 4 to 6 months, that infants start to display higher-order voluntary attention orienting systems (Mulder *et al.*, 2009; Reynolds and Romano, 2016), such as disengagement and increase speed of shifting to other stimuli of interest (Amso and Scerif, 2015). It is these processes that are targeted and assessed in the 'Gap-Overlap task' used in the PDP battery.

The Gap-overlap paradigm (gap task) is frequently utilised amongst investigations of children with autism. The gap task measures the time to disengage from a centrally presented stimulus in order to re-direct gaze to one peripherally located. Being able to orientate to surrounding stimuli and shift flexibly between areas of visual focus have been linked to social communication abilities; an area that is typically compromised in children on the autistic spectrum (Bryson *et al.*, 2004; Elsabbagh *et al.*, 2009). The speculation within the autism literature is that the social difficulties are in part due to deficits in visual attention (van der Geest *et al.*, 2001). In support of this, impairments have been reported in the disengagement responses to the gap task in this population (Wainwright-Sharp and Bryson, 1993), and in populations at high risk for autism (Elsabbagh *et al.*, 2009). In these investigations, impairments in disengagement manifest as slower response times.

There is evidence to suggest that ex-preterm populations are more prone to social interaction difficulties than term born peers and often show signs of introversion and neuroticism (Johnson and Marlow, 2016). These children are also reported to be at a higher risk for ASD (Johnson and Marlow, 2016). Given this, it is could be postulated that the findings reported in studies of children with ASD could be similar to those of the PDP cohort in response to the Gap-overlap paradigm. To date, the author is unaware of any preterm research to utilise the Gap task to investigate visual attention development over the first two years after birth.

Within this chapter, data from Gap-overlap paradigm collected across 3 time points of the PDP are explored. The data are investigated both cross-sectionally and longitudinally to fully explore the differences and changes of attentional disengagement behaviours of the infants involved.

5.1 Methods

The GAP task utilised eye-tracking technology to record speed of eye saccades to stimuli presented on a visual display unit (VDU). The infant was presented with a stimulus in the centre of the screen in the first instance; a second stimulus was then presented on the periphery of the screen, to which the infant was required to look. Stimuli were presented in a range of formats in attempts to assess visual attention, specifically the speed of disengagement of attention and voluntary visual attention shifting. For example, central stimulus either disappeared before the presentation of the secondary stimulus, or it remained on screen; the latter termed an overlap trial and challenged the infants' disengagement abilities.

Longitudinal measures of orienting visual attention were achieved with administration of the Gap task at 6, 12 and 30 months of age (Wass, Porayska-Pomsta and Johnson, 2011). The 3 time points allowed for the developmental trajectories of the visual system to be explored across the two populations. Should any differences in speed processing or performance be observed either group, the multiple time points would hope to determine when these problems begin to emerge.

5.1.1 Apparatus

Those within their first year were placed in a travel seat 65cm from the Visual Display Unit (VDU) see Figure 5-1. For those at 30 months of age, the child sat on a parent or carers lap on a beanbag ensuring they were at an average distance of 65cm from the VDU, with their eye line in the centre of the unit.

Beneath the VDU was a 60 Hz Tobii 1750 eye tracker (1024 x 768 pixels monitor). The eye-tracker was calibrated, incorporating all dimensions of the VDU, please see Figure 5-1 for measurements and calibration calculations used.

The GAP task stimuli were presented using the Talk2Tobii toolbox and custom-written MATLAB scripts shared with the PDP from the Birkbeck Babylab, originally written by Sam Wass (Wass, Porayska-Pomsta and Johnson, 2011).

5.1.2 Procedure

The task began with 6 calibration points across the screen. Each point was presented independently, using an animated shape of numerous colours in combination with sound effects to draw the child's attention to each area of the screen. Quality of the recording was reported after all calibration points were shown. A pause in the script allowed the experimenter to either accept or decline the calibration. If calibration had not been achieved, the initial script was repeated.

The premise of the task was to identify the infant's ability to disengage from a more visually pleasing, central stimulus to a less detailed, peripheral image. This was achieved by the use of a colourful and detailed cartoon clock as the central stimulus, and a simple cartoon cloud as the peripheral stimulus, both shown in Figure 5-2. Movement of gaze from the central stimulus to the peripheral stimulus was rewarded by the disappearance of the cloud, revealing a different and more interesting animation and sound effect. The speed of relocating gaze to the peripheral stimulus varied dependent upon the different test conditions.

3 test conditions were included within the paradigm, please see Figure 5-2.

1. The baseline condition provided a measure of saccadic reaction time. During these trials, the central stimulus was presented for a total of 200ms and

disappeared simultaneously as the peripheral stimulus appeared, removing the need to disengage from the central stimulus.

2. The Gap condition displayed a blank screen between the central stimulus disappearance and the presentation of the peripheral stimulus. The speed to engage with the peripheral stimulus was expected to be the fastest within these trials.
3. Lastly, the overlap condition saw the central clock stimulus presented for 200ms, before the peripheral stimulus appeared simultaneously to this, with both stimuli remaining on screen until the infant looked to less visually pleasing cloud stimulus, challenging the infant's disengagement ability.

The script comprised of a random presentation of the 3 test conditions across 3 separate blocks. Each block was separated by two short videos to maintain the infant's attention throughout the task. Progression through the task was controlled by the infant's looking behaviour. Each trial required the infant to engage with the central stimulus before the trial could commence, ensuring that the gaze was not biased to one side of the screen. Upon centralisation of the gaze, as depicted in Figure 5-2 each condition had 200ms of the central stimulus before deviation into the different condition settings. In any condition, the infant had to look to the peripheral stimulus within 3000ms of its presentation. The number of trials where the infant's attention remained centralised in the presence of a peripheral image was defined here as inability to disengage, sometimes termed 'sticky attention'.

Once the infant had met all criteria stated above to achieve a valid trial, determined by online criteria written into the MATLAB script, the paradigm was terminated after 12 baseline, 12 gap and 16 overlap valid trials.

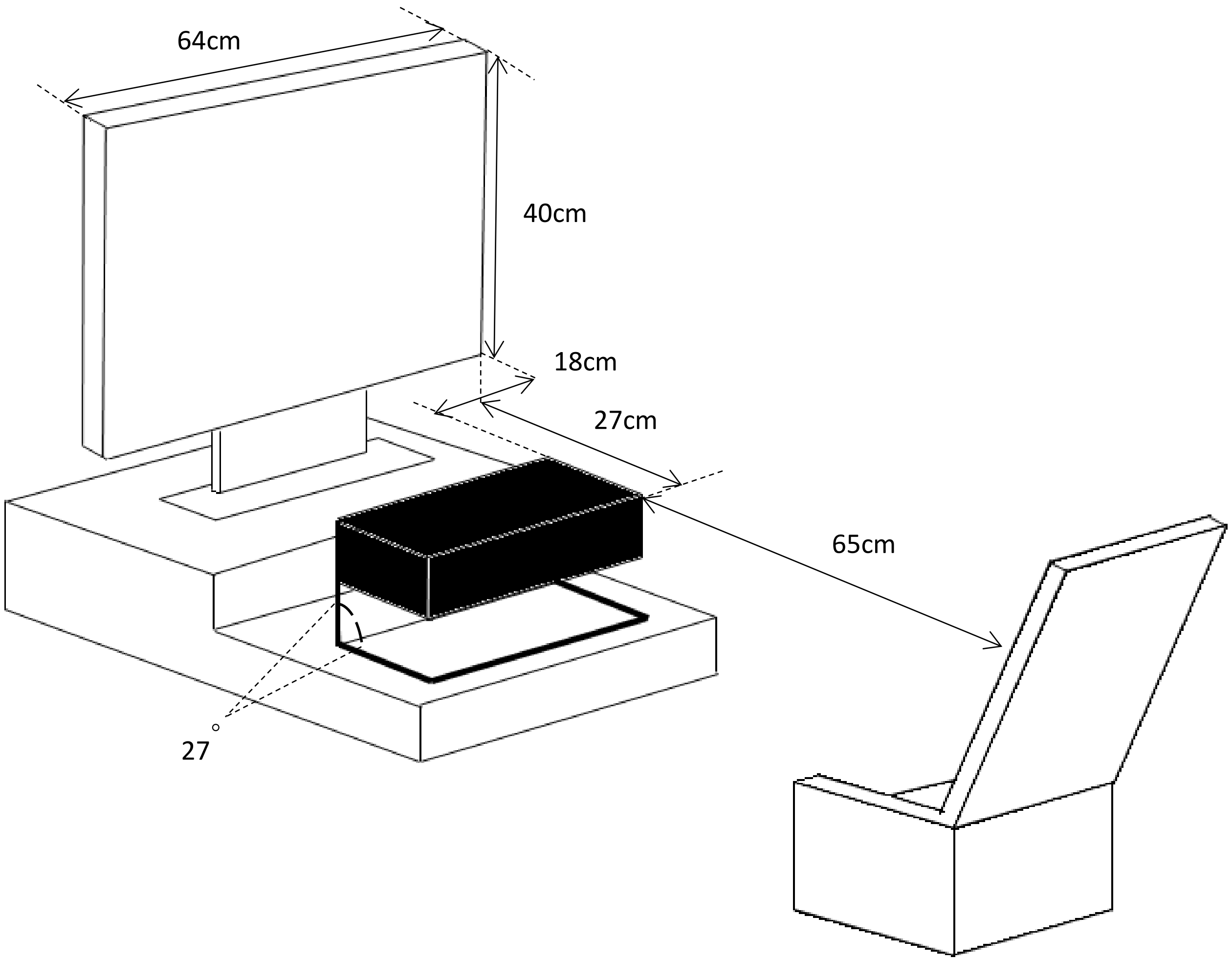
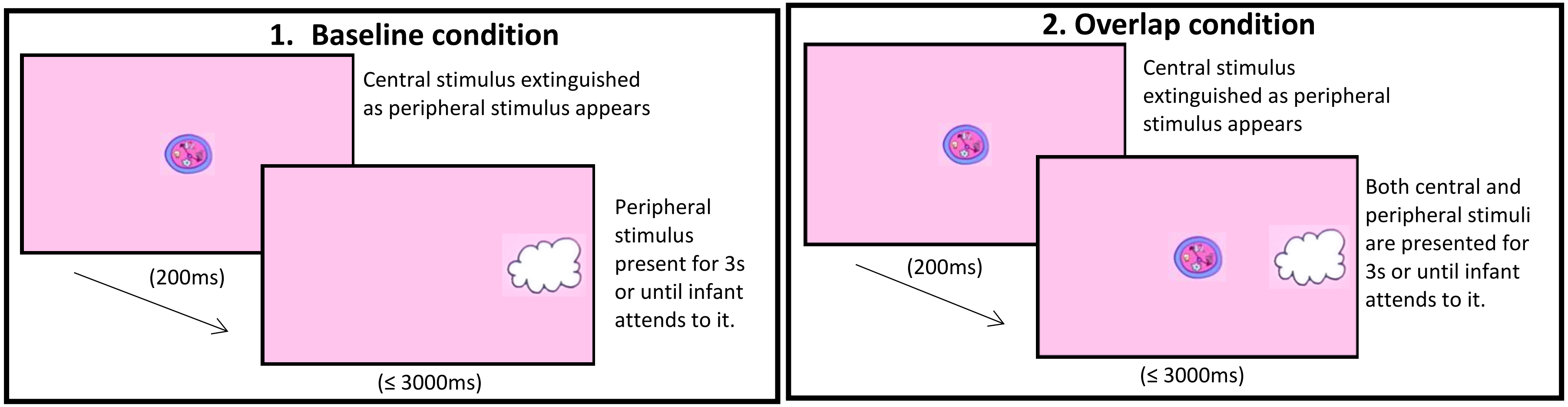


Figure 5-1. Set up and dimensions of visual display unit with eye-tracker; travel car seat used to correctly position the infants at 6 and 12months of age. During the 30m time point, the travel seat was removed and the child sat on a parent or carers lap at the same distance.



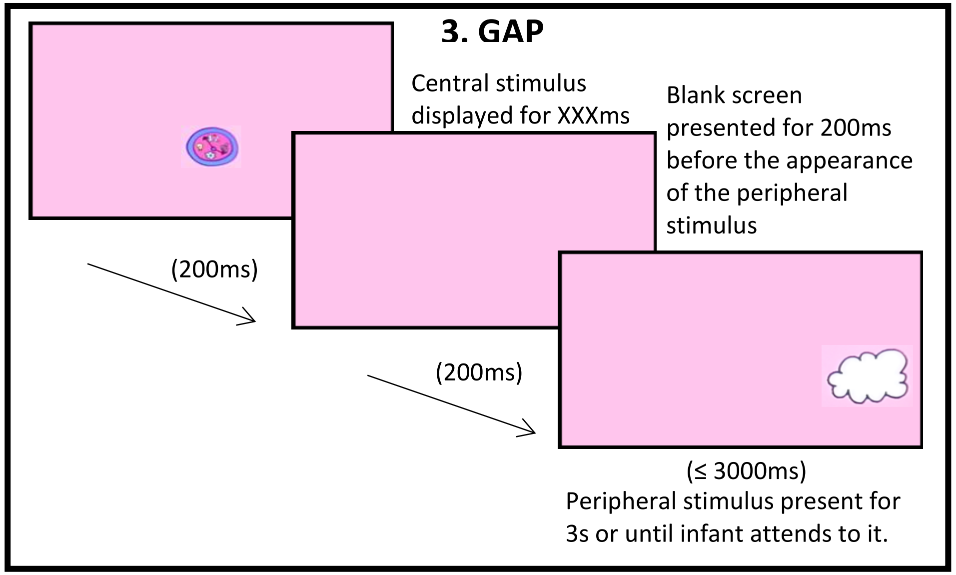


Figure 5-2. Representation of the visual displays and timing within the 3 conditions of the GAP task. All conditions start with a central stimulus, and end with a less visually interesting peripheral stimulus.

5.1.3 Coding

After data collection, we had concerns about the accuracy of the internal criteria used by the programme to determine trial validity. Data were re-processed offline with the help of Dr Emily Jones, Birkbeck University, which led to a different number of total trials for each participant. Post hoc we used the number of valid trials for exclusion purposes only: we excluded infants if they did not achieve a minimum of 5 valid trials within each condition.

The paradigm produced 3 primary response time (RT) variables, one for each condition. From these response times, a 'time to disengage' was calculated by subtracting the baseline RT from the Overlap RT, and the 'gap-effect' was calculated by subtracting the Gap RT from Overlap RT. As mentioned above, trials where disengagement did not occur will also be reported.

5.1.4 Analysis

The statistical process was as described in section 2.3, with each assessment age explored independently in the first instance, followed by a longitudinal look at the data. The longitudinal relationships were explored in the following order: 6 and 12 months; 12 and 30 months. This section is concluded by looking across all three time points. The 3 way longitudinal model was primarily explored only including participants with all 3 observations, however the study numbers were small with this approach. The final model includes infants that presented with a minimum of two observations across the 3 assessment ages.

The variable selected a priori to best reflect the performance on the Gap task was the 'disengagement' response time. This variable was selected as it accounts for the baseline performance variability between individuals, and includes the response to the overlap trials. The overlap trials were hypothesised to create the greatest variability in gaze shift response patterns.

The relationship between study group, male sex, IMD quintile, cognitive z-score and disengagement RT were explored with the use of regression models to determine the predictably of these characteristics on the gap-task performance. In all cases, two models were performed, the first reported includes the first three predictors and the second included the adjustment for the Bayley-III cognitive z-score. Six and 12 month scores were adjusted using the 12m Bayley cognitive score and 30 month scores using the 2 year cognitive score.

For the longitudinal analysis, rather than exploring all response times to all variables, only the disengagement RTs and the gap-effect were explored as both values accounted for individual variation. Multilevel mixed effects regression models were explored for each subgroup of longitudinal measurements thereby accounting for the repeated measures across assessment ages. Likelihood ratio (LR) tests were used to determine whether the inclusion of interaction terms were necessary in each model. The coefficients and 95% confidence intervals reported are from the Mixed-effect models.

Within the longitudinal data, participants were included in the models if they achieved >5 trials in each condition and at each age. This significantly reduced the numbers within each study group for the full three level model. An additional analysis was therefore included that allowed infants with valid data at any two time points to be included in the 3 way model, as an exploratory sensitivity analysis.

Task prediction:

If social communication difficulties are in part influenced by visual attention, as reported within the autism literature (Elsabbagh *et al.*, 2009) it is hypothesized that VP infants in the current study will display slower disengagement and associated responses time patterns to the Gap-overlap task.

The main research questions are:

1. Do the disengagement response times (RTs) differ between the two study groups?
2. Do differences in RTs between groups vary across trial conditions?
3. Do differences in saccade RTs in response to the Gap trials vary between study groups?
4. Do VP infants demonstrate non-disengagement more frequently compared to term infants at any age?
5. Do differences in RTs between groups vary with age of assessment?

5.2 Results – cross-sectional data

5.2.1 6 months

During the 6 month assessment phase, 36 term and 22 VP infants completed the Gap-overlap task (Table 5-1). No differences were observed between the two study groups in the response times to the 3 task conditions: baseline, gap and overlap. The gap condition produced the fastest shift in gaze (Table 5-2; Figure 5-3). No differences were seen in the speed to disengage from the central stimulus between the two study groups (overlap trial minus baseline; Figure 5-4); nor were there any differences in the gap effect (overlap minus gap; Figure 5-5). These results were consistent after adjusting for cognitive z-score collected at 12 months.

		Term (n=36)	VP (n=22)
Gestational age	*Median (range); weeks^{+d}*	40^{+0} (37^{+1} – 41^{+5})	26^{+0} (23^{+4} – 31^{+4})
Male sex		17 (47.22%)	15 (68.18%)
IMD Quintile	1	6	2
	2	8	4
	3	3	8
	4	11	5
	5	8	3
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=34; VP=15)	108.68 (12.20)	100.00 (8.66)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=34; VP=15)	.07 (1.03)	-.66 (.73)

Table 5-1. Demographic details of infants in Gap-overlap task at 6 month assessment.

Table 5-2 displays the means and SDs of the outcome variables of the GAP task at the 6 month assessment age. No significant differences were observed between the two groups in any of the response times at this age. When running a repeated measures ANOVA, a significant main effect of condition was observed (F(1.15, 64.53) = 390.92, p<.001), with Gap trials producing the fastest response and the overlap the slowest, as seen from the means reported in Table 5-2. No main effects of study group were seen nor were there any significant interaction terms.

GAP task Variable		*Term (N=36; excluded as <5 = 6)*	*Preterm (N=22; excluded as <5 = 6)*	*p*
Baseline RT (ms)	Mean	386. 16	388.78	.86
	(SD)	50.85	59.66	
GAP RT (ms)	Mean	281.41	274.16	.27
	(SD)	22.10	26.94	
Overlap RT (ms)	Mean	675.50	667.81	.84
	(SD)	145.14	120.12	
Disengagement RT (ms)	Mean	289.34	279.03	.76
	(SD)	127.68	123.41	
Gap-effect (ms)	Mean	394.09	393.65	.99
	(SD)	137.43	118.54	
Non-disengagement	N	14	10	.42

Table 5-2. Response times (ms) to each condition of the Gap-Overlap task and the total number of trials where disengagement from the central stimulus did not occur during the 6 month assessment by study group.

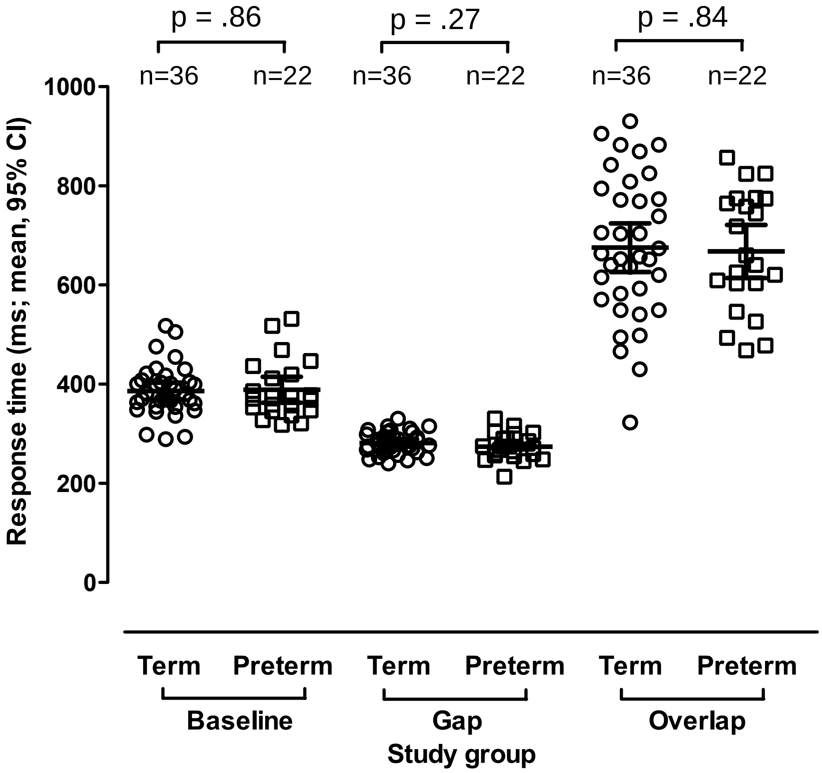


Figure 5-3. Response times (ms) of each condition within the Gap overlap task at 6 months of age by study group.

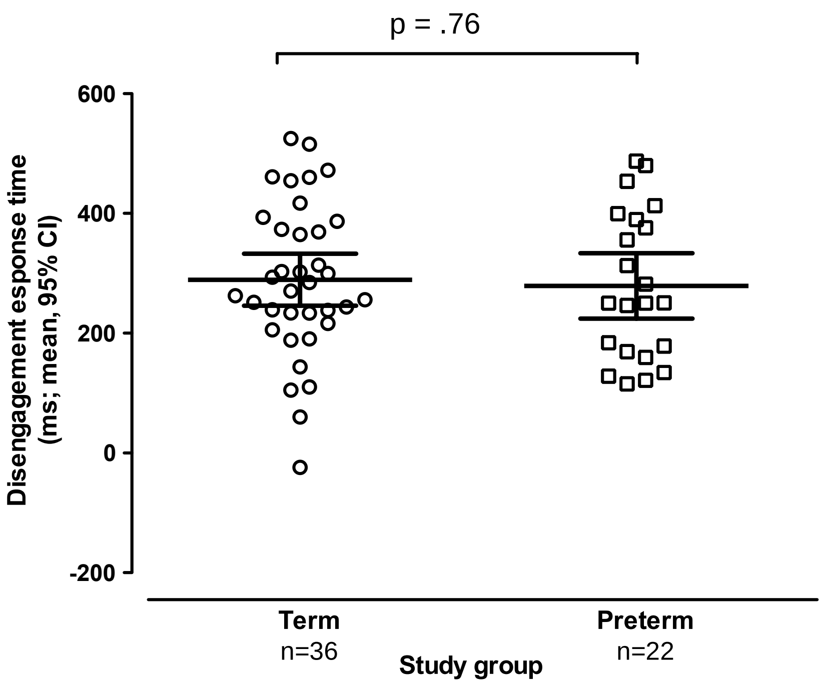


Figure 5-4. Disengagement response times (ms) of term and very preterm infants during the Gap-overlap task at 6 months of age.

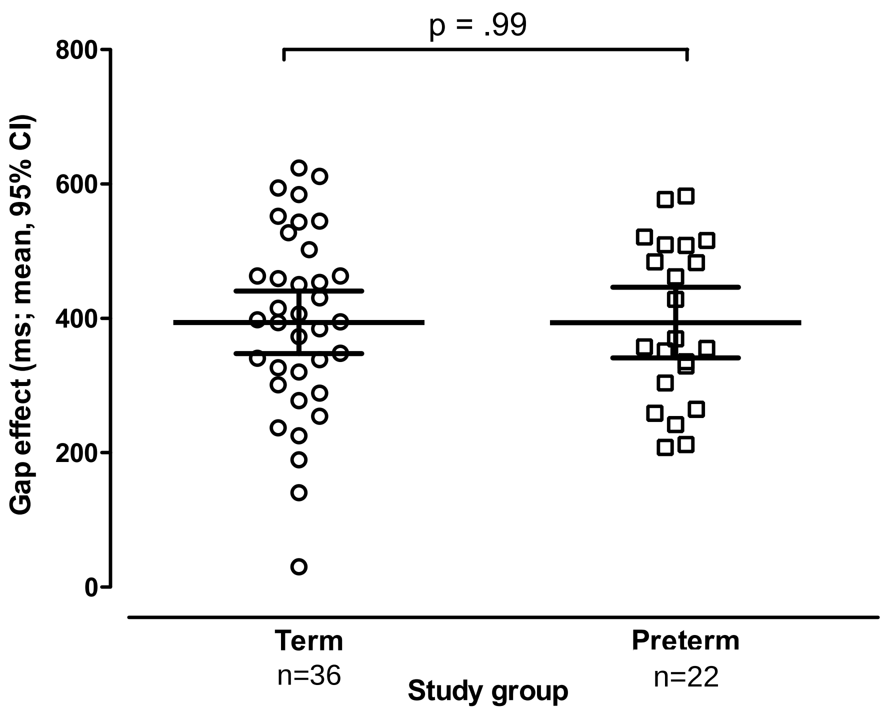
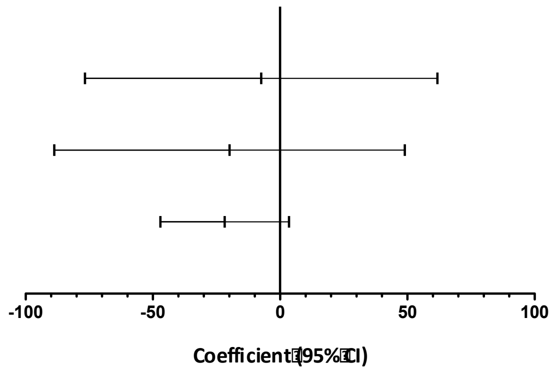


Figure 5-5. Gap-effect (ms) of term and very preterm infants during the Gap-overlap task at the 6 months of age.

The main outcome, disengagement RT, was explored in two regression models (Table 5-3 and Table 5-4). None of the predictive variables included have a significant predictive effect on the disengagement RT during the 6 month assessment. This result was maintained when adjusting for the Bayley-III z-scores at 12 months.

Table 5-3. Linear regression model of disengagement RT during the 6 month GAP task (F(3,54) = 1.05, p = .38). The baseline group are term-born females with an IMD quintile of 1.

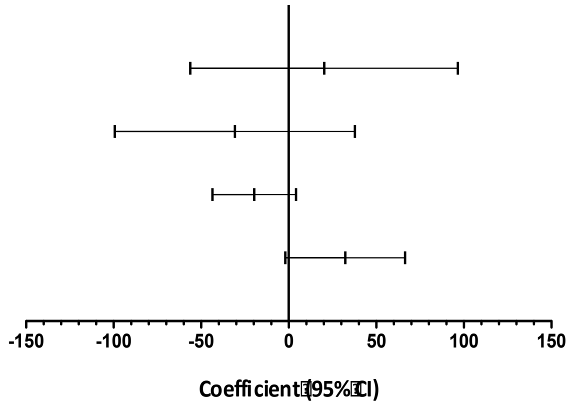
Predictor	Term (n=36)	Preterm (n=22)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	40^{+0} (37^{+2} – 41^{+5})	26^{+0} (23^{+4} – 31^{+4})	-7.42	-76.68 – 61.84	.83
Male sex	17 (47.22%)	15 (68.18%)	-19.87	-88.77 – 49.03	.57
IMD Quintile	4 (1-5)	3 (1-5)	-21.77	-47.04 – 3.49	.09
Disengag RT(Const.)	289.34 (127.68)	279.03 (123.41)	368.28	266.58 – 469.99	.000



Overall model fit was R^2 = 0.05

Table 5-4 Linear regression model of disengagement RT during the 6 month GAP task including the adjustment of the 12m Bayley-III z-score ($F(4, 44) = 1.62$, $p = .17$). The baseline group are term-born females with an IMD quintile of 1.

Predictor	Term (n=34)	Preterm (n=15)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	40^{+0} (37^{+2} – 41^{+5})	26^{+2} (24^{+3} – 29^{+4})	20.20	-56.24 – 96.63	.60
Male sex	17 (50)	12 (80)	-30.75	-99.35 – 37.81	.37
IMD Quintile	2.5 (1-5)	3 (1-5)	-19.68	-43.56 – 4.19	.10
12m cog. Z-score	.71 (1.03)	-0.66 (.73)	32.27	-1.91 – 66.45	.06
Disengag RT(Const.)	289.34 (127.68)	279.03 (123.41)	379.76	284.46 – 475.05	.000



Overall model fit was $R^2 = 0.13$

5.2.2 12 months

During the 12 month assessment phase, 31 term born and 23 VP infants completed the Gap-overlap task (Table 5-5). No differences were observed between the two study groups in the response times to the 3 task conditions: baseline, gap and overlap. The gap condition produced the fastest shift in gaze (Table 5-6; Figure 5-6). No differences were seen in the speed to disengage from the central stimulus (overlap trial minus baseline; Figure 5-7) between the two study groups; nor were there any differences in gap effect (overlap minus gap; Figure 5-8). These results were consistent after adjusting for cognitive z-score collected at 12 months.

		Term (n=31)	***VP (n=23)***
Gestational age	*Median (range); weeks^{+d}*	40^{+0} (37^{+1} – 42^{+0})	26^{+0} (23^{+6} – 31^{+4})
Male sex		14 (45.16%)	14 (60.87%)
IMD Quintile	1	3	3
	2	6	5
	3	6	8
	4	11	5
	5	5	2
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=31; VP=18)	107.26 (12.51)	99.17 (11.41)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=31; VP=18)	-.05 (1.05)	-.73 (.96)

Table 5-5. Demographic details of infants in Gap-overlap task at 12 month assessment.

Table 5-6 displays the means and SDs of the outcome variables of the GAP task at the 12 month assessment age. No significant differences were observed between the two groups in any of the response times at this age. Again, a significant main effect of condition was observed, with the Gap trials producing the fastest response ($F(1.18, 61.24) = 619.83$, $p<.001$), but with no main effect of study group or interaction terms were observed.

GAP task Variable		*Term (n=31; excluded as <5 = 1)*	*Preterm (n=23; excluded as <5 = 7)*	*p*
Baseline RT	Mean	378.70	378.24	.97
	(SD)	45.51	38.28	
GAP RT	Mean	264.12	272.83	.30
	(SD)	28.19	33.08	
Overlap RT	Mean	702.56	696.02	.81
	(SD)	92.08	111.09	
Disengagement RT	Mean	323.87	317.78	.84
	(SD)	112.11	108.90	
Gap-effect RT	Mean	438.44	423.19	.61
	(SD)	103.87	116.25	
Non-disengagement	N	11	9	.17

Table 5-6. Response times (ms) to each condition of the Gap-Overlap task and the total number of trials where disengagement from the central stimulus did not occur during the 12 month assessment by study group.

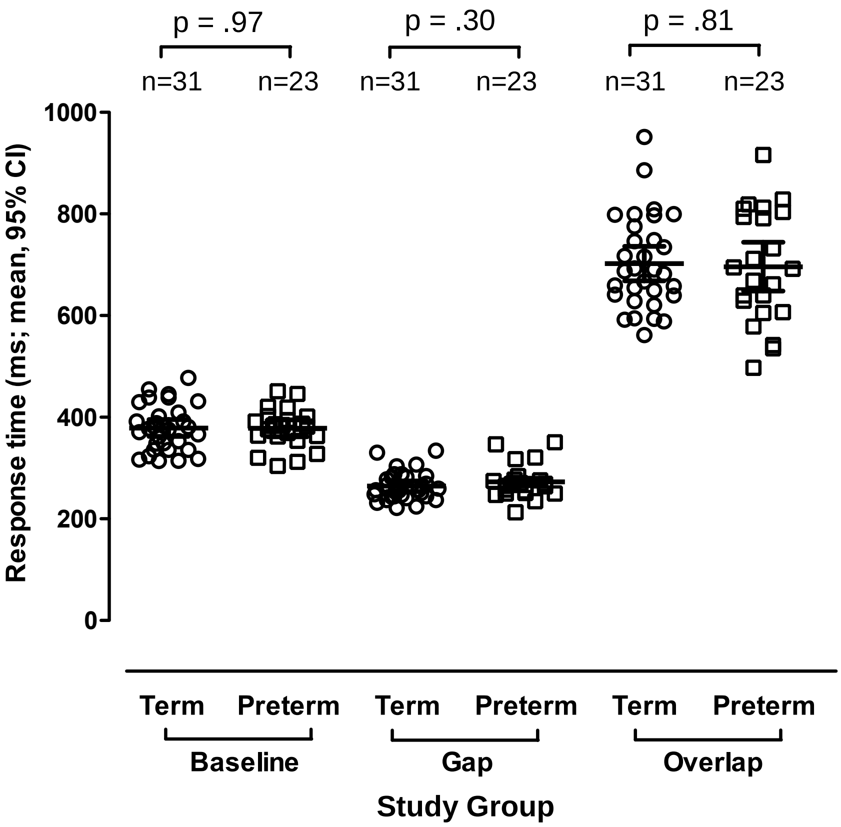


Figure 5-6. Response times (ms) of each condition within the Gap overlap task at 12 months of age by study group.

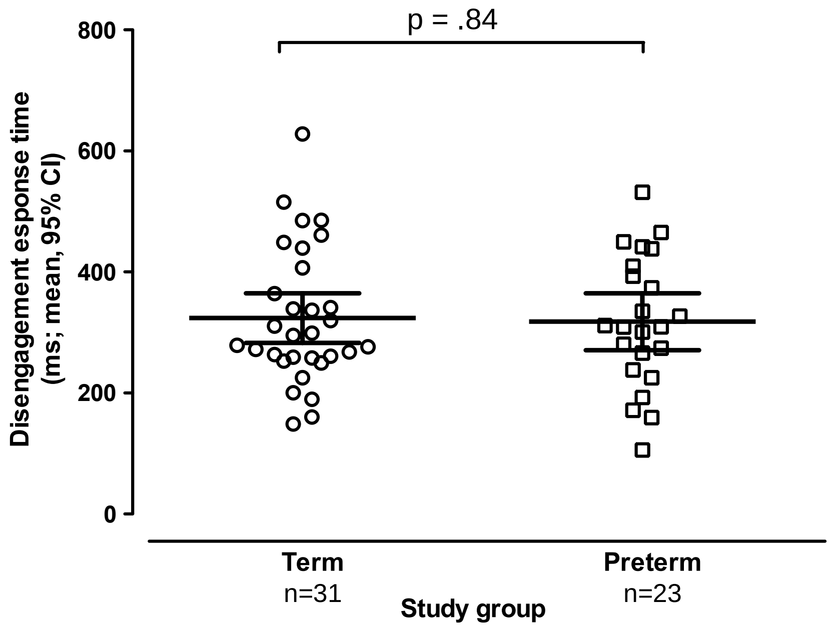


Figure 5-7. Disengagement response time (ms) of term and very preterm infants during the Gap-overlap task at 12 months of age.

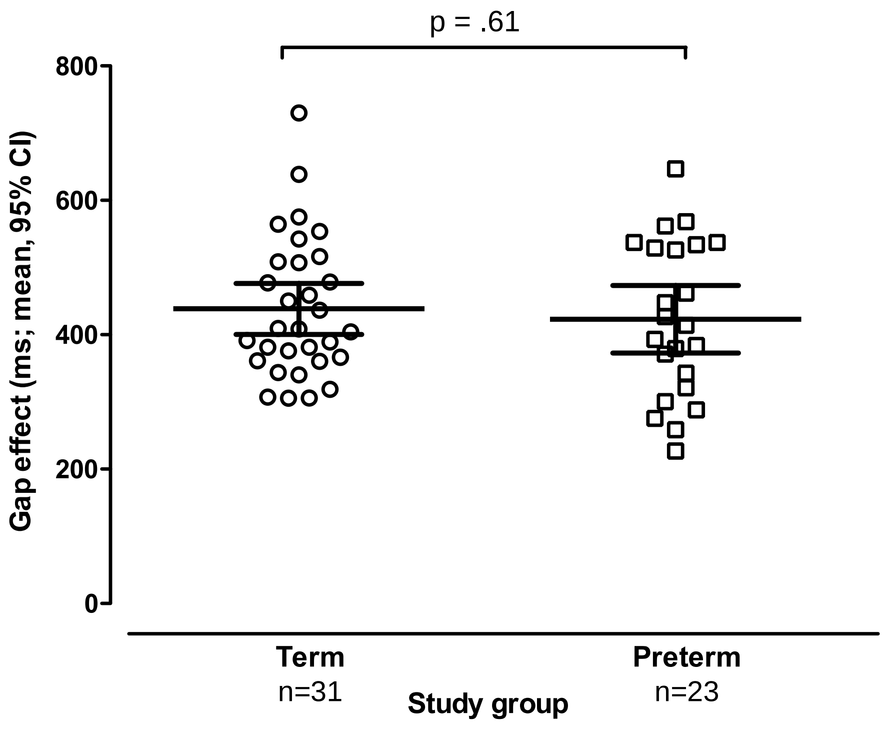
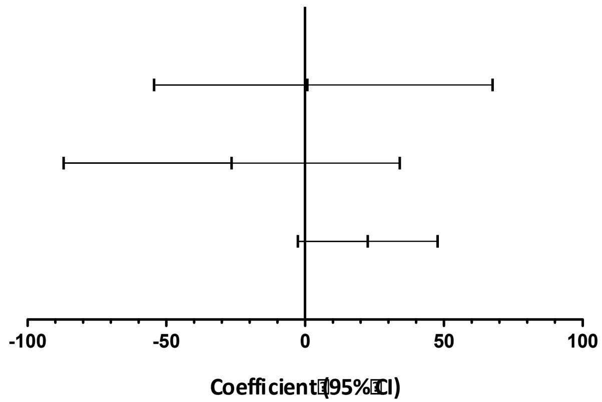


Figure 5-8. Gap-effect (ms) of term and very preterm infants during the Gap-overlap task at 12 months of age.

The main outcome, disengagement RT, was explored in two linear regression models (Table 5-7 and Table 5-8). The variables included within the models did not have significant predictive effects on the disengagement RT during the 12 month assessment. This result was maintained when adjusting for the Bayley-III z-scores at 12 months (Table 5-8), although a borderline effect of IMD quintile was seen (β = 25.49, p=.06) suggesting a possible relationship with a higher deprivation score and slower disengagement behaviour in the Gap task.

Table 5-7. Linear regression model of disengagement RT during the 12 month GAP task (F(3, 50) = 1.55, p = .21). The baseline group are term-born females with an IMD quintile of 1.

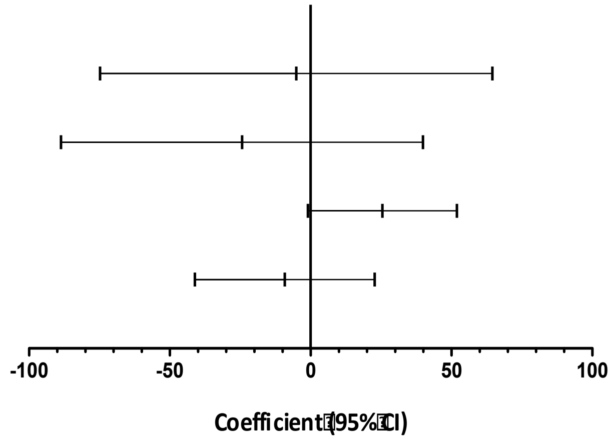
Predictor	Term (n=31)	Preterm (n=23)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	39^{+6} (37^{+1} – 42^{+0})	26^{+4} (23^{+6} – 31^{+4})	.83	-54.4 – 67.60	.83
Male Sex	14 (45.16)	14 (60.87)	-26.45	-87.01 – 34.11	.39
IMD Quintile	4 (1-5)	3 (1-5)	22.61	-2.55 – 47.77	.08
Disengag RT(Const.)	323.87 (112.11)	317.78 (108.90)	261.43	162.35 – 360.50	.000



Overall model fit was R^2 = 0.09

Table 5-8 Linear regression model of disengagement RT during the 12 month GAP task including the adjustment of the 12m Bayley-III z-score ($F(3, 50) = 1.55$, $p = .21$). The baseline group are term-born females with an IMD quintile of 1.

Predictor	Term (n=31)	Preterm (n=18)			
	Median (Range)	Median (Range)	Coef	95%CI	P
Study Group	39^{+6} (37^{+1} – 42^{+0})	26^{+4} (24^{+0} – 29^{+4})	-5.11	-74.75 – 64.53	.88
Male Sex	14 (45.16)	12 (66.67)	-24.38	-88.69 – 39.93	.45
IMD Quintile	4 (1-5)	3 (1-5)	25.49	-.99 – 51.97	.06
12m cog. Z-score	-.04 (1.05)	-.73 (.96)	-9.18	-41.11 – 22.74	.57
Disengag RT(Const.)	323.87 (112.11)	317.78 (108.90)	250.58	146.91 – 354.24	.000



Overall model fit was $R^2 = 0.13$

5.2.3 30 months

During the 30 month assessment phase, 17 term and 13 VP infants completed the Gap-overlap task (Table 5-9). No differences were observed between the two study groups in the response times to the 3 task conditions: baseline, gap and overlap. The gap condition again produced the fastest shift in gaze (Table 5-10; Figure 5-9). No differences were seen in the speed to disengage from the central stimulus (overlap trial minus baseline; Figure 5-10) between the two study groups; nor were there any differences in gap effect (overlap minus gap; Figure 5-11). In this instance, when adjusting for the cognitive z-scores collected at 2 years, VP infants were on average 110 ms faster than the term born infants in their disengagement response times.

		Term (n=17)	***VP (n=13)***
Gestational age	*Median (range); weeks^{+d}*	40^{+0} (37^{+0} – 42^{+1})	26^{+2} (23^{+6} – 29^{+4})
Male sex		10 (58.82%)	11 (84.62%)
IMD Quintile	1	5	4
	2	2	0
	3	2	4
	4	4	3
	5	4	2
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=13; VP=12)	109.62 (12.33)	102.92 (12.15)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=13; VP=12)	-.15 (.99)	-.39 (.98)

Table 5-9. Demographic details of infants in Gap-overlap task at 30 month assessment.

Table5-10 displays the means and SDs of the outcome variables of the GAP task at the 30 month assessment age. No significant differences were observed between the two groups in any of the response times at this age. Again, a significant main effect of condition was observed ($F(1.44, 40.39) = 290.50$, $p<.001$), but with no main effect of study group or interaction terms were reported.

GAP task Variable		*Term (n=17; excluded as <5 = 7)*	*Preterm (n=13; excluded as <5 = 7)*	*p*
Baseline RT	Mean	347.68	380.94	.11
	(SD)	37.67	71.72	
GAP RT	Mean	252.90	269.95	.20
	(SD)	31.81	39.64	
Overlap RT	Mean	593.83	581.32	.74
	(SD)	101.68	98.61	
Disengagement RT	Mean	246.16	200.38	.14
	(SD)	82.73	81.92	
Gap-effect RT	Mean	340.93	311.37	.38
	(SD)	98.43	74.87	
Non-disengagement	n	15	8	.75

Table 5-10. Response times (ms) to each condition of the Gap-Overlap task and the total number of trials where disengagement from the central stimulus did not occur during the 30 month assessment by study group.

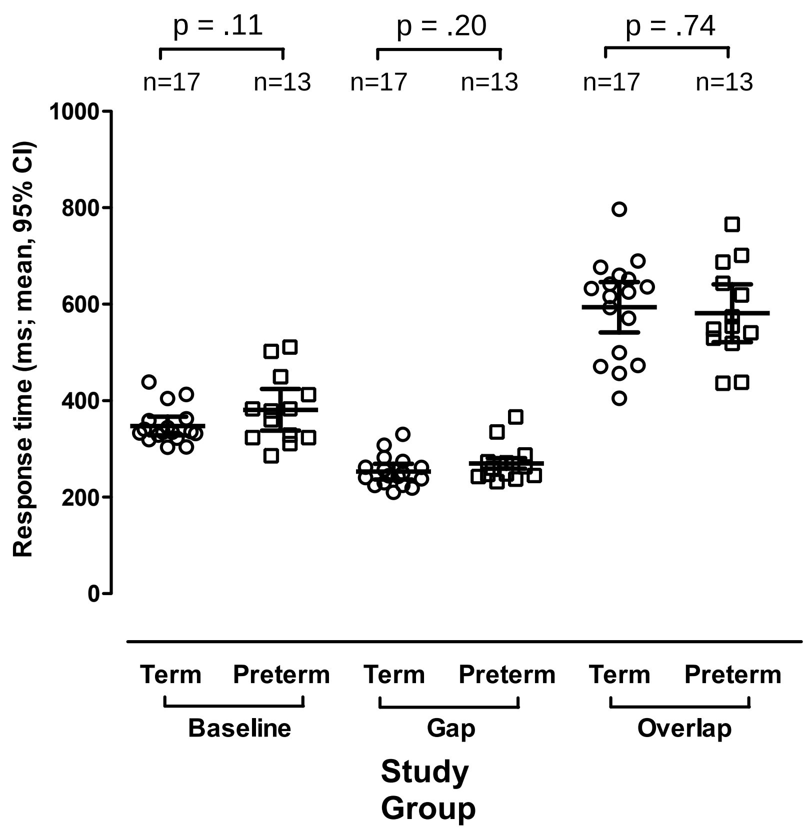


Figure 5-9. Response times (ms) of each condition within the Gap overlap task at 30 months of age by study group.

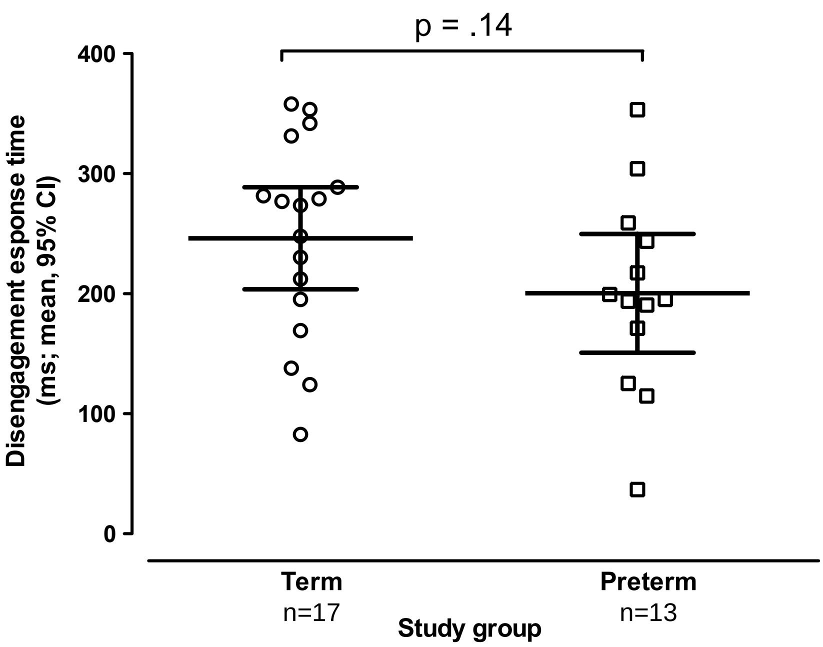


Figure 5-10. Disengagement response time (ms) of term and very preterm toddlers during the Gap-overlap task at 30 months of age.

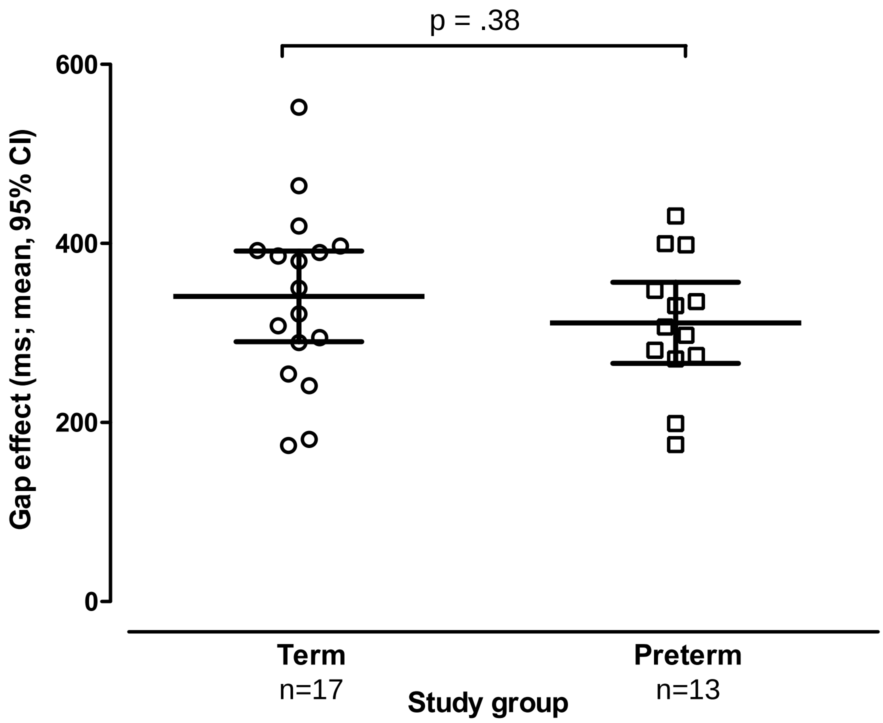
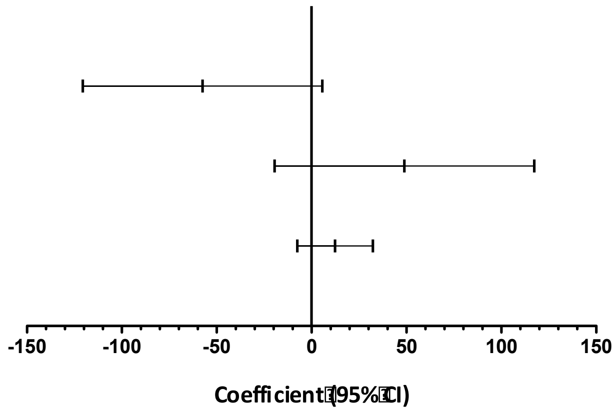


Figure 5-11. Gap-effect (ms) of term and very preterm toddlers during the Gap-overlap task at the`30 months of age.

Disengagement RT was modelled again at this age point in two linear regression models (Table 5-11 and Table 5-12). In the primary model, none of the predictors appear to have significant effects on the disengagement RT during the 30 month assessment. However, when adding the adjustment for cognitive score at this age, a significant predictive effect of study group was observed (Table 5-12). The β-coefficient for the study group variable indicates that the VP children are responding faster than the term-born children when adjusting for their scores on the Bayley-III at 30 months ($\beta = -112.43$, $p = .001$).

Table 5-11. Linear regression model of disengagement RT during the 30 month GAP task (F(3, 26) = 2.00, p = .14). The baseline group are term-born females with an IMD quintile of 1.

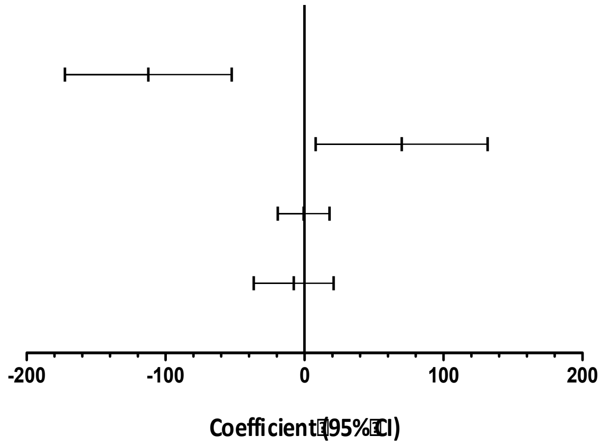
Predictor	Term (n=17)	Preterm (n=13)			
	Median (Range)	Median (Range)	Coef	95%CI	P
Study Group	39^{+5} (37^{+2} – 42^{+1})	26^{+4} (23^{+6} – 29^{+4})	-57.45	-120.64 – 5.74	.07
Male Sex	10 (58.82)	11 (84.62)	48.97	-19.48 – 117.41	.15
IMD Quintile	3 (1-5)	3(3)	12.39	-7.48 – 32.26	.21
Disengag RT(Const.)	246.16 (82.73)	200.38 (81.92)	180.18	96.23 – 264.12	.000



Overall model fit was R^2 = 0.19

Table 5-12 Linear regression model of disengagement RT during the 30 month GAP task including the adjustment of the 30m Bayley-III z-score ($F(4, 20) = 4.31$, $p = .01$) . The baseline group are term-born females with an IMD quintile of 1.

Predictor	Term (n=13)	Preterm (n=12)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	40^{+1} (38^{+5} – 42^{+1})	26^{+5} (23^{+6} – 29^{+4})	-112.43	-172.48 - -52.38	.001
Male Sex	7 (53.85)	10 (83.33)	69.92	8.07 – 131.77	.03
IMD Quintile	4 (1-5)	3(1-5)	-.63	-19.312 – 18.06	.95
30m cog. Z-score	.15 (.99)	-.39 (.98)	-7.81	-36.61 – 20.99	.58
Disengag RT(Const.)	246.16 (82.73)	200.38 (81.92)		167.61 – 321.80	.000



Overall model fit was $R^2 = 0.46$

5.3 Results – longitudinal data

5.3.1 6 to 12 months

At 6 and 12 months, 20 term-born and 16 VP infants had usable data sets (Table 5-13). Both the groups displayed a small increase in time to disengage from the central stimulus at 12 months, and the gap between stimulus presentations created a slightly longer delay in gaze shift (Table 5-14); however neither measure were significantly different from the times at 6 months (Figure 5-12 and Figure 5-13). These results were not influenced by male sex, IMD quintile or after adjustment for cognitive score at 12 months.

		Term (n=20)	***VP (n=16)***
Gestational age	*Median (range); weeks^{+d}*	39^{+6} (37^{+1} – 41^{+2})	26^{+2} (23^{+6} – 31^{+4})
Male sex		7 (35%)	10 (62.50%)
IMD Quintile	1	2	2
	2	6	2
	3	3	7
	4	7	4
	5	2	1
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=34; VP=15)	108.00 (13.02)	101.25 (8.56)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=34; VP=15)	.01 (1.10)	-.55 (.72)

Table 5-13. Demographic details of infants in Gap-overlap task at 6 and 12 months of age.

A repeated measure ANOVA on both the gap-effect and disengagement response times at 6 to 12 months, showed no main effects of age or group and no significant interactions. When fitting a multi-level mixed-effects model to the data, none of the predefined variables significantly predicted either outcome variable. The full models for each outcome, including the adjustment for Bayley-III z-scores at 12 months, are displayed in Table 5-15 and Table 5-16.

GAP task Variable			***Term (n=20)***	***Preterm (n=16)***
Disengagement RT	6m	Mean (ms)	318.95	281.24
		(SD)	121.43	123.82
	12m	Mean (ms)	329.2	326.04
		(SD)	91.07	115.63
	Mean diff		10.25	44.8
	(95%CI)		-58.46 – 78.96	-41.70 – 131.30
Gap-effect	6m	Mean (ms)	427.52	395.01
		(SD)	126.99	125.57
	12m	Mean (ms)	434.69	439.23
		(SD)	93.51	112.46
	Mean diff		7.17	44.22
	(95%CI)		-64.22 – 78.56	-41.85 – 130.29

Table 5-14. Response time to disengage (ms) and speed of gaze shift for gap-effect in the Gap-Overlap task at 6 and 12 months, including mean difference in response across the two time points.

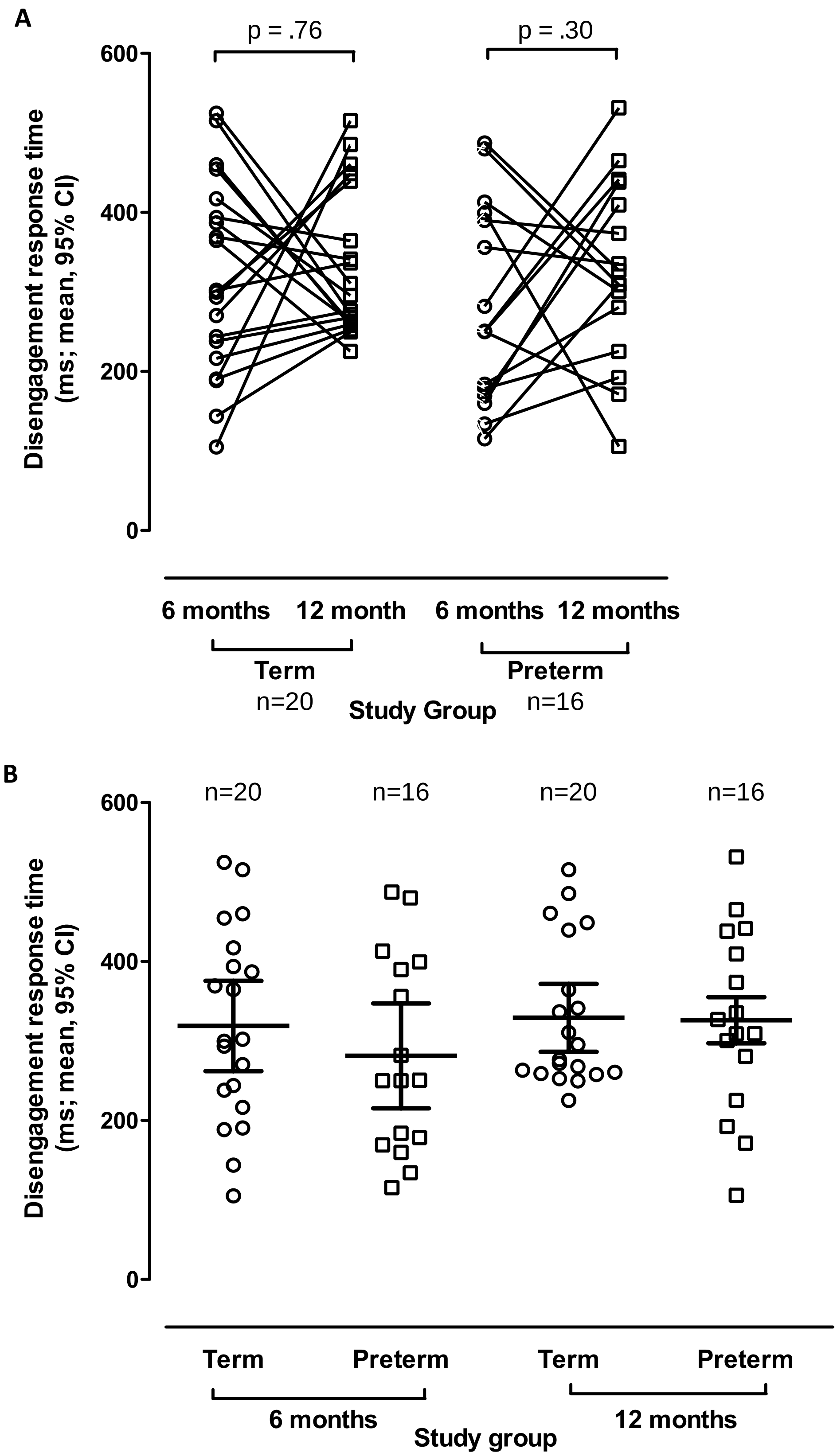


Figure 5-12. Two displays of the change of disengagement response times (ms) over the 6 to 12 months (A) illustrates the individual changes over time in each study group; (B) illustrates group difference by age to visually display how the mean responses change over time

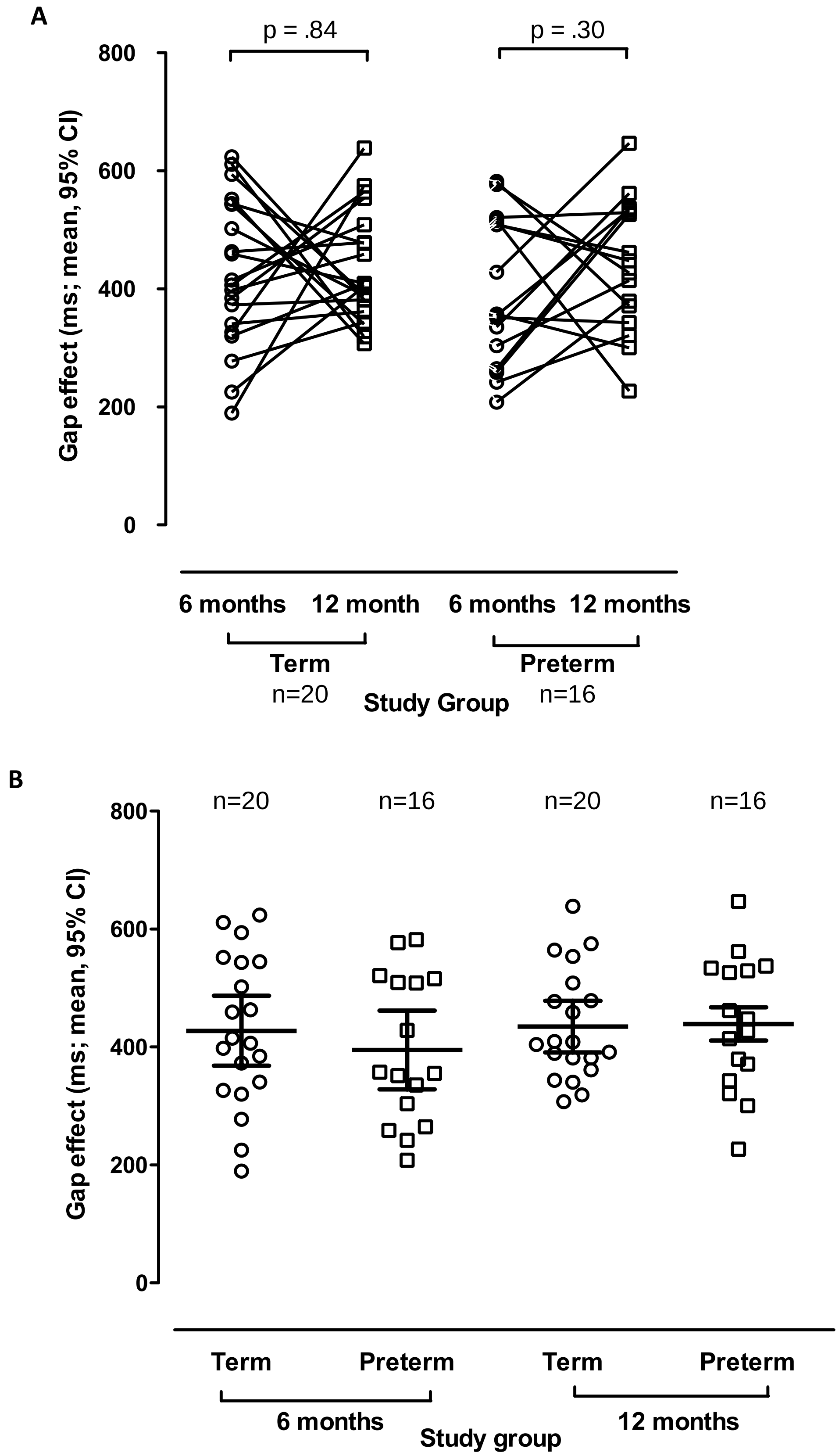
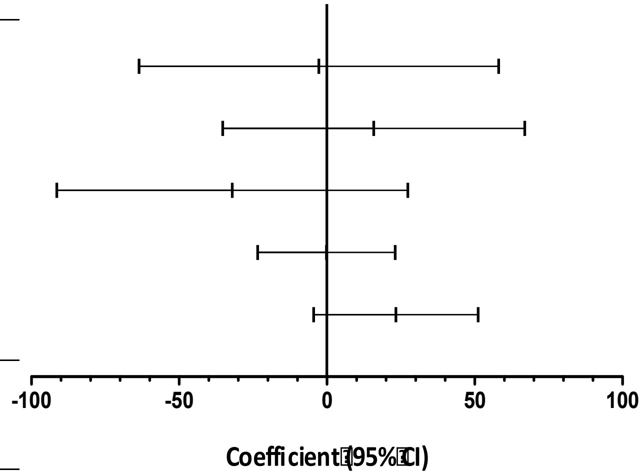


Figure 5-13. Two displays of the change of gap-effect (ms) over the 6 to 12 month assessment (A) illustrates the individual changes over time in each study group; (B) illustrates group difference by age to visually display how the mean responses change over time

Table 5-15. Multi-level mixed effects regression model of disengagement RT across 6 and 12 month assessment including the adjustment of the 12m Bayley-III z-score. The baseline group are term-born females at 6 months with an IMD quintile of 1.

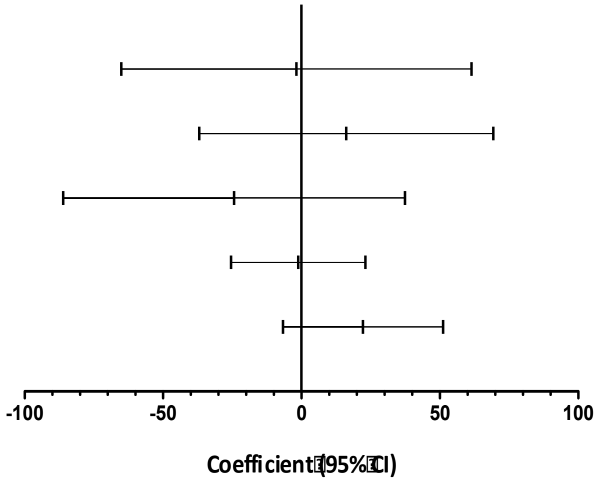
Predictor	Term (n=20)	Preterm (n=16)			
	Median (Range)	Median (Range)	Coef	95%CI	P
Study Group	39^{+5} (37^{+0} – 41^{+2})	26^{+5} (24^{+4} – 29^{+4})	-2.73	-63.56 – 58.09	.93
Age (12)	329.2 (91.07)	307.87 (109.36)	15.85	-35.23 – 66.94	.54
Male sex	7 (35)	9 (75)	-32.0	-91.37 – 27.37	.29
IMD Quintile	3 (1-5)	3 (2)	-.15	-23.46 – 23.15	.99
12m cog. Z-score	.01 (1.10)	-.55 (.72)	23.38	-4.45 – 51.21	.10
Gap-effect RT(Const.)	-	-	327.47	237.63 – 417.32	.000



Wald chi^2 = 2.96 (p = .56)

Table 5-16. Multi-level mixed effects regression model of gap-effect across 6 and 12 month assessment including the adjustment of the 12m Bayley-III z-score. The baseline group are term-born females at 6 months with an IMD quintile of 1.

	Term (n=20)	Preterm (n=12)			
Predictor	Median (Range)	Median (Range)	Coef	95%CI	P
Study Group	39^{+5} (37^{+0} – 41^{+2})	26^{+5} (24^{+4} – 29^{+4})	-1.80	-65.06 – 61.47	.96
Age (12)	434.69 (93.51)	422.79 (106.26)	16.19	-36.94 – 69.32	.55
Male sex	7 (35)	9 (75)	-24.33	-86.08 – 37.42	.44
IMD Quintile	3 (2)	3 (2)	-1.13	-25.37 – 23.11	.93
12m cog. Z-score	.01 (1.10)	-.55 (.72)	22.26	-6.69 – 51.20	.13
Gap-effect RT(Const.)	-	-	434.65	341.21 – 528.10	.000



Wald chi^2 = 4.90 (p = .43)

5.3.2 12 to 30 months

At 12 to 30 months, 7 term-born and 10 VP infants had usable data sets (Table 5-11). Both the groups displayed a decrease in time to disengage from the central stimulus at 30 months (Table 5-18; Figure 5-14). The VP toddlers increased the speed they disengaged on average by 155ms, compared to the term-born infants whose speed decreased on averaged by 23ms. The gap between stimulus presentations also had a smaller impact in the gaze shift speed with the VP infants displaying a greater increase in speed compared to the term born (Table 5-18; Figure 5-15). These results were not influenced by male sex, IMD quintile or adjustment for cognitive score at 30 months.

		Term (n=7)	***VP (n=10)***
Gestational age	*Median (range); weeks^{+d}*	39^{+4} (38^{+5} – 40^{+2})	26^{+2} (23^{+6} – 29^{+4})
Male sex		3 (42.86%)	8 (80%)
IMD Quintile	1	3	3
	2	2	4
	3	1	0
	4	0	2
	5	1	1
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=6; VP=9)	105.83 (8.61)	105.00 (13.23)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=6; VP=9)	-.15 (.69)	-.22 (1.07)

Table 5-17. Demographic details of infants in Gap-overlap task at 12 and 30 months of age.

A repeated measure ANOVA on the disengagement RT across the 12 to 30 months assessments showed a main effect of age ($F(1, 15) = 7.04$, $p=.018$), and a borderline interaction between age and group ($F(1, 15) = 3.85$, $p<.07$), with the preterm born participants showing a greater decrease in RT from 12 to 30 months compared to the term born participants ($t(9)= 3.01$, $p = .02$). When exploring the gap-effect over the 12 and 30 month time points, there was a significant main effect of age in an repeated measures ANOVA ($F(1,15) = 15.05$, $p = .001$). There were no interactions effects or main effect of study group.

GAP task Variable			***Term (n=7)***	***Preterm (n=10)***
Disengagement RT	12m	Mean (ms)	269.63	336.63
		(SD)	49.82	135.97
	30m	Mean (ms)	246.26	180.84
		(SD)	70.51	75.60
	Mean diff		-23.37	-155.79 **
	(95%CI)		-94.47 – 47.73	-259.15 – -52.43
Gap-effect	12m	Mean (ms)	386.61	443.61
		(SD)	40.29	141.04
	30m	Mean (ms)	322.80	298.46
		(SD)	80.64	74.98
	Mean diff		-63.81	-145.15**
	(95%CI)		-138.05 – 10.43	-251.27 – -39.03

Table 5-18. Mean and SDs of disengagement RTs and gap-effect variables across the 12 and 30m assessment ages within the two groups; *p<.05; **p<.01; ***p≤.001

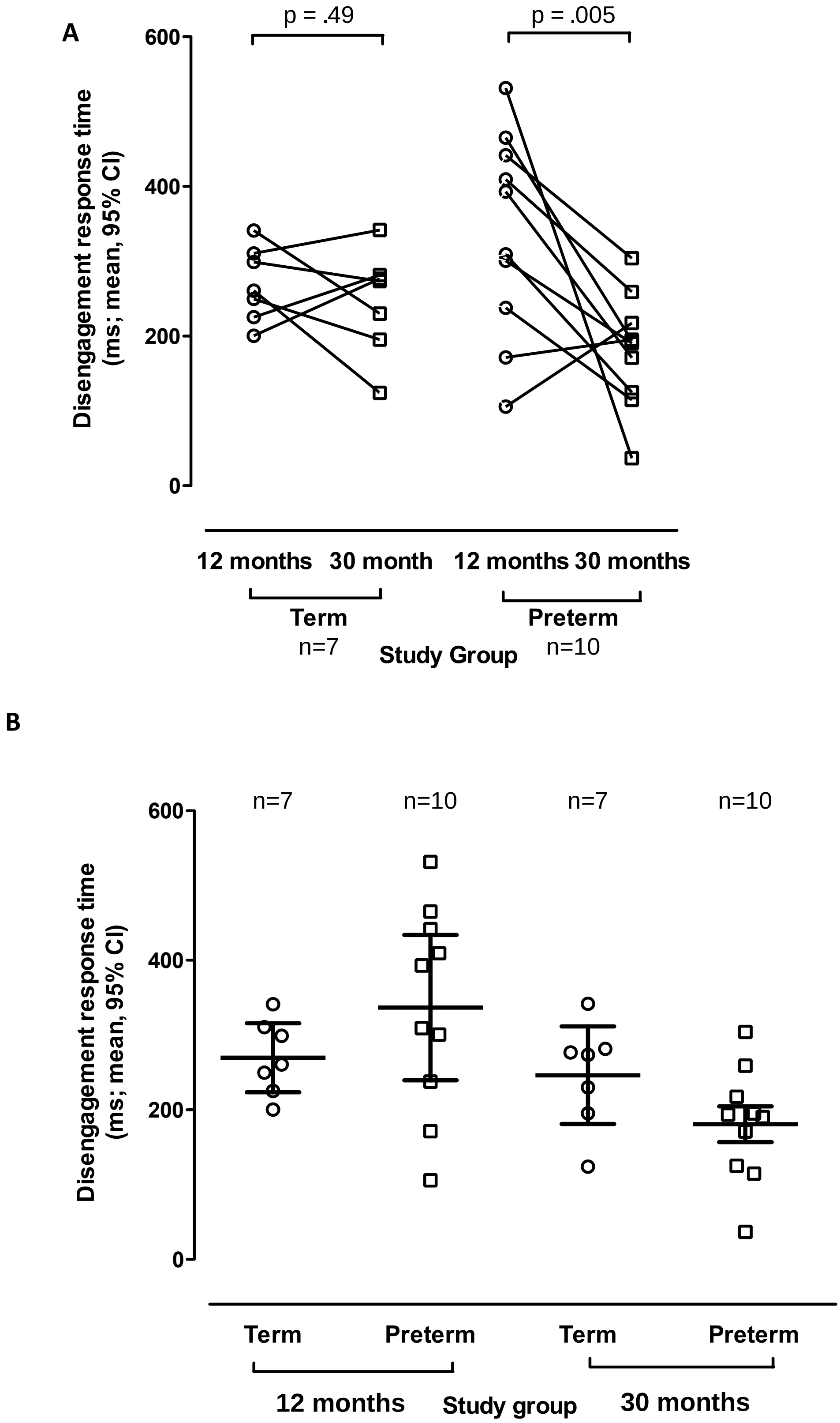
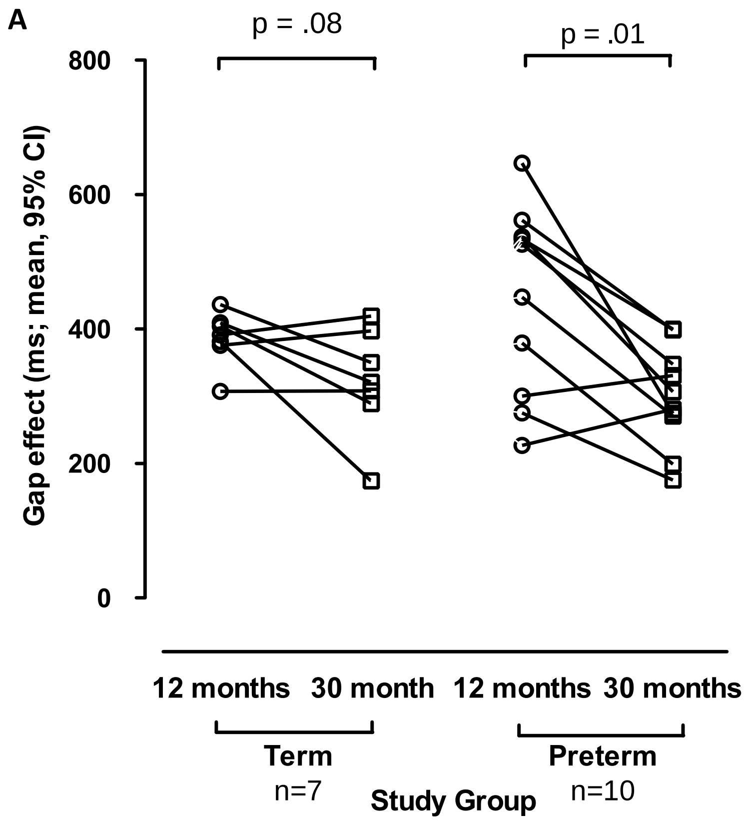


Figure 5-14. Two displays of the change of disengagement response times (ms) over the 12 to 30 month (A) illustrates the individual changes over time in each study group; (B) illustrates group difference by age to visually display how the mean responses change over time.



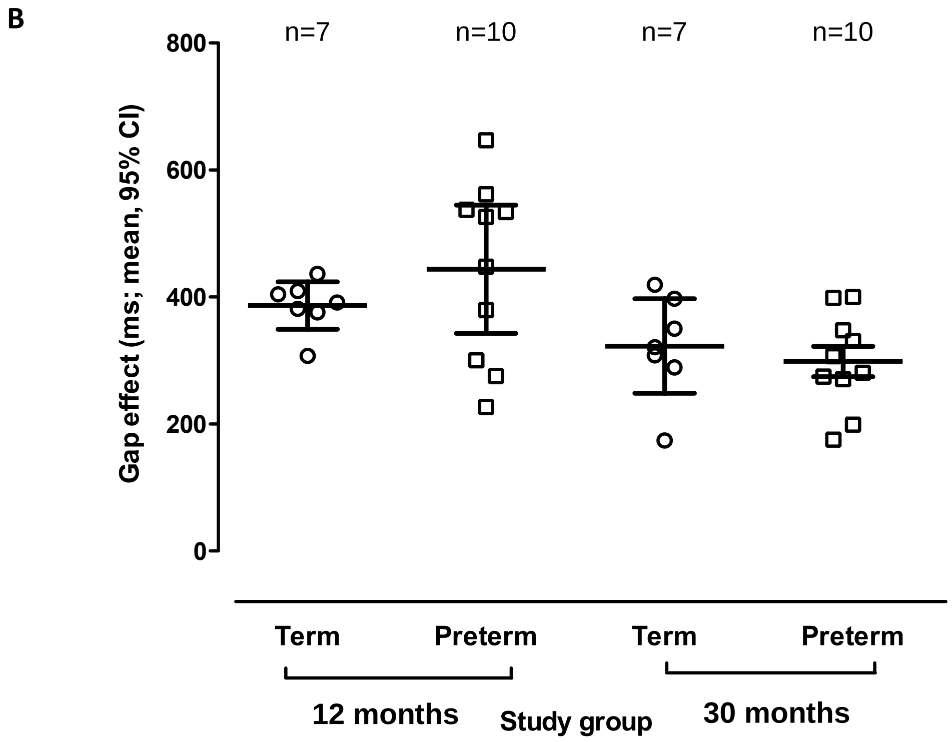


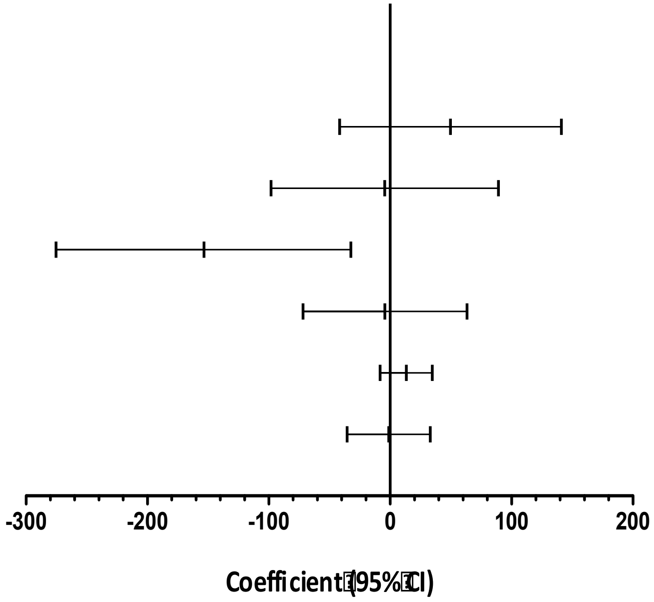
Figure 5-15. Two displays of the change in gap-effect (ms) over the 12 to 30 month assessments (A) illustrates the individual changes over time in each study group; (B) illustrates group difference by age to visually display how the mean responses change over time.

The first multi-level mixed effects model explored the disengagement RTs from 12 to 30 months. When adjusting for study group, male sex, IMD quintile, age of assessment and an interaction term between age and study group, a significant effect of prematurity at the 30 month time point was observed and continued to effect the model when additionally adjusting for 30m Bayley-III z-scores ((*-132.41, 95% CI (-248.67–--16.16), p = .03*) and (*-153.30, 95% CI (-275.26–-32.35), p = .01*) respectively). The more complex model including the adjustment for cognitive scores is reported in Table 5-19.

The second multilevel mixed-effects model explored the Gap-effect over the 12 and 30 month assessments. When adjusting for the same variables as the previous model, the interaction term did not have a significant impact on the gap-effect outcome and was therefore omitted. The gap-effect significantly decreased with age before and after adjustment for cognitive score ((*-111.65, 95% CI (-170.15–-53.155), p < .001*) and (*-103.78, 95% CI (-165.97–-41.60), p = .001*) with cognitive score adjustment). IMD quintile had a significant impact on gap-effect outcome, with a higher deprivation status predicting a longer gap-effect (*23.64, 95% CI (1.71–45.56), p = .04*), however, this effect is no longer significant at a 95% confidence level when adjusting for 30m cognitive scores (*21.32, 95% CI (-1.26–43.91), p = .06*). The more complex model including the adjustment for cognitive scores is reported in Table 5-20.

Table 5-19. Multi-level fixed effects regression model of disengagement RT across the 12 and 30 month assessments including the adjustment of the 30m Bayley-III z-score. The baseline group are term-born females at 12 months with an IMD quintile of 1.

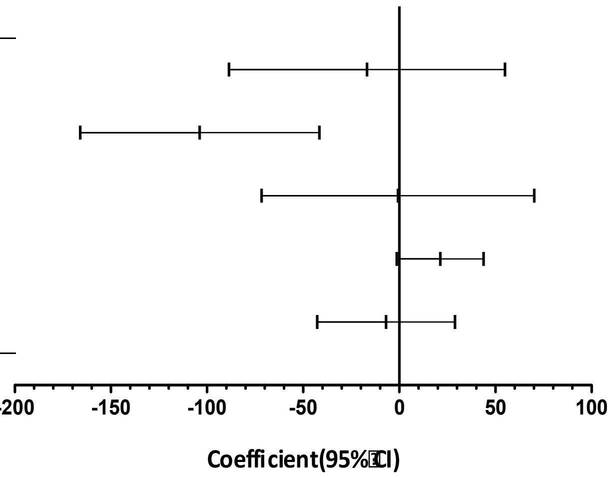
Predictor	Term (n=7)	Preterm (n=10)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	39^{+3} (38^{+5} – 40^{+2})	26^{+3} (23^{+6} – 29^{+4})	49.73	-41.57 – 141.04	.29
Age (30m)	266.60 (49.92)	167.14 (65.70)	-4.51	-98.20 – 89.18	.93
Study Group# 30m	-	-	-153.30	-274.25 – -32.35	.01
Male sex	2 (33.33)	7 (77.78)	-4.23	-71.82 – 63.37	.90
IMD Quintile	2 (1-5)	3 (1-5)	13.30	-8.22 – 34.82	.23
30m cog. Z-score	-.15 (.69)	-.22 (1.07)	-1.16	-35.34 – 33.01	.95
Disengag RT(Const.)	-	-	241.30	154.34 – 328.27	.000



Wald chi^2 = 18.40 (p>chi2 = .005)

Table 5-20 Multi-level mixed effects regression model of gap-effect RT across 12 and 30 month assessment including the adjustment of the 12m Bayley-III z-score. The baseline group are term-born females at 12 months with an IMD quintile of 1.

Predictor	Term (n=6)	Preterm (n=9)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	39^{+3} (38^{+5} – 40^{+2})	26^{+3} (23^{+6} – 29^{+4})	-16.76	-88.55 – 55.03	.65
Age (30m)	266.60 (49.92)	167.14 (65.70)	-103.78	-165.97 – -41.60	.001
Male sex	2 (33.33)	7 (77.78)	-.67	-71.61 – 70.27	.99
IMD Quintile	2 (1-5)	3 (1-5)	21.32	-1.26 – 43.91	.06
30m cog. Z-score	-.15 (.69)	-.22 (1.07)	-6.88	-42.74 – 28.99	.70
Disengag RT(Const.)	-	-	368.82	285.87 – 451.76	.000



Wald chi^2 = 14.20 (p>chi2 = .01)

5.3.3 6, 12 and 30 months

Infants with data in all three time points included 5 term-born and 8 VP infants (Table 5-21). Overall, there was a significant decrease in disengagement response time with age, once accounting for study group, male sex, SES and cognitive score at 30 months. Within the two study groups, the term born cohort displayed a general decrease in time to disengage with age; compared to the VP cohort who displayed an initial increase in RT to disengage at 12 months, before decreasing again at 30 month (Table 5-22; Figure 5-16). A similar pattern was displayed in the gap-effect (Table 5-22; Figure 5-17).

		Term (n=5)	***VP (n=8)***
Gestational age	*Median (range); weeks^{+d}*	39^{+1} (38^{+5} – 40^{+0})	26^{+2} (23^{+6} – 29^{+4})
Male sex		2 (40%)	7 (87.5%)
IMD Quintile	1	2	2
	2	2	0
	3	1	4
	4	0	2
	5	0	0
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=4; VP=7)	103.75 (6.29)	108.57 (12.15)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=4; VP=7)	-.32 (.51)	.07 (.98)

Table 5-21. Demographic details of infants in Gap-overlap task at 6, 12 and 30 months of age.

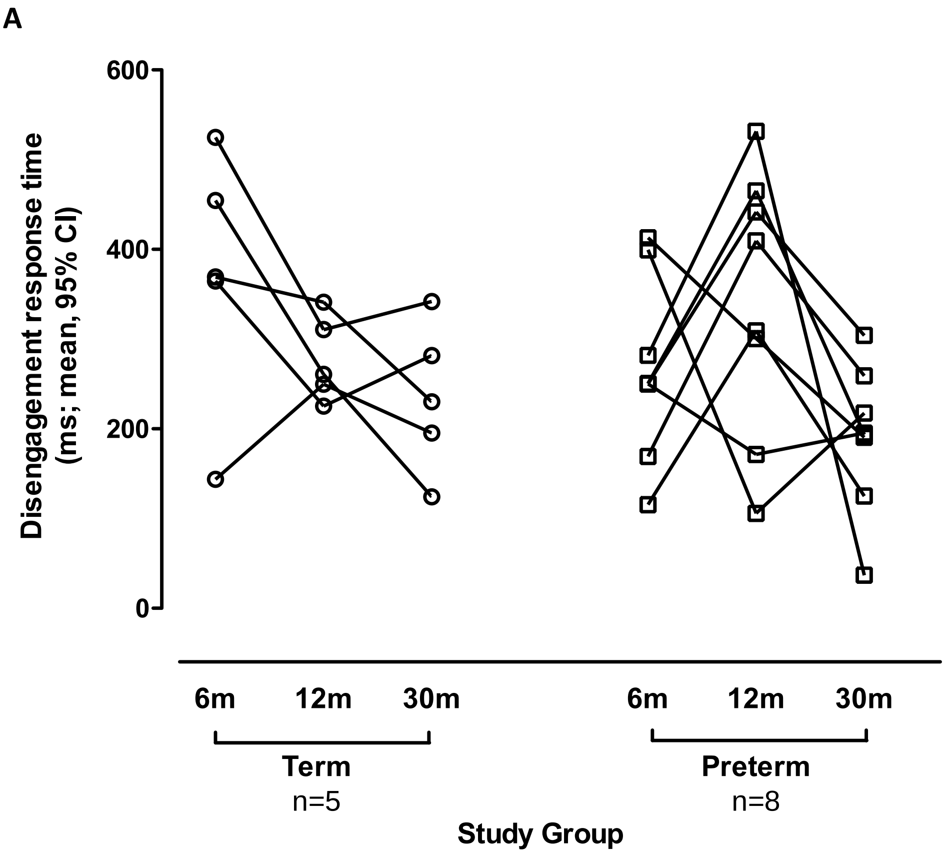
The final longitudinal investigation looked at the data across all three time points. For a direct comparison to the previous longitudinal data sets in the first instance, infants were excluded if they had <5 trials in each condition across any of the time points, and no imputations were carried out. This lead to a small sample size and the results in this section should only be considered exploratory (Table 5-22).

GAP task Variable			Term (n=5)	Preterm (n=8)
Disengagement RT	6m	Mean (ms)	371.42	266.24
		(SD)	143.51	101.77
	12m	Mean (ms)	277.6	341.88
		(SD)	47.25	147.95
	30m	Mean (ms)	234.67	190.29
		(SD)	82.87	81.31
Gap-effect	6m	Mean (ms)	490.92	371.18
		(SD)	161.07	109.84
	12m	Mean (ms)	378.79	452.91
		(SD)	41.33	142.06
	30m	Mean (ms)	302.45	312.79
		(SD)	87.53	69.41

Table 5-22. Mean and SDs of disengagement RTs and gap-effect variables across the 12 and 30m assessment ages within the two groups.

A repeated measures ANOVA for disengagement RTs over the 3 time points suggested a borderline main effect of age ($F(1.75, 19.27) = 3.42$, $p = .06$). There were no main effects of study group, nor any interaction effects.

When exploring the longitudinal relationship of the gap-effect outcome variable, a repeated measures ANOVA again produced a main effect of age ($F(1.59, 17.51) = 4.47$, $p = .03$). No main effect of group or any interactive effects were observed.



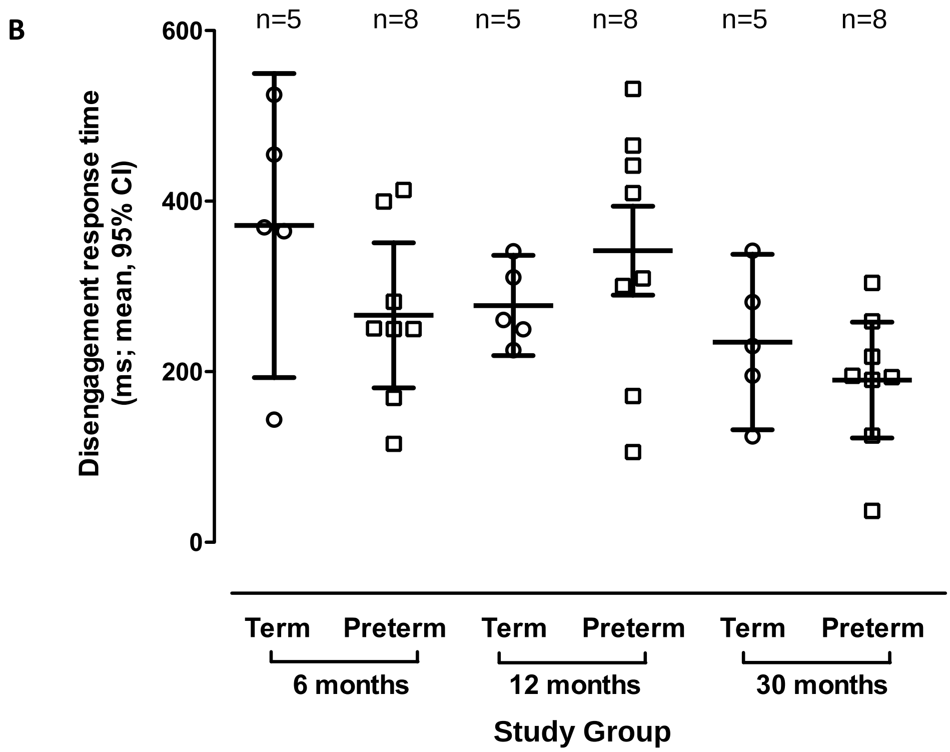
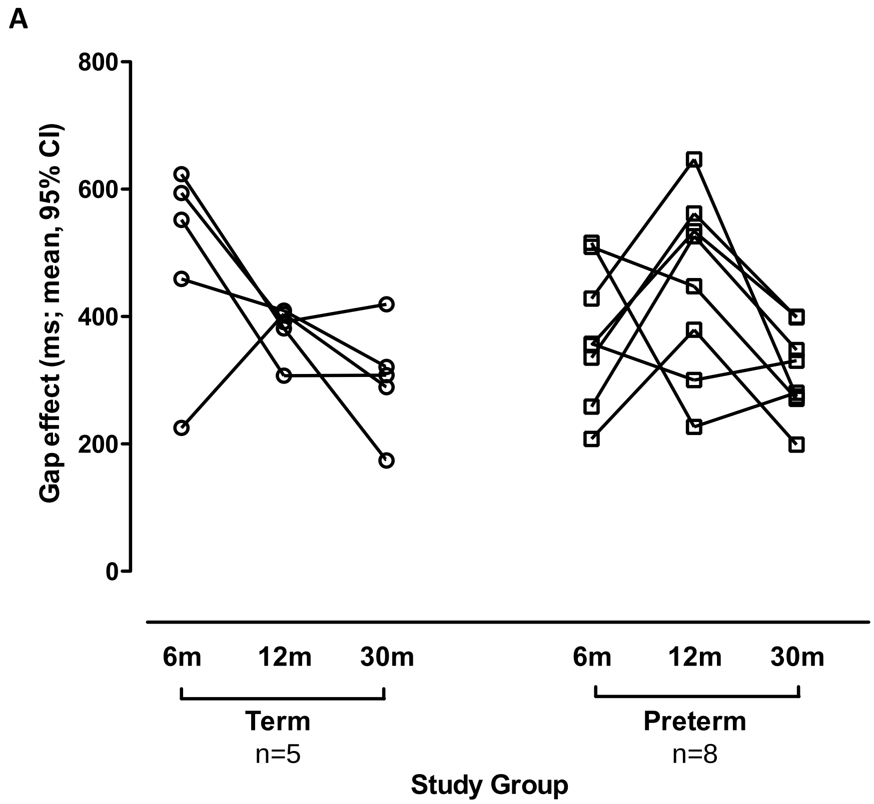


Figure 5-16. Two displays of the change of disengagement response times (ms) over the 3 assessments (A) individual changes over time in each study group (B) grouped by age to visually display how the mean responses change over time.



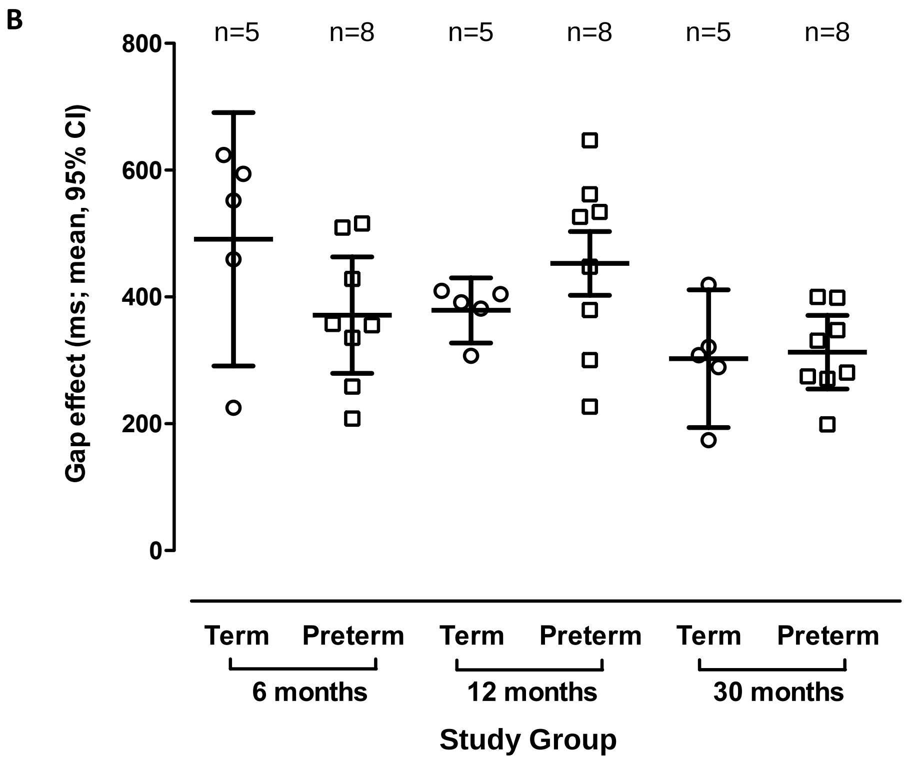


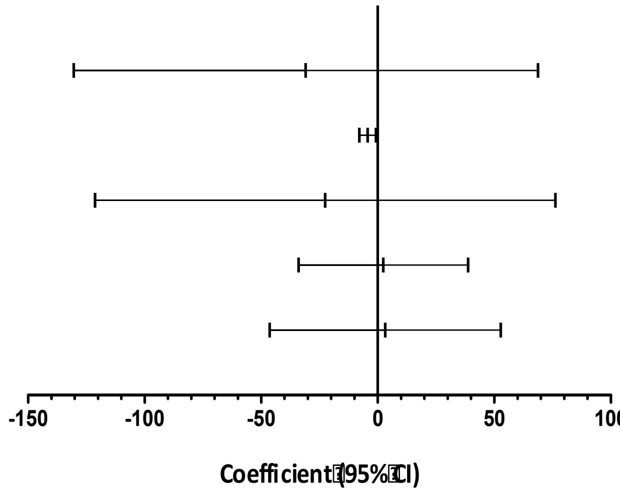
Figure 5-17. Two displays of the change in gap-effect (ms) over the 3 assessments (A) illustrates the individual changes over time in each study group; (B) illustrates group difference by age to visually display how the mean responses change over time.

The multilevel mixed effects models fitted to the disengagement RTs across the 3 time points, found a significant main effect of age when entered as a continuous variable before and after adjustment for cognitive performance ((*-4.59, 95% CI (-7.87–-1.31), p < .01*) and (*-4.30, 95% CI (-7.87–-.74), p < .02*) respectively). When treating age as a categorical variable, and setting the 6 month time point as baseline, there was a significant effect at of age at 30 months, again before and after adjustment for cognitive score ((*-99.33, 95% CI (-180.38–-18.29), p < .02*) and (-92.30*, 95% CI (-180.37–-4.22), p = .04*) respectively). The most complex model is again displayed in Table 5-23.

The multilevel mixed effects model for the gap-effect over the 3 time points, found a significant effect of age when treated as a continuous ((*-4.96, 95% CI (-8.31–-1.61), p < .01*) and (*-4.28, 95% CI (-7.80–-.76), p < .02*) with the adjustment of cognitive scores). When treating age as a categorical variable, it was again the 30 month time point that had a significant predictive effect on the gap-effect outcome (before adjustment (*-108.42, 95% CI (-191.27–-25.57), p = .01*) and after adjustment (*-91.32, 95% CI (-178.21–-4.42), p = .03*) respectively). The more complex model is displayed in Table 5-24.

Table 5-23. Multi-level mixed effects regression model of disengagement response time across 6, 12 and 30 month assessments including the adjustment of the 30m Bayley-III z-score in the primary longitudinal sample with infants that had complete 3 time point samples. The baseline group are term-born females at 6 months with an IMD quintile of 1.

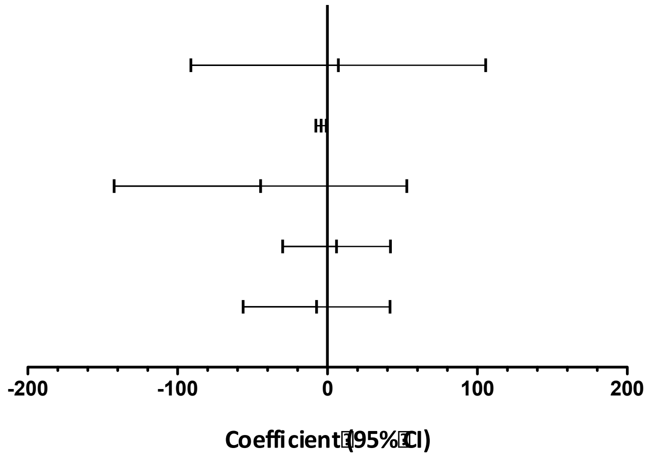
Predictor	Term (n=5)	Preterm (n=8)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	39+0 (38+5 – 39+4)	26+2 (23+6 – 29+4)	-30.85	-130.46 – 68.76	.54
Age	-	-	-4.30	-7.87 - -.74	.02
Male sex	3 (25)	18 (85.71)	-22.50	-121.25 – 76.25	.66
IMD Quintile	2 (1)	3 (1.5)	2.43	-33.95 – 38.81	.89
30m cog. Z-score	-.22 (.6)	.18 (.8)	3.24	-46.34 – 52.83	.90
Disengag RT(Const.)	-	-	368.88	250.28 – 487.48	.000



Wald chi^2 = 6.98 (p>chi2 = .22)

Table 5-24. Multi-level mixed effects regression model of gap-effect RT across 6, 12 and 30 month assessments including the adjustment of the 12m Bayley-III z-score in the primary longitudinal sample with infants that had complete 3 time point samples. The baseline group are term-born females at 6 months with an IMD quintile of 1.

Predictor	Term (n=5)	Preterm (n=8)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	39+0 (38+5 – 39+4)	26+2 (23+6 – 29+4)	7.22	-91.18 – 105.63	.89
Age	-	-	-4.28	-7.80 - -.76	.02
Male sex	1 (25)	6 (85.71)	-44.67	-142.22 – 52.89	.37
IMD Quintile	2 (1)	3 (1.5)	6.05	-29.89 – 41.98	.74
30m cog. Z-score	-.22 (.6)	.18 (.8)	-7.24	-56.23 – 41.75	.77
Disengag RT(Const.)	-	-	455.33	338.17 – 572.50	.000



Wald chi^2 = 7.12 (p>chi2 = .21)

The nature of multilevel models means missing data can be handled relatively efficiently. A second sample was therefore explored, including all infants who completed >5 trials per condition over a minimum of 2 assessment ages. The summary of this cohort can be found in Table 5-25.

GAP task Variable			***Term***	***Preterm***
Disengagement RT	6m	N	26	17
		Mean (SD)	312.38 (119.97)	279.20 (120.19)
	12m	N	22	18
		Mean (SD)	321.97 (91.01)	324.88 (111.88)
	30m	N	13	11
		Mean (SD)	249.58 (81.20)	182.53 (71.94)
Gap-effect RT	6m	N	26	17
		Mean (SD)	419.44 (120.38)	398.96 (122.66)
	12m	N	22	18
		Mean (SD)	432.10 (89.83)	435.58 (115.28)
	30m	N	13	11
		Mean (SD)	342.74 (100.27)	301.78 (71.98)

Table 5-25. Means of Disengagement RTs and Gap-effect in longitudinal sample of infants with 2 or more valid datasets.

A significant predictive effect of age at 30 months was observed in a mixed effects linear regression model for disengagement RT before adjustment for cognitive performance (*-73.67, 95% CI (-8.31–-1.61), p < .01*). Upon adjustment, an interaction term between study group and age had a significant effect on the model, with preterm birth at 12 months displaying a greater disengagement RT (110.36*, 95% CI (4.45–216.28), p = .04*). This model also suggested a main effect of prematurity (*-87.66, 95% CI (-162.81–-12.52), p = .02*), see Table 5-26.

Due to interaction term the model presented in Table 5-26 this becomes hard to interpret; therefore for clarity, the main effect of prematurity allowed for two separate models to be fitted to the data between the two study groups. Within the preterm model, a borderline effect of age at both 12 (70.78*, 95% CI (-1.97–-143.53), p < .06)* and 30 months (*-80.67, 95% CI (-160.65–-.68), p < .05)* was observed with the infants showing an increase in disengagement RTs from 6 month to 12 months, before decreasing at 30 months (Table 5-27). This effect of age was not observed within the term born infants (Table 5-28).

When exploring the gap-effect within this second cohort before when controlling for cognitive performance, a main effect of age was observed (*-3.67, 95% CI (-5.87–-1.47), p = .001).* A shorter gap effect at 30 months creating this age effect (*-81.44, 95% CI (-133.55–-29.34), p = .002).* Within this model, IMD quintile also had a significant predictive effect on the gap-effect outcome where greater deprivation status related to long gap effects (*15.11, 95% CI (.28*–29.95)*, p <.05)*. Upon adjustment for the cognitive performance, no interaction terms had a significant impact on the model parameters, however, the effect of age still remained (*-3.41, 95% CI (-5.82*– 1.00.)*, p = .006)*, again with this specifically driven with the shorter in gap-effect RT at 30 months (*-72.72, 95% CI (-131.10*–-14.34)*, p = .015)*. The more complex model is reported in Table 5-29.

All longitudinal models reported in the chapter were not significantly different from a standard regression model, indicating there are no group-level random effects.

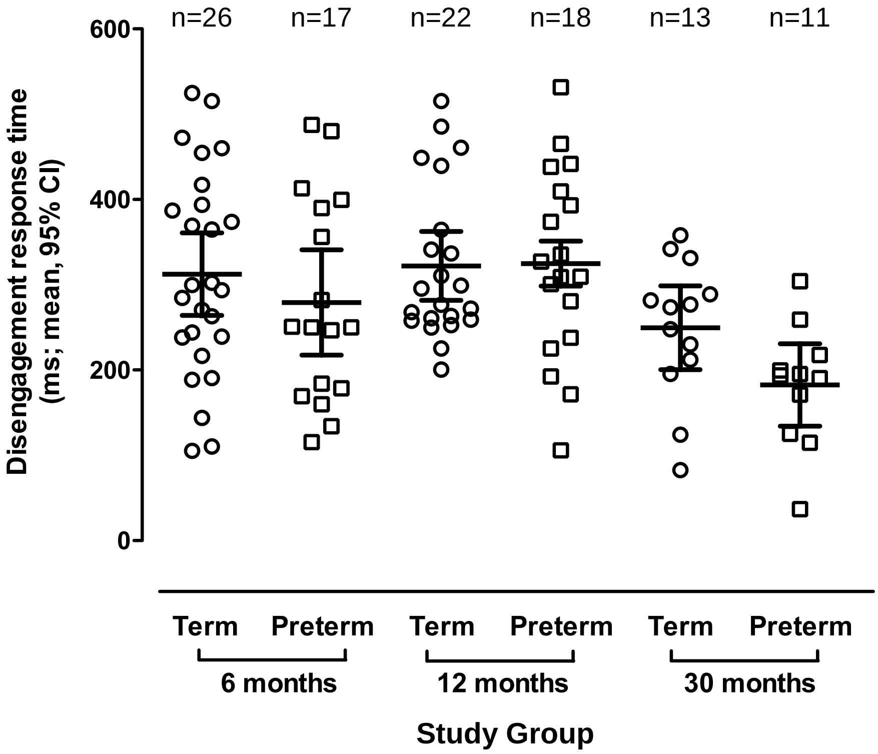


Figure 5-18. Disengagement response time (ms) of term and very preterm toddlers during the Gap-overlap task across the 3 assessments including infants with a minimum of 2 sets of data.

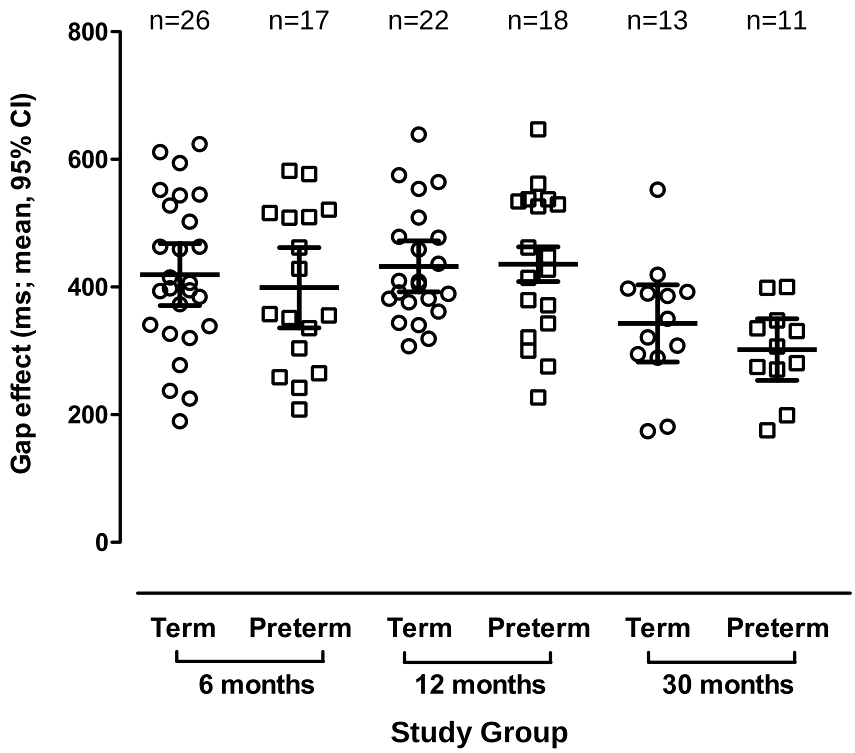
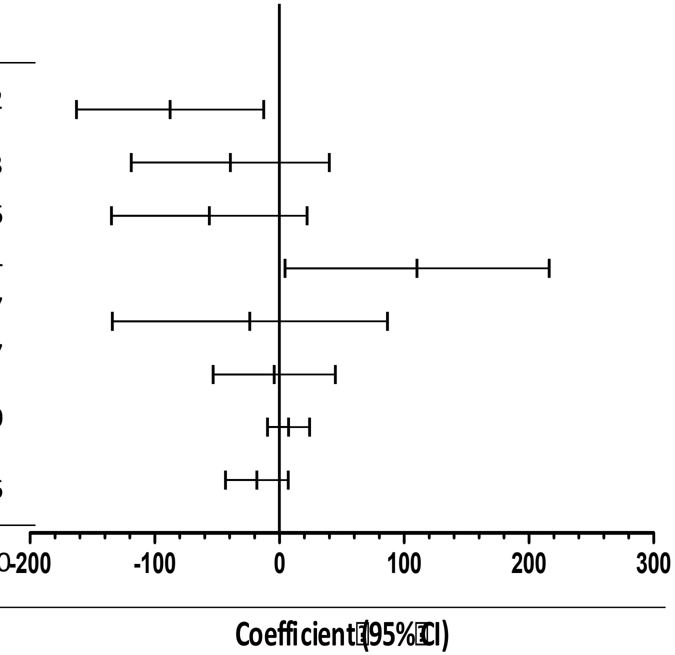


Figure 5-19. Gap-effect (ms) of term and very preterm toddlers during the Gap-overlap task across the 3 assessments, including infants with a minimum of 2 sets of data

Table 5-26. Multi-level mixed effects regression model of disengagement RT across 6, 12 and 30 month assessments including the adjustment of the 30m Bayley-III z-score in the second longitudinal sample with infants that had ≥2 valid time points, including interaction term between study group and age. The baseline group are term-born females at 6 months with an IMD quintile of 1.

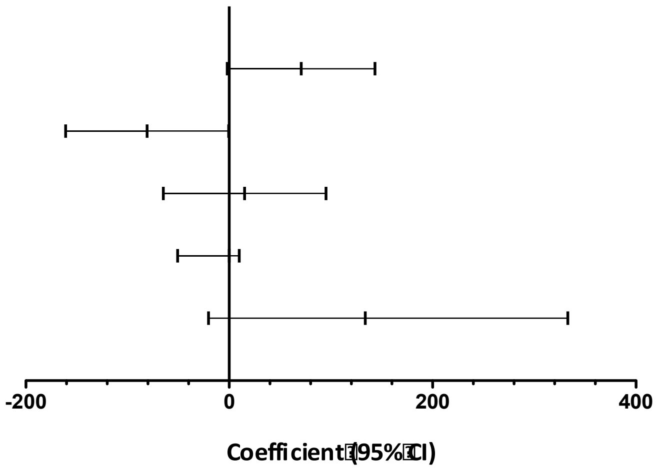
Predictor	Term (n=14)	Preterm (n=15)			
	Median (range)	Median (range)	Coef	95%CI	P
Study Group (SG)	40+1 (38+5 – 41+5)	26+2 (23+6 – 29+4)	-87.66	-162.81 - -12.52	.02
Age – 12m	-	-	-39.35	-118.72 – 40.01	.33
– 30m	-	-	-56.24	-134.71 – 22.22	.16
12m#SG	-	-	110.36	4.45 – 216.28	.04
30m#SG	-	-	-23.68	-133.95 – 86.60	.67
Male sex	7 (50)	11 (73)	-4.20	-53.19 – 44.80	.87
IMD Quintile	3.5 (3)	3 (2)	7.28	-9.56 – 24.13	.40
30m cog. Z- score	-.22 (.8)	.18 (1.2)	-18.11	-43.27 – 7.06	.16
Disengag RT(Const.)	-	-	314.77	235.84 – 393.70	.000



Wald chi^2 = 25.7 (p>chi2 = .001)

Table 5-27. Multi-level mixed effects regression model of disengagement RT across 6, 12 and 30 month assessments including the adjustment of the 12m Bayley-III z-score in the second longitudinal sample with infants that had ≥2 valid time points within the very preterm group only. The baseline group are females at 6 months with an IMD quintile of 1.

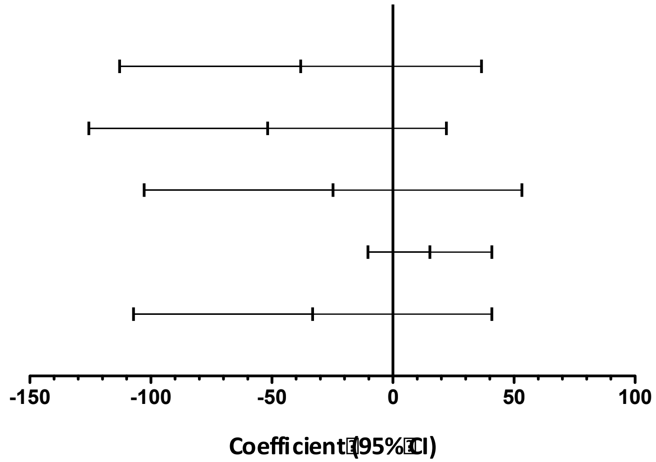
Predictor	Preterm (n=15)			
	Median (range)	Coef	95%CI	P
Age – 12m	-	70.78	-1.97 – 143.53	.06
– 30m	-	-80.67	-160.65 - -.68	.05
Male sex	11 (73)	15.22	-64.82 – 95.27	.71
IMD Quintile	3 (2)	.02	-50.57 – 9.92	1.0
30m cog. Z-score	.18 (1.2)	-20.33	133.96 – 333.09	.19
Disengag RT(Const.)	-	233.52	133.96 – 333.09	.000



Wald chi^2 = 15.41 (p>chi2 = .009)

Table 5-28. Multi-level mixed effects regression model of disengagement RT across 6, 12 and 30 month assessments including the adjustment of the 12m Bayley-III z-score in the second longitudinal sample with infants that had ≥2 valid time points within the term group only. The baseline group are females at 6 months with an IMD quintile of 1.

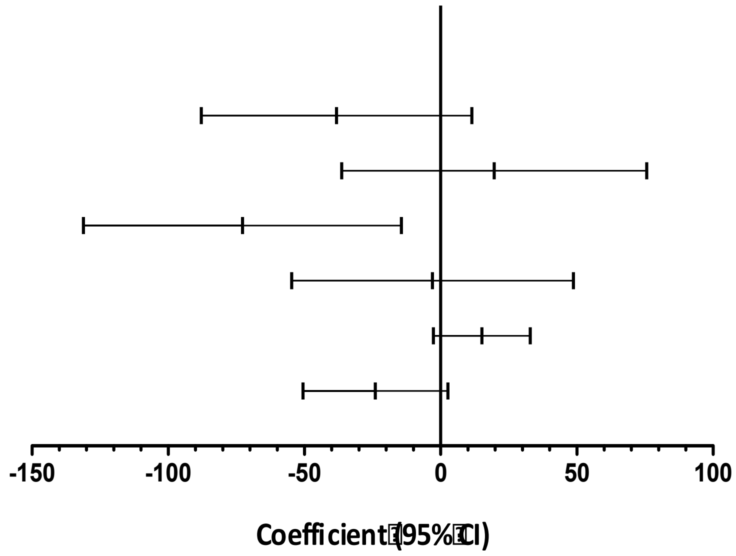
Predictor	Term (n=14)	Coef	95%CI	P
	Median (range)			
Age – 12m	-	-38.13	-112.90 – 36.64	.32
– 30m	-	-51.72	-125.57 – 22.13	.17
Male sex	7 (50)	-24.72	-102.78 – 53.34	.54
IMD Quintile	3.5 (3)	15.32	-10.26 – 40.90	.24
30m cog. Z-score	-.22 (.8)	-33.12	-107.10 – 40.87	.38
Disengag RT(Const.)	-	294.92	195.12 – 394.72	.000



Wald chi^2 = 4.13 (p>chi2 = .53)

Table 5-29. Multi-level mixed effects regression model of gap-effect RT across 6, 12 and 30 month assessments including the adjustment of the 30m Bayley-III z-score in the second longitudinal sample with infants that had ≥2 valid time points. The baseline group are term-females at 6 months with an IMD quintile of 1.

Predictor	Term (n=14) Median (Range)	Preterm (n=15) Median (Range)	Coef	95%CI	P
Study Group	40+1 (38+5 – 41+5)	26+2 (23+6 – 29+4)	-38.20	-87.87 – 11.47	.13
Age – 12m	-	-	19.69	-36.32 – 75.69	.49
– 30m	-	-	-72.72	-131.10 - -14.34	.02
Male sex	7 (50)	11 (73)	-2.97	-54.72 – 48.79	.91
IMD Quintile	3.5 (3)	3 (2)	15.16	-2.66 – 32.97	.10
30m cog. Z-score	-.22 (.8)	.18 (1.2)	-23.91	-50.47 – 2.65	.08
Disengag RT(Const.)			377.93	302.86 – 452.99	.000



Wald chi^2 = 16.80 (p>chi2 = .01)

5.4 Discussion

This chapter investigated differences in visual attention between term and preterm born infants at 6, 12 and 30 months of age. The Gap-overlap paradigm administered utilised eye-tracking technology to track the speed of ocular saccades to a peripheral target from a central stimulus. The paradigm comprised of 3 conditions: the gap condition, where the target appeared after a short interval following the central stimulus disappearance; the baseline condition, where the target appeared simultaneously to the central stimulus disappearance; and finally the overlap condition, where both the central and peripheral targets were presented concurrently. Previous research has found an effect of condition on the saccadic response times within the paradigm, with the gap condition eliciting the fastest responses, and the overlap creating the greatest delay (Hood and Atkinson, 1993). This was replicated in the current investigation, with this pattern presenting at all time points within the current study.

The overall reduction in speed over time to all conditions is consistent with previous findings within the literature. However, cross-sectional investigations did not produce any main effects of group or any group by condition interactions, suggesting no clear sign of visual attention dysfunction within the VP infants at any age. When exploring the longitudinal relationship across the assessment ages to explore how performance changes with time on the task, no significant effects of age were observed within the first model, exploring the 6 to 12 month RTs. However, when looking across the 12 to 30 month data, there was an interaction observed between the study groups and age of the disengagement RTs; the VP infants showed a greater reduction in RT from 12 to 30 months. When modelling the full 3 time points and including all infants that had 2 or more valid data sets, it was the VP infants who displayed significant changes across assessments in their disengagement behaviour. A significant increase in response time at 12 was observed before a decrease again at 30 months within the VP group. This difference in study group performance is only apparent once controlling for cognitive performance. Upon adjustment, the VP infants displayed the greatest variation in

RTs over the assessments compared to the term infants who remained relatively consistent.

When looking at the gap-effect variable (the difference between the fastest and slowest trial responses), the effect of age was a main effect, showing a decrease in response across both groups between 12 and 30 months, and again when looking across all 3 time points. There were no interaction effects when exploring this variable suggesting the effect of the gap, even if not significantly so within each age point, was slower in the response of the VP infants, leading to no group differentiation in the gap-effect variable.

The main variable investigated throughout was the disengagement reaction time (RT). This variable was calculated with the subtraction of the baseline responses from the overlap trial responses, adjusting for individual variation and producing a measure of the extent the overlap trial challenges the infants' attentional system. The overlap trials create attentional competition as the infant is required to look away from the more visually pleasing central stimulus to the target in the periphery. Failure to disengage from this stimulus has been interpreted as a deficit in young children (Hood and Atkinson, 1993), and has been compared to prolonged fixation behaviour, termed 'obligatory attention' reportedly found in adults (Stechler and Latz, 1966). The delay created from the overlap trials is thought to be caused by the initial engagement with the central stimulus before the presentation of the peripheral target. This engagement prevents the ocular movement to the periphery and it has been suggested that attentional processes are required to be disengaged before the shift in gaze can be made. Following the saccade, the visual attention is then reengaged with the peripheral target (Posner and Petersen, 1990; Fischer and Weber, 1993). Removal of the central stimulus increases the speed of saccades to the periphery, as this automatically disengages the attention and allows saccades to the next visual movement on the screen. These theories correspond to the RTs produced previously and within the current study to the 3 conditions in the Gap-overlap paradigm.

In children with autism, attentional deficits, specifically visual attention and attentional shifting, have been reported and are thought to correlate with the social-communication difficulties typically observed in this population (Elsabbagh *et al.*, 2009). A large study, BASIS, investigating young autistic populations and those at high-risk of the disorder, used the Gap-overlap paradigm and have repeatedly found longer RTs to the disengagement within the high risk and vulnerable groups (Elsabbagh *et al.*, 2009; Green *et al.*, 2015). When this group trialled intervention procedures to help with later known problems within this disorder, the disengagement RTs decreases, speculated to be a sign of positive response from the intervention (Green *et al.*, 2015). This pattern of results has also been observed in health controls following executive training (Wass, Porayska-Pomsta and Johnson, 2011). The suggestion within these studies is that a slower disengagement response is reflective of later atypical attention, and appears to be the strongest theory within the literature.

In terms of the exploration of preterm attentional abilities, literature searches did not produce any studies utilising the Gap-overlap paradigm within this population. However, it has been shown on a number of occasions that preterm infants are at risk of later attentional difficulties (Mulder *et al.*, 2009; Anderson, 2014). These difficulties have been associated with lower achievements later in life (Sigman *et al.*, 1991; Rose, Feldman and Jankowski, 2001).

A number of studies have found disengagement times to be slower in high-risk preterm cohorts (Landry *et al.*, 1985; Rose, Feldman and Jankowski, 2001, 2002), whilst others, investigating of low risk preterm cohorts, report response times to be faster than that of term controls within the first year of life (Foreman *et al.*, 1991; Hunnius *et al.*, 2008). Studies that have found these reduced RTs theorise that the early exposure to visual stimuli created by premature birth leads to faster maturation of the visual system in the infants that are not as high risk (Hunnius *et al.*, 2008). The perinatal complications are likely to be higher in high risk cohorts and may confound these developments in high risk infants (Hitzert *et al.*, 2014). Preterm infants have also been reported to have different patterns of exploratory behaviour in terms of investigating their environment. Previous literature has suggested these infants can spend less time examining objects around them (Landry and Chapieski, 1988), with this potentially having knock on effects to later cognitive development (de Haan *et al.*, 2000).

Although it is commonly accepted that the shorter the look, the more efficient the information processing, as seen in habituation paradigms and in accordance to the theories above (Rose, Feldman and Jankowski, 2001, 2002), an alternative theory has been suggested in the case of disengagement times. In a study by van der Geest, *et al.,*(2001) the gap-overlap paradigm was utilised to explore again, the visual attentional difficulties in children on the autistic spectrum at the age of 10 years (van der Geest *et al.*, 2001). This group found that autistic children displayed faster reaction times in the gap-effect to the gap-overlap task compared to healthy controls in contrast to previous findings. The authors concluded that instead of an inability to disengage, these children displayed a poorer engagement of attention to the central stimulus in the first instance, and therefore the disengagement response is faster and saccades quicker than that of the term controls (van der Geest *et al.*, 2001). They propose this lack of engagement could offer a possible explanation for high saccadic frequency, previously reported in autistic individuals (Kemner *et al.*, 1998). However, this difference was not observed within the disengagement RTs in previous publications and the results could be reflective of slower saccadic reactions to the gap trials.

The literature presents with conflicting theories regarding the explanation of saccadic responses to visual attention tasks. When reflecting on the results of this longitudinal observation of a high risk VP cohort, it is possible that both theories have merit. If considering the first year of assessments, although not significantly different from the term controls at each age point, the VP cohort presented with a significant increase at 12months from the 6 month in the gap-effect and disengagement RTs. This could be an indication of latter attentional problems as seen within the studies of autistic populations discussed above. However, at 30 months, the preterm response decreases significantly in line with term controls.

This could be a possible indication of engagement deficits beginning to emerge within the cohort, as suggested by van der Geest et al. (2001). Given the support in the literature for slow disengagement as an earlier indicator of later cognitive deficits, the data collected at the 12 month assessment may be a better identifier of later attentional deficits.

Overall the responses of visual attentional system assessed with the use of the gap-overlap paradigm appear to be more variable within the preterm cohort compared to term born controls. Although both groups consistently responded within statistically similar timeframes across the 3 time points, the changes from one time point to the next were greater within the preterm infants when controlling for global cognitive abilities. The overall performance of the children on the task was consistent with previous research, replicating the response time patterns to each condition at each age. In terms of predictive validity of the assessment, it will be crucial to correlate the data collected with cognitive abilities and visual attentional performance later in life.

Chapter 6 Information Processing

Processing speeds have been found to be associated, if not fundamental, to EF abilities in both typical and atypical populations (Kail and Salthouse, 1994; Mulder, Pitchford and Marlow, 2010). A historical view within the literature associates efficient processing to superior cognitive domain performance over time, as it is has been interpreted as a marker of domain experience; the more experienced a domain, the more established the neural links and thereby the faster and better the performance (Chi, 1977; Kail and Salthouse, 1994). An alternative, but not unrelated theory, considers a more global perspective reflective of the developmental changes in the brain, such as myelination and synaptic pruning (Luna, 2009). These models allude to the same outcome however, both highlighting the importance of processing speed in the interpretation of global cognitive abilities.

Developmental improvements in processing speeds have been proposed to be significant in infancy and early childhood, continue to improve in middle to late childhood and become 'asymptotic' in adolescence (Kail and Salthouse, 1994; Kail and Ferrer, 2007), seemingly a non-linear trajectory (Kail and Ferrer, 2007). In infancy, information processing is often assessed using visual paradigms with 'length of look' evaluated; visual recognition memory and habituation tasks are frequently utilised (McCall and Carriger, 1993; Kavsek and Bornstein, 2010; Rose, Feldman and Jankowski, 2012). It is theorised that if an infant spends a prolonged time looking at an image, their visual processing is slower than that of a child with a shorter look (Kavsek and Bornstein, 2010). It has been proposed that habituation reflects both memory consolidation and learning, with speed of processing of information clearly linked to both abilities (Colombo and Mitchell, 2009). An alternative paradigm to demonstrate learning within this literature is that of the conjugate reinforcement paradigm by Rovee and Rovee (1969) and will be further explored later in this chapter (Heathcock *et al.*, 2004).

In childhood and into adolescence, processing speed is typically evaluated using manual response time tasks. For example, a match-to-sample paradigm, a visual search task whereby the participant must find a match to a target and performance is defined according to the quantity found within 1 minute (Rose *et al.*, 2012) has been used to investigate processing speeds in 11 year olds. However, assessments such as these can be confounded with fine motor performance difficulties particularly in early childhood but also more generally (Kyllonen and Zu, 2016; Ebaid *et al.*, 2017). A relatively new approach to limitations such as these, is the use of touchscreen technology (Pitchford and Outhwaite, 2016). This medium eliminates the need for precise motor abilities and those as young as two have been shown to have the acquired motor abilities for this assessment structure (Nacher *et al.*, 2015). This medium is explored with the preterm cohort later in the chapter.

Associations have been made between reduced processing speeds and the academic difficulties and behavioural patterns observed following preterm birth (Mulder, Pitchford and Marlow, 2011a). These observations associated to poor academic achievement have been primarily from later childhood studies (Mulder, Pitchford and Marlow, 2010, 2011b; Aarnoudse-Moens *et al.*, 2013). Although there have been reports of reduced processing speed in infancy, it is not clear whether problems observed in the early years are directly translated to the later observations (Rose, Feldman and Jankowski, 2009; Rose *et al.*, 2012). As with many cognitive areas the assessment of processing speed in the toddler years is limited (Rose, Feldman and Jankowski, 2009). The work by Rose *et al.* (2009) begun to address this gap by exploring the processing speeds difficulties reported in ex-preterm populations, and found an element of continuity in the problems reported. The group later proposed a model that incorporated processing speed, memory, attention and representational competence that predicted global outcomes at 11 years (Rose *et al.*, 2012). Although minimal, this provides evidence to suggest it may be possible to identify deficits as early as infancy that later predicts performance outcomes.

Preceding conscious cognitive effort is the speed in which the brain handles sensory information (Escera *et al.*, 2000). Information from sensory systems is initially required to be processed and collated. This is considered a ‘lower-order’ cognitive process that underlines ‘higher-order’ functions such as EF (Demetriou *et al.*, 2002; Aarnoudse-Moens, Smidts, *et al.*, 2009). In preterm infants, the white matter tract abnormalities often reported have been associated to slow processing of information across brain regions, leading to poorer cognitive performances and behavioural difficulties (Aarnoudse-Moens, Smidts, *et al.*, 2009). It is therefore possible that detection of slower neural processing speeds could be a predictor of later cognitive function.

Event-Related Potentials or ERPs have long been considered the most temporally accurate method of measuring the speed of neural processes. ERPs are electroencephalography (EEG) waves time-locked to a specific stimulus. The oscillating electrical impulses naturally produced by the brain at rest are interrupted when presented with new sensory information, such as a short sound burst, generating a change the electrical activity across the brain (EEG waves). When the stimulus presentation is time-locked to the electrical activity (ERPs) the speed of transfer of information across multiple brain networks can be evaluated. This technique allows the speed of transient changes within the brain to be compared across populations (Woodman, 2010; Sokhadze *et al.*, 2017).

Involuntary attention is the initial detection of sensory information by the brain and can be observed in ERP responses. For auditory stimuli, a component is elicited approximately 100ms after a sound is played, reflecting this involuntary processing (Escera *et al.*, 2000). ERP paradigms are often considered the most feasible method for assessing neurological function in infants and young children due to the versatility of the technique. Exploration of the neural response to sound has been used to assess processing speeds in infants as it does not require active focus. In a preterm cohort at 12 months of age, correlations have been reported between auditory ERP components differences and later developmental scores compared to infants born at term (Fellman *et al.*, 2004). This group later observed differences in the same cohort at 5 years of age, leading them to conclude that the primary auditory processing could be impaired within the cohort (Mikkola *et al.*, 2007). A trend is emerging within the preterm literature, with the implication that the auditory delays being observed could be impacting language development and may explain the delays frequently reported in the population (Jansson-Verkasalo *et al.*, 2003, 2010; Hövel *et al.*, 2014). As alluded to in Chapter 1, section 1.4, language delays have been correlated to reduced neural volume in regions connecting the auditory cortices (Northam *et al.*, 2012). By further investigating the auditory processing in terms of auditory attention, it could shed further light on this topic.

Both behavioural and neuropsychological techniques were implemented during the PDP to investigate performance speed differences in both EF and attentional measures within the two cohorts.

6.1 Behavioural processing speeds

The behavioural techniques used within the PDP battery investigate information processing alongside EF abilities. Both the speed of response and overall performance was investigated. The measure from the first year utilised the conjugate reinforcement paradigm to investigate the speed of learning. In the second year, the toddlers were assessed on a newly developed touchscreen base application, the BabyScreen application, and an established paradigm from the literature, the Multi-location Multi-step task.

The classic mobile paradigm (Rovee and Rovee, 1969) was selected as it has previously been used to explore learning in preterm cohorts. Preterm infants have demonstrated a reduced level of learning compared to that of term born controls in a number of studies (Heathcock *et al.*, 2004, 2005; Haley *et al.*, 2008). The criterion used in Haley *et al.,* (2008) was used to determine whether the infants learnt during the task, in order to make the work comparable to results found previously in the field. The speed in which the infants achieved this criterion was also investigated by averaging the number of kicks per minute over the duration of the task.

The BabyScreen app, designed by Deirdre Murray and colleagues (Twomey *et al.*, In press) used a touchscreen tablet to assess multiple domains of executive function, in combination with the speed in which toddlers could process and respond to visual stimuli. The BabyScreen application used multiple touch responses to assess learning and speed of processing within the different EF domains through a novel app that the children had not seen before. The instructions for the app were minimal and very little verbal communication was given during administration. The premise of the design prompted the infant to 'discover' what to do as they interacted with the task. It was theorised that the process of discovery demonstrated a greater representation of naturalistic EF. The task also did not require highly precise movements and recorded touch response type and time automatically. The rationale behind the design was to remove any biases caused by receptive language delay and fine motor difficulties, the common limitations to traditional cognitive tests in use today (Twomey *et al.*, In press).

The 'multi-location multi-step task' (Zelazo, Reznick and Spinazzola, 1998) is an advanced version of the A-not-B task (Piaget, 1954) and was carried out at the 30 month time point. Zelazo et al., modified the typical A*B* task, to see if perseverative errors often observed with the first year after birth continue at the age of 2 years (Zelazo, Reznick and Spinazzola, 1998). In the study in 1998, perseverative errors were observed, postulated to be a sign of cognitive inflexibility in relation to motor learning. In a similar structure to the A*B* paradigm, the MLMS apparatus included multiple wells in which the snack or toy was hidden. However, rather than a simple request to the child to find the snack or toy, multiple actions had to be completed before the snack or toy could be obtained. This task has been analysed based on time to correct response, with this reflecting information processing capacity within the EF task. Given the inclusion of the AB paradigm at the assessment at 12 months of age, this task was a logical continuum when exploring EF performance at 30 months of age.

6.1.1 Methods

6.1.1.1 Conjugate Mobile Reinforcement Paradigm: Apparatus and procedure

The infants were placed supine, at a 45 degree angle, into a Graco Travel Cot, with the dimensions 80cm x 26cm x 26cm. Attention of the infant was directed to the mobile by covering the sides and base of the cot in white sheets to remove any distractions from outside. The mobile was positioned in the bottom left hand corner of the cot, and was suspended at a height of 50cms above the cot. The infants' legs were in line with a mobile in the bottom left hand corner and their head in the top right hand corner.

Two mobiles were used during the task, the order of which was counterbalanced equally within the two cohorts. The first of the two mobiles was approximately 34cm in length, with four three dimensional soft toys with bells hanging from a circular hoop which was suspended from the mobile stand. The second mobile was 42cm long and composed of 3 strands of two dimensional cardboard shapes of multiple colours with bells that hung from the bottom. The two mobiles were intentionally visually very distinctively as in this version of the task the mobile is changed in the third phase to test the infants understanding of their learnt skill.

The task comprised 3 test phases. The first, the baseline condition, was 2 minutes in length, and involved attaching a stationary ribbon to the infants' right ankle. The ribbon was considered stationary as it was attached to the cot at the base of the mobile stand. This phase recorded the frequency of natural leg kicks produced per minute by the infant. The second phase, the training phase, was 6 minutes in length and required switching to a ribbon attached to the top of the mobile. A kick from the infant now caused a displacement of the mobile. This ribbon was maintained for the final phase of the task, the generalisation phase. This condition required the removal of the mobile that had been in place for the baseline and training conditions and was replaced with the second visually distinctive mobile. This final phase was 2 minutes in duration. The ribbon length was adjusted to ensure it remained taut for all phases of the task.

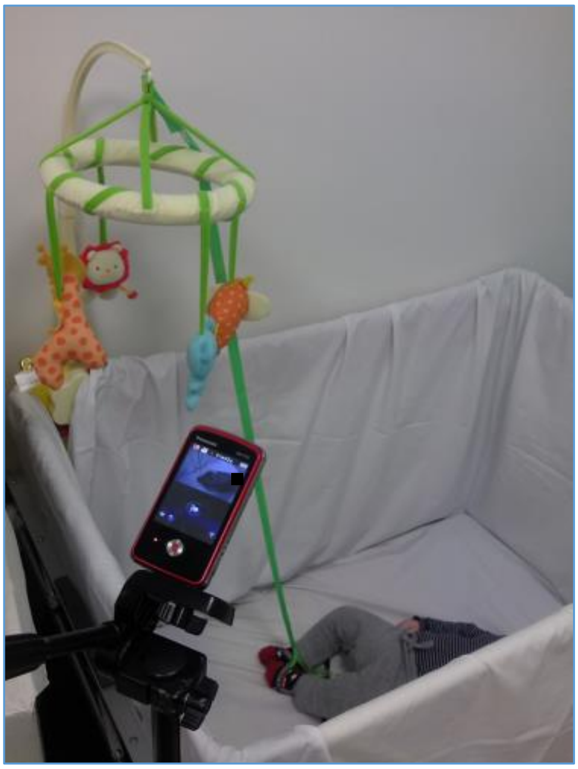


Figure 6-1. Photograph of the conjugate mobile reinforcement paradigm lab set up.

A video camera (model: Panasonic HDMI HM-TA2) was positioned in the bottom right hand corner of the cot, and was adjusted to ensure the infant was in full view during the recording.

Coding and Analysis:

Each phase of the task was video recorded; offline coding recorded the number of independent and simultaneous leg kicks per 30s intervals within each phase of the task. Although recorded, the simultaneous leg kicks were not considered in the analysis. A kick was defined as an extension or flexion of the hip and/or knee joints (with enough force to create movement on the ribbon), which returns along the route from which it came (Rovee and Rovee, 1969).

To account for the variation in natural baseline movements, a relative response ratio (RRR) was calculated for each minute of the training phase (training leg kick response per minute/ average baseline leg kick response per minute; see Figure 6-2). In order to categorise performance within the mobile reinforcement paradigm, other studies within the literature have defined ‘learning behaviour’ by calculating the RRR and employing a criterion alongside this that infant has to meet to be considered to have learnt during the paradigm (Haley *et al.*, 2008). The criterion states the infant must display a ratio of 1.5 (50% increase from baseline response) in two consecutive minutes of training to achieve a ‘learner’ status (Haley *et al.*, 2008). If this criterion is not met the infant is considered a non-learner in this task. To ensure the data is comparable to results from previous studies the same criteria was employed here. Accordingly, the infants were therefore categorised into 4 groups in the analysis of this paradigm: term learner (TL); term non-learner (TNL); preterm learner (PL); preterm non learner (PNL).

$$\text{RRR} = \frac{\text{Number of ribbon leg kicks during each minute of training/generalisation Phase}}{\text{Average ribbon leg kicks during baseline}}$$

Definition of learning:
Learner = RRR ≥ 1.5 for 2 consecutive minutes of the training phase

Figure 6-2. Relative response ratio calculation, used to remove baseline leg kick bias. The RRR was calculated for each minute of the training phase and of the generalisation phase. The learning definition was defined only by performance within the training phase.

Interrater reliability was performed and produced a reliability coefficient of .97 when double coding 25% of the data.

RRR has been criticised following its use in previous studies. Those with higher nature baseline kick frequencies are potentially at a disadvantage when needing to meeting this criterion as they are required to kick at an even greater frequency to achieve the specified 1.5 ratio (Millar and Weir, 2015). Furthermore, preterm infants have previously been reported to display a higher spontaneous kick frequency compared to term controls (Heathcock *et al.*, 2005). This suggests the disadvantage discussed by Millar and Weir (2005) could bias the groups that we are investigating here. In response to this, baseline kick rates were investigated by study group to confirm no significant differences were present dependent on gestation at birth. Baseline data was not normally distributed and could not be corrected with transformation, this was therefore investigated using a Mann-Whitney U-test.

An odds ratio was used to investigate relationship between learner status and study group (term vs preterm and learner vs non-learner). Speed of learning was determined by comparing response rates between the different learning status groups (learner vs non-learner) in 2 minute intervals across the training phase. A mixed effects model was also fitted to the data to further explore the relationship between the groups including the previously stated confound variables determined in chapter 2. Initially a single, mixed effects model was fitted to the data including: the 'study group/learning status' 4-way between-subject grouping variable; time (repeated within-subject predictor, detailing responses from the 2 minutes of baseline, and the training phase split into 3 epochs of 2 minutes) and the other stated predictors: male sex and IMD quintile. However, the best model fit included an interaction between the group status and time, as determined by a likelihood ratio test. This produced a complex 4-way interaction that was difficult to interpret, therefore for clarity, the learning status groups were separated and differences in study group were explored within each model. Reported in the subsequent sections are two mixed effect models for each learning category, adjusting for study group; male sex; IMD quintile; and time. As previously, these relationships were explored before the additional adjustment of cognitive z-scores at 12 months. Post-hoc t-tests were then run to examine any interactions.

To investigate whether those that met the learning criterion displayed a specific increase in kick rate on the leg attached to the mobile, leg kicks were analysed independently across the 4 groups (term-learner; term-non-learner; preterm-learner; preterm-non-learner). Kick rate difference was computed by subtracting the non-ribbon leg from the ribbon leg, with a positive kick rate difference indicating a higher level of kicking on the ribbon leg. The kick rate difference at baseline was compared to the difference in the last two minutes of training. As before, the two learning status groups (learner vs non-learner) were analysed independently to investigate the relationship between study groups. Mixed linear regression models were fitted to both groups adjusting for study group, time (kick rate difference at baseline and at the end of training), male sex and IMD quintile. Independent models were then run with the additional adjustment of cognitive z-scores.

RRR were calculated again during the generalisation phase. Employing the same criterion as before, infants who displayed a RRR of 1.5 for the 2 minutes of the generalisation phase were considered to have transferred their learnt behaviour from the training phase. Again, separate mixed-effect linear regression models were fitted to the two learning status groups (learner vs non-learner) both adjusting for study group, time (RRR at baseline and during the generalisation phase), male sex and IMD quintile. Lastly, the models were repeated, adjusting for cognitive z-scores.

The variable selected a priori to best reflect the performance on this task was the binary learning status outcome.

6.1.1.2 BabyScreen application

During the 30 month assessment, cognitive processing was tested using a newly developed touchscreen application. For the administration, the child was seated at a table, with the touchscreen device place flat on the table in front of them. The experimenter and parent in the room remained in the room but the parent was instructed not to talk during the assessment. The application was administered on an Apple iPad Air 2, with the volume set at 70% of the maximum capacity. Before starting, the parent was asked to estimate their child’s touchscreen exposure.

The BabyScreen software version 1.5 (Hello Games Ltd, UK) assesses 4 areas associated with EF, termed ‘constructs’, being selective attention, working memory, hidden object retrieval, and object permanence. Each construct comprised a varying number of items. Overall 18 different items were administered. The rationale and design of each construct was developed utilising basics concepts from established experimental designs, including the A*B* paradigm and Multi-Location Multi-step task, both of which were used in the current study.

The paradigm starts in demonstration mode. The experimenter says ‘let me show you’, and completes the action required to complete the item. This is the only verbal instruction given during the session. The application utilises multiple different touchscreen techniques, which at the age of 30 months, may or may not be familiar to the child, depending upon prior touchscreen exposure. Throughout the different test items, it was expected that the child would make trial and error responses in order to determine the correct touch response to complete it.

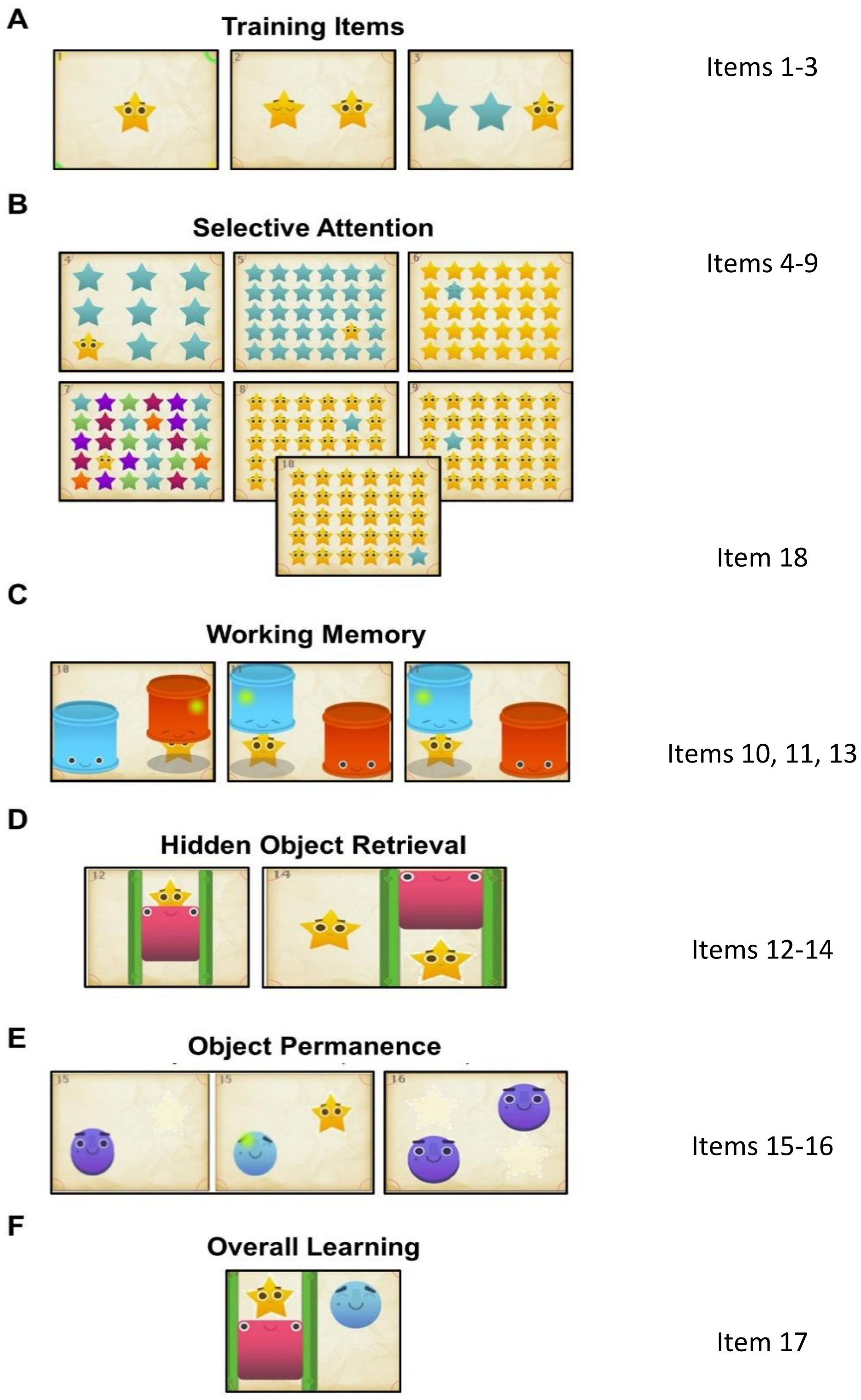


Figure 6-3. Visual representation of the BabyScreen items, separated into the 4 EF constructs, with the addition of training and overall learning.

The application starts with 3 training items, to familiarise the child with the concept of the paradigm, please see Figure 6-3 (A). The main theme running through the application is the target image. The child is required to press or repeatedly tap the gold star with a face until it disappears. The length of press or repeated tapping motion is set such that it is unlikely the child achieves the required level of touch by accident. This ensures the behaviour to obtain the target is intentional. When reaching the required touch level on the target, the star disappears, accompanied by a musical sound signifying this achievement and the app progresses to the next item.

Selective attention is assessed in items 4 through 9, and in an additional item at the very end of the paradigm, item 18, which is a repeat of items 8 and 9 (B). This is theorised to assess selective attention as there are multiple distractor stimuli alongside the target, which the child had to inhibit to pass the item. Progression through the 6 items of this construct gradually increased in difficulty, for instance, additional distractors and dimensions (such as colour), and ended with a rule change on item 8, where the target changed to the blue star with no face.

The working memory was assessed in items 10, 11 and 13 (C). These items presented the target initially, but was then cover by one of two cups that fell from the top of the screen. The child had to swipe the correct cup in an upward motion to reveal the target star, the cups then disappeared and the target was then pressed/tapped until it disappeared, as learnt previously.

For the hidden object construct, the child had to move the red box to reveal the target star, as shown in image D. Items 12 and 14 assessed this construct, with item 14 adding an additional target star.

Object permanence, assessed by items 15 and 16, displayed a blue button with a smiley face. The child was required to press and hold the button to display the target star. The target was only achieved when the child simultaneously held the button and tapped/pressed the star to complete the item. The second of these items had two buttons and two stars.

Item 17 (F), is an amalgamation of touch responses utilised over the application to obtain an overall measure of learning over the task. Tapping, holding, swiping and pressing were all needed to achieve the items within the paradigm and item 17 required all of these motions.

The structure of each item was the same. Upon the initiation the child was given 30 seconds to successfully complete the item by making the target star disappear. If the child was not able to complete the item in this timeframe, the screen would go into 'demo mode'. At this point the experiment said 'watch me', and successful completed the item. The child was then given a further 30 seconds to complete the item. If it was apparent that the child was not going to accomplish the required action to achieve a pass on that particular item, at the end of the second block, the experimenter could choose to skip the item. The task would then move on to the next item.

Coding and Analysis:

The key variables produced by the BabyScreen app were: number of items completed (without and with a demonstration), the response time of the child to each item, and finally, in those where distractor stimuli were present, accuracy of the child's first press was also analysed. Each item was investigated separately, and performance on each construct as a whole were analysed.

Irrespective of whether the child required a demonstration, the total amount of time taken was summated from the first and second attempts at each item, as required. In cases where the items were skipped, this was either due to an inability to complete the item or, if circumstances indicated the child was not going to be able to achieve a pass, if a previously similar item was not passed, or if the child had become distressed. Due to this, when exploring the response times to all items within the paradigm, these RTs could have skewed findings and not truly reflect a child's difficulties with a particular item. Accordingly, the highest response time in the cohort was applied to all cases where the item was skipped and the item was considered to be a 'failed' attempt.

It was only possible to measure accuracy, as mentioned, on items where distractor stimuli were present; items 4-11, 13 and 18. This score was calculated based on which attempt successful completion of the item occurred and whether the first touch on the screen due the successful attempt was an interaction with the distractor stimulus or the target. The scale for accuracy was from 0 to 4. A score of 4 indicated the highest accuracy, where the target was obtained in the first attempt and the target was the first interaction. Scores of 1 indicated the poorest accuracy where success on the item was achieved during the second attempt and an interaction with a distractor stimulus was the first interaction. Task failure gave an accuracy score of 0.

Statistical analysis approach was as detailed in chapter 2. Total task performance by study group was first explored with a particular interest in the proportion of items completed without the need of a demonstration. Each construct was then explored, investigating overall performance on the construct and performance within the individual items. Whole constructs and individual items were explored in terms of performance (pass/fail), response time (RT) and, where applicable, accuracy score (ACC) between the two study groups.

Multiple linear regression models were fitted to explore the relationship of predictor variables with the overall number of items completed without demo and with the RTs of the main learning measure, item 17. These models were fitted with and without the adjustment using 30 month Bayley-III cognitive z-scores.

The variable selected a priori to best reflect the performance on this task was the response times to overall learning measure, task 17.

6.1.1.3 MLMS

The Multi-location Multi-step paradigm was an adaptation on the experimental design reported by Zelazo et al., in 1998. The paradigm explores various aspects of EF, with a particular focus on working memory skills and cognitive flexibility. The MLMS task was designed to be undertaken at 30 months as an extension to the A not B paradigm at 12 months. The MLMS was primarily designed to provoke subjects into making preservative errors, as observed in the 12 month A*B* paradigm. During data collection, it was clear that the children within this cohort were not making traditional errors however, and a ceiling effect was being observed. Nevertheless, the subjects started to perform motor responses indicative of preservative errors, but modified their responses before the errors were made, considered perseverative hesitations. Although it was not possible to accurately code this hesitation, it was accounted for by measuring the response time. It is theorised that this visible updating of working memory, is reflective of slower information processing speeds.

The MLMS apparatus (Figure 6-4) comprised a set of horizontal drawers attached to a cardboard base, with a long rectangular cardboard box and cloth. There were 5 spaces but only 3 drawers were used for this task, the central and two end spaces. Attached to each draw was a string which extended to a spot of Velcro 13cm away on the cardboard base. 4 symbols were used during the task; a green triangle, red square, blue circle and yellow star. The triangle was used for the training phase and the other 3 for the experimental phases.

The task begun with the experimenter hiding a snack in the central draw, pointing to the green triangle attached to the equivalent string to indicate the food is in that draw, the experiment then covered the green triangle with the cardboard box and placed the cloth over the whole apparatus. The experimenter then demonstrated the 3 steps the child had to follow to locate the snack. The first step required the lifting of the cloth to reveal the apparatus, followed by the removal of the cardboard box to reveal the symbols attached to the strings. Lastly, the appropriate string is pulled to reveal the snack. The training phase is initiated following the demonstration of the task procedure. On each trial, the experimenter hides and covers the apparatus, and asks ‘can you find the snack?’. On the first attempt, prompts were offered where necessary to ensure the child executed the 3 steps correctly. The same verbal direction was given in each instance.

The training phase comprised of 3 trials or continued until the child could follow the 3 steps without any significant verbal prompts. The green triangle was then removed, and replaced with the 3 other symbols. The yellow star was always in the centre, but the red square and the blue circle were counterbalanced between the two sides. The pre-switch phase was then initiated with the snack being hidden in one of the two far side positions. The same procedure for each trial was followed as described previously, but with a small modification. After the experimenter had hidden the snack, they stated ‘the snack is in this one’, whilst simultaneously pointing to the equivalent symbol. The experimenter then opened the middle draw and stated ‘it is not in this one’ whilst pointing to the star, ‘and not in this one’, again simultaneously pointing to the final symbol. The child is lastly reminded of the foods location again by opening the draw it is in and pointing to the corresponding symbol. The apparatus is then covered in the same procedure as the training phase. The pre-switch phase was completed when the child located the food in 3 consecutive trials.

Upon completion of the pre-switch phase, the post-switch phase was initiated. The same trial procedure was followed, but the location of the snack was switched to the draw at the opposite end of the apparatus. The child then was asked to locate the snack in the new location. The time for the child to locate the final snack location was then measured.

Coding and Analysis:

The MLMS task was video recorded and coding performed offline. Time to retrieve the snack was recorded for each trial. The timer was started once the experiment asks ‘can you find the snack?’, and stopped when the subject selected the correct draw. Any perseverative errors made during the task were additionally recorded.

The time taken during training trials was averaged and used to adjust responses in the pre-switch and post-switch phases, in order to account for individual variation. Baseline responses were compared to ensure no fundamental differences were observed between study groups.

The child’s performance during the task dictated the number of pre-switch trials they completed. As stated the child had to correctly identify the location of the snack on 3 consecutive trials to reach the post-switch phase. If an error was made following 2 correct trials, the child had to pass another 3 additional trials consecutively, completing 6 pre-switch trials in total before the post-switch phase was initiated. The maximum number of pre-switch trials was set at 8, if a child was unable to correctly locate the snack in 3 consecutive trials after 8 attempts, the child was considered to have failed the task. The total number of pre-switch trials was compared between study groups as a whole, and within those that went on the meet the criterion. The trial theorised to display the highest error rate within the pre-switch phase was the first after training. The differences between groups were therefore explored within the first pre-switch trial. Subsequently, the response times to the three consecutively correct trials were then explored across study groups. The response times for each trial were calculated as a proportion of time taken to complete the training trials, referred to as ‘adjusted RT’. This adjustment accounted for individual variation during the training (or baseline) condition.

To further investigate the speed of processing over the task, a multi-level mixed effect model was produced to investigate how the time taken to complete the trials varied across the task. It was hypothesized that time taken would decrease over the pre-switch period as the child practiced the multiple step routine to locate the snack. An increase in response time was then hypothesised during the post-switch trial due to the higher working memory load required to modify the prepotent response established during the preswitch phase. The data was transformed into long format to conduct this analysis, including the adjusted training variable as the baseline group, the adjusted trial times for the 3 consecutive pre-switch trials and the adjusted post-switch trial response time as the main outcome variable. Post-hoc investigations then independently explored the change in time taken from one trial to the next throughout the paradigm.

The number of errors made during the post-switch trial and the change in adjusted time taken from the final pre-switch to the post-switch trial was investigated. The relationship between the adjusted RT for the post-switch trial and study group, male sex and IMD quintile were then explored using a multiple regression model, before lastly adjusted for 30 month cognitive z-scores in a separate model, to determine the extent to which performance was dependent on overall cognitive ability.

The variable selected a priori to best reflect the performance on this task was the adjusted response time for the post-switch trial.

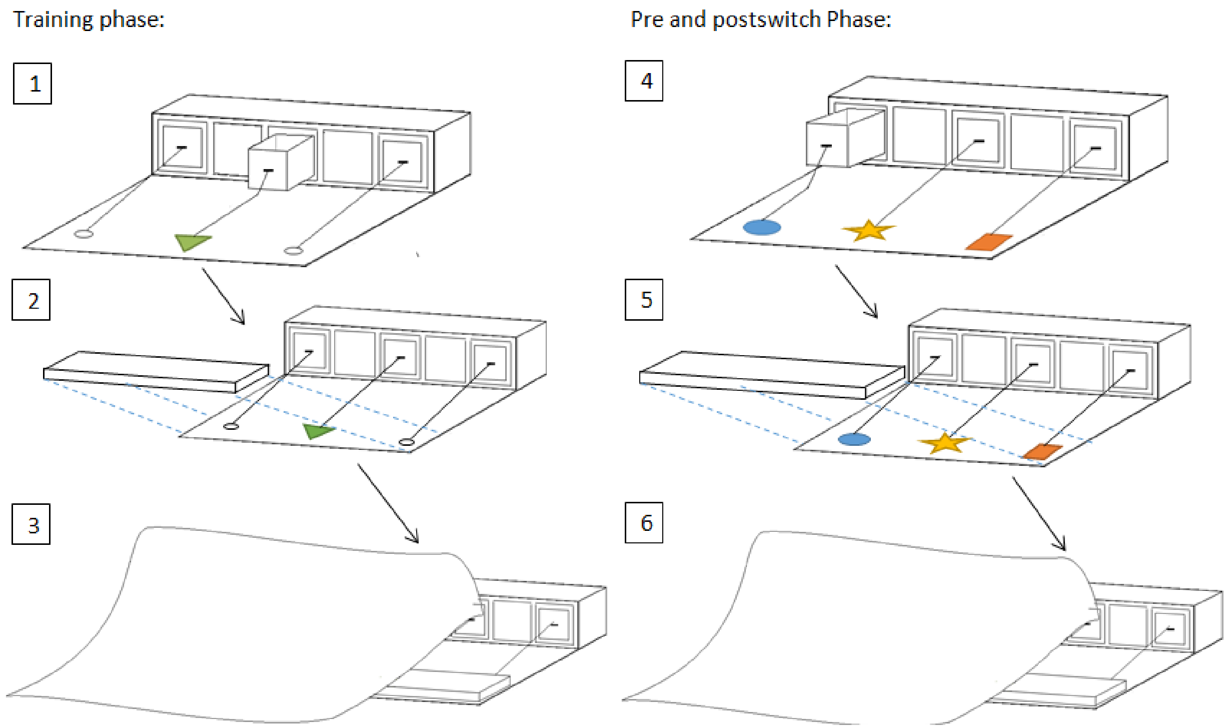


Figure 6-4. MLMS apparatus show the multi-step procedure children were asked to imitate. 1- shows the training phase where only one symbol was attached to the base board, the food was hidden, the drawer closed and the cardboard cover placed over the symbol line (2). 3 illustrates the cloth being placed over the apparatus. 4 shows the pre- and post-switch set up with the additional symbols. The food is hidden in one of the end drawers and the same procedure is followed in images 5 and 6.

6.1.2 Results

6.1.2.1 Mobile task results

At 3 months, results for the conjugate mobile reinforcement paradigm were available for 71 infants (Table 6-1). Fourteen of 31 VP infants (45%) met criterion for learning compared to 25 (61%) of 41 term infants (odds ratio: .41 (*95% CI .12, 1.10), p = .15,* when adjusting for 12 month cognitive z-scores, male sex, and IMD quintile). Overall, both term and VP learners displayed a gradual increase in kick rate in response to the mobile throughout the training phase, in contrast to the non-learners. The learner groups additionally displayed a specific increase in effector leg within this phase. In the generalisation phase, 75% of the term learners and 71% of the VP learners maintained the kick response to the new mobile.

		Term (n=41)	***VP (n=31)***
Gestational age	*Median (range); weeks^{+d}*	40^{+0} (37^{+1} – 42^{+0})	26^{+0} (23^{+4} – 30^{+2})
Male sex		20 (48.78)	21 (67.74)
IMD Quintile	1	5	4
	2	6	6
	3	5	11
	4	17	6
	5	8	4
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=35; VP=26)	109.14 (.11)	98.27 (9.15)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=35; VP=26)	.11 (1.01)	-.81 (.77)

Table 6-1. Demographic details of infants in mobile conjugate task during the 3 month assessment.

Baseline kick rates were similarly distributed between study groups, with term infants displaying a median of 5.25 kicks per minute (IQR: 2.75–7.5) and VP infants a median of 5.75 kicks per minute (IQR: 2.75 – 10.5); z = -.64, p = .52). Baseline kicks were also similar rates between learners and non-learners (Table 6-2; Figure 6-5).

	Baseline (Median kick rate in 2 minutes (IQR))
Term (n=41)	**5.25** (2.75–7.5)
Term Learner (n=25)	3.75 (2 – 7.5)
Term Non-learner (n=16)	5.75 (4 – 10.63)
Preterm (n=31)	**5.75** (2.75 – 10.5)
Preterm Learner (n=14)	3.63 (2 – 6.5)
Preterm Non-learner (n=17)	6.5 (5.5 – 12.25)

Table 6-2. Median and IQR of baseline kick rate per minute in Mobile Task by study group and further subdivided in to learner and non-learners according to the study criterion (Figure 6-2).

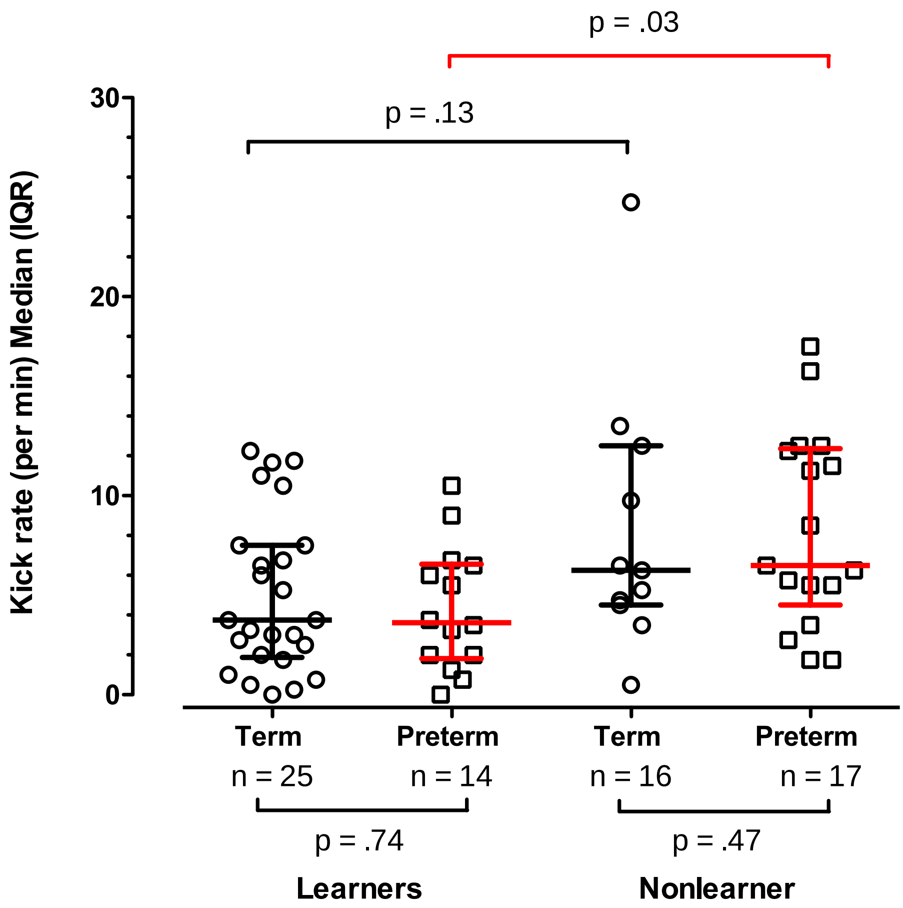


Figure 6-5. Baseline kick rate per min across the Term and Very Preterm infants to complete the Mobile Task at 3 months of age.

From baseline, non-learners showed no change over the three training phases. In contrast, learners showed progressive increase in kick rates over the first two epochs (Table 6-3; Figure 6-6).

	Baseline RRR	*Training phase 0-2 minutes (Median RRR (IQR))*	*Training phase 3-4 minutes (Median RRR (IQR))*	*Training phase 5-6 minutes (Median RRR (IQR))*
Term (n=41)	**1**	**1.17** (.8 – 2)	**1.61** (1.14 – 1.61)	**1.56** (.91 – 2.13)
Term Learner (n=25)	1	1.93 (1.23 – 2.67)	2.64 (1.88 – 3.5)	2 (1.57 – 4.69)
Term Non-learner (n=16)	1	.69 (.32 - .93)	.94 (.61 – 1.35)	.88 (.58 – 1.23)
Preterm (n=31)	**1**	**1.14** (.74 – 1.71)	**1.23** (.8 – 2.19)	**1.35** (.73 – 3.13)
Preterm Learner (n=14)	1	1.65 (1.32 – 1.88)	2.22 (1.62 – 4.6)	3.31 (2.08 – 5.5)
Preterm Non-learner (n=17)	1	.9 (.69 – 1.14)	.88 (.76 – 1.22)	.76 (.65 – 1.15)

Table 6-3. Median kick rate for baseline phase, and median RRR for each 2 min epoch of the training phase and generalisation phase of the mobile reinforcement paradigm.

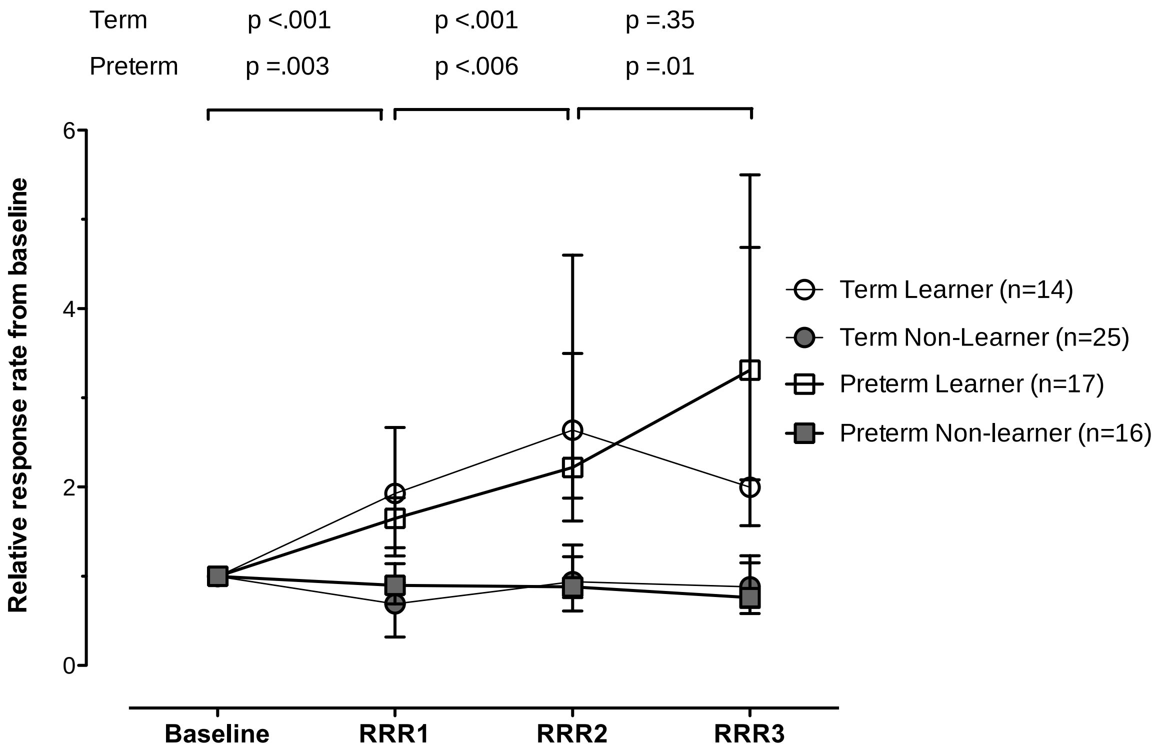


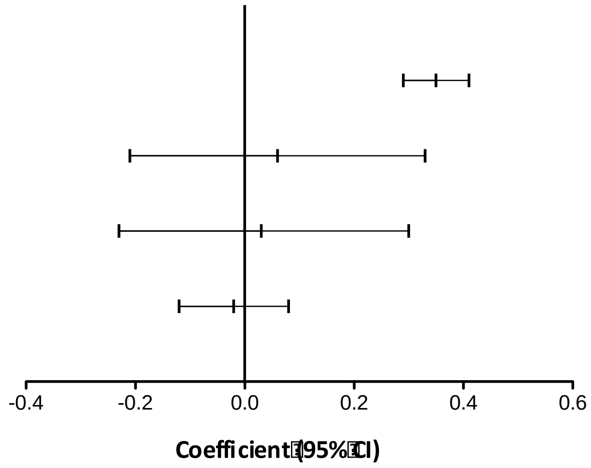
Figure 6-6. Changes from baseline in relative response rates in term and Very Preterm infants by learning status

Both learning groups showed a significant rise in kick rates from baseline to time 1 and then to time 2 (Figure 6-6). The kick rates for term infants plateaued at this point, whereas those for VP infants continued to increase ($p=.35$ and $p=.01$, respectively).

In a mixed effects model among learners, when adjusting for study group, male sex and IMD quintile, a significant main effect of time was observed (baseline to RRR3) with an increase of .35 in the RRR for each phase of the task (95% CI (.29 – .41), $p < .001$; Table 6-4). This effect persisted after further adjustment for the 12 month cognitive z-scores ($\beta = .34$, 95% CI (.28 – .40), $p < .001$; Table 6-5). No effects were seen for non-learners (not shown).

Table 6-4. Multilevel mixed effects model with the learner category modelling the RRR across the training phase. The baseline group was the baseline RRR in term-born females learners with an IMD quintile of 1.

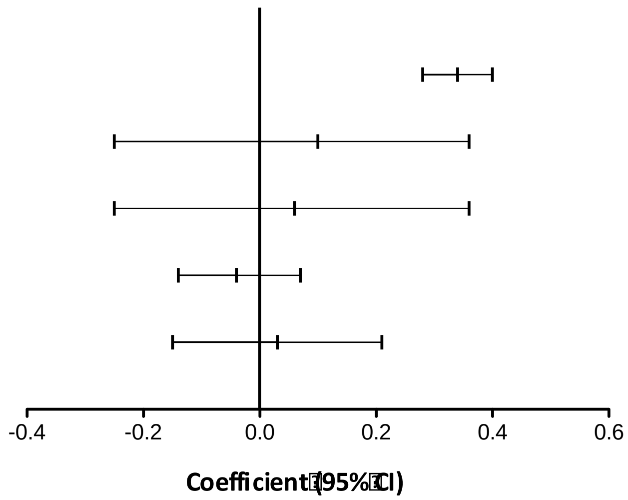
Predictor	Term (n=25)	Preterm (n=14)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Time	-	-	.35	.29 - .41	.000
Study group	39^{+6} (37^{+1} – 42^{+0})	26^{+2} (23^{+4} – 29^{+3})	.06	-.21 - .33	.68
Male sex	15 (60%)	9 (64%)	.03	-.23 - .30	.80
IMD quintile	4(1-5)	4(1-5)	-.02	-.12 - .08	.68
(Const.)	-	-	.14	-.26 - .55	.48



Wald chi^2 = 152.11 (p<.001)

Table 6-5. Multilevel mixed effects model with the learner category modelling the RRR across the training phase additionally adjusting for 12 month cognitive z-scores. The baseline group was the baseline RRR in term-born females learners with an IMD quintile of 1.

Predictor	Term (n=25)	Preterm (n=14)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Time	-	-	.34	.28 - .40	.000
Study group	40^{+0} (37^{+1} – 42^{+0})	25^{+5} (24^{+4} – 29^{+3})	.10	-.25 - .44	.59
Male sex	15 (71 %)	7 (58%)	.06	-.25 - .36	.72
IMD quintile	4(1-5)	4(1-5)	-.04	-.14 - .07	.49
12m cog z-score	.06 (.94)	-.87 (.46)	.03	-.15 - .21	.74
(Const.)	-	-	.18	-.29 - .65	.45



Wald chi^2 = 126.00 (p<.001)

To determine if leg kicks were lateralised to the effector leg, the leg kick difference was compared between the learning and study groups (Table 6-6; Figure 6-7). Overall, the term born infants displayed an increase from baseline in the effector leg (z = 2.65, p = .008); the VP infants did not (z = 1.00, p = .31). However, when separating into learner and non-learner groups, there was lateralisation of the learnt response in the learners (Terms p<.005; VP p<.05), which persisted after adjustment for the study confounds (Table 6-7; Table 6-8), in contrast to no lateralisation in either of the non-learner groups was observed.

	Median leg difference during baseline (IQR)	***Mean leg kick difference during last 2 minutes of training (SD)***	***MWU; p***
Term (n=41)	**.5 (-.5 – 1.5)**	**1.89 (3.01)**	2.65; .008
Term Learner (n=25)	.25 (-.5 - 1)	2.49 (3.32)	2.83; .005
Term Non-learner (n=16)	.63 (.25 – 1.88)	1.03 (2.35)	.29; .78
Preterm (n=31)	**-.25 (-1.5 - .75)**	**.65 (3.50)**	1.00; .31
Preterm Learner (n=14)	-.63 (-1.25 - .25)	1.66 (3.43)	1.98; .05
Preterm Non-learner (n=17)	0 (-1.5 – 2.75)	-.3 (3.41)	-.63; .53

Table 6-6. Differences between kick rates of each leg from baseline to the end of training; Mann Whitney U (MWU) tests were used to compare the difference from baseline to the end of training in each study group; then by learning status (exact p values reported; Significance following Bonferroni correction: p<.008)

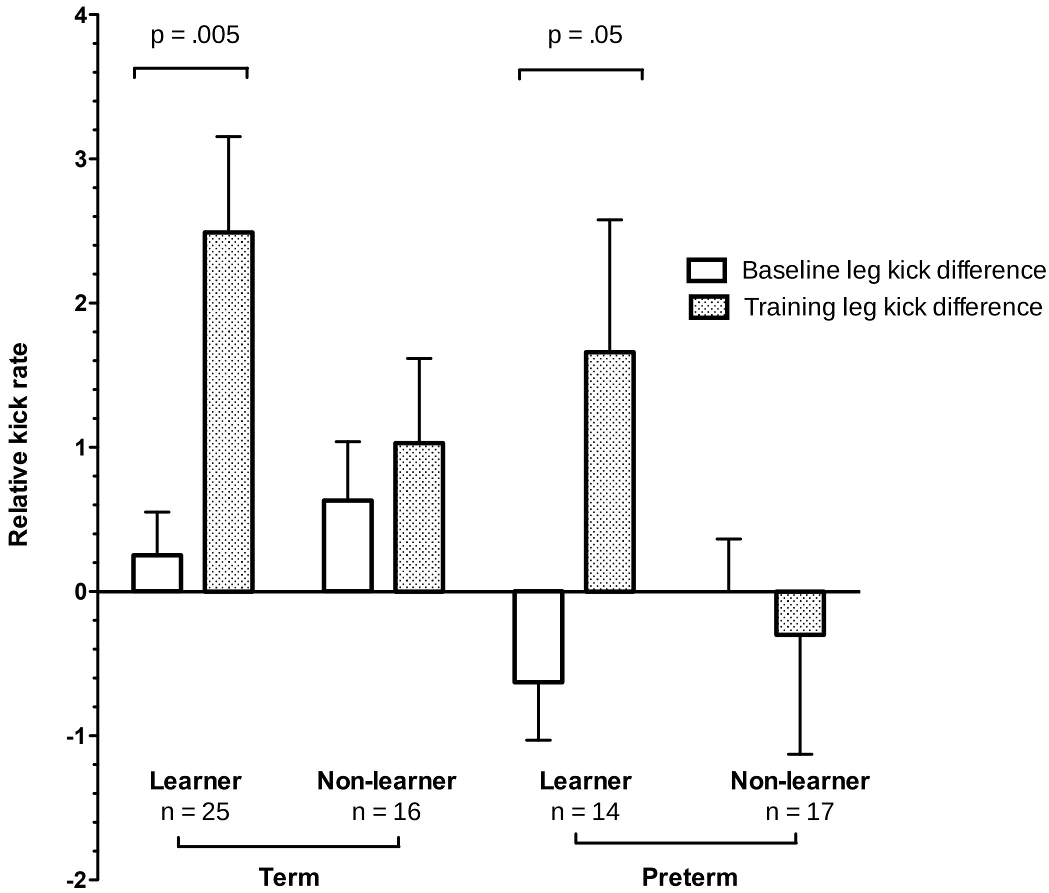
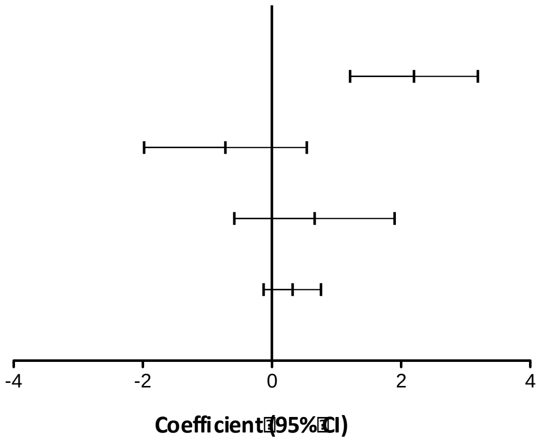


Figure 6-7. Difference in kick rate of each leg at baseline and the end of the training phase by study group and learner status; Mann Whitney U tests used to compare between groups.

Table 6-7. Multilevel mixed effects model with the learner category modelling leg kick difference from baseline to the end of training. The baseline group was the baseline leg kick difference in term-born female learners with an IMD quintile of 1.

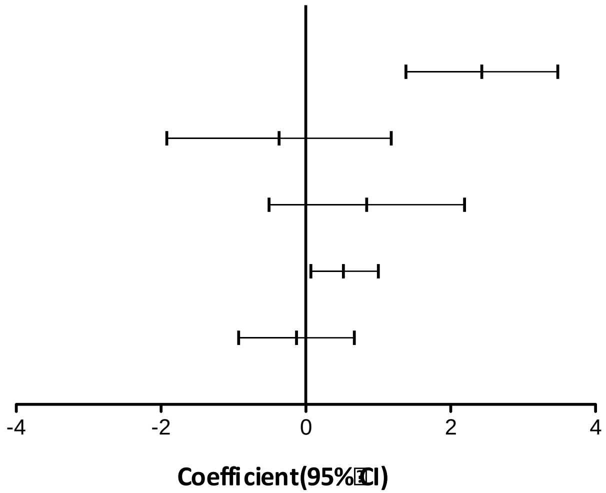
Predictor	Term (n=25)	Preterm (n=14)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Time	-	-	2.20	1.21 – 3.19	.000
Study group	39^{+6} (37^{+1} – 42^{+0})	26^{+2} (23^{+4} – 29^{+3})	-.72	-1.98 - .54	.26
Male sex	15 (60%)	9 (64%)	.66	-.58 – 1.90	.30
IMD quintile	4(1-5)	4(1-5)	.32	-.13 - .76	.16
(Const.)	-	-	-1.16	-3.07 - .76	.24



Wald chi^2 = 23.37 (p<.000)

Table 6-8. Multilevel mixed effects model with the learner category modelling the leg kick difference from baseline to the end of training additionally adjusting for 12 month cognitive z-scores. The baseline group is the baseline leg kick difference in term-born female learners with an IMD quintile of 1.

Predictor	Term (n=25) Median (Range)	Preterm (n=14) Median (Range)	Coef	95%CI	P
Time	-	-	2.43	1.38 – 3.48	.000
Study group	40^{+0} (37^{+1} – 42^{+0})	25^{+5} (24^{+4} – 29^{+3})	-.37	-1.92 – 1.18	.64
Male sex	15 (71%)	7 (58%)	.84	-.51 – 2.19	.23
IMD quintile	4(1-5)	4(1-5)	.53	.07 – 1.00	.02
12 m cog z-score	.06 (.94)	-.87 (.46)	-.13	-.93 - .67	.75
(Const.)	-	-	-2.00	-4.11 - .11	.06



Wald chi^2 = 27.67 (p<.001)

During the generalisation phase (Table 6-9; Figure 6-8), 18 of the 24 term-learners (75%) and 10 of the 14 VP learners (71%) displayed RRR of 1.5 or greater, a significant increase from baseline (z = 5.11; p < .001) compared the non-learners (z = 1.07; p < .29). This effect of time persisted within the learning category after adjustment for the study demographic variables (Table 6-10 and Table 6-11).

	Baseline RRR = 1	*Generalisation phase 0-2 minutes Median RRR (IQR))*
Term (n=40)	**1**	**1.7** (.95 – 3.00)
Term Learner (n=24)	1	2.38 (1.70 – 4.03)
Term Non-learner (n=16)	1	1.05 (.55 – 1.46)
Preterm (n=30)	**1**	**1.45** (.62 – 2.74)
Preterm Learner (n=14)	1	2.91 (1.79 – 6.93)
Preterm Non-learner (n=16)	1	.65 (.34 – 1.05)

Table 6-9. Median and IQR relative response ratio during the generalisation response across study group then subdivided into learning status. Not all infants that completed the training phase completed the additional generalisation phase and therefore were omitted from this analysis.

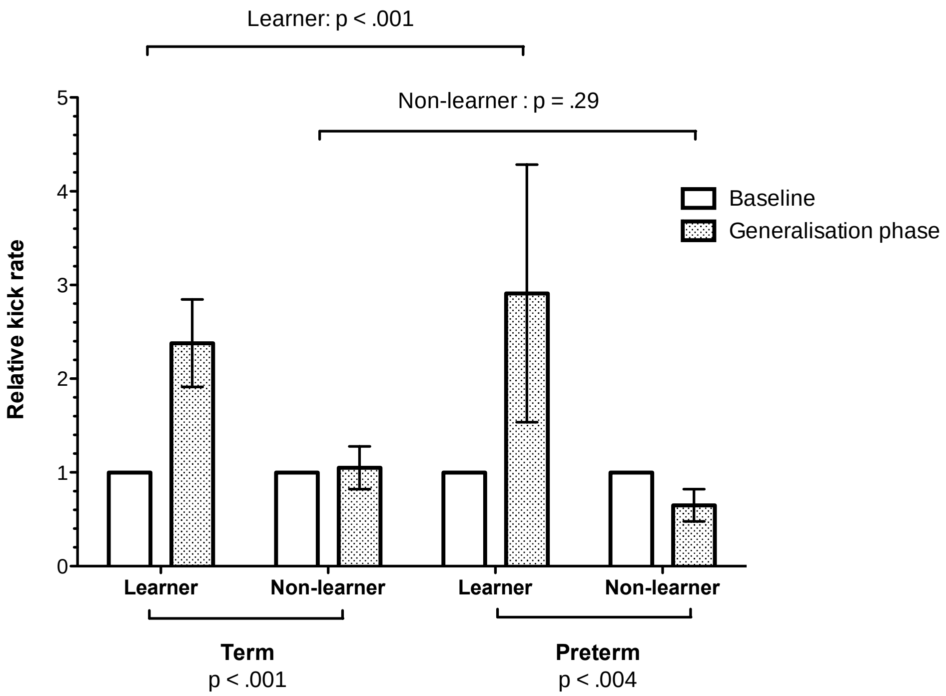
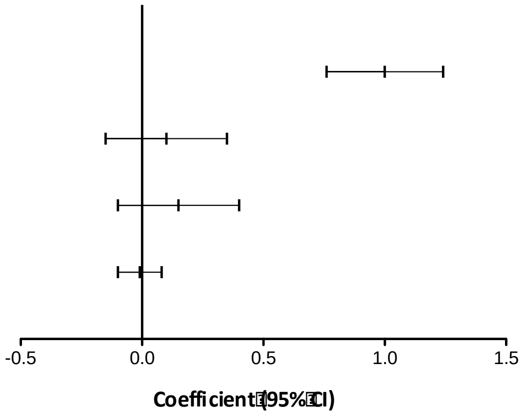


Figure 6-8. Relative response ratio from baseline to the generalisation phase across study group and learning status; Mann Whitney U tests used to compare between groups and learning status.

Table 6-10. Multilevel mixed effects model with the learner category modelling the RRR from baseline to the generalisation phase. The baseline group was the baseline RRR in term-born female learners with an IMD quintile of 1.

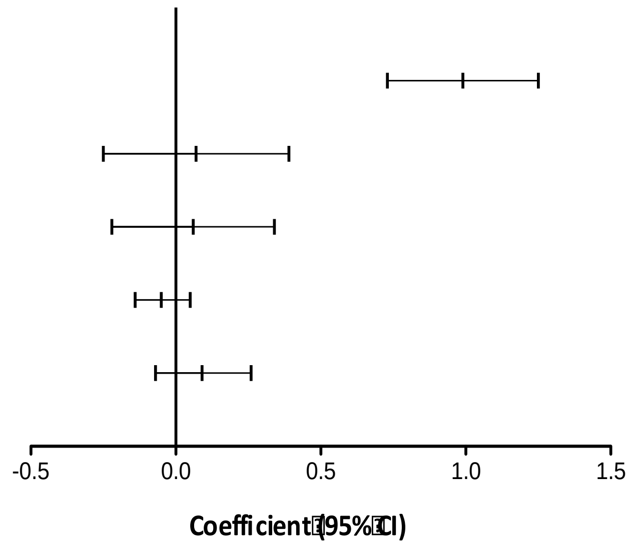
Predictor	Term (n=24) Median (range)	Preterm (n=14) Median (range)	Coef	95%CI	P
Time	-	-	1.00	.76 – 1.24	.000
Study group	39^{+6} (37^{+1} – 42^{+0})	26^{+2} (23^{+4} – 29^{+3})	.10	-.15 - .35	.43
Male sex	14 (58%)	9 (64%)	.15	-.10 – .40	.24
IMD quintile	4(1-5)	4(1-5)	-.01	-.10 - .08	.75
(Const.)	-	-	-.8	-.47 - .31	.68



Wald chi^2 = 68.67 (p<.000)

Table 6-11. Multilevel mixed effects model with the learner category modelling the RRR from baseline to the generalisation phase additionally adjusting for 12 month cognitive z-scores. The baseline group is the baseline RRR in term-born female learners with an IMD quintile of 1.

Predictor	Term (n=24)	Preterm (n=14)	Coef	95%CI	P
	Median (range)	Median (range)			
Time	-	-	.99	.73 – 1.25	.000
Study group	40^{+0} (37^{+1} – 42^{+0})	25^{+5} (24^{+4} – 29^{+3})	.07	-.25 - .39	.66
Male sex	14 (70 %)	7 (58%)	.06	-.22 - .34	.67
IMD quintile	4(1-5)	2.5(1-5)	-.05	-.14 - .05	.32
12 month Cog z-score	.06 (.94)	-.87 (.46)	.09	-.07 - .26	.28
(Const.)	-	-	.10	-.33 - .54	.64



Wald chi^2 = 58.93 (p<.001)

6.1.2.2 BabyScreen Application Results

Eighteen term-born children and seventeen VP children completed the BabyScreen programme (Table 6-12). Overall 45% of the children had daily use of a touchscreen, similarly distributed between the two groups. VP children completed a median 16 of the 18 items (range 14-18) compared to a median of 18 for the Term children (r: 15-18; z = 2.39, p<.02). VP children completed fewer items overall without a demonstration (med: 13; r: 10-17) compared to term children (med: 16; r: 8-17; z = 2.54, p=.01; Figure 6-9). The number of trials completed showed no correlation with the frequency of touchscreen exposure or the composite cognitive z-scores at 30 months.

		Term (n=18)	***VP (n=17)***
Gestational age	*Median (range); weeks^{+d}*	40^{+2} (37^{+0} –42^{+1})	26^{+3} (23^{+6} – 29^{+4})
Male sex		9 (50)	14 (82.35)
IMD Quintile	1	2	3
	2	2	0
	3	3	7
	4	7	5
	5	4	2
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=16; VP=17)	109.38 (13.77)	101.47 (11.29)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=16; VP=17)	.13 (1.11)	-.50 (.91)

Table 6-12. Demographic details of infants to complete the BabyScreen Application

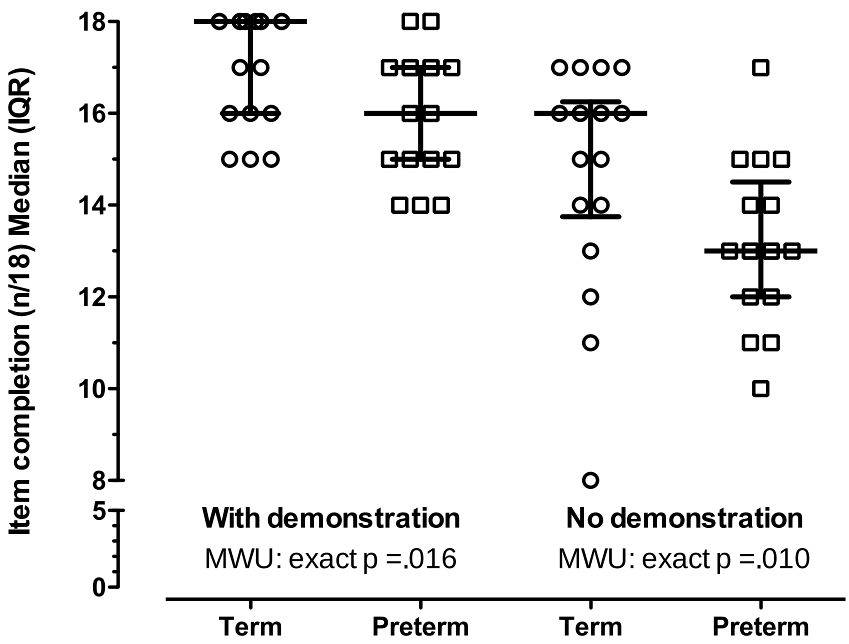


Figure 6-9. Total number of items completed without and with demonstration in the BabyScreen application by study group at 30 months of age. Mann Whitney U (MWU) tests were used to compare the number completed in each study group (exact p reported to account for tied values).

When exploring the number of children to complete each item (Table 6-13), no item had significantly more failures with or without Bonferroni adjustment (p=.05/18 or .003); however, 65% and 53% VP children failed tasks 15 and 16 respectively, the object permanence construct (Figure 6-10).

Response times (RTs) in seconds were then compared for each item (Table 6-14). No differences were observed in the RTs to each item before or after Bonferroni adjustment (p=.05/18 or p=.003) up to item 14. In the object permanence items, VP infants were slower in their responses, significantly so in the second item before Bonferroni correction (Figure 6-10; item 15: z = -1.89, p = .06; item 16: z = -2.34, p = .02).

Item Number/ Construct	*Completed (No Demo - %)*	*Completed (Demo - %)*	*Failed (%)*	Chi^2_{trend}	*P*
Task 1 (Training Item)					
Term	16 (88.9%)	2 (27.3%)	0 (0%)		
Preterm	11 (64.7%)	5 (29.4%)	1 (5.9%)	3.19	0.20
Task 2 (Training Item)					
Term	14 (77.8%)	3 (16.7%)	1 (5.6%)		
Preterm	12 (70.6%)	5 (29.4%)	0 (0%)	1.63	0.44
Task 3 (Training Item)					
Term	16 (88.9%)	2 (11.1%)	0 (0%)		
Preterm	16 (94.1%)	1 (5.9%)	0 (0%)	.31	0.58
Task 4 (Selective Attention)					
Term	18 (100%)	0 (0%)	0 (0%)		
Preterm	16 (94.1%)	1 (5.9%)	0 (0%)	1.09	0.30
Task 5 (Selective Attention)					
Term	17 (94.4%)	1 (5.6%)	0 (0%)		
Preterm	17 (100%)	0 (0%)	0 (0%)	0.97	0.32
Task 6 (Selective Attention)					
Term	18 (100%)	0 (0%)	0 (0%)		
Preterm	16 (94.1%)	1 (5.9%)	0 (0%)	1.09	0.30
Task 7 (Selective Attention)					
Term	17 (94.4%)	1 (5.6%)	0 (0%)		
Preterm	17 (100%)	0 (0%)	0 (0%)	0.97	0.32
Task 8 (Selective Attention)					
Term	14 (77.8%)	3 (16.7%)	1 (5.6%)		
Preterm	11 (64.7%)	4 (23.5%)	2 (11.8%)	0.81	0.67
Task 9 (Selective Attention)					
Term	16 (88.9%)	1 (5.6%)	1 (5.6%)		
Preterm	12 (70.6%)	5 (29.4%)	0 (0%)	4.21	.12
Task 10 (Working Memory)					
Term	11 (61.1%)	5 (27.8%)	2 (11.1%)		
Preterm	5 (29.4%)	10 (58.8%)	2 (11.8%)	3.89	0.14
Task 11 (Working Memory)					
Term	17 (94.4%)	1 (5.6%)	0 (0%)		
Preterm	16 (94.1%)	1 (5.9%)	0 (0%)	0.00	0.97
Task 12 (Hidden Object Retrieval)					
Term	15 (83.3%)	3 (16.7%)	0 (0%)		
Preterm	17 (100%)	0 (0%)	0 (0%)	3.10	.08
Task 13 (Working Memory)					
Term	18 (100%)	0 (0%)	0 (0%)		
Preterm	16 (94.1%)	1 (5.9%)	0 (0%)	1.09	0.30
Task 14 (Hidden Object Retrieval)					
Term	15 (83.3%)	3 (16.7%)	0 (0%)		
Preterm	16 (94.1%)	0 (0%)	1 (5.9%)	4.01	0.14
Task 15 (Object Permanence)					
Term	6 (33.3%)	7 (38.9%)	5 (27.8%)		
Preterm	2 (11.8%)	4 (23.5%)	11(64.7%)	5.04	0.08
Task 16 (Object Permanence)					
Term	9 (50.0%)	6 (33.3%)	3 (16.7%)		
Preterm	4 (23.5%)	4 (23.5%)	9 (52.9%)	5.30	0.07
Task 17 (Combined Learning)					
Term	12 (66.7%)	2 (11.1%)	4 (22.2%)		
Preterm	5 (29.4%)	5 (29.4%)	7 (41.2%)	4.96	0.08
Task 18 (Selective Attention)					
Term	17 (94.4%)	1 (5.6%)	0 (0.0%)		
Preterm	15 (88.2%)	1 (5.9%)	1 (5.9%)	1.10	0.58

Table 6-13. Total number of children to complete each item with and without a demonstration.

Construct	*N*	*Median RT*	*IQR*	*Min - max*
Task 1 (Training Item)				
Term	18	10.15	3.117 - 17.716	.73 - 33.367
Preterm	17	13.30	2.30- 39.719	1.23- 90
Task 2 (Training Item)				
Term	18	9.03	3.58 - 16.58	1.567- 90
Preterm	17	14.316	4.93- 37.85	1.467-68.28
Task 3 (Training Item)				
Term	18	3.43	1.4 -7.63	.93- 66.27
Preterm	17	7.017	2.35-10.00	.75-34.935
Task 4 (Selective Attention)				
Term	18	3.28	1.25-6.05	.55-27.035
Preterm	17	3.917	1.5- 8.916	1.033 -41.118
Task 5 (Selective Attention)				
Term	18	2.62	1.43-10.05	.717-53.05
Preterm	17	3.117	2.13-5.43	.717- 22.53
Task 6 (Selective Attention)				
Term	18	3.48	1.53-5.55	.649-22.95
Preterm	17	3.766	2.33-9.15	1-31.749
Task 7 (Selective Attention)				
Term	18	3.225	1.38-6.88	.916-51.768
Preterm	17	4.65	1.649-6.917	.732- 14.23
Task 8 (Selective Attention)				
Term	18	17.68	10.42-26.12	1.716-90
Preterm	17	16.5	5.349-37.95	1.3-90
Task 9 (Selective Attention)				
Term	18	5.375	2.916-13.45	.583-90
Preterm	17	11.7	4.467-38.37	.716-50.25
Task 10 (Working Memory)				
Term	18	26.90	10.73-44.15	4.48-90
Preterm	17	41.82	26.80-52.02	2.38-90
Task 11 (Working Memory)				
Term	18	9.217	4.35-12.75	2.218-52.268
Preterm	17	7.05	4.449-10.03	1.267-54.419
Task 12 (Hidden Object Retrieval)				
Term	18	7.28	5.032-19.117	3.917-60.435
Preterm	17	6.13	4.53-13.918	2.017 - 22.58
Task 13 (Working Memory)				
Term	18	8.14	4.28-10.83	.88 - 15.55
Preterm	17	7.78	5.53-12.18	1.685-34.23
Task 14 (Hidden Object Retrieval)				
Term	18	10.08	7.033-20.418	4.083-56.335
Preterm	17	9.28	8.73-13.666	7.117 - 90
Task 15 (Object Permanence)				
Term	18	35.81	17.685 -90	12.13-90
Preterm	17	90	42.75-90	5.835 - 90
Task 16 (Object Permanence)				
Term	18	31.30	15.45-60.638	6.65-90
Preterm	17	90	48.169 -90	11.05-90
Task 17 (Combined Learning)				
Term	18	23.09	13.23 - 52.586	5.968 – 90
Preterm	17	61.25	25.735-90	8.3-90
Task 18 (Selective Attention)				
Term	18	5.24	2.6-7.699	1.28-31.75
Preterm	17	4.03	2.717-23.15	1.18-90

Table 6-14. Median response times to each trial for the two study groups.

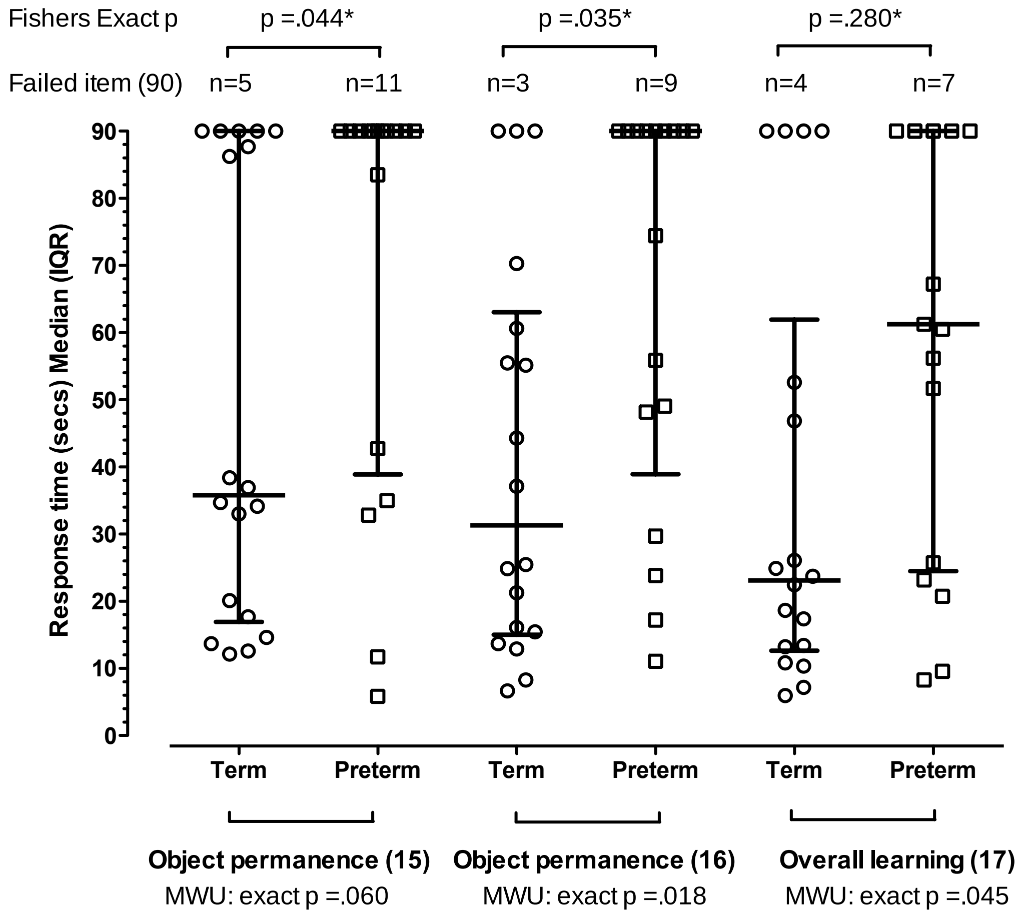


Figure 6-10. Response times on BabyScreen on object permanence and overall learning items completed by Term and Very Preterm children at 30 months of age. Fisher's exact test was used to compare the proportions in each group that failed each item; Mann Whitney U (MWU) tests were used to compare the response times in each study group (exact p to account for tied values).

To explore overall RT performance for each construct, the RTs were averaged across items (Table 6-15). The constructs were compared across study groups (Bonferroni adjustment: p = .05/6 or p = .008). In the RTs of the 6 overall constructs, VP children were significantly slower on the object permanence construct (Figure 6-10; z = -2.30, p = .02) and the overall learning measure (z = -2.30, p = .05).

	Term (n=18) *Median (range)*	*VP (n=17)* *Median (range)*	*z; P*
Training Construct	11.50 (1.52 – 55.63)	14.60 (1.62 – 45.10)	-.86; .39
Selective Attention Construct	8.20 (2.22 – 30.51)	11.39 (3.30 – 21.89)	-1.45; .15
Working Memory Construct	13.20 (3.21 – 50.65)	18.00 (4.42 – 55.77)	-1.06; .29
Hidden Object Constuct	11.97 (4.32 – 47.79)	7.59 (4.57 – 48.88)	.96; .34
Object Permanence Construct	35.48 (13.15 – 90)	72.93 (21.94 – 90)	-2.30; .02
Overall Learning	23.09 (5.97 – 90)	61.25 (8.3 – 90)	-2.01; .05

Table 6-15. The median and range response time of each construct

The overall constructs were then correlated with frequency of touchscreen use and the Bayley-III cognitive z-score (Table 6-16). There were no correlations before or after Bonferroni adjustment for the touchscreen use, and only a weak correlation with the Bayley-III in the overall learning construct within the VP group (Figure 6-11; r^2 = -.47, p = .06).

A: *Term (n=18)*	***Touchscreen use; r^2 (p)***	***Cognitive z-score ; r^2 (p)***
Training Construct	.09 (.73)	.13 (.13)
Selective Attention Construct	.07 (.8)	.10 (.72)
Working Memory Construct	-.01 (.98)	.41 (.12)
Hidden Object Constuct	-.04 (.89)	.41 (.12)
Object Permanence Construct	.19 (.48)	-.14 (.61)
Overall Learning	.32 (.23)	.16 (.56)

B: *VP (n=18)*	***Touchscreen use; r^2 (p)***	***Cognitive z-score; r^2 (p)***
Training Construct	.11 (.68)	.53 (.03)
Selective Attention Construct	.30 (.24)	-.36 (.16)
Working Memory Construct	.34 (.18)	-.15 (.57)
Hidden Object Construct	.04 (.89)	.27 (.29)
Object Permanence Construct	-.01 (.96)	.19 (.46)
Overall Learning	-.07 (.79)	-.47 (.06)

Table 6-16. Correlations between frequency of touchscreen use and Bayley-III cognitive z-score at 2 years of age for A) term born toddlers and B) Very Preterm toddlers.

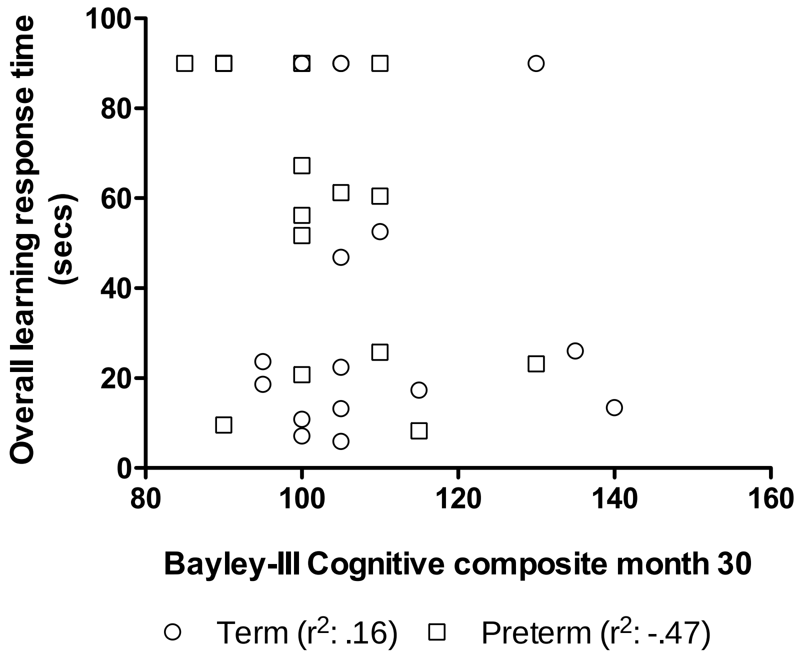


Figure 6-11. Correlation between the Bayley-III cognitive z-scores from the 2 year assessment.

Next the BabyScreen application accuracy scores were compared (Table 6-17). The only item to indicate a significant difference between study group in terms of accuracy was in response to the first working memory trial (Figure 6-12; z = 2.12, p = .03), where the term children demonstrated a higher level of accuracy. Overall mean accuracy (on a 4 point scale) for VP children was 2.97 (SEM: .08) compared to 3.08 for the term born children (SEM: .07; p=.32).

		Term (n=18) Median (range)	VP (n=17) Median (range)	z; P
Selective attention	**Item 4** (n: T=18; VP=17)	3 (3 – 4)	4 (1 – 4)	-.98; .33
	Item 5 (n: T=18; VP=17)	4 (1 – 4)	4 (3 – 4)	-.32; .75
	Item 6 (n: T=18; VP=17)	3 (3)	3 (1 – 3)	1.03; .30
	Item 7 (n: T=18; VP=17)	3.5 (2 – 4)	3 (3 – 4)	.00; 1.00
	Item 8 (n: T=17; VP=15)	3 (1 – 3)	3 (1 – 3)	.61; .54
	Item 9 (n: T=17; VP=17)	3 (1 – 3)	3 (1 – 3)	1.77; .08
	Item 18 (n: T=18; VP=16)	3 (1 – 3)	3 (1 – 3)	.09; .93
Working memory	**Item 10** (n: T=16; VP=15)	3 (1 – 4)	2 (1 – 4)	2.12; .04
	Item 11 (n: T=18; VP=17)	3.5 (2 – 4)	4 (2 – 4)	-.47; .64
	Item 13 (n: T=18; VP=17)	3 (3 – 4)	3 (1 – 4)	-.88; .38

Table 6-17. Median and range of accuracy scores (on a 4 point scale) for Term and VP children on all correctly passed items with distractor stimuli on the BabyScreen Application

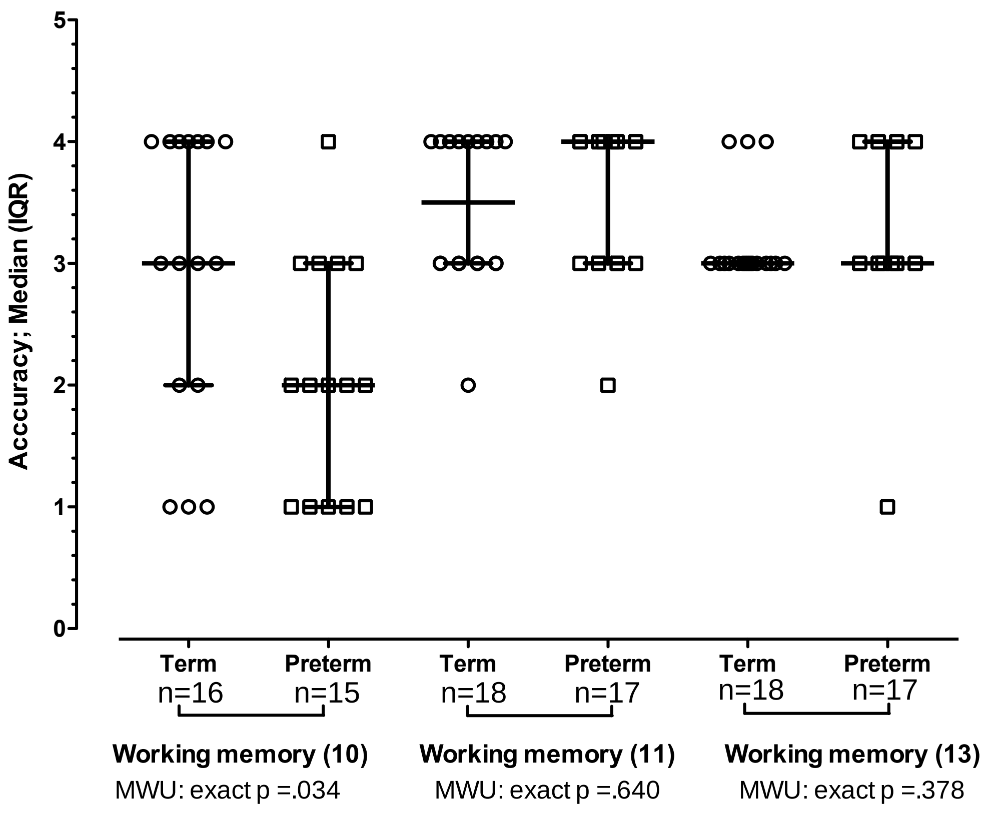
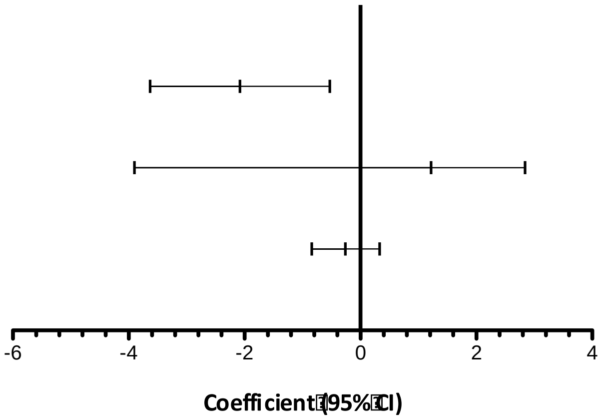


Figure 6-12. Accuracy scores on BabyScreen application for correctly passed working memory items completed by Term and Very Preterm children at 30 months of age. Mann Whitney U (MWU) tests were used to compare the accuracy scores in each study group (exact p to account for tied values).

Finally, the relationships between the two main outcome variables, total number of trials completed without a demo (shown in Figure 6-9), and the speed of completion of the main learning measure, item 17 (shown in Figure 6-10) were explored. Using multiple linear regression models, the group differences observed in these two outcomes remained after adjusting for the predetermined study confounds: male sex, IMD quintile and Bayley-III cognitive z-score at 2 years (Table 6-18 and 6-19).

Table 6-18. Linear multiple regression model for the total number of trials completed without a demonstration from the experimenter (F(3, 31) = 2.77, p = .06). Baseline group were term born females with an IMD quintile of 1.

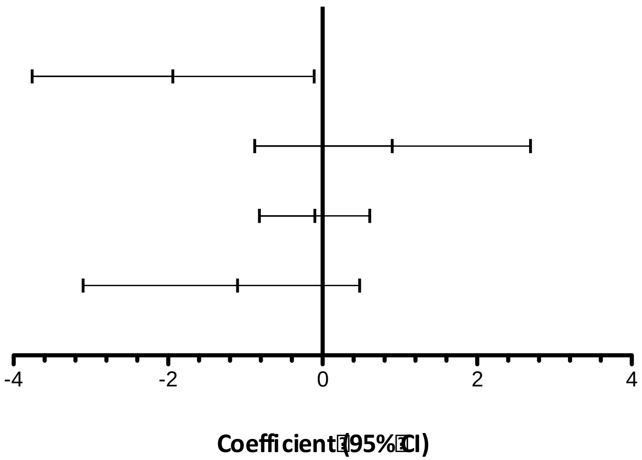
Predictor	Term (n=16)	Preterm (n=17)	Coef.	95%CI	P
	Median (Range)	Median (Range)			
Study group	39^{+6} (37^{+0} – 42^{+1})	26^{+3} (23^{+6} – 29^{+7})	-2.08	-3.63 - -.53	.01
Male sex	9 (50%)	14 (82%)	1.22	-.39 – 2.84	.13
IMD quintile	4 (1-5)	3 (1-5)	-.26	-.84 - .33	.38
Const	-	-	15.06	12.64 – 17.48	.000



Overall model fit was R^2 = .21

Table 6-19. Linear multiple regression model for the total number of trials completed without a demonstration from the experimenter with the adjustment of the 30 month cognitive z-scores (F(4, 28) = 1.27, p = .31). Baseline group were term born females with an IMD quintile of 1.

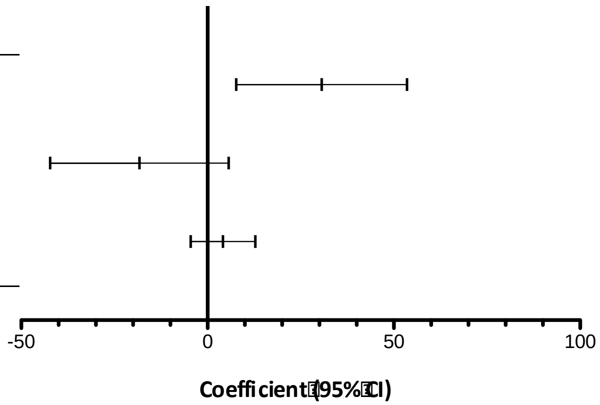
Predictor	Term (n=16) Median (Range)	Preterm (n=17) Median (Range)	Coef.	95%CI	P
Study group	40^{+1} (37^{+0} – 42^{+1})	26^{+2} (26^{+3} – 29^{+4})	-1.94	-3.76 - -.11	.04
Male sex	7 (43%)	14 (82%)	.90	-.88 – 2.69	.31
IMD quintile	4 (1.5)	3(1)	-.10	-.82 - .61	.77
2 year cog z-score	.13 (1.11)	-.50 (.91)	-.31	-1.10 - .48	.43
Const	-	-	14.54	11.61 – 17.47	.000



Overall model fit was R^2 = .15

Table 6-20. . Linear multiple regression model for the time for completion of overall learning measure, trial 17 (F(3, 31) = 2.78, p = .06). Baseline group were term born females with an IMD quintile of 1.

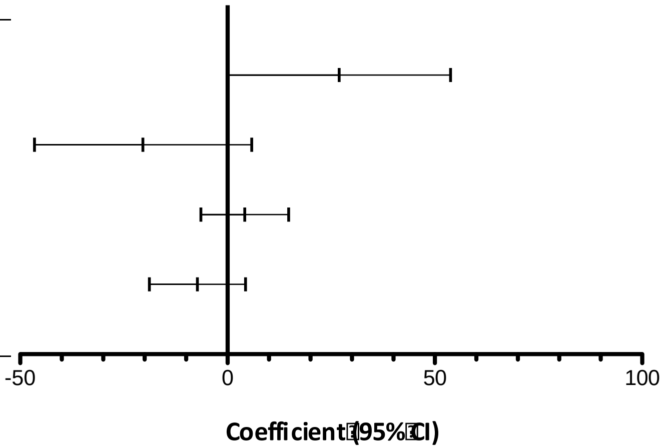
Predictor	Term (n=16) Median (Range)	Preterm (n=17) Median (Range)	Coef.	95%CI	P
Study group	39^{+6} (37^{+0} – 42^{+1})	26^{+3} (23^{+6} – 29^{+7})	30.61	7.70 – 53.53	.01
Male sex	9 (50%)	14 (82%)	-18.29	-42.23 – 5.66	.13
IMD quintile	4 (1)	3 (1)	4.15	-4.53 – 12.84	.34
Const	-	-	30.93	-4.97 – 66.82	.000



Overall model fit was R^2 =.21

Table 6-21. . Linear multiple regression model for the total number of trials completed without a demonstration from the experimenter with the adjustment of the 30 month cognitive z-scores ($F(4, 28) = 1.93$, $p = .13$). Baseline group were term born females with an IMD quintile of 1.

Predictor	Term (n=16)	Preterm (n=17)	Coef.	95%CI	P
	Median (Range)	Median (Range)			
Study group	40^{+1} (37^{+0} – 42^{+1})	26^{+2} (26^{+3} – 29^{+4})	26.91	.07 – 53.76	.05
Male sex	7 (43%)	14 (82%)	-20.37	-46.57 – 5.84	.12
IMD quintile	4 (1.5)	3(1)	4.18	-6.39 – 14.74	.43
2 year cog z-score	.13 (1.11)	-.50 (.91)	-7.25	-18.86 – 4.36	.21
Const	-	-	32.60	-10.55 – 75.75	



Overall model fit was $R^2 = .22$

6.1.2.3 Multi-Location Multi-Step (MLMS) results

Twenty-five term-born and 24 VP toddlers completed the Multi-Location Multi-Step paradigm (Table 6-22). Twenty-one infants from each study group passed the pre-switch criterion (Table 6-23). Those that did not pass the pre-switch criterion were excluded from the subsequent analyses. During data collection, physical hesitations were observed where the toddlers often paused over incorrect locations without making a physical error. By using the 'time to correct response', delays in motor responses were reflected in the response time.

Overall there were no differences observed between the term and VP responses to the MLMS task. However, term born infants appeared to show a greater increase in speed of response from the training trials the first of the consecutive pre-switch trials, compared to the VP infants that displayed a greater decrease to the second pre-switch trial.

		Term (n=25)	***VP (n=24)***
Gestational age	*Median (range); weeks^{+d}*	40^{+2} (37^{+0} –42^{+0})	26^{+2} (23^{+4} – 31^{+4})
Male sex		11 (44%)	17 (70.83%)
IMD Quintile	1	3	5
	2	4	1
	3	3	9
	4	9	6
	5	6	3
Bayley-III cognitive composite score at 12m	Mean (SD) (n:T=23; VP=22)	106.52 (11.22)	101.81 (10.18)
Bayley-III Cognitive z-score at 12m	Mean (SD) (n:T=23; VP=22)	-.10 (.90)	-.47 (.82)

Table 6-22. Demographic details of toddlers to complete the Multi-Location Multi-Step Paradigm.

	Term	*Preterm*
Criterion reached	21	21
Criterion not-reached	4	3
		Fisher's exact = 1.00

Table 6-23. Proportion of Term and Very Preterm children to pass the pre-switch criterion

Term born infants took a median of 8.64 seconds to complete the training phase compared to the VP who took 9.45 (z = -.29, p = .77). The first pre-switch trial created the greatest number of errors, 6 within the term group (24%) and 4 within the VP group (19%). The children had to correctly locate the snack in 3 consecutive tasks during the pre-switch phase, 17 of the 21 term-born children achieved this in 3 trials compared to 18 of the VP children (z = .33, p = .74; Table 6-24).

Number of Preswitch trials	***Term (n = 21)***	***Preterm (n = 21)***
3	17	18
4	2	1
5	2	1
6	0	1

MWU: z = .33, p = .74

Table 6-24. Total number of trials to reach pre-switch criterion by study group.

The three consecutive pre-switch trials for each child were then explored. The response times (RTs) were calculated as a proportion of the training RT, with term born children showing a greater decrease from the training phase in the first pre-switch trial compared to the VP infants (z = -2.73, p < .006; -1.55, p = .12 respectively). The VP infants then display a greater decrease from trial one to trial two (z = -1.20, p = .23), whereas the term- born children plateau (z = -3.46, p < .001; Table 6-25). From the final pre-switch trial, the VP children displayed a greater increase in their relative response to the post-switch trial but not significantly so (Figure 6-13)

	Term *Median RTs (IQR)*	*Preterm* *Median RTs (IQR)*
Training phase (Adjusted training phase = 1)	8.64 (6.56 – 10.28)	9.45 (6.30 – 14.18)
Preswitch trial 1 (adjusted)	.74 (.62 – 1.29)	.88 (.65 – 1.06)
Consecutive preswitch trial 1 (adjusted)	.73 (.62 – .88)	.79 (.65 – 1.02)
Consecutive preswitch trial 2 (adjusted)	.77 (.48 – .96)	.61 (.46 – .79)
Consecutive preswitch trial 3 (adjusted)	.76 (.48 – 1.00)	.66 (.45 – .94)
Post-switch (adjusted)	.71 (.53 – 1.15)	.76 (.59 – 1.49)

Table 6-25. Median trial times for MLMS task including children that met the task criterion.

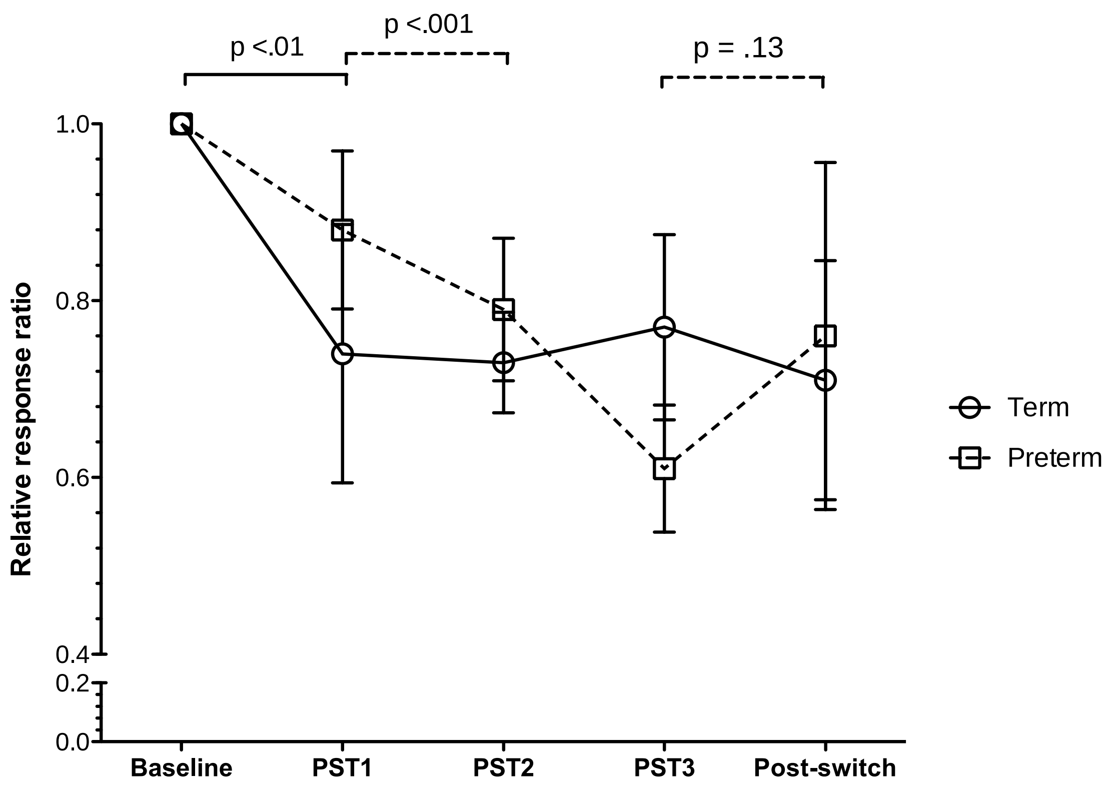
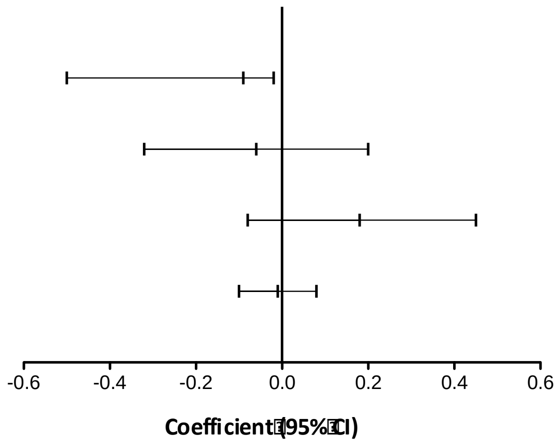


Figure 6-13. Time taken relative to training phase across the duration of the Multi Location Multi Step task; Mann Whitney U was used to assess relative response times from the previous time points; adjusted according to the training RT (baseline) through the pre-switch trials (PST1–3) and post-switch trial.

When controlling for study group, male sex and IMD quintile, a decrease in relative RT of .05 across each phase of the trial was observed using a mixed effects models (β = -.05, 95% CI (-.09 – -.015), p = .006; Table 6-26). This effect persisted after the additional adjustment of the 2 year cognitive z-scores (Table 6-27). When inputting each time point into the model as an independent predictor, each trial was significantly lower than the baseline, with the greatest decrease being observed over the pre-switch phase. There was no significant relationship between time and study group.

Table 6-26. Multilevel mixed effects model of the time taken to complete the MLMS trials from baseline to the post-switch phase. The baseline group was the average time taken during the training trials in term-born females with an IMD quintile of 1.

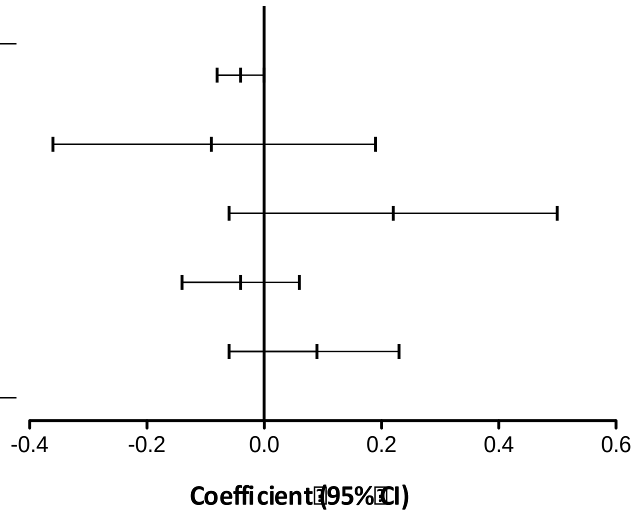
Predictor	Term (n=21)	Preterm (n=21)	Coef	95%CI	P
	Median (range)	Median (range)			
Time	-	-	-.05	-.09 - -.02	.006
Study group	39+6 (37+0 – 42+0)	26+5 (23+4 – 31+4)	-.06	-.32 - .20	.67
Male sex	8 (38%)	16 (76%)	.18	-.08 - .45	.18
IMD quintile	4 (1-5)	3 (1-5)	-.01	-.10 - .08	.78
(Const.)	-	-	-.20	-.58 - .18	.30



Wald chi^2 = 9.77 (p=.04)

Table 6-27. Multilevel mixed effects model of the time taken to complete the MLMS trials from baseline to the post-switch phase additionally adjusting for 12 month cognitive z-scores. The baseline group was the average time taken during the training trials in term-born females with an IMD quintile of 1.

Predictor	Term (n=19)	Preterm (n=20)	Coef	95%CI	P
	Median (range)	Median (range)			
Time	-	-	-.04	-.08 - -.00	.04
Study group	39+6 (37+0 – 42+0)	26+3 (23+4 – 29+4)	-.09	-.36 - .19	.53
Male sex	7 (37%)	16 (80%)	.22	-.06 - .50	.12
IMD quintile	4 (1-5)	3 (1-5)	-.04	-.14 - .06	.42
2 year Cog z-score	-.03 (.93)	-.38 (.80)	.09	-.06 - .23	.26
Time from baseline (Const.)	-	-	-.10	-.49 - .29	.60



Wald chi^2 = 7.97 (p<.16)

6.1.3 Discussion of behavioural Information processing measures

In the current study, 3 behavioural paradigms were used to investigate the processing speed difference. Firstly, at 3 months of age, the mobile reinforcement paradigm was used to investigate the speed of learning within the cohort.

In the current investigation a proportion of both term born and VP infants learnt to adapt their movements in response to the mobile during the conjugate reinforcement paradigm as measured by the relative response ratio previously utilised by Hayley *et* al., (2008). The speed at which the infants detected the mobile movement following the initiation of the training phase was fundamental to assessing speed of processing within the paradigm. During this phase, the term born infants appeared to plateau in their kick response to the mobile in contrast to the VP infants that showed a continual increase. This could be reflective of slower information processing by the VP infants, as the term born infants appeared to display a level of habituation to the mobile which was not displayed by the VP infants in the time provided. Slower habituation has been speculated to indicate less effective visual processing and could be suggestive of possible cognitive delays later in life (Slater, 1997). Previous reports have suggested that preterm infants displayed a reduced level of learning behaviour using similar paradigms to the mobile reinforcement at 3 months of age (Heathcock *et al.*, 2004; Haley, Weinberg and Grunau, 2006; Haley *et al.*, 2008; Lobo and Galloway, 2013). However, as there were no significant differences in the proportion of infants from each study group to learn at the task, it is not possible to conclude the same from the current cohort.

There are concerns within the literature regarding the use of the RRR due to the dependency of the measure on the natural leg kick frequency, particularly within a preterm cohort, who have been previously reported to display a higher kick rate than term born infants naturally (Heathcock *et al.*, 2005). In those with a high baseline, the infant would be expected to kick at a much faster rate than those with a lower kick rate (Millar and Weir, 2015). However, in this investigation, there was no significant difference between the term and preterm baseline leg kick rate.

Another confirmation of the validity of this measure was the significant increase in the effector leg within both learner groups during the training phase, with no association within the non-learner groups. This is in contrast to a previous report by Heathcock et al. (2005) who found poor lateralisation of response leg response in preterm infants. In the absence of longer term outcomes, Heathcock et al. (2005), speculated that failure to lateralise might identify infants in need of developmental support. This may still be a valid prediction, as the infants within the current study do not fall within the clinical cut off for developmental delay, at 12 or 30 months of age. The lateralisation observed here in combination with the higher proportion of VP infants to learn at the paradigm compared to previous investigations could therefore be an indication of higher cognitive performance within this cohort of preterm infants. The main effects of the investigation remained following adjustment for global cognitive performance as measured by the Bayley-III 12 month z-scores.

The infants within the learning groups also went on to generalise their response to the new mobile in the final phase. Both the lateralisation of leg kick response and the maintained RRR during the generalisation phase gives support to the tasks validity. However, the term and VP infants that did not learn during the task could be reflective of one or more of the paradigm limitations. Although infants that were clearly distressed and/or tired were excluded from the analysis, the paradigm was the last to be run in an hour long assessment and therefore, although appeared reactive, infants may have performed differently on a different day. Other limitations include possible vision difficulties, particularly within the preterm infants. Notes were kept on reported problems but more subtle difficulties may not have been detected in the timeframe of the study. Both mobiles, however, produced noise upon movement, and it was agreed that the sound response from the mobile could stimulate those with poor vision in a similar manner to the visual movement.

The second assessment explored within this chapter was a pilot investigation of the BabyScreen Application using touchscreen technology. The results suggested a difference in overall performance according to study group, with the term born children completing more items, with fewer instances of assistance required in the form of item demonstrations, compared to the VP children. This result was still present after adjusting for cognitive scores, suggesting that the application was accessing an aspect of cognitive performance that was not accounted for by the Bayley-III cognitive scale. The initiation of the demonstration occurred automatically at 30 seconds, suggesting that the VP infants required more time to complete each item overall, even though not significantly so on an item by item basis.

When looking within the specific constructs within the task, VP children were significantly less accurate that their term born counterparts on the first item designed to assess working memory. The nature of this trial required the child to recognise the location of the target and to use that information before initiating their response. It could be speculated that the VP children did not process this information fast enough to update their motor responses accordingly before starting their search for the target. This accuracy improved in the second working memory item. Speed of processing and working memory capacity have long since been associated within the preterm population in relation to later cognitive difficulties (Rose and Feldman, 1996; Mulder, Pitchford and Marlow, 2010). The difficulties in this item may reflect this.

The final item structure of the paradigm could be argued to be of greatest difficulty and therefore most informative. Item 17 included all touch processes used within the paradigm. The preterm children were significantly slower to complete this item than their term born counterparts, and although the response times were significantly but weakly correlated with the z-scores on the Bayley-III, there was still a significant effect of study group following global score adjustments. This suggests the paradigm may access difficulties within the VP children that the Bayley-III scale is not detecting.

The technology focused world we live in today means that children are exposed to different types of information and information platforms that have not been investigated fully in the developmental literature. The use of touchscreens in our everyday society has become unavoidable and children have been shown to attend well to such devices (Lovato and Waxman, 2016). Traditional developmental assessments in early childhood are known to be difficult to administer and arduous (Marlow, 2013), utilising the attraction of this technology on 2 year olds as a method of assessing EF skills hoped to overcome some of these issues.

Recent studies exploring the use of such technology have found favourable results suggesting its reliability and validity of testing psychometric profiles within young populations (Pitchford and Outhwaite, 2016). The application design used in the current investigation was developed to reduce the amount of verbal communication required to assess a child's cognitive performance profile and relied on the application recording the learning process displayed by the child's interactions with the touchscreen. Although primarily designed to assess the cognitive profile of individuals, the time-restricted element throughout the paradigm provided a measures of processing speed in association with the main cognitive outcomes.

Overall the performance on the paradigm was high, as the pass rate on a lot of the items was at ceiling. This could suggest that the initial items were too simple to detect any difference in constructs such as selective attention within this cohort. As the application is still under development, these results are exploratory in nature. Evidence is limited on how well the different constructs are targeting the domains proposed. However, although only a weak correlation, the association of the final item with the cognitive performance on the Bayley-III is encouraging for the task developers. The presence of a study group effect after cognitive score adjustment also supports theories that the Bayley-III scale is missing aspects of cognitive performance that could be reflective of later difficulties (Anderson and Burnett, 2017). Previous studies have reported the proportion of variance accounted for by cognitive performance equates to that of information processing speed performance (Rose, Feldman and Jankowski, 2009; Mulder, Pitchford and Marlow, 2010). The association of speed of information processing and cognitive abilities here is consistent with previous research within the preterm literature linking information processing difficulties with later difficulties in IQ.

The final behavioural task utilised, was the Multi-Location Multi-Step. Typically used to assess working memory and cognitive flexibility, following observations of hesitation behaviours during data collection, it was concluded that the paradigm was potentially reflecting processing speed difficulties. On a number of trials, the children would pause over an incorrect location before changing to correct one, almost visibly demonstrating an update in working memory. By measuring the speed of locating the snack within the task proposed to incorporate the hesitation behaviour as reliably categorising behaviours was not possible. The results, although not significantly different between groups, indicate that the term children performed at a more consistent level throughout the pre-switch phase of the paradigm and were able to update their motor responses sufficiently following the switch of snack location in contrast to the VP children. A gradual decrease in response times to the multi-step procedure during the pre-switch phase was observed within the VP cohort and when comparing responses from the last trial of the pre-switch phase to that of the post-switch, performance appeared to slow in the completion of the final trial following the switch. This delay could be picking up on the hesitation behaviours observed during data collection and could reflect a slower relay of information required to update the motor responses in order to make the correct box identification in the final phase.

Previous investigations have not utilised this task to assess processing speed, but rather working memory and cognitive flexibility. The study by Pozzetti *et al.*, (2014) utilised the performance on this task in combination with other experimental EF measures (spin the pots; reverse categorisation) to perform an exploratory factor analysis to determine which of the EF domain differentiated a cohort of preterm infants from a control group. Of the 3 factors they generated from the multiple paradigms, cognitive flexibility discriminated between the two groups. This factor was in part provided by performance on the Multi location Multi step paradigm (Pozzetti *et al.*, 2014). Perseverative errors within this paradigm have been observed in previous investigations of 2 year olds. The developers of the MLMS Zelazo *et al.,* (1998) reported preservative errors in a cohort of 2 years olds with a mean age of 24 months. When the children were asked to physically search during a pre-switch phase, they display perseveration in the post-switch trials. The absences of these behaviours in the current study could be an effect of maturity, with the children able to modify their responses sufficiently to not produce these errors by 30 months of age.

Although not traditional measures of processing speed, the tasks included within this chapter were posited to tap into this ability. Delays in information processing speeds have been commonly reported within the preterm literature (Rose, Feldman and Jankowski, 2009; Mulder, Pitchford and Marlow, 2010) and many suggest it is a primary cause of global cognitive impairments (Rose, Feldman and Jankowski, 2009) and the academic difficulties seen later in life (Mulder, Pitchford and Marlow, 2010). Information processing abilities are vital to later cognitive performance, and have been shown to mediate performance on EF tasks in ex-preterm cohorts at 10 years of age (Mulder, Pitchford and Marlow, 2011b). In contrast Aarnoudse-Moens *et al.*(2009) observed differences in inhibition and shifting behaviours that were still apparent following adjustment of processing speeds. Few studies have investigated processing speed in very preterm infants within the second year on life. However from those that have, difficulties in processing speeds appear to effect the performance of various cognitive domains throughout the early years (Rose, Feldman and Jankowski, 2009).

Overall the differences in speed of information processing within these two populations are subtle. Although there are some clear differences in performance throughout the different tasks, the results are not conclusive of processing difficulties. One explanation might be the relative high cognitive scores of our VP sample. The processing speed differences are more apparent within the 30 month tasks compared to the performance at 3 months. This is perhaps suggestive that the

3 month time point is too early to display any significant processing speed differences; or it could be indicative of developmental changes within the brain over the 2 years that is impacting processing performance. If to consider the interactive specialisation framework proposed by Johnson (2000), interactions between cortical regions and networks sharpen with age, becoming more refined and specialised. The reduced processing speeds observed at 30 months could therefore be indicative of abnormal connectivity development between networks. A more detailed investigation of processing speed advancements through toddlerhood was run by Rose *et al.* (2009). A consistency of processing speed delays within a group of preterm infants was reported from 7 months to 3 years in this investigation (Rose, Feldman and Jankowski, 2009). Within the PDP assessment, supplementary tasks were performed that could have provided additional measures of processing speed, a gaze-shifting paradigm for example. However, these tasks contained social components and it was concluded that this additional element could bring confusion to conclusions in the current context.

6.2 Neural Measures: Auditory processing

Arguably the most effective way to investigate information processing speeds is with the use of electroencephalography (EEG) recordings, specifically with Event Related Potentials (ERPs) that illustrate the brains response to a specific stimulus. The temporal accuracy of this technique lends to the exploration of individual and group variability in speed of information processing.

Highlighted in section 1.4, MR research by Northam et al., postulated an association between reduced interhemispheric fibres connecting the two auditory cortices and the language delays typically reported in preterm children (Northam *et al.*, 2012). These findings in combination with the multiple studies implying auditory discrimination difficulties within preterm cohorts (Jansson-Verkasalo *et al.*, 2003, 2004, 2010; Therien *et al.*, 2004; Mikkola *et al.*, 2007; Ortiz-Mantilla *et al.*, 2008) led to the development of an ERP paradigm designed to investigate the auditory attentional response between the two cohorts. The research question posed was whether any temporal delays across the hemispheres of the brain could be observed in preterm toddlers, indicative of auditory information processing difficulties. The primary objective of this investigation was to simply look at the speed of processing of auditory sounds in the cohort at 2 years.

The neural pathway of auditory information following initial detection is not conclusively understood. Unlike the clear hemi-decussation of the ophthalmic pathways (Winawer and Horiguchi, 2015), the auditory pathway has multiple points where the two ears converge. This aids with spatial orientation and sound location interpretation (Scott and Wise, 2004). However, there does appear to be a consensus in the literature that suggests the dominant pathway runs contralaterally from the ear of stimulation (Scott and Wise, 2004; Stefanatos *et al.*, 2008). The information then is required to be passed between the two hemispheres following the long-standing accepted view that the two auditory cortices are responsible for different aspects of sound interpretation and understanding (Nicholls, 1996). The structure responsible for relaying this information is the interhemispheric pathways via the Corpus Callosum; specifically the splenium (Bamiou *et al.*, 2007). This structure has been identified as small in size within preterm populations (Northam *et al.*, 2012; Thompson *et al.*, 2012; Nosarti *et al.*, 2014; Pannek *et al.*, 2014), and it is speculated that it is the inefficient transfer of information across this structure that impacts language development (Chiara Nosarti *et al.*, 2004; Northam *et al.*, 2012; Bruneau *et al.*, 2015).

With this as a focus, an auditory ERP paradigm aimed to assess the speed of auditory processing across hemispheres. To achieve this, sounds were presented monaurally in an oddball structure, employing the rationale that sounds are first processed contralaterally to the ear of presentation and following interhemispheric communication, are secondarily processed in the ipsilateral hemisphere. There are numerous theories surrounding the trajectory of a sound with speculations as to whether a sound is always required to be bisymmetrically processed (direct access model by Zaidel (1986)). A detailed review of this literature is provided in Appendix 4, explaining the current theoretical auditory pathway and the different models proposed for the path trajectory of a sound following detection. In any regard, auditory attention will be the primary focus of the current investigation. The following components were hypothesised to be observed: the involuntary attentional response (typically termed the N1 component) and the later more analytical component for an auditory sound (P3). These predicted components are characteristically observed within auditory ERP investigations (Escera *et al.*, 2000), however, it is important to consider the age of the participants.

Developmental changes of auditory ERP components are complex as they are reflective of cortical folding throughout the early years and are known to flip polarities as maturation occurs (Kushnerenko, Van den Bergh and Winkler, 2013). Typically, a large amount of research has been conducted within the first year of life, and then later at approximately 5 years (Hövel *et al.*, 2014). A few studies have investigated ERP components after the first year (Choudhury and Benasich, 2011; Putkinen *et al.*, 2012), however, the predominant focus has been on the deviant response during auditory oddball paradigms with the components of interest peaking in the fronto-central region (Mismatch Negativity or MMN). Due to the ambition of the current investigation to target possible dysfunction within the interhemispheric transfer of auditory information, the MMN response was not the main focus of the investigation. Although the N1 and P3 components are hypothesised to be observed, due to the attempt to address the question of reduced speed of transcallosal transfer, the monaurally presentation proposed an additional complication when predicting the location of these components. Although a number of monaural ERP paradigms have been conducted within the literature (Gilmore, Clementz and Berg, 2009; Bruneau *et al.*, 2015), few have been run in children, and to the authors' knowledge, none in a 30 month old population. A gap is therefore inherently present within the literature and due to this, component analysis was primarily data driven.

The paradigm comprised two simple speech sounds, presented monaurally in an oddball design. Due to the age of the infants, a paradigm that required prolonged active attention in order to obtain sufficient data was not going to be possible. The paradigm was therefore designed to be passive, to allow for maximum data collection. Hemispheric differences in ERP responses were investigated to compare possible transcallosal delays. To the authors' knowledge, no other ERP study has explored the neural response described, in a preterm population at 2 years of age.

6.2.1 Auditory Oddball Paradigm - methods

The two monosyllabic, computer-synthesized consonant-vowel sounds were utilised: da/-/ga/ continuum. The two 100ms phoneme variants were chosen as both were voiced sounds with consistent voice onset time (VOT), however, were distinctly different due to the articulation location of the consonant sounds. /da/ is an alveolar sound, as the tongue is placed behind the alveolar ridge at the front of the mouth to produce the sound, whereas, /ga/ a velar sound, is produced by placing the tongue further back on the soft palate at the roof of the back of the mouth. The two sounds were clearly distinguishable but comparative in the wave structure for use in this paradigm. Amplitude and frequency of both sounds were explored to ensure the sounds did not deviate too greatly from one another, see

Figure 6-14 – 6-16. An initial sound burst was present to a greater extent in the /ga/ sound, this had the potential to cause timing issues if the sounds were counterbalanced between the frequent and infrequent stimuli between subjects. The sounds were therefore kept consistent to avoid any timing differences. Both were shared with the PDP study by a fellow research group within the ICH Developmental Cognitive Neurosciences and Neuropsychiatry section, headed by Professor Baldeweg.

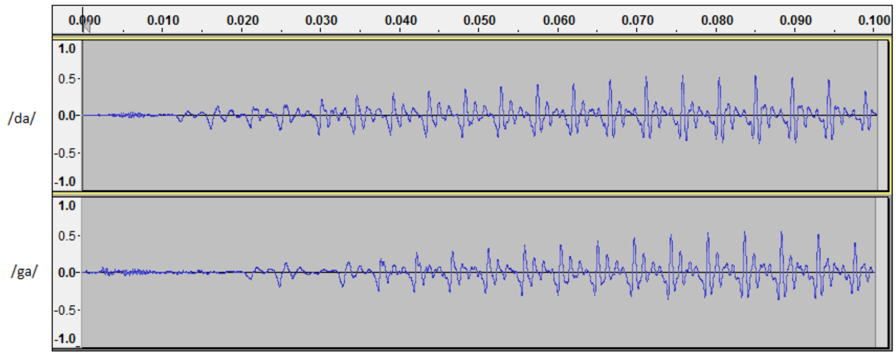


Figure 6-14. Waveform display of both stimuli used in the auditory oddball paradigm (/da/ and /ga/ sounds respectively).

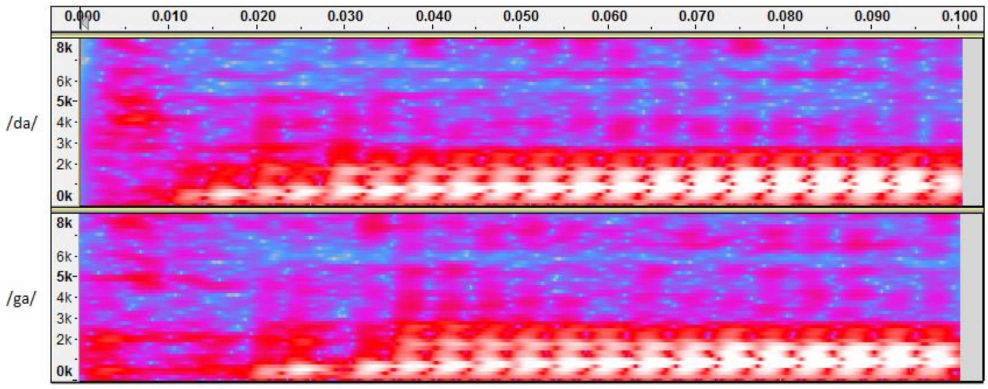
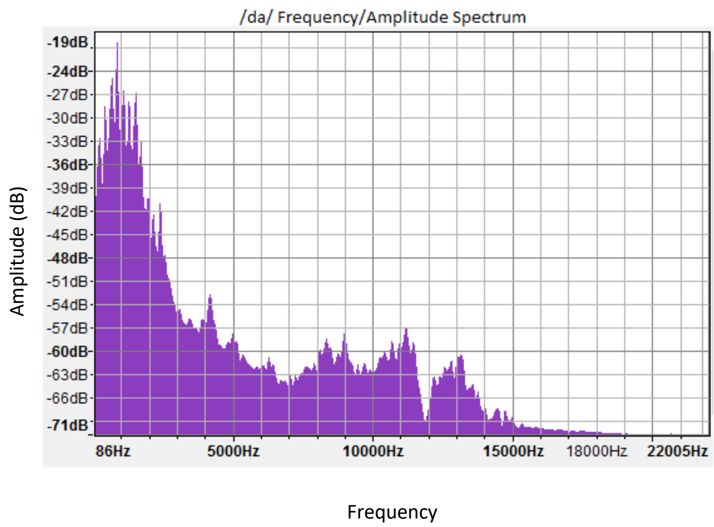


Figure 6-15. Spectral view of /da/ and /ga/ sounds respectively.



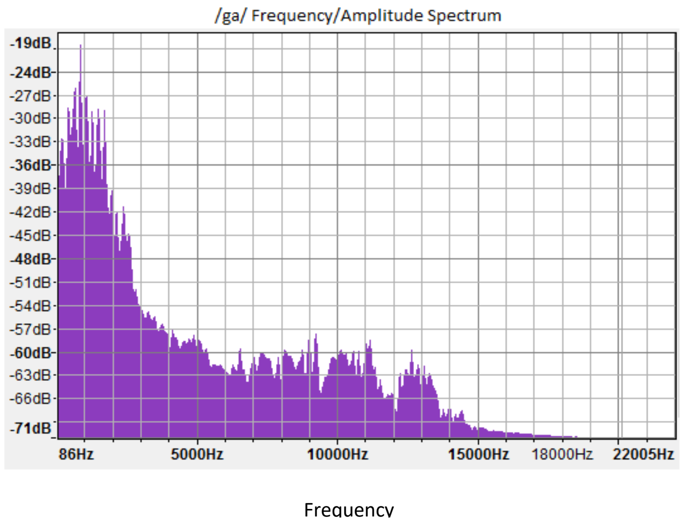


Figure 6-16. Frequency/amplitude spectrum plots for /da/ and /ga/ sounds respectively.

The paradigm was structured in an oddball design within MATLAB. The frequency of the two phonemes was presented in a 70:30 ratio. /da/ was presented as the deviant sound and /ga/ as the standard. The sounds were presented monaurally in 4 blocks of 100 trials with a random inter-stimulus interval of between 0.7-1s. Each block alternated between the left and right ear presentation, with the first ear of presentation counterbalanced between subjects.

A note here on the paradigm design. Due to the age of the participants, the maximum number of trials tolerated was 400 trials. This gave 280 standard sounds and 120 deviant sounds. The ratio of the frequent to infrequent was set marginally higher than that of traditional oddball paradigms in an attempt to increase the number of deviant trials without losing the novelty response. This was a risk in the design, however, the novelty response was not the predominant response of interest, as explained above. Due to this, the deviant trials were compared to the standard trials to explore the novelty response, but the standard trials are the primary focus of the analyses.

The geodesic EGI system was used for the recordings. The toddlers were either sat on a parents lap or on a small chair in front of the VDU. Distance from the VDU was not measured as it was not fundamental to the task. The child was positioned so that the video camera could see the child throughout the task however. The net was placed using the procedure described below. Lightweight children's headphones, supplied by Urbanz, were then gently placed over the EGI net, taking care to not rest on the reference electrode.

To reduce movement artefacts the toddlers watched either a popular children's cartoon without sound on the VDU or were able to watch a video of choice on an iPad, again without sound. The lights were switched off and experimenters and parents did not speak throughout the administration.

6.2.1.1 ERP Pre-processing

The EEG was acquired using a 32-channel EGI Geodesic Array Sensor net against an online vertex reference. Netstation 4.5.1 was used to record the data (Electrical Geodesic, Eugene, Oregon), at a sampling frequency of 250 Hz and impedences kept below 80 kohm.

Participant head circumference was measured prior to the initiation of the EEG session and the appropriate sensor net was prepared accordingly. The nets were soaked for a total of 5 minutes in 1000mls of lukewarm water, with 5ml of baby shampoo and 10g of Potassium Chloride. When placing the net, the central electrode, 18, was positioned in line with the nasion, and the reference vertex electrode centred between the ears and in line with the inion. Once positioned, the impedances were measured, corrected and saved before the start of each recording.

Post-recording, the data was passed through a bandpass filter offline of 0.1-30Hz. The data was segmented into 1000ms epochs; 200ms pre-stimulus onset and 800ms post-stimulus. The artefact detection tool was programmed to mark channels as bad according to the following parameters within each epoch or trial: if the maximum to minimum peak amplitude exceeded 100µV, accounting for eye-movement and eye-blink artefacts; if ≥ 5 channels were marked bad within one trial, the entire trial was excluded; and lastly, if the >4 channel were in the same topographic position the trial was marked as bad. A participant was excluded from further analysis if less than 10 good trials per condition resulted from this processing tool. The automated process was check manually, ensuring that all bad channels had been correctly identified.

Following artefact detection, the bad channel replacement tool used a pre-programmed algorithm to estimate the voltages of the channels marked bad by using the voltage patterns of surrounding electrodes. The data was then averaged across each participant and re-referenced to the average reference, excluding the EOG channels, before being baseline corrected, 100ms pre-stimulus onset.

Grand average waveforms were created for the two groups, term and preterm. The waveforms for each condition were compared within and between groups. Maximum and minimum amplitudes were calculated for each component within the waveforms, using a maximum and minimum peaks and time-window analysis.

The steps for data processing were as follows: segmentation, artefact detection, bad channel replacement, averaging, averaged referencing, and baseline correction before all participant data was compiled into one grand average file.

6.2.1.2 ERP Analysis

Butterfly plots were used to determine the components present within the data and for electrode selection. The plots were explored across all individuals, with the electrodes and time window of the component peaks noted from each. From these plots, two components were identified: a negative peaking component around 100ms (will be termed N1) with a window of 80-200ms, with the most prominent electrode grouping as: 15, 23, 25 (left hemisphere); and 16, 24, 26 (right hemisphere) according to the EGI 32 channel geodesic map (see Figure 6-17); and positive peaking component around 300ms (will be termed P3) with a window of 235-395ms, with the most prominent electrode grouping as: 13, 15, 23, 25 (left hemisphere); and 14, 16, 24, 26 (right hemisphere), according to the EGI 32 channel geodesic map (see Figure 6-18).

The paradigm therefore produced responses to: standard and deviant phonemes, responses according to the ear of presentation (sounds presented to the left or right ear), and each component was investigated in the left and right hemisphere independently. Therefore the standard and deviant sounds were handled separately in the first instance for clarity, with the standard sounds the primary focus. For each component, amplitude and latency measures were investigate using a repeated measures ANOVA with the between-subject factor as study group and with within-subject measures of ear of stimulation and hemisphere. This was repeated within the deviant sounds.

For the N1 component, the minimum amplitude was reported for the standard phonemes along with the latency measures, however for the P3 component, the mean amplitude was investigated and no latency measures were investigated due to the broader nature of the component. For the deviant tones, due to the fewer number of trials available, the mean amplitudes for both components were investigated. Finally, differences in the two components were compared between the standard and deviant sounds. Although the novelty response is often centralised, there was little evidence available to formulate a prediction for where the novelty response might be observed in monoaural presentations as reports are mixed (Gilmore, Clementz and Berg, 2009). As such, it was decided to persist with a data lead approach and for these analyses the data were collapsed across the ear of presentation and hemisphere, to explore the novelty effect of the deviant phonemes within the N1 and P3 components. Both the minimum and mean peaks and peak latencies were explored for the N1, and the only the mean peak investigate for the P3 component. The electrode selection remained consistent with the independent trial type analysis as the butterfly plots did not elect a strong centralised response in the central region.

The variable selected a priori to best reflect the performance on this task was the left hemisphere N1 latency measure, collapsed across ear of sound presentation and standard and deviant tones. This was selected as the left hemisphere is considered the predominant hemisphere for speech processing therefore hypothesised to display the greatest response. As difference in speed processing was hypothesized, a latency response was selected as the most appropriate reflection of a speed difference. Evaluation of the sound is most likely to be reflected in the P3 component; however, the more prominent response was hypothesised to be the N1 due to the automaticity of detecting a sound and was therefore selected on this basis.

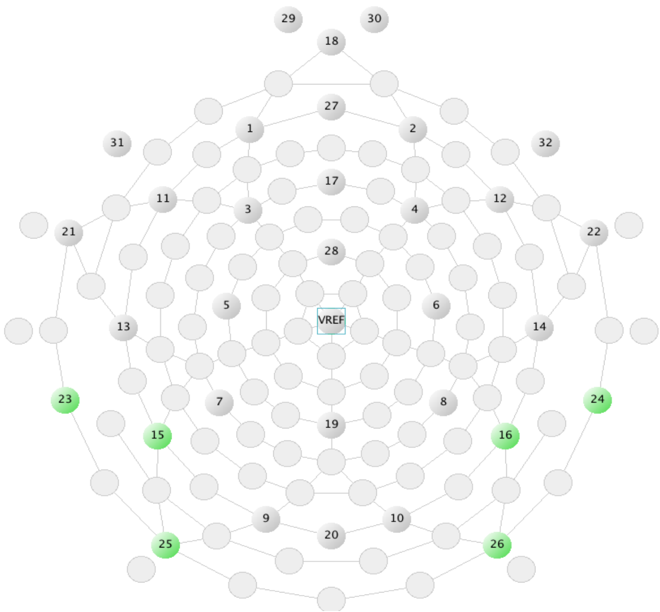


Figure 6-17. EGI 32 channel Geodesic map display electrode selection for N1 component

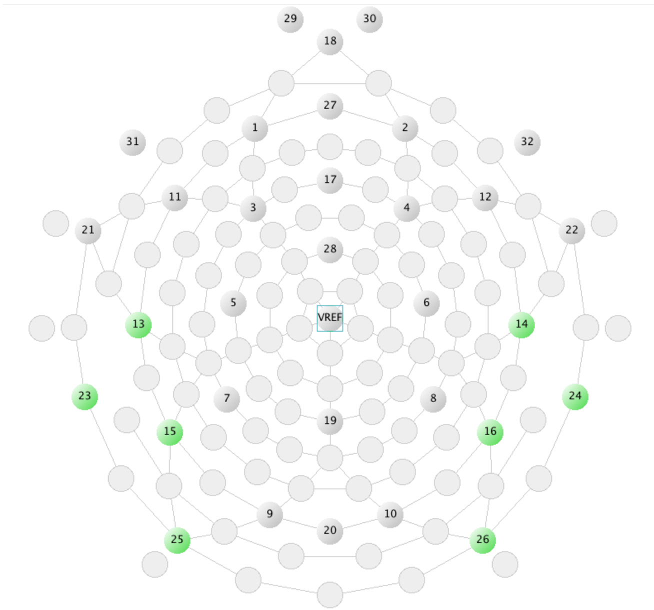


Figure 6-18. EGI 32 channel Geodesic map display electrode selection for N1 component

6.2.2 Results

Fourteen term born and fourteen VP infants were included in the auditory oddball ERP analysis (Table 6-28).

		Term (n=14)	***VP (n=14)***
Gestational age	*Median (range); weeks^{+d}*	40^{+4} (37^{+2} – 42^{+1})	24^{+6} (23^{+6} – 29^{+4})
Male sex		7 (50%)	10 (71.43%)
IMD Quintile	1	1	2
	2	1	1
	3	3	6
	4	5	3
	5	4	2
Bayley-III cognitive composite score at 2 years	Mean (SD) (n:T=13; VP=14)	110.77 (14.98)	101.07 (12.28)
Bayley-III Cognitive z-score at 2 years	Mean (SD) (n:T=13; VP=14)	.25 (1.21)	-.56
Bayley-III language composite score at 2 years	Mean (SD) (n:T=11; VP=14)	121.18 (10.83)	96.35 (18.39)
Bayley-III language z-score at 2 years	Mean (SD) (n:T=11; VP=14)	.17 (1.04)	-2.22 (1.77)

Table 6-28. Demographic details of toddlers to complete the auditory oddball paradigm

6.2.2.1 Standard Phonemes trials

In the exploration of the N1 component, the term born children produced a greater negative response to a left ear sound across both hemispheres compared to the VP infants (Figure 6-19; main effect of study group: $F(2,26) = 3.46$, $p = .074$; with post hoc tests suggesting greater negativity in the terms with Bonferroni correction: $t(26) = -2.01$, $p = .06$). No notable differences were observed with sounds to the right ear between the study groups or across hemispheres.

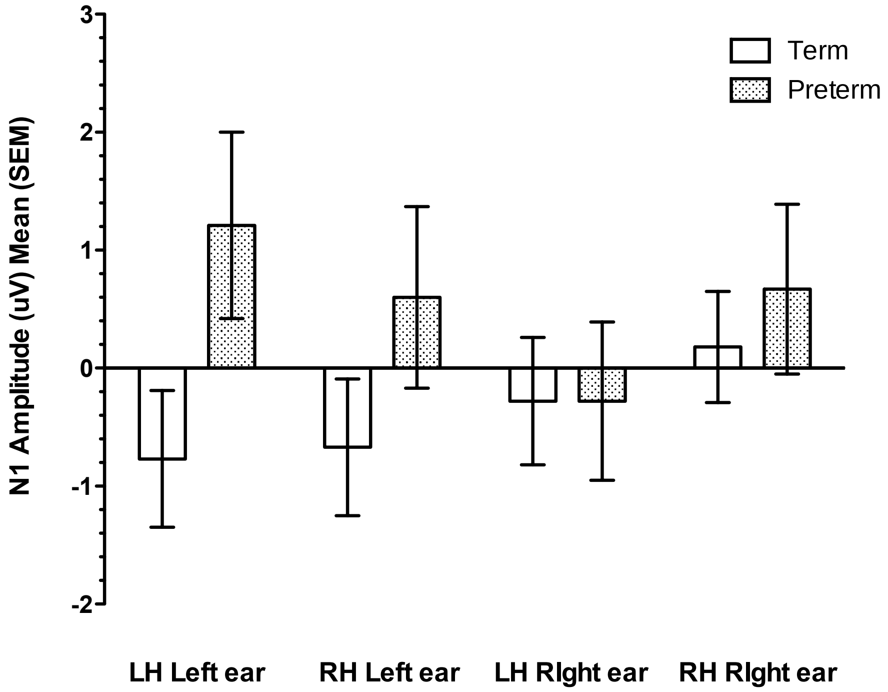


Figure 6-19. Mean amplitude response of the N1 component across the left (LH) and right (RH) hemispheres according to ear of presentation for standard sounds between term and very preterm toddlers.

For the P3 component, sounds to the left ear elicited a greater mean amplitude response in the right (contralateral) hemisphere for the term born infants compared to the VP infants (Figure 6-20; hemisphere by study group interaction: ($F(2,26) = 12.81$, $p = .001$); with post hoc tests suggesting a greater response in the terms: $z = 2.25$, $p = 0.02$). In addition, a significant effect of hemisphere was observed within groups, with terms showing a greater amplitude in the right (contralateral to the ear of stimulation; $z = -2.54$, $p = 0.01$); and the preterms showing a greater response in the left hemisphere (ipsilateral to the ear of stimulation; z = 2.04, p = 0.04).

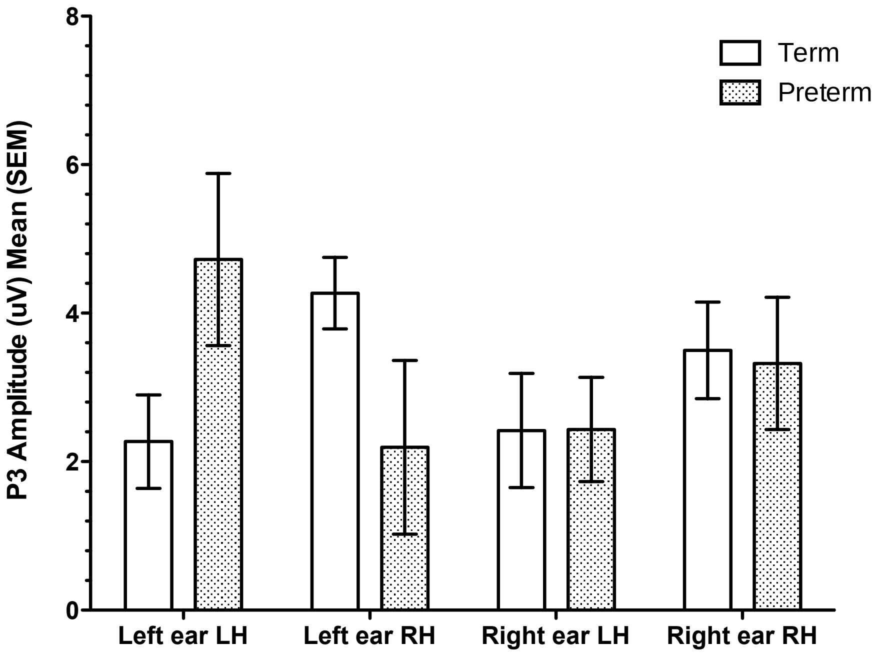


Figure 6-20. Mean amplitude response of the P3 component across the left (LH) and right (RH) hemispheres according to ear of presentation for standard sounds between term and very preterm toddlers.

		Term (n=14)		***Preterm (n=14)***	
Component	**Ear of stimulation; measure**	**Left hem (SD)**	**Right hem (SD)**	**Left hem (SD)**	**Right hem (SD)**
N1	Left Ear; mean amp	-.77 (2.16)	-.68 (2.18)	1.21 (2.97)	.60 (2.88)
	Right Ear; mean amp	-.28 (2.02)	.18 (1.76)	-.28 (1.42)	.67 (2.69)
	Left Ear; latency	126.29 (28.90)	120.57 (26.81)	132.67 (30.38)	133.43 (31.31)
	Right Ear; latency	130.57 (24.45)	118.67 (21.40)	124.67 (18.89)	120.95 (15.82)
P3	Left Ear; mean amp	2.27 (2.35)	4.27 (1.79)	4.72 (4.32)	2.19 (4.39)
	Right Ear; mean amp	2.40 (2.88)	3.50 (2.42)	2.43 (2.64)	3.32 (3.33)

Table 6-29. Auditory ERP N1 and P3 component mean amplitudes and latencies for the standard sounds across study groups.

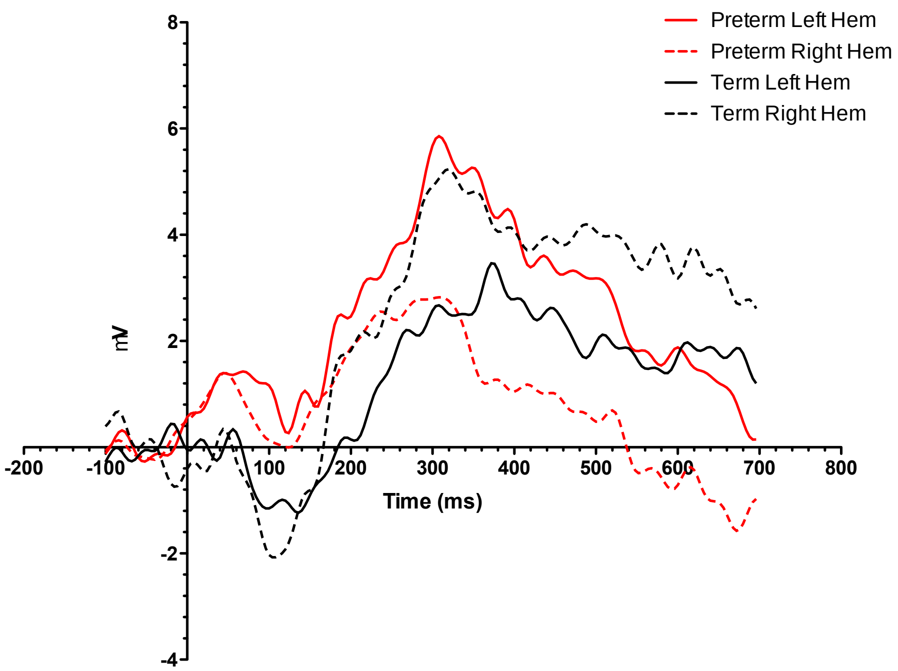


Figure 6-21. Auditory ERP waveform of the N1 and P3 component response to left ear stimulation across hemispheres with standard sounds; the N1 electrode grouping was used for this illustration.

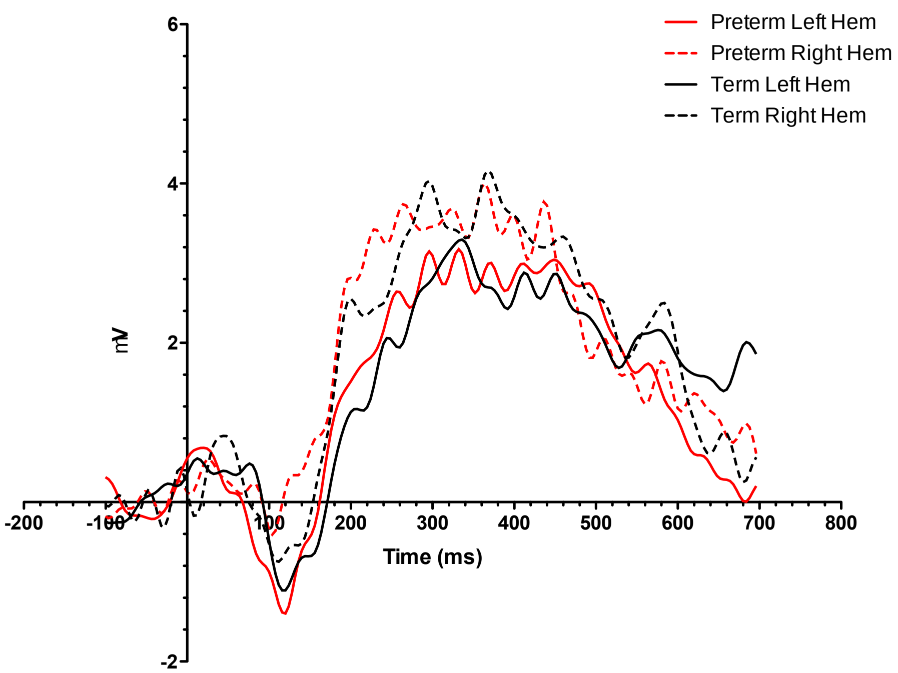


Figure 6-22. Auditory ERP waveform of the N1 and P3 component response to right ear stimulation across hemispheres with standard sounds; the N1 electrode grouping was used for this illustration.

6.2.2.2 Deviant Phonemes trials

For the N1 component in response to the deviant tones, the contralateral hemisphere displayed a greater negativity in both study groups when sounds were presented to the right ear (Figure 6-23; $F(2,26) = 7.51$, $p = .011$). However, the N1 component peaked faster in the ipsilateral hemisphere (latency main effect of hemisphere: $F(2,26) = 7.78$, $p = .01$; post-hoc exploration indicates a faster response in the right hemisphere: $z = 2.29$, $p = .022$; Figure 6-25). No differences were observed following deviant tones to the left ear (Figure 6-24).

For the P3 component, there were no significant main effects or interactions between the groups in either ear of presentation.

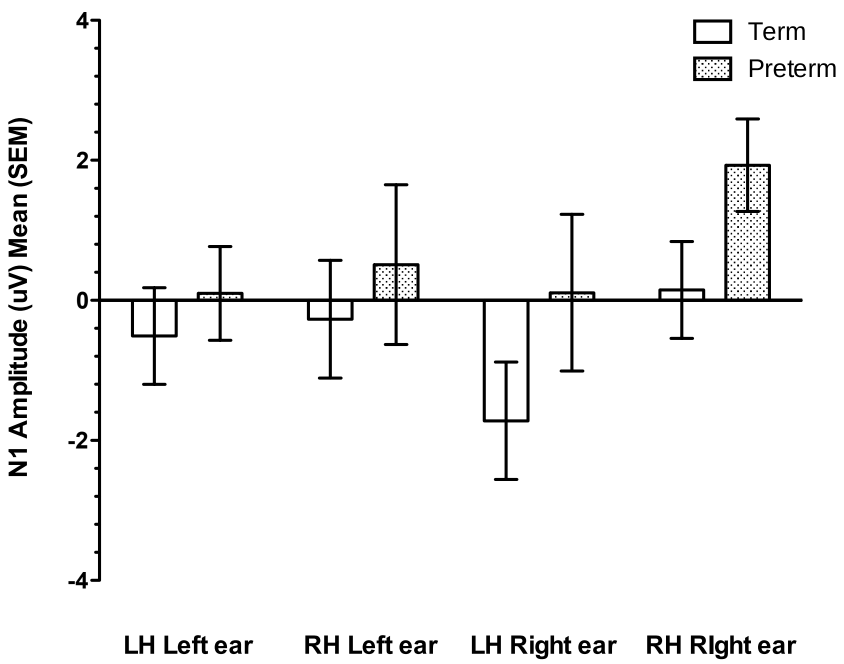


Figure 6-23. Mean amplitude response of the N1 component across the left (LH) and right (RH) hemispheres according to ear of presentation for deviant sounds between term and very preterm toddlers.

		Term (n=14)		*Preterm (n=14)*	
Component	**Side of sound presentation; measure**	**Left hem (SD)**	**Right hem (SD)**	**Left hem (SD)**	**Right hem (SD)**
N1	Left Ear; mean amp	-.51 (2.59)	-.27 (3.16)	.10 (2.52)	.51 (4.26)
	Right Ear; mean amp	-1.72 (3.13)	.15 (2.57)	.11 (4.18)	1.93 (2.46)
	Left Ear; latency	126.76 (25.63)	121.10 (33.09)	135.71 (30.16)	125.71 (41.14)
	Right Ear; latency	132.57 (23.75)	119.43 (22.36)	130.67 (16.78)	110.48 (27.41)
P3	Left Ear; mean amp	2.46 (2.48)	4.42 (5.28)	3.38 (4.58)	2.39 (5.27)
	Right Ear; mean amp	1.57 (2.43)	2.90 (2.21)	2.05 (5.68)	4.97 (4.88)

Table 6-30. N1 and P3 components mean amplitudes and latencies for deviant sounds from the auditory ERP oddball paradigm

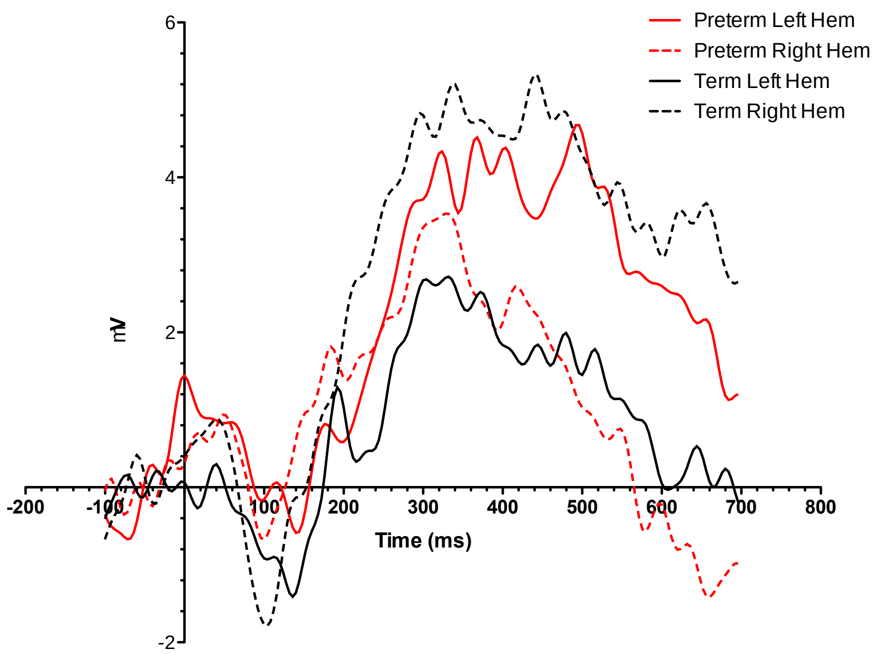


Figure 6-24. Waveform of the N1 and P3 response across hemispheres to the deviant sounds following left ear presentation. The electrode grouping used for this illustration was those for the N1 analyses.

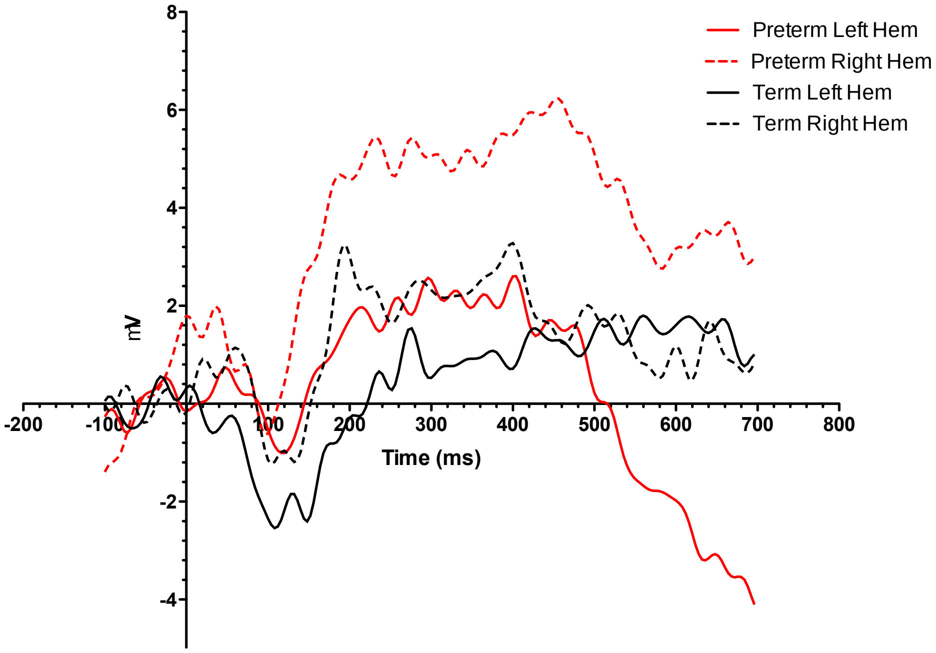


Figure 6-25. Waveform of the N1 and P3 response across hemispheres to the deviant sounds following right ear presentation. The electrode grouping used for this illustration was those for the N1 analyses.

6.2.2.3 Standard vs Deviant phoneme response

When comparing the differences in response between the deviant and standard phonemes, the responses were collapsed across hemisphere and ear of presentation as there were no expectations that one hemisphere or a particular ear of presentation would elect a greater response to a novel sound over the other (Table 6-31; Figure 6-26).

		Term (n=14)		*Preterm (n=14)*	
Component	**Measure**	**Standard**	**Deviant**	**Standard**	**Deviant**
N1	Mean amp	-1.2 (3.11)	-1.46 (4.16)	1.53 (4.76)	.90 (5.49)
	Min amp	-6.61 (3.48)	-8.44 (5.12)	-4.01 (5.10)	-7.32 (5.88)
	Latency	124.02 (16.48)	125.19 (17.67)	127.93 (17.39)	125.64 (20.23)
	Combined latency	124.61 (16.16)		126.79 (15.47)	
	Combined latency, left hemisphere only*	129.05 (19.58)		130.93 (16.45)	
P3	Mean amp	6.48 (3.32)	6.18 (4.89)	7.86 (7.11)	6.84 (7.76)
	Max amp	12.17 (3.79)	13.17 (5.04)	14.55 (7.54)	14.69 (7.85)

Table 6-31. N1 and P3 components mean, minimum and max amplitudes and latencies for standard and deviant sounds from the auditory ERP oddball paradigm. *Variable selected a priori as most informative reflection of performance: left hemisphere response collapsed across ear of sound presentation and sound frequency.

The deviant sounds elicited a greater minimum N1 response in both study groups ($F(2,26) = 5.61$, $p = .02$) compared to the standard tones; although this effect was weakened when averaging the peak N1 response ($F(2,26) = .225$, $p = .64$). However, the term mean amplitude was larger in response to a deviant tone compared to the VP infants ($F(1,26) = 3.31$, $p = .08$).

The P3 component did not appear to significantly differ between the two trial types, when exploring the mean amplitude of the component.

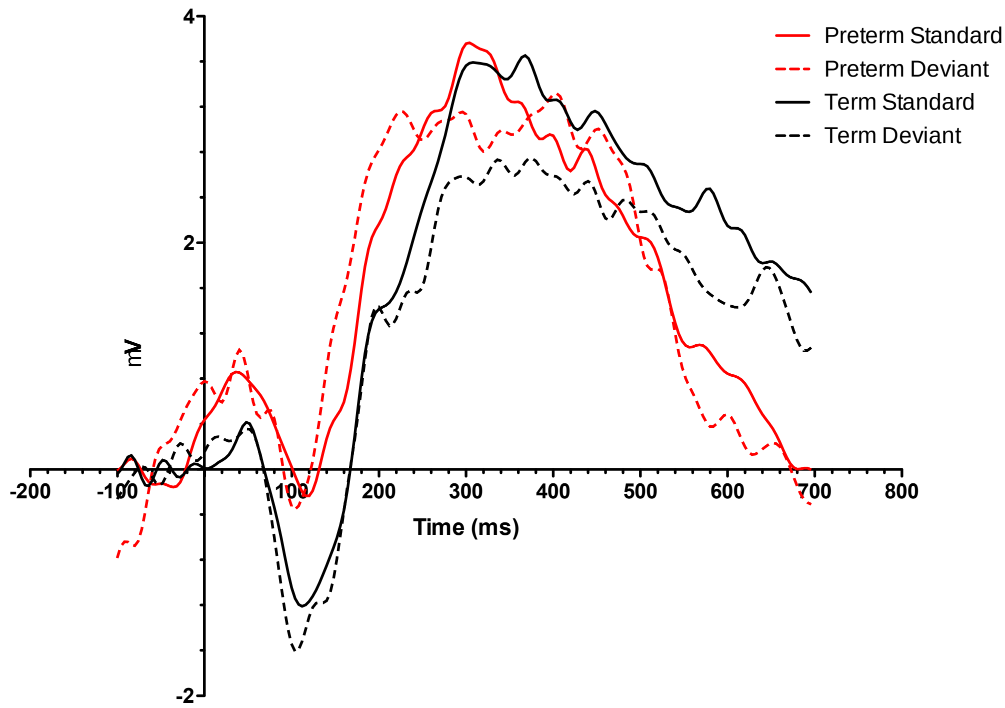
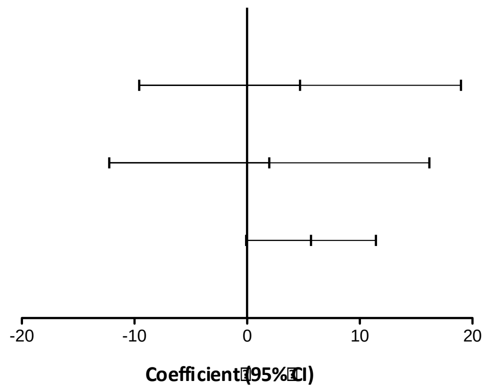


Figure 6-26. Auditory ERP waveform of the N1 and P3 components in response to the standard and deviant tones, averaged across ear of sound presentation and hemisphere; the N1 electrode grouping was used for this illustration.

The variable selected a priori was the N1 latency response from the left hemisphere collapsed across the ear of presentation and standard and deviant responses. Although no latency differences were observed, this was variable was further explored in relation to the study confound variables. IMD quintile had a significant impact on the first model, before adjustment for Bayley-III cognitive score, with greater deprivation suggesting a slower latency response to these sounds (Table 6-32). This however did not withstand adjustment for cognitive score (Table 6-33). This additional model also included the language z-score from the Bayley-III due to the speech sounds used within the paradigm. No significant relationships were highlighted (Table 6-32 and 6-33).

Table 6-32. Linear regression model of left hemisphere response to speech sounds in auditory oddball paradigm, collapsed over ear and frequency of sound presentation (F(3, 24) = 1.44, p = .25). The baseline group are term-born females at 12 months with an IMD quintile of 1.

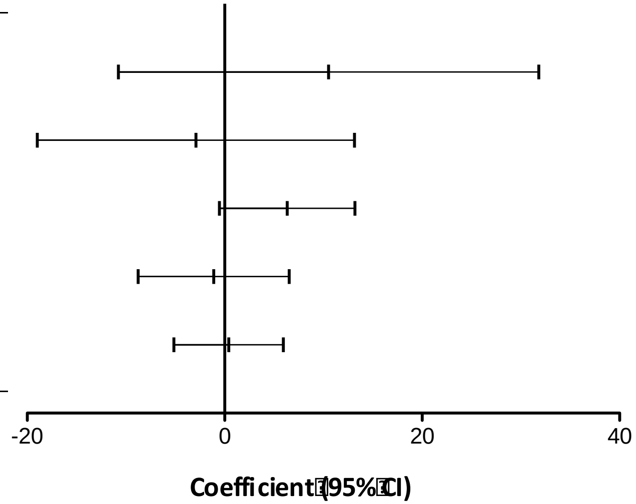
Predictor	Term (n=14)	Preterm (n=14)	Coef	95%CI	P
	Median (Range)	Median (Range)			
Study Group	40^{+4} (37^{+2} – 42^{+1})	24^{+6} (23^{+6} – 29^{+4})	4.70	-9.58 – 18.99	.50
Male sex	7 (50%)	10 (71.43%)	1.96	-12.25 – 16.17	.78
IMD Quintile	4 (1–5)	3 (1–5)	5.67	-.09 – 11.43	.05
(Const.)	-	-	106.99	82.69 – 131.28	.000



Overall model fit R^2 = 0.15

Table 6-33. Linear regression model of left hemisphere response to speech sounds in auditory oddball paradigm, collapsed over ear and frequency of sound presentation including the adjustment of the 12m Bayley-III cognitive and language z-score ($F(5, 19) = .91$, $p = .50$). The baseline group are term-born females at 12 months with an IMD quintile of 1.

Predictor	Term (n=11)	Preterm (n=14)	Coef.	95%CI	P
	Median (Range)	Median (Range)			
Study group	39^{+6} (37^{+0} – 42^{+1})	26^{+3} (23^{+6} – 29^{+7})	10.51	-10.78 – 31.82	.31
Male sex	4 (36.36%)	10 (71.43%)	-2.92	-18.97 – 13.14	.71
IMD quintile	4 (1-5)	3 (1-5)	6.33	-.53 – 13.19	.07
2 year cog z-score	.44 (1.21)	-.54 (.99)	-1.11	-8.76 –6.53	.76
2 year lang z-score	.17 (1.04)	-2.22 (1.77)	.40	-5.15 – 5.94	.88
Const	-	-	102.89	74.14 – 131.64	.000



Overall model fit was R^2 = .19

6.2.3 Discussion of Neural Information processing measures: Auditory ERP

The ERP auditory oddball paradigm elicited some very interesting differences between the term and VP toddler responses. The waveforms in Figure 6-19 and 6-20 highlight the main differences observed during the paradigm. Two components were identified: the N1 at approximately 100ms post stimulus presentation, and the P3 at approximately 300ms post-stimulus presentation. The N1 is typically regarded as the brains involuntary attention response and the P3 as conscious processing of the stimulus (Escera *et al.*, 2000).

During standard tone presentation to the left ear the term born infants displayed a greater N1 response compared to the VP infants and the P3 was largest in the term born right hemisphere, contralateral to the ear stimulated. In contrast the VP group displayed a stronger ipsilateral P3 response to a left ear standard tone. There were no distinguishable differences to the right ear sound presentations and no latency differences were observed to either ear presentation.

Multiple cerebral processes are speculated to contribute to the generation of the N1 component, with the primary process being the synchronisation of the primary and secondary auditory cortices after sensory perception. The N1 is classically recorded in the front-central location (Tomé *et al.*, 2014). Due to the monaural presentation in the current paradigm, the N1 was recorded over the temporo-parietal areas. An investigation conducted previously utilised an auditory oddball paradigm with sounds presented monaurally and similarly detected the N1 component in the temporo-parietal regions (Gilmore, Clementz and Berg, 2009). This investigation into interhemispheric lateralisation found a stronger response detected in the contralateral hemisphere to the ear of stimulation (Gilmore, Clementz and Berg, 2009). Although there were no significant differences observed across hemispheres, the N1 amplitudes in the current investigation suggest a greater contralateral response by both study groups to the ear of presentation (see Figure 6-19). This supports the theory that the contralateral pathways through the brain stem from the ear of stimulation are dominant in the processing of a sound in the first instance (Scott and Wise, 2004).

The P3 is classically observed in central locations and is typically produced in response to conscious processing of a sound (van Dinteren *et al.*, 2014). When a sound is presented monaurally, disagreements over the laterality of the P3 response are present within the literature (for a review please see (Gilmore, Clementz and Berg, 2009). It is plausible that the P3 observed in the current paradigm is not the classic component often reported in the literature, but it may reflect the auditory evaluation process due to the timing of the component (Escera *et al.*, 2000).

When comparing the P3 amplitudes of term and VP toddlers, the responses were very different. Using standard sounds, a hemisphere by group interaction was observed in response to left ear stimulation. The term born toddlers displayed a greater right hemisphere response; and the VP infants a greater left hemisphere response. The response of the term group reflects previous reports implying that the contralateral pathway is the dominant trajectory of a sound and upon detection is processed primarily in the contralateral hemisphere (Gilmore, Clementz and Berg, 2009). The VP toddlers however displayed a greater ipsilateral response. This difference is intriguing.

The broad nature of the P3 component in this instance prohibited the exploration of peak latency comparisons across hemispheres. On visual inspection of the waveforms following left ear stimulation in the term group (Figure 6-21) the latency of the contralateral response peaked before the ipsilateral response. This may reflect interhemispheric transmission of the sounds from the right to the left hemisphere. The sounds used in the paradigm were speech-based. This information may require transference to the left auditory cortex in order to be processed, as the left is typically considered to be the language dominant hemisphere. In contrast, the VP group response shows the reverse, with the ipsilateral response displaying greater amplitude and a considerably depressed left hemisphere response. This suggests that the VP group is handling the speech information differently to the term group.

There are multiple theories regarding when and how speech is then processed in ipsilateral hemisphere (Musiek, 1986; Musiek *et al.*, 1989; Bamiou *et al.*, 2007). The consensus appear to suggests that for speech sounds to be processed efficiently, they must be transferred to the speech dominant cortex, classically the left auditory cortex (Bamiou *et al.*, 2007). When a sound is presented to the left ear, in order to be processed in the left hemisphere, the theory dictates the information must be transferred through the transcallosal fibres. The greater ipsilateral response in the VP toddlers in response to left ear stimulation may suggest the VP children have stronger ipsilateral connections as a result of poorer communication across hemispheres via the transcallosal pathway. This interpretation is supported by the MR studies that have reported poorer tract formation and reduced speed of transfer of speech information across the corpus callosum (Northam *et al.*, 2012). The language scores are lower in the current cohort suggesting some language impairments may be present. It was not possible to fully explore the language association with the P3 component as it was considered too broad in this instance. However, this greater response in the left ipsilateral hemisphere to a left ear sound, could suggest that compensatory mechanisms have changed how the brain handles speech stimuli within this cohort, probably due to difficulties in interhemispheric communication, thereby ensuring direct processing to the language dominant hemisphere of the brain. In imaging studies of acallosal patients, the number of ipsilateral fibres exceeds that of those crossing contralaterally, suggesting the reduction or absence of interhemispheric fibres creates compensatory connections (Nowicka and Tacikowski, 2011). These findings are consistent with this explanation.

These data are interpreted with caution for several reasons. Firstly, the mean amplitude was the only possible measure for the P3 component and therefore the latencies could not be properly investigated. Handedness was not identified during the assessments due to inconsistent reports in the accuracy of handedness measures in individuals so young (for review see Scharoun and Bryden, 2014). Hand preference is often associated with the language domain hemisphere and it is possible that some within the cohort have right sided language dominance.

Typically in the literature, the deviant sounds within an oddball paradigm are detected by voltage deflections in the P3 response. The P3a component is often considered as an indication of conscious auditory attention, and reported in paradigms where the participants are asked to actively focus on the deviant sounds (van Dinteren *et al.*, 2014). This was not the case in the current investigation in which the N1 amplitude was significantly greater in response to the deviant sounds across both study groups, when collapsing across ear of sound presentation and hemisphere. A number of factors could explain this discrepancy. Firstly, due to the monaural presentation, the P3 was not detected in the typical fronto-central location. Therefore it is conceivable the subtlety of the voltage deflection caused by deviant sounds is lost due to current topography of the P3 component measured (Gilmore, Clementz and Berg, 2009). Secondly, the deflection in the P3 may reflect the active conscious acknowledgment of the sound deviation (Escera *et al.*, 2000). The current paradigm was conducted passively without the toddlers actively attending to the sounds, therefore was not asked to actively attend to the deviant. Thirdly, the response may have been lost due to the higher deviant to standard phoneme ratio (30:70) as the greater the difference the larger the response (Näätänen and Winkler, 1999). The length of the paradigm had to be shorter due to the age of the participants, therefore a higher ratio was set in order to acquire more deviant sounds in a shorter paradigm. However, this could have weakened the response to the deviant sounds as they occurred at a greater frequency. A final explanation could be the phonemes used were too subtle in structure to elect this response within the toddler brain. Nevertheless, the difference observed in N1 amplitude is likely to reflect the involuntary processing of the brain detecting the deviant sound and suggests the components detected are reflective of active processing of the sounds (Tomé *et al.*, 2014).

Future work to strengthen these findings would require increasing the number of trial within each condition in order to clarify the P3 peak and obtain latency measures. It would also be preferable to reduce the number of deviant stimuli and to administer the paradigm binaurally; jointly this would increase the likelihood of the detecting a novelty response and allow for better predictions on the component location. The final modification would be to include simple tones to investigate whether there is a difference in how the brain responds to the tones over the speech sounds. This has the potential to provide the necessary evidence to qualify the theory above regarding how the preterm brain processes speech sounds.

This paradigm has ultimately not demonstrated a difference in speed of neural processing due to the absence and inability to measure the latency response between term and preterm infants. It does provide evidence to suggest the VP infants are processing auditory information utilising different neural mechanisms and may indicate functional connectivity differences within the VP brain that have not before been observed in an ERP investigation. These differences are consistent with my original predictions.

Chapter 7 Prediction of global cognitive performance at 12 months and 2 years from earlier measures of EF, IP and Attention.

Established paradigms assessing Executive function (EF), information processing (IP) speeds and attentional abilities have identified deficits that emerge around 2 years of age in very preterm children, with particular difficulties in the EF measures. These difficulties are evident after adjusting for cognitive performance. The question still remains to what extent the variation in the Bayley-III cognitive score is accounted for by EF performance. The aim of the current chapter is to explore: the relationship of the EF, IP and attentional measures to the Bayley-III cognitive scores at both the 12 months and 2 year time points, and to explore the relationship between the EF, IP and attentional scores to the overall learning construct from the BabyScreen application as a comparative measure for global cognitive performance.

Although no study group differences were observed in the first year experimental EF, IP and attention tasks, it is possible the variation in these measures reflects the variation in cognitive scores. Although first year EF performance is not correlated with the second year EF task scores as explored in Chapter 4, it is possible the variation across scores in the first year could still be, in part, reflective of the variability in cognitive scores at 2 years.

In the second year, differences were observed in the Bayley-III cognitive scale performance between the term and VP children. These scores only partly correlated with the EF, IP and attentional measures. It is plausible that collectively the variation across the EF, IP and attentional measures could account for a greater amount of variation in the Bayley-III cognitive scores at 2 years.

The BabyScreen application differentiated EF performance whilst incorporating a measure of IP speed. The overall learning construct from this application was designed to be a measure of broad cognitive ability and may be considered to be targeting similar abilities as the Bayley-III cognitive scale. The variation in this measure will be explored in relation to the other EF, IP and attentional scores from the 30 month assessment, as an additional insight as to how the experimental scores reflect variability in this alternative measure of global cognitive performance.

The research questions for the follow chapter are as follows:

1) Will the targeted assessments of EF, IP and attention within the first year relate to the cognitive z-scores at:
   a) 12 months?
   b) 2 years?
2) To what extent do the EF, IP and attentional variables from the 30 month assessment:
   a) Explain the variation in cognitive z-score of the Bayley-III at 2 years?
   b) Account for variation in the BabyScreen overall EF construct?

7.1 Methodology

Due to study numbers it was not practical to perform the preferred multi-level cluster analysis on the data collected within this longitudinal study. Given this, the relationship between the data collected at each time point of the study was addressed by a series of sequential linear regressions.

Cognitive z-scores of the Bayley-III at both 12 months and 2 years were utilised as the outcome measures. All the experimental predictor variables utilised were those selected a priori in previous chapters: Learning status (Mobile paradigm), proportion of trials correct (DRT), disengagement RT at 6, 12 and 30 months (GAP), level of A*B* error (A*B* paradigm), highest level achieved (DCCS), post-switch trial RT (MLMS), overall learning construct RT (BabyScreen), and Left hemisphere N1 latency measure (ERP auditory oddball). All continuous data were reduced to z-scores.

The predetermined demographic factors: IMD quintile, male sex and study group, are also included within the models. The R-squared values will be reported for each model, reflective of the proportion of the variation in the dependent variable accounted for by the predictors within each model.

For the 12 month and 2 year Bayley-III cognitive score data, all predictor variables from the first year were added incrementally according to age of assessment.

For the 2 year Bayley-III cognitive score data, 4 models were fitted to fully investigate the contributions of the 30 month variables to the data. Model 1 included the well-established 30 month EF and attentional task variables from the literature (DCCS; MLMS; and GAP disengagement RT) that have direct theoretical links with the first year tasks. Model 2 included the more experimental variables from the 30 month assessment, the overall learning construct RTs from the BabyScreen, and the left hemisphere N1 latency response collapsed over ear of presentation and standard and deviant tones from the auditory oddball paradigm. Model 3 includes all measures from the 30 month assessment phase. The final model, model 4, included the demographic variables stated above, and the 12 month Bayley-III cognitive z-score.

A final set of sequential regression models were performed with the overall learning RT construct from the BabyScreen as the dependent variable. As for the 2 year Bayley-III data, the established EF and attention measures were included primarily; the auditory oddball was explored as a solitary experimental predictor in the second model and the final model included all 30 month measures.

7.2 Results

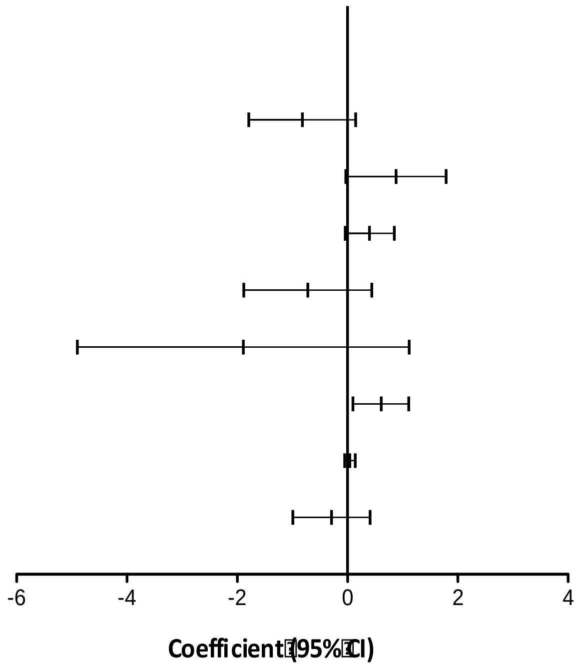
7.2.1 Predictive effects of the first year EF, IP and attentional performances on the cognitive score of the Bayley-III at 12 months of age

Overall the first model, including just the 3 month mobile task variable and the study demographic variables accounted for only 26% of the variation observed in the 12 month Bayley-III cognitive z-score. This increases to 44% with the addition of the 6 month variables and to 62% for the full model at 12 months. In both the 6 month and final models the inclusion of the attentional measure from the GAP task at 6 months was a significant influence on the model fit (Table 7-1).

The forest plot in Table 7-1 displays the model with all predictor variables from the first year.

Table 7-1. Sequential linear regression modelling the cognitive z-score of the Bayley-III at 12 months utilising EF, IP and attentional variables from the first year of assessments. The baseline group was term-born females.

Predictor	3 months - Coeff (95%CI)	6 months - Coeff (95%CI)	12 months - Coeff (95%CI)
N (T/PT)	61 (35 / 26)	33 (22 / 11)	22 (15 / 7)
Study Group	-.85 (-1.34 – -.35) ***	-1.10 (-1.78 – -.41)**	-.82 (-1.79 – .15)
Male	.09 (-.40 – .58)	.51 (-.14 – 1.17)	.88 (-.03 – 1.79)
IMD Quintile	.19 (.00 – .37) *	.19 (-.10 – .48)	.40 (-.04 – .85)
Mobile	-.05 (-.54 – .43)	-.29 (-.96 – .37)	-.72 (-1.88 – .44)
DRT	-	-1.78 (-4.29 – .74)	-1.89 (-4.90 – 1.12)
Gap at 6m	-	.45 (.09 – .82)*	.61 (.10 – 1.11)*
A-not-B	-	-	.04 (-.05 – .14)
Gap at 12m	-	-	-.29 (-.99 – .41)
Bayley-III cog at 12m (Const.)	-.53 (-1.37 – .31)	.23 (-1.56 – 2.02)	-.56 (-2.77 – 1.64)
R^2	**.26**	**.44**	**.62**



***p<.05; **p<.01; ***p≤.001**

7.2.2 Predictive effects of the first year EF, IP and attentional performances on the cognitive score of the Bayley-III at 2 years of age

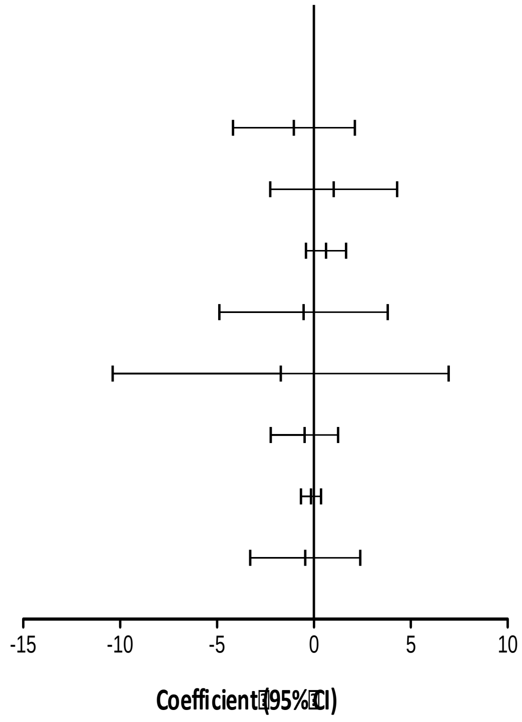
Table 7-2 presents the 3 sequential linear regression models fitted to the 2 year Bayley-III cognitive z-score and a fourth that included only the 12 month Bayley-III cognitive z-score and demographic variables as predictors. As in Table 7-1, the progressive addition of age related variables allowed the level of variability accounted for by each to be explored. Again the model that included all first year variables had the greatest predictive effect on the 2 year Bayley-III cognitive score outcome, with 48% of the variance accounted for by these measures. This level of variance accounted for by the experimental predictors is substantially greater than that of the model that included solely the Bayley-III cognitive scores at 12 months as a predictor; with only 26% of the variance at 2 years accounted for.

Unlike in the models for the 12 month Bayley-III cognitive scores, the attentional measures from the GAP task are no longer significantly impacting the models for the 2 year outcome.

The forest plot presented in Table 7-2 represents the full model with all first year variables.

Table 7-2. Sequential linear regression modelling the cognitive z-score of the Bayley-III at 30 months. The baseline group was term-born females.

Predictor	**3 months - Coeff (95%CI)**	**6 months - Coeff (95%CI)**	**12 months - Coeff (95%CI)**	**Bayley (95% CI)**
N (T/PT)	35 (14 / 21)	21 (11 / 10)	13 (7 / 6)	40 (18 / 22)
Study Group	-.61 (-1.27 – .05)	-.42 (-1.39 – .55)	-1.03 (-4.17 – 2.12)	-.02 (-.75 – .71)
Male	-.12 (-.80 – .56)	.18 (-.77 – 1.12)	1.02 (-2.25 – 4.29)	-.36 (-1.12 – .39)
IMD Quintile	.03 (-.21 – .21)	.12 (-.25 – .49)	.63 (-.40 – 1.66)	-.08 (-.16 – .33)
Mobile	-.43 (-1.08 – .21)	-.34 (-1.33 – .65)	-.53 (-4.87 – 3.81)	-
DRT	-	.93 (-3.05 – 4.92)	-1.71 (-10.38 – 6.96)	-
Gap at 6m	-	-.10 (-.59 – .39)	-.48 (-2.22 – 1.25)	-
A-not-B	-	-	-.15 (-.66 – .37)	-
Gap at 12m	-	-	-.44 (-3.28 – 2.39)	-
Bayley-III 12m z-score	-	-	-	.31 (-1.19 – .81)
Bayley-III cog at 30m (Const.)	.12 (-.87 – 1.12)	-.91 (-3.30 – 1.48)	-.71 (-5.81 – 4.39)	-.19 (-1.19 – .86)
R^2	.15	.13	.48	.26



***p<.05; **p<.01; ***p≤.001**

7.2.3 Predictive effects of the second year EF, IP and attentional performances on the cognitive score of the Bayley-III at 2 years

Table 7-3 explores to what extent the variance in the 2 year Bayley-III cognitive z-score is accounted for by the 30 month experimental variables. The first of the three models included the established EF and attentional measures from the literature and accounted for 11% of the Bayley-III cognitive z-scores.

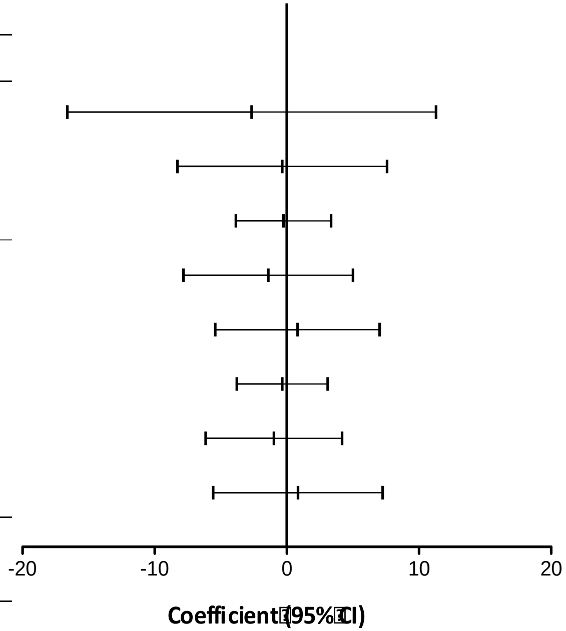
The second model included the more experimental measures from the 30 month assessment stage, the Babyscreen RTs to the overall EF construct and the ERP auditory attentional measure. This model accounted for 32% of the variation in Bayley-III cognitive performance at 24/30 months, accounting for a greater level of variance than those from well-established EF tasks.

The final model incorporated all measures to explore the total variance accounted for by the 30 month experimental variable, with 38% of the variance accounted for. None of the predictor variables significantly impacted any of the models suggesting equal contribution of each predictor to the variance accounted for by the models.

Table 7-4 displays additional regression models fitted to the overall EF construct from the BabyScreen paradigm. Model 3 includes all 30 month experimental variables and indicates 76% of the variation in this outcome is accounted for within this model. The forest plot represents the largest model.

Table 7-3. 3 Linear regression models with the cognitive z-score of the Bayley-III at 30 months as the outcome variable. The baseline group was term-born females.

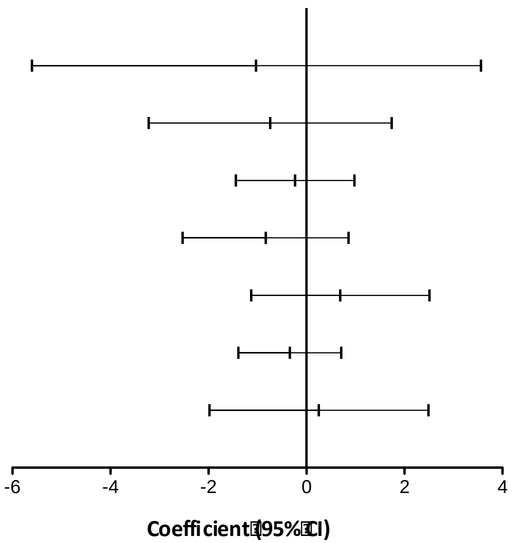
Predictor	Model 1 (95%CI)	Model 2 (95%CI)	Model 3 (95%CI)
N (T/PT)	24 (12 / 12)	23 (12 / 11)	11 (6 / 5)
Study Group	-.09 (-1.42 – 1.25)	-.15 (-1.45 – 1.15)	-2.66 (-16.61 – 11.28)
Male	-.22 (1.22 – .78)	-.52 (-1.70 – .66)	-.34 (-8.27 – 7.59)
IMD Quintile	.15 (-.13 – .43)	.02 (-.49 – .53)	-.24 (-3.85 – 3.36)
DCCS	.06 (-.43 – .54)	-	-1.41 (-7.82 – 5.00)
MLMS	-.13 (-.71 – .45)	-	.81 (-5.42 – 7.03)
Gap at 30m	-.05 (-.51 –.41)	-	-.35 (-3.79 – 3.09)
Babyscreen (task 17)	-	-.49 (-1.06 – .07)	-.98 (-6.14 – 4.18)
Auditory ERP	-	-.23 (-.81 – .36)	.84 (-5.56 – 7.25)
Bayley-III cog at 30m (Const.)	-.49 (-2.62 – 1.64)	.46 (-1.64 – 2.56)	6.45 (-25.54 – 38.44)
R^2	.11	.32	.38



***p<.05; **p<.01; ***p≤.001**

Table 7-4. 3 Linear regression models fitted on the Babyscreen overall EF construct at 30 months as the outcome variable. The baseline group was term-born females.

Predictor	Model 1 (95%CI)	Model 2 (95%CI)	Model 3 (95%CI)
N (T/PT)	16 (7 / 9)	24 (13 / 11)	11 (6 / 5)
Study Group	.99 (-1.09 – 3.07)	.94 (.02 – 1.86)*	-1.03 (-5.6 – 3.56)
Male	-1.03 (-2.46 – .40)	-.52 (-1.45 – .42)	-.74 (-3.22 – 1.74)
IMD Quintile	.27 (-.30 – .74)	.12 (-.26 –.50)	-.23 (-1.44 – .98)
DCCS	-.16 (-.82 – .51)	-	-.83 (-2.53 – .86)
MLMS	.11 (-.63 – .85)	-	.69 (-1.13 –2.51)
Gap at 30m	-.05 (-.58 – .48)	-	-.34 (-1.39 – .71)
Auditory ERP	-	-.18 (-.66 – .31)	.25 (-1.98 – 2.49)
Babyscreen: Overall EF construct (Const.)	.45 (-2.66 – 3.56)	-.16 (-1.73 – 1.41)	3.93 (-4.88 – 12.74)
R^2	.48	.21	.76



***p<.05; **p<.01; ***p≤.001**

7.3 Discussion

Through the use of sequential regression analyses, the variances in the Bayley-III cognitive scores at both 12 months and 2 years of age were explored in relation to the EF, IP and attentional measures acquired during the PDP battery.

For the Bayley-III cognitive scores at 12 months of age, 62% of the variance was accounted for by the performances on the EF, IP and attentional measures. The same measures accounted for 48% of the Bayley-III cognitive score variation at 2 years. This is in contrast to the Bayley-III cognitive scale itself, where the 12 month scores only accounted for 24% of the variation at 2 years.

The EF, IP and attentional measures at 30 months were poor predictors of the contemporaneous Bayley-III cognitive score, only accounting for 38% of the variance. Contrastingly, these measures at the 30 months predicted 76% of the variance in the overall learning construct from the BabyScreen application. These results are discussed in the subsequent section as they leave interesting questions regarding the specificity and sensitivity of the Bayley-III.

7.3.1 Predictive nature of the EF, IP and attentional measures from the first year of the PDP on Bayley-III performance variability at 12 months and 2 years

The EF, IP or attentional measures nor cognitive composite scores differentiate VP and term groups in the first year, but they appear to be accessing the same areas of performance, explaining 62% of the variance in cognitive scores at 12 months. It is less clear what accounts for the remaining variance of Bayley-III scores however. The cognitive scale may access other areas that contribute to the final score and the relative importance of these can only be shown by looking at how they predict later outcomes.

It is plausible that the experimental assessments included within the first year do not provide sensitive enough measures of EF abilities to account for the additional variation in the Bayley-III outcome. It is equally likely that the Bayley-III is not accurately assessing EF abilities within the cognitive scale, as seen in the second year assessments with EF deficits still present after adjustment for the Bayley-III scores.

An interesting observation within these analyses is level of variation accounted for by the first year EF, IP and attention measures in the Bayley-III cognitive scores at 2 years. The predictors that accounted for 62% of the variability in the Bayley-III cognitive scores at 12 months accounted for 48% of the variability in the cognitive scores at 2 years. This is in contrast to the Bayley-III itself, which only accounted for 26% of the 2 year scores variability. This poor predictive validity of the Bayley-III cognitive scores across the two time points was observed in Chapter 3 and suggests the EF, IP and attentional measures from the first year are accessing abilities that give a greater indication of later cognitive performance in line with the Bayley-III cognitive scores at 2 years. However, this is not to say that the Bayley-III cognitive scale at 2 is assessing the necessary skills to detect deficits later in life. It is likely that the EF difficulties observed in VP populations are more subtle and therefore more specific assessments of EF, IP speeds and attentional abilities are required in the early years in order to detect these deficits.

The Gap task measure at 6 months had a significant impact on the model fitted to the 12 month Bayley-III cognitive score. The positive relationship of the 6 month Gap disengagement RTs within the models suggests the greater the time to disengage during the gap task, the better the performance on the Bayley-III. The implication from these findings is that the attentional ability targeted by the Gap task, specifically attention orientation, is having a greater influence in the Bayley-III cognitive performance at 12 months. Within Chapter 5, an increase in disengagement RT is seen from the 6 to the 12m time point within the preterm cohort. One theory proposed to explain this relationship between the Gap measure at 6 months and the Bayley-III score at 12, is a faster disengagement time reflects poorer engagement with the primary stimulus (van der Geest *et al.*, 2001). If an infant is displaying poor focus during the Bayley-III, the scores are likely to be reflective of this. However, when including the 12 month Gap disengagement RT, this measure no longer significantly impacts the model. The preterm infants' disengagement response was slower at 12 months and was in line with the term response; as the GAP variable is not significant at 12 months, it is unlikely that the attentional focus is a responsible for the poorer performance in the 12 month Bayley-III cognitive score data. This result is hard to interpret without the use of a multilevel analysis, and is likely the result is confounded by the repeated measures aspect of the data within the 12 month model. Inclusion of both Gap measures when not accounting for repeated effects could be misleading and therefore should be disregarded in the larger model.

In summary, the first year of assessments of EF, IP and attention appear better predictors of later Bayley-III cognitive scale performance in contrast to the Bayley-III itself. The question that remains is how effective is the Bayley-III cognitive score at 2 years at detecting later problems.

7.3.2 Predictive validity of the experimental measures from the 30 month assessment on the Bayley-III performance variability at 24/30 months

When exploring the relationship between the experimental variables and the Bayley-III scores in the second year, it is clear that there is a disagreement between the two batteries as to what skills are being measured. When using all predictor variables available, only 38% of the variation in Bayley-III cognitive score was accounted for.

Although the Bayley-III has been shown previously to detect cognitive impairment in numerous populations, it is frequently reported to miss those with mild cognitive impairments later in life. In the current cohort, none of the infants were considered to fall within the mild to moderately delayed category according to the Bayley-III cognitive score at 2 years of age. This result, in combination with the differences observed between the study groups in the EF tasks at 30 months of age, is likely suggestive that a proportion of the cohort will later present with cognitive impairments and/or academic difficulties. This lack of correlation between the study predictors and the amount of variation accounted for in the Bayley-III is therefore not surprising. What remains to be seen, is whether the EF tasks administered during the PDP are a better reflection of later abilities over the Bayley-III cognitive scale. This question unfortunately cannot be answered in the current investigation without subsequent follow-ups. However, by looking at the relationship of the experimental variables with an alternative measure that reflects global cognitive ability, additional evidence could support the experimental EF approach.

The Babyscreen application, reported in Chapter 6, is still under development. However, the task has been designed around well-established measures in the developmental literature that assess EF performance. The overall learning construct was selected as the best measure from the PDP battery to explore as an alternative measure for more general cognitive performance. The overall learning construct incorporated both EF and processing speed elements from the task. 76% of the variation of this construct was accounted for by the EF, IP and attentional measure collected at 30 months of age. This high percentage suggests that this construct is likely tapping into the same abilities as the other well established measures utilised in the 30 month assessment phase; unlike the Bayley-III cognitive scale at this time point.

Another interesting point to consider is the 48% of variation accounted for by the first year EF, IP and attentional measures to the 2 year Bayley-III cognitive scores. As this variability is not reflected in the EF, IP and attentional measures in the 30 month assessment, this result could suggest that the Bayley-III is measuring more immature EF abilities which are more reflective of abilities in the first year of development. The emergence of EF and its sub-domains is highly theorised, however, a common opinion many researchers adopt is the differentiation of EF sub domains over time (Miyake *et al.*, 2000). During the second year, although considered in this thesis as a unified construct in terms of assessing EF, it is not disregarded that the different EF domains are likely to be emerging. Given what is understood by later preterm studies, specific EF domain deficits are often reported to be the probable cause of impairments and academic difficulties. If these domains are beginning to emerge in the second year, assessments that better differentiate the domains are likely to be better suited at detecting the emerging deficits. Although the EF tasks included in the current PDP 30 month assessment were not specifically targeting each of the different EF sub-domains, each task does have a predominant domain focus. Likewise, the BabyScreen application is subdivided into constructs that overlap with the different EF subskills. The experimental tasks at 30 months could therefore be considered to be more sensitive to the different EF domains that, perhaps, the Bayley-III cognitive scale is not. This theory is supported by the predictive nature of the more simplistic first year task structures to the 2 year cognitive scores.

7.4 Concluding comments

The first year of experimental assessments are better predictors of later Bayley-III scores; however, it appears that the Bayley-III may not be targeting the same abilities as those assessed in EF tasks at 30 months. Although it cannot be concluded that the EF tasks from the 30 month assessment are a better indication of later EF abilities, there are difficulties observed in the preterm population that are not being detected in the Bayley-III at 2 years. Making a prediction in the first year of life about later cognitive performance is therefore not simple. A possible explanation as to why the EF abilities in the first year are not reflective of the abilities in the second could be down to an evolving structure of EF into a differentiated construct. Due to the broad array of difficulties observed following preterm birth and the different levels of impairments observed in later life, it could be proposed that cognitive development in different children will take different trajectories, and therefore a linear relationship of EF function from the first year into the next is unlikely. Equally, these difficulties could be present from birth, but detecting these difficulties may not possible until later in development when the different EF domains can be assessed.

Another area highlighted in the preterm literature as having a modulating effect on the later EF performances is speed of processing. The Bayley-III is not particularly sensitive to processing speeds as it is predominantly experimenter lead and it therefore not likely to be assessing the same skills as those targeted in the EF, IP and attentional measures. This is in contrast to the BabyScreen application that is regarded to incorporate both in the overall learning construct. The EF, IP and attentional measures account for a much greater proportion of the variation in the overall learning construct of the BabyScreen and although this is not yet an established measure of EF or general cognitive function, it highly probable from these observations that it is tapping into the same cognitive abilities as those measured by classic EF tasks. The self-driven aspect of the application allows for processing speed abilities of the infant to be incorporated into the evaluation of performance in an EF measure.

Chapter 8 Discussion

Cognitive impairments are typically reported in later childhood of children born preterm. Current assessments conducted within the early years appear to miss children in the mild to moderate impairment range and it is often not until the children reach school age that these problems are recognised. Once at school, the reported difficulties impact transition into school and have lasting effects on academic attainment. This highlights the importance to improve early identification methodologies in order to formulate targeted intervention schemes to aid the developmental trajectories of preterm children in the early years and optimise developmental outcomes.

Within this thesis, a cohort of very preterm and term born infants were followed from infancy through to 30 months of age in a longitudinal design incorporating 4 assessment stages at 3, 6, 12 and 30 months of age. The main study objectives were:

1. To explore differences in EF, information processing speed and attention at 3 time points within the first year and additionally at 2 years of age between the term and very preterm cohorts before and after adjustment for global cognitive score. This sought to identify the emergence of any EF difficulties not accounted for by global cognitive performance.
2. The variation in first year measures of EF, IP and attention were compared to the variation in cognitive scores of the Bayley-III at both:
    a. 12 months
    b. 2 years

    This sought to explore the extent to which performance in EF tasks in the first year predicted the variation on the Bayley-III cognitive scores and examined the effectiveness of the cognitive scale at detecting variation in EF abilities within a very preterm cohort.

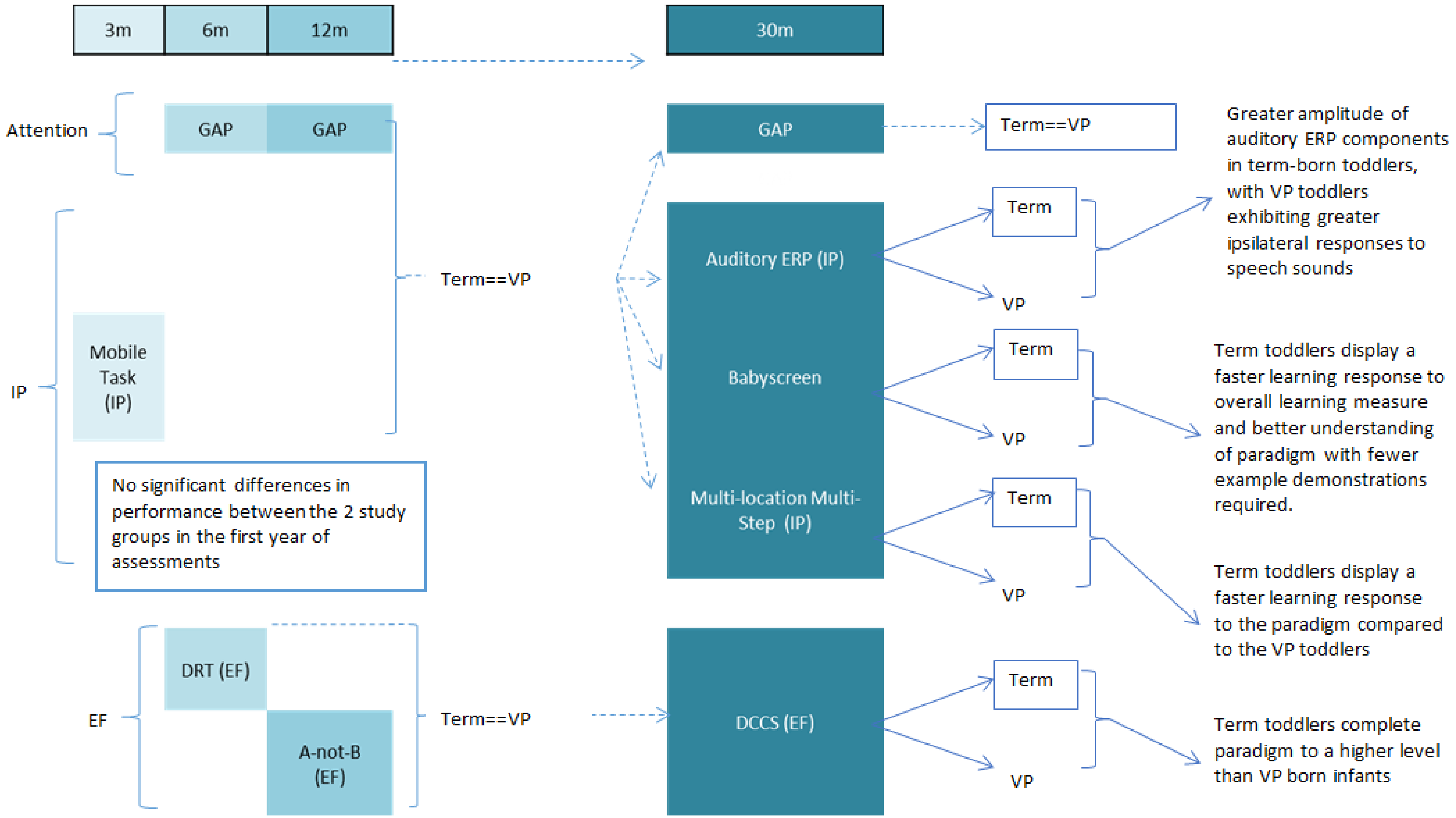
3. Finally, the EF, IP and attention scores at 2 years were used to investigate the variation of the Bayley-III at 2 years.

8.1 Objective 1: The exploration of Executive Function, Information Processing speed and Attention over the first 2 and a half years

In middle to late childhood, ex-preterm infants are often observed to suffer from a range of cognitive impairments and attentional difficulties which later lead to reduced academic attainment that is typically not accounted for by IQ scores. The cognitive impairments described in ex-preterm populations are unusual as they do not appear to be global, but rather specific difficulties in EF sub-skills, namely, working memory, in combination with reduced information processing speeds and attention deficits (Johnson, 2007; Mulder, Pitchford and Marlow, 2010, 2011b; Rose *et al.*, 2012). The area that is of fundamental importance to the preterm developmental literature is identification of measures that can highlight deficits within the population early in life, before they impact academic attainment and other outcomes.

In the current investigation, no group differences were detected between the two study groups within the first year of assessments. At 30 months of age particular differences were observed in EF and IP related measures. The attentional measures from the Gap task did not suggest conclusive differences between the two groups although the VP responses were less consistent over time compared to the term born responses. Overall, the VP toddlers displayed a poorer performance in the DCCS task completing fewer levels and passing a fewer number of trials, completed fewer items on the BabyScreen application with more experimental demonstrations required to aid completion and overall displayed slower responses to the overall learning construct of the task, and although not significantly different, the VP toddlers displayed a greater response time following the post-switch trial of the MLMS task. The auditory oddball paradigm additionally indicated possible differential mechanisms regarding the processing of speech sounds between the two study groups. Overall study finding are summarised in Figure 8-1. Following adjustment for global cognitive score, the group differences observed were maintained across the battery.

Figure 8-1. Diagrammatic representation summarising the divergence of term and preterm abilities at the 30 month assessment stage with the prominent findings from each task.



Investigations in infancy and early childhood have sought to identify early signs of difficulties in specific EF sub-domains in preterm populations due to the difficulties reported in middle childhood. Upon review of the literature, it is clear EF abilities differentiate into sub-domains later in life, specifically: working memory, inhibition and cognitive flexibility (Diamond, 2013). How the different domains interrelate early on is a matter debate (Barkley, 1997; Miyake *et al.*, 2000; Anderson, 2002) and there is limited evidence to suggest a differentiated construct within the first two years (Garon, Bryson and Smith, 2008). It is plausible that the emergence of the sub-domains begins within the first two years after birth. However, it has been proposed that it is more likely basic EF skills emerge within the early years and only later differentiate into the different sub domains (Garon, Bryson and Smith, 2008).

The results of this longitudinal investigation suggest that EF difficulties, independent of global cognitive abilities, being to emerge within this VP population at 2 years of age. This is in line with previous investigations (Rose, Feldman and Jankowski, 2009; Pozzetti *et al.*, 2014), where the populations reported could be considered lower risk to those in the current investigation, with older gestational ages (Pozzetti *et al.*, 2014) and higher birth weights (Rose, Feldman and Jankowski, 2009). These differences observed previously were primarily within the EF sub-domains. To the author's knowledge, this current study is the first detailed investigation incorporating the performances of EF, IP and attentional measures through the first year with a follow up within the second year.

Within the current battery of assessments, although viewed to assess EF as a unified construct, each task has previously been associated with a specific EF domain, as detailed in the relevant chapters. It is unlikely one domain can be assessed without the influence of another (Pozzetti *et al.*, 2014). However, should these tasks assess specific domains, it could be postulated from the cognitive profile in school-aged ex-preterm children, that VP infants would present with a more prominent difficulty in the paradigms assessing working memory, for example the A not B task. This was not the case in the current investigation as the VP toddlers displayed a more general difficulty across all EF measures.

The gradual emergence of EF difficulties within VP populations could be explained by the neuroconstructive approach to cognitive development, whereby the neuronal connectivity patterns later on in development are the product of abnormal developmental trajectories, rather than an impaired origin (Oliver *et al.*, 2000). This approach would provide further clarity to the absence of performance differences in the first year observations. Alternatively or in combination with, the developmental trajectory of these abilities could reflect the interactive specialisation theory proposed by Johnson (2000). This theory considers that the refining of cortical activity during development and the strengthening of inter-regional communications reflect the emerging behaviour profiles observed in the early years. The strengthening of neuronal networks could offer an explanation for the commonly reported association of processing speeds difficulties and EF performances. It is likely a combination of these two frameworks could offer explanations regarding the difficulties differentiating EF sub-domains within the early years, and the association of EF performances with information processing abilities. Nevertheless, a commonality across EF domains is likely in the early years and although the various EF domains may be emerging in the second year, the current investigation suggests targeting specific difficulties may not be possible in children so young. We do have to consider that specific EF sub-domain differences may not be observed within the current cohort later in childhood. More general EF difficulties reported at 30 months of age in this current investigation may be a true reflection of later cognitive performance differences. Until future follow ups are conducted, this cannot be confirmed.

EF performance differences were not present in the first year of the current investigation. However, it is plausible the measures utilised, although established EF tasks in the literature, may not be sensitive enough to detect subtle performance difference at this early stage of development. For example, it was clear the delayed response task at 6 months did not surpass chance in this instance, but the parameters may have been too challenging for infants this young. Previous investigations have reported performance differences in VP populations in the first year. Sun *et al*. (2009) reported EF difficulties at 8 months of age in a VP population of similar medical risk to those in the current investigation. Sun *et al.,* (2009) utilised a modified A not B task that provided inhibition and working memory measures, although the classic A*B* switch was not incorporated and therefore could be considered simply a delayed response task. The differences observed in the 8 month cohort could therefore reflect the development of EF abilities from 6 to 8 months and supports reports from the literature that delayed response tasks are more reliable at assessing EF abilities at 8 months (Diamond, 1990). The classic A*B* paradigm incorporates the inhibition of prepotent responses during the A*B* switch trials and has be argued to be a greater EF challenge than the simple DRT format (Diamond, 1991). The lack of performance differences in the A*B* task in the current investigation could therefore suggest differences in EF between term and VP infants are subtle and are lost in the additional challenge created by the inhibition of prepotent response. Additional research into paradigm sensitivity could find subtle differences in EF are detectable within the first year.

One important factor that must also be considered in the interpretation of these results is the use of age adjustment of the infants within the investigation. As highlighted within Chapter 1, correcting for gestational age at birth is commonly employed before the age of 2 years. After this age, there is disagreement within the literature as to whether corrected age or chronological age is more appropriate. Many research studies will employ corrected age so as not to disadvantage the individuals, however, at school these children are typically compared to their year group. VP infants could be considered older than the term born participants within the current investigation based on dates of birth, and differences in cohort performance would therefore be even more pronounced in this case. Using chronological age rather than corrected age will identify more VP children with problems as they are assessed against ‘older’ children’s standards. This may identify children in whom later EF problems appear better than conventional age correction (Wilson and Cradock, 2004). Later outcomes are required to assess this issue.

In this study, EF difficulties are present independent of global cognitive difficulties. In preterm cohorts, it is has been reported that speed of IP and attentional abilities similarly impact global cognitive performances and therefore are also important factors to consider when exploring the variability in Bayley-III cognitive scores. The Bayley-III cognitive scale is a measure of developmental progression. Although the measure is relied upon to identify children at high risk for later difficulties it was not originally designed to assess individual executive processes. It is therefore important to explore to what extent each measure contributes to the cognitive scores in order to understand how a generalised measure such as the Bayley-III could be adapted to formulate more efficient earlier identification methods.

8.2 Objective 2: To what extent do Executive Function, Information Processing speed and Attention measures from the first year account for the variation in cognitive scores in the Bayley-III at 12months and 2 years

Identification of early measures that can highlight deficits within the preterm populations is of importance if we are to establish targeted interventions to help improve developmental outcomes. Deficits although may be subtle during the first two years, could be the first signs of a deviant developmental trajectory that leads to the difficulties in academic progress identified in studies of school-aged children born preterm (Johnson, Fawke, *et al.*, 2009). Currently, it is believed that the Bayley-III is not providing the necessary sensitivity for this level of early identification (Johnson, Moore and Marlow, 2014; Spencer-Smith *et al.*, 2015). Irrespective of whether EF differences were observed within the current investigation, it was important to explore the level of variation accounted for by the EF measures utilised within the battery in relation to a standardised assessment commonly used to identify later cognitive delays.

At 12 months of age, the Bayley-III cognitive score suggested the VP infants were performing on average 7 points lower than the term born infants in the cognitive scale. There were no differences observed in the EF, IP nor attentional measures within the first year and adjustment for the Bayley-III 12 month cognitive score had no effect on the overall study group differences. Although no significant differences were observed between the study groups, the variation within the EF, IP and attention scores may have accounted for a level of variation in the Bayley-III cognitive score at 12 months.

Sequential regression analyses in Chapter 7 revealed that with each additional EF measure through the first year, progressively more variation in the Bayley-III cognitive score at 12 months was explained by the models. The final model including all measure of EF, IP and attention measures from the first year accounted for 62% of the variation in the Bayley-III at 12 months. Previously, the Bayley-III cognitive scale has been criticized for the poor predictive validity of later cognitive impairments in preterm populations (Lobo and Galloway, 2013). From these results it could concluded that a moderately high proportion of the Bayley-III is targeting EF abilities within the first year. Rather than the measure being observed as having poor predictive validity, instead it could be reflective of the significant developmental changes in EF abilities during the first year. If this was the case it could be expected that the EF, IP and attentional measures from the first year and the 12 month cognitive scores should account for the same amount of variation in the second year Bayley-III cognitive scores. However, this was not evident and gives rise to speculation regarding the different skills the developmental test is assessing at the two ages. As discussed, the Bayley-III was designed to evaluate developmental progression, therefore assesses much broader areas than the EF tasks utilised in the current investigation (Bayley, 2006).

The variability in first year EF, IP and attentional measures were then explored in relation to the variability of the Bayley-III cognitive scores at 2 years. In chapter 3, the Bayley-III scores were explored independently at the first and second year assessments and followed up with a longitudinal exploration of the scores. This demonstrated a poor relationship between the measures at the two time points, with the 12 month scores only accounting for 23% of the variation at 2 years. In contrast, the EF, IP and attentional measures from the first year accounted for 48% of the variability in the second year assessment. This could be considered a high proportion of variability to be accounted for when the measure are taken up to two years prior.

The Bayley-III cognitive scores could therefore be interpreted to be reflective of EF, IP and attentional abilities from the first year; however, in the second, the measure appears to be targeting a different set of abilities. Before commenting on the implications of these findings, objective 3 also needs to be considered. The final question that remained in the current investigation was whether the EF, IP and attentional abilities at 30 months were reflective of the 2 years Bayley-III cognitive scores.

8.3 Objective 3: To what extent do Executive Function, Information Processing speed and Attention measures from the 30 month assessment account for the variation in cognitive scores in the Bayley-III at 2 years

The Bayley-III has been shown to display a level of predictive validity at 2 years within the literature (Bode *et al.*, 2014). This measure reliably identifies those at risk of severe impairments, but mild to moderate delays are often missed (Aylward, 2002; Hack *et al.*, 2005). To better understand why the measure is not identifying these individuals, the variation in the EF, IP and attentional scores were compared to the variation in Bayley-III 2 year cognitive scores. Only 38% of the variability in Bayley-III scores was accounted for by these measures. Given the difference observed within the measures obtained at 30 months of age, it could be strongly argued that a proportion of the infants within the current study will go on to develop cognitive difficulties later in life. Future follow ups will explore this relationship to investigate whether better predictor variables are apparent in the current study over the traditional measures.

Upon review of the relationship between the Bayley-III scores and the experimental data, two interpretations could be made. Firstly, it could be proposed that the Bayley-III is targeting more immature measures of EF abilities. In the first year it is likely the Bayley-III cognitive score is sufficient for assessing EF performance as only basic EF abilities have developed. However, during the second year, the developmental changes within the EF construct are such that these Bayley-III cognitive measures may not be sensitive enough to detect subtle differences beginning to emerge in the preterm population, differences that could be reflective of later academic performance.

An alternative view considers the developmental trajectory of cognition as a whole. Cognitive assessments typically incorporate additional functions above EF, IP and attention, for example long-term memory, which is not assessed in the current investigation. As the variation in the Bayley-III is well explained by the EF, IP and attentional measures within the first year, this could be reflective of an undifferentiated state of cognition within the first year. This could again reflect the interactive specialisation theory whereby immature connections reflect less specialised behavioural profiles. However, as these connections become more refined and interregional communication becomes more rehearsed with experiences in the second year, this could be indicative of differentiation of global cognition in the same vein as the EF sub-domains. As such, the Bayley-III may be targeting broader aspects of cognition that the EF, IP and attentional measures do not account for, explaining the reduction in variation accounted for by these measures. Of these two interpretations however, if to take the view the Bayley-III cognitive scale is targeting additional aspects of cognition, a more accurate reflections of later abilities might be expected, it is therefore likely there is an element of overlap across these two interpretations.

Thus I conclude that EF and IP difficulties are present within this current VP cohort at 2 years of age that are not fully identified by the cognitive scale of the current edition of the Bayley Scales of Infant and Toddler development. These findings provide further evidence for the need to improve identification methods used in standard clinical practice before interventions schemes can be established.

Typically children born preterm present with a greater need for special educational support once at school (special educational needs, SEN) (Bowen et al., 2002; Johnson et al., 2009a; Saavalainen et al., 2008). A strong positive correlation of SEN with increasing gestational age at birth has been observed (Mackay *et al.*, 2010); extremely preterm cohorts (<26 weeks of gestation) are up to 13 times more likely to need additional support than term born infants (Johnson, Fawke, *et al.*, 2009). The rise in preterm survival rates has led to a substantial increase in pressure on the educational systems and its resources (Marlow, 2004). Earlier detection of cognitive impairments would enable the development of targeted interventions with the objective to elevate some of these economic pressures as well as improving the transition from nursery to school for these children (Johnson, Wolke and Marlow, 2008). Infants in need of support may benefit most from schemes that are started early, targeting the cognitive functions as they develop, although current attempts have not produced long term changes in outcome (Spittle *et al.*, 2007).

This research provides additional evidence to the rich and varied literature. A significant proportion of studies referenced within this thesis have very different approaches, from broad recruitment criteria in terms of gestational age, to neonatal factors, to specific changes in the methodologies of particular paradigms. The nature of current research requires the scientific community to explore new areas in order to extend prior knowledge. However, in areas wishing to determine patterns of performance in order to achieve interventions, such as preterm research, a level of consistency is required to advance scientific understanding. These discrepancies in the literature highlight the need for longitudinal observations such as this one to remove individual variability to purely focus on age dependent changes (Garon, Bryson and Smith, 2008). The unique longitudinal and broad focus of tests utilised here are important if evolving deficits in EF are to be understood in this group.

8.4 Clinical implications

The EF tasks utilised in the current investigation show performance differences between the term and VP groups that are not explained by the Bayley-III cognitive scores. These results suggest that the Bayley-III cognitive scale at 2 years may not be sensitive enough to detect the specific delays typically observed in the VP population and therefore may be missing children that later go on to present with mild cognitive delays. Although the current population need to be explored at a later time point to confirm these suspicions, these results may be of interest to practicing clinicians that assess infants in clinics following preterm birth.

Similarly, parents of children born very preterm should be aware of the sensitivity of the measures used within current clinical practice. Although it must be emphasized that the Bayley-III has been shown to reliably detect those likely to present with severe cognitive impairments (Aylward, 2002; Hack *et al.*, 2005), the current findings suggest this measure may not have the necessary sensitivity for early identification of milder impairments and is in agreement with previous research (Johnson, Moore and Marlow, 2014; Spencer-Smith *et al.*, 2015). Further follow-ups are required to draw definitive conclusions on the accuracy of these current findings.

8.5 Limitations

In terms of cohort selection bias, the study groups were unbalanced in relation to the child's sex and the VP group had an excess of males. Bias could be introduced as males are known to have poorer cognitive and motor outcomes when assessed at later ages (Peacock *et al.*, 2012). This was addressed by controlling for sex differences in all secondary analyses.

Similar levels of education were found between groups, but 87% of mothers had a university degree suggesting selection bias. This may be reflected in the higher Bayley-III cognitive scores for VP children than expected from other published studies (Christian, Morrison and Bryant, 1998; Spencer-Smith *et al.*, 2015). Additionally, both groups displayed a variety of ethnic backgrounds. It could be argued that a marginally greater level of diversity was present within the preterm cohort. This could reflect a recruitment bias within the term born cohort, however all ethnic groups were given equal opportunity to take part in study during antenatal class recruitment sessions. By utilising scores from the Index of Multiple Deprivation (a nationally available score related to postcode (NPEU, 2013)), a range of values were shown but again did not vary by quintile.

Recruitment from the tertiary level neonatal unit at UCH led to a cohort mean gestational age of approximately 26 weeks, with a ratio of 17:33 very preterm (<32 weeks of gestation) to extremely preterm infants (EP; <27 weeks of gestation) respectively. Poorer prognosis is often reported in infants born before 27 weeks in terms of academic success. It was not possible to stratify the early performance outcomes assessed in this current investigation by gestational weeks due to the ratio of VP to EP infants. Again, by controlling for cognitive z-scores on the Bayley-III in the secondary analyses, any bias created by this excess of extremely preterm children was addressed.

Attrition occurred at each level of the study. This was typically due to participating families moving away from the London area and therefore not able to attend within the month timeframe set in the study protocol. Travel to the hospital was necessary for the assessments as some of the study equipment was not transportable. It was therefore possible some families declined the study follow-up invitations due to transportation difficulties into the hospital for the visit. Reimbursement for travel and taxis were offered to minimise this potential deterrent in participation.

Within each phase of the study, each task detailed within this thesis has a marginally different sub-cohort for each task analysed. This was due to attrition within the assessments; although the family attended the follow-up, the child may not have completed all aspects of the assessment. This could reflect a potential bias within the results, as children that are underperforming may have been more likely to not to complete all tasks. However, when exploring the attrition rates at each time point for each task reported in this thesis, the greatest difference in proportion of term and VP children not to complete the task was within the Bayley-III scores at 12 months where 30% of the VP infants' scores are missing. The VP scores were collected by the hospital and therefore outside of the control of this study. When exploring the data collected within the remit of the study, there were no significant differences between the proportion of term born infants and VP infants in each task reported. Although there may have been bias introduced by those that did not complete each task, and this could be explored in future works, there was not a significant difference between the study groups and delays were still being observed at 30 months independent of any task attrition.

The distribution of cognitive scores on the Bayley-III within the current VP group was unusual as previously preterm cohorts have been reported to score lower than the standardised mean, which was not evident in the current cohort (Spencer-Smith *et al.*, 2015). The current VP cohort did not include any children with severe disabilities as defined by cognitive scores <-2SD below the mean, suggesting the scores are reflective of previous VP populations once those with severe disabilities are excluded. It should also be noted that scores increased by .8SD on the Bayley-II over the 11 years prior to the updated Bayley-III release (Moore, Johnson, *et al.*, 2012). These higher performances could therefore be indicative of a continued gradual increase in performance over time.

Within the term born cohort, there was a level of selection bias in the scores acquired for the Bayley-III scales at both the 12 month and 30 month follow up. This was due to the volume of data acquired during the assessment timeframe. The focus was on the Bayley-III cognitive scale as this was fundamental to the research question framing this thesis. If the infants and/or toddlers were capable of completing the additional language and motor scales of the Bayley then these were completed. In many circumstances this was too much for the participants and had fatigued by this stage of the follow-up. It is therefore it is likely that the higher performing infants completed the language and motor scales and explains the unusually high term scores on page 87. It was for this reason that the language and motor scores were not exclusively taken into account at any point within the thesis.

The moderate cohort sample size reported in this thesis was reflective of the challenges faced by longitudinal research studies. The number of children to complete the 4 assessment phases was not as high as desired, but achieving a detailed longitudinal investigation with multiple follow-ups was going to face challenges with group numbers. Although a multivariate cluster analysis was not possible due to missing data across the multiple paradigms performed, this was a consequence of the highly detail assessment batteries for each assessment. A primary objective of the study was to obtain a detailed understanding of the emergence of EF, IP and attentional abilities across the first 2 years. In order to achieve this, detailed batteries were necessary. Performance bias was addressed by administering each assessment day in the same order for each subject, however, occasions arose where data collection was not possible, primarily due to ill health and temperament of the infant which could not be accounted for. Overall, the attrition rate across the 4 time points was low and the sample size at the final time point was reflective of the time constraints of a PhD.

The potential bias from un-blinded assessments was addressed by scoring offline where possible and using blinded secondary scorers. Bayley training was accredited by the person leading the team doing the clinical testing.

8.6 Future work

From the current investigation both the ERP auditory oddball paradigm and the BabyScreen Application would benefit from additional follow-up work. The opposing VP and term brain responses to the monaural speech sound within the ERP paradigm are intriguing. To further explore these differences, inclusion of simple tones to investigate whether the neural response changes to tones over the speech sounds would be of particular interest. Similarly, the ceiling effects produced by the early items in the BabyScreen application indicate the tool needs refining for children of this age. The results reported here make it an interesting pilot investigation. The child driven aspect of the task could be beneficial to future investigations, as it provides more accurate measures of processing speed in association to EF abilities and removes administrative bias. The removal of language dependent instructions also holds considerable promise for the utility of this tool with other clinical populations where communication is a particular challenge.

The current investigation has provided the initial steps in exploring EF, IP and attention in a VP population suggesting differences are first detectable after 12 months of age. Further research is required before these results can be utilised in clinical practice and translated into interventions. The logical next step to the current study would be to investigate the relationships observed in a larger sample of VP children. This would allow for a clear range of EF performances to be identified in response to the tasks administered in the second year within the VP group.

Although differences were not observed within first year of assessments, EF, IP and attentional differences could still be apparent from birth using more sensitive tests. The question remains as to whether detecting these differences would be possible. Future work could refine these measures. The challenge to overcome behavioural capabilities to reach the level of complexity required to accurately identify the subtle performance differences however may prove challenging.

The next assessment of this longitudinal cohort has been initiated to follow up on the findings in the current investigation. I have discussed the problems in accepting early assessments as outcomes for the tests performed in infancy. It is vital to follow this group into school age and characterise the middle childhood outcome further. Only then can the predictive validity of these early tests be accurately assessed. Cognitive performance on global measures is fairly constant from early school IQ assessments through to 19 years (Linsell *et al.*, In Press) or young adult life (Breeman *et al.*, 2015) so associations between early testing and middle childhood findings will be important indicators of the stability of cognitive processes.

To date, a number of intervention studies have been conducted in an attempt to improve cognitive outcomes in preterm populations post-discharge from hospital, but evidence is limited, and only short-term benefits have been observed (for reviews see Spittle *et al.*, 2007; Guralnick, 2012). The implication from the current findings could suggest formulating targeted interventions is not appropriate in early childhood for VP populations. If EF is to be accepted as an undifferentiated construct in the early years, targeting interventions to specific domains difficulties identified in later life may not be appropriate or feasible in early childhood (Aarnoudse-Moens, Weisglas-Kuperus, *et al.*, 2009; Mulder, Pitchford and Marlow, 2010; Aarnoudse-Moens *et al.*, 2011). A more generalised approach aimed at collectively improving memory, flexible thinking, attentional abilities, inhibitory behaviours and processing speeds, may be more advantageous in the first two years after birth in those that indicate signs of difficulty. The schemes may then need to adapt with age, in line with the emergence of the EF sub-domains. If the changes in impairments within VP populations can be mapped though childhood, ideally with the use of longitudinal observations such as the one reported here, a clearer picture of the developmental trajectory could be obtained. EF sub-domains are distinguishable within middle childhood (Brydges *et al.*, 2014) with specific difficulties in IP and working memory reported in VP populations (Mulder, Pitchford and Marlow, 2011a, 2011b). Whether this differentiation occurs within the third year after birth is still to be seen (Howard, Okely and Ellis, 2015). By establishing this trajectory, more targeted interventions can be developed and applied as soon as they are applicable to the child's stage of development. This would likely provide the greatest chances of success in improving the developmental outcomes of VP children. The current research suggests differences in developmental abilities are apparent in the early years and therefore interventions should be put in place to aid development and reduce the gap that appears between the VP populations and term born peers. The development and assessment of interventions need to be tracked through to middle childhood, as long term effects need to be observed before clinical programmes can be established. The future focus of VP research should primarily look to explore the differentiation of EF after the 2 year time point reported here, before intervention schemes can be developed accordingly.

When considering the future of EF assessment, ease of administration should also be of high importance. As highlighted in the introductory chapter, the current clinical recommendations are not consistently achieved due to various constraints within the hospital trusts. A more efficient and easily attainable assessment would ease of detection of developmental delay within preterm and other clinical populations requiring prolonged follow-up. An example of this could be the development of the BabyScreen application for clinical use. An application based programme would reduce administrator error and could be completed within a shorter timeframe compared to current clinical follow-up. This therefore could be a valuable direction for future EF research.

More detailed observations exploring the underlining causes of the impairments reported in this population are additionally required. An area that has shown promise but requires further clarity is the use of neuroimaging techniques in the identification of later developmental outcomes. White matter tract abnormalities have been reported within ex-preterm populations in adulthood (Nosarti *et al.*, 2014) and in infancy (Counsell *et al.*, 2008), with relationships between microstructural abnormalities and cognitive (Counsell *et al.*, 2008; Thompson *et al.*, 2012) and language performances observed (Northam *et al.*, 2012). However, regions of interest are still being explored and biomarkers to later cognitive outcomes are yet to be identified. Within the wider PDP study neonatal MRI scans were collected. These are currently being utilised to explore structure function relationships within the current cohort.

In addition, medical factors during the perinatal period are likely to impact later cognitive development, however further epidemiological work is required in order to classify the risks. Although there is limited evidence to suggest the perinatal complications consistently influence outcomes, it is likely this is due to the poor categorisation of risk severity within the neonatal period across the literature. A recent investigation has observed significant correlations between gestational age, maternal steroids and number of surgeries in relation to later EF performance (Duvall *et al.*, 2015). By expanding on studies such as these and including neuroimaging measures, whilst continuing to explore the developmental trajectory of cognitive abilities in ex-preterm children, a more detail picture of how and why these problems emerge would be clarified.

8.7 Concluding remarks

Advances in neonatal care have continued to progress over the past decades and babies are surviving from much lower gestations than ever before. In comparison, the early developmental profile of these children following preterm birth is yet to be fully understood. The current investigation has provided a detailed longitudinal observation of the developmental trajectories of EF, IP and attentional measures in a cohort of very preterm children through the first two years after birth. Overall, VP scores in the first year assessments tend to reflect the scores of term infants. Once aged 2 years, VP performance to assessments show similar patterns to those expected at early school age but further follow up is required to confirm whether the low scores at 2 are identifiable of later challenges. The poorer performances observed at 2 were not accounted for by global cognitive score in the Bayley-III. Moreover, performances in the EF, IP and attentional measures are poor predictors of the Bayley-III at 2 years. Alternative general EF measures including the BabyScreen application may prove to be the best way of identifying individuals at risk in future.

Experimental measures appear to be able to target specific difficulties in cognitive performance in the early years. Currently no clinical assessments are able to accurately identify proportional levels of cognitive impairment likely to be observed later in life. Currently, developmental assessments such as the Bayley-III cognitive tests assess much broader areas of development and therefore do not purely target EF abilities. Clinically these tests are relied upon to identify those who are not following typical developmental trajectories. The principles of these developmental measures are therefore not designed to detect the subtle variability in EF performances and it is not surprising the identification of later cognitive impairments is poor. The question that remains is whether it is in fact possible to accurately identify cognitive impairments within the first year of life and whether general measures can be adapted to predict these difficulties. Multiple factors currently stand in the way of achieving this objective; predominantly the evolution of EF abilities. From observations in the current study, in addition to evidence from the developmental literature, EF abilities appear to develop dramatically over the first few years of life. This makes detecting impairments in this domain challenging, if not impossible. Although the aim of the developmental literature is to identify difficulties within the first year, this may not be feasible given the developmental trajectory of the construct. The current investigation provided important information regarding the emergence of these cognitive abilities across the first two years of life in children born preterm. It is of fundamental importance to follow-up on these findings to clarify the stability of these performances in relations to the bigger developmental picture.

References

Aarnoudse-Moens, C. S. H., Duivenvoorden, H. J., Weisglas-Kuperus, N., Van Goudoever, J. B. and Oosterlaan, J. (2011) 'The profile of executive function in very preterm children at 4 to 12 years', *Developmental Medicine and Child Neurology*, 54(3), pp. 247–253. doi: 10.1111/j.1469-8749.2011.04150.x.

Aarnoudse-Moens, C. S. H., Smidts, D. P., Oosterlaan, J., Duivenvoorden, H. J. and Weisglas-Kuperus, N. (2009) 'Executive function in very preterm children at early school age', *J Abnorm Child Psychol*. 2009/06/03, 37(7), pp. 981–993. doi: 10.1007/s10802-009-9327-z.

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., Duivenvoorden, H. J., van Goudoever, J. B. and Oosterlaan, J. (2013) 'Executive Function and IQ Predict Mathematical and Attention Problems in Very Preterm Children', *PLoS ONE*. Edited by A. V. García, 8(2), p. e55994. doi: 10.1371/journal.pone.0055994.

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B. and Oosterlaan, J. (2009) 'Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children', *Pediatrics*, 124(2), pp. 717–728. doi: 10.1542/peds.2008-2816.

Akshoomoff, N. (2002) 'Selective attention and active engagement in young children.', *Developmental neuropsychology*, 22(3), pp. 625–642. doi: 10.1207/S15326942DN2203_4.

Allen, L. and Kelly, B. B. (2015) *Transforming the Workforce for Children Birth Through Age 8*, *Transforming the Workforce for Children Birth Through Age 8: A Unifying Foundation*. National Academies Press (US). doi: 10.17226/19401.

Alvarez, J. A. and Emory, E. (2006) 'Executive Function and the Frontal Lobes: A Meta-Analytic Review', *Neuropsychology Review*. Springer US, 16(1), pp. 17–42. doi: 10.1007/s11065-006-9002-x.

Amso, D. and Scerif, G. (2015) 'The attentive brain: insights from developmental cognitive neuroscience', *Nature Reviews Neuroscience*. Nature Research, 16(10), pp. 606–619. doi: 10.1038/nrn4025.

Anderson, P. (2002) 'Assessment and development of executive function (EF) during childhood', *Child Neuropsychol*. 2003/03/15, 8(2), pp. 71–82. doi: 10.1076/chin.8.2.71.8724.

Anderson, P. J. (2014) 'Neuropsychological outcomes of children born very preterm', *Seminars in Fetal and Neonatal Medicine*, 19(2), pp. 90–96. doi: 10.1016/j.siny.2013.11.012.

Anderson, P. J. and Burnett, A. (2017) ‘Assessing developmental delay in early childhood — concerns with the Bayley-III scales’, *The Clinical Neuropsychologist*. Routledge, 31(2), pp. 371–381. doi: 10.1080/13854046.2016.1216518.

Anderson, P. J. and Doyle, L. W. (2008) ‘Cognitive and Educational Deficits in Children Born Extremely Preterm’, *Seminars in Perinatology*, pp. 51–58. doi: 10.1053/j.semperi.2007.12.009.

Atkinson, J., Hood, B., Wattam-Bell, J. and Braddick, O. (1992) ‘Changes in Infants’ Ability to Switch Visual Attention in the First Three Months of Life’, *Perception*. SAGE PublicationsSage UK: London, England, 21(5), pp. 643–653. doi: 10.1068/p210643.

Aylward, G. P. (2002) ‘Cognitive and neuropsychological outcomes: More than IQ scores’, *Mental Retardation and Developmental Disabilities Research Reviews*, pp. 234–240. doi: 10.1002/mrdd.10043.

Aylward, G. P. (2005) ‘Neurodevelopmental Outcomes of Infants Born Prematurely’, *Journal of Developmental & Behavioral Pediatrics*, 35(6), pp. 394–407. doi: 10.1097/01.DBP.0000452240.39511.d4.

Aylward, G. P. (2013) ‘Continuing Issues With the Bayley-III’, *Journal of Developmental & Behavioral Pediatrics*, 34(9), pp. 697–701. doi: 10.1097/DBP.0000000000000000.

Bamiou, D. E., Sisodiya, S., Musiek, F. E. and Luxon, L. M. (2007) ‘The role of the interhemispheric pathway in hearing’, *Brain Res Rev*. 2007/08/21, 56(1), pp. 170–182. doi: 10.1016/j.brainresrev.2007.07.003.

Barkley, R. A. (1997) ‘Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD’, *Psychological Bulletin*, 121(1), pp. 65–94. doi: 10.1037/0033-2909.121.1.65.

Barkley, R. A. (2001) ‘The Executive Functions and Self-Regulation: An Evolutionary Neuropsychological Perspective’, *Neuropsychology Review*. Kluwer Academic Publishers-Plenum Publishers, 11(1), pp. 1–29. doi: 10.1023/A:1009085417776.

Baron, I. S., Erickson, K., Ahronovich, M. D., Litman, F. R. and Brandt, J. (2010) ‘Spatial location memory discriminates children born at extremely low birth weight and late-preterm at age three.’, *Neuropsychology*, 24(6), pp. 787–794. doi: 10.1037/a0020382.

Bayley, N. (1969) ‘Bayley Scales of Infant Development’, *New York, NY: The Psychological Corporation*.

Bayley, N. (2006) 'Bayley Scales of Infant and Toddler Development', *Harcourt Assessment, San Antonio, TX*, Third Edit.

Beaino, G., Khoshnood, B., Kaminski, M., Marret, S., Pierrat, V., Vieux, R., Thiriez, G., Matis, J., Picaud, J. C., Rozé, J. C., Alberge, C., Larroque, B., Bréart, G. and Ancel, P. Y. (2011) 'Predictors of the risk of cognitive deficiency in very preterm infants: The EPIPAGE prospective cohort', *Acta Paediatrica, International Journal of Paediatrics*, 100(3), pp. 370–378. doi: 10.1111/j.1651-2227.2010.02064.x.

Beauchamp, M. H., Thompson, D. K., Howard, K., Doyle, L. W., Egan, G. F., Inder, T. E. and Anderson, P. J. (2008) 'Preterm infant hippocampal volumes correlate with later working memory deficits', *Brain*. 2008/09/19, 131(Pt 11), pp. 2986–2994. doi: 10.1093/brain/awn227.

Best, J. R. and Miller, P. H. (2010) 'A developmental perspective on executive function.', *Child development*. NIH Public Access, 81(6), pp. 1641–60. doi: 10.1111/j.1467-8624.2010.01499.x.

Beveridge, M., Jarrold, C. and Pettit, E. (2002) 'An Experimental Approach to Executive Fingerprinting in Young Children', *Infant and Child Development*, 11(2), pp. 107–123. doi: 10.1002/icd.300.

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M. and Anand, K. J. S. (2002) 'Cognitive and behavioural outcomes of school-aged children who were born preterm', *Journal of American Medical Association*, 288, pp. 728–737.

Blair, C. and Razza, R. P. (2007) 'Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten', *Child Development*. Blackwell Publishing Inc, 78(2), pp. 647–663. doi: 10.1111/j.1467-8624.2007.01019.x.

Bode, M. M., D'Eugenio, D. B., Mettelman, B. B. and Gross, S. J. (2014) 'Predictive validity of the bayley, third edition at 2 years for intelligence quotient at 4 years in preterm infants', *J Dev Behav Pediatr*, 35(9), pp. 570–575. doi: 10.1097/dbp.0000000000000110.

Bohm, B., Smedler, A. C. and Forssberg, H. (2004) 'Impulse control, working memory and other executive functions in preterm children when starting school', *Acta Paediatr*. 2004/10/27, 93(10), pp. 1363–1371.

Botting, N., Powls, A., Cooke, R. W. I. and Marlow, N. (1998) 'Cognitive and educational outcome of very-low-birthweight children in early adolescence', *Developmental Medicine and Child Neurology*, 40(10), pp. 652–660. doi: 10.1111/j.1469-8749.1998.tb12324.x.

Bouslama, M., Adla-Biassette, H., Ramanantsoa, N., Bourgeois, T., Bollen, B., Brissaud, O., Matrot, B., Gressens, P. and Gallego, J. (2015) ‘Protective effects of intermittent hypoxia on brain and memory in a mouse model of apnea of prematurity.’, *Frontiers in physiology*. Frontiers Media SA, 6, p. 313. doi: 10.3389/fphys.2015.00313.

Bowen, J. R., Gibson, F. L. and Hand, P. J. (2002) ‘Educational outcome at 8 years for children who were born extremely prematurely: A controlled study’, *Journal of Paediatrics and Child Health*, 38(5), pp. 438–444.

Breeman, L. D., Jaekel, J., Baumann, N., Bartmann, P. and Wolke, D. (2015) ‘Preterm Cognitive Function Into Adulthood’, *Pediatrics*, 136(3), pp. 415–423. doi: 10.1542/peds.2015-0608.

Breeman, L. D., Jaekel, J., Baumann, N., Bartmann, P. and Wolke, D. (2016) ‘Attention problems in very preterm children from childhood to adulthood: the Bavarian Longitudinal Study’, *Journal of Child Psychology and Psychiatry*, 57(2), pp. 132–140. doi: 10.1111/jcpp.12456.

Bruneau, N., Bidet-Caulet, A., Roux, S., Bonnet-Brilhault, F. and Gomot, M. (2015) ‘Asymmetry of temporal auditory T-complex: Right ear-left hemisphere advantage in Tb timing in children’, *International Journal of Psychophysiology*, 95(2), pp. 94–100. doi: 10.1016/j.ijpsycho.2014.07.012.

Brydges, C. R., Fox, A. M., Reid, C. L. and Anderson, M. (2014) ‘The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis’, *Intelligence*, 47, pp. 34–43. doi: 10.1016/j.intell.2014.08.010.

Bryson, S. E., Landry, R., Czapinski, P., McConnell, B., Rombough, V. and Wainwright, A. (2004) ‘Autistic spectrum disorders: Causal mechanisms and recent findings on attention and emotion’, *International Journal of Special Education*, 9(1), pp. 14–22.

Bull, R., Espy, K. A. and Wiebe, S. a (2008) ‘Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years.’, *Developmental Neuropsychology*, 33(3), pp. 205–228. doi: 10.1080/87565640801982312.

Bull, R. and Johnston, R. S. (1997) ‘Children’s arithmetical difficulties: contributions from processing speed, item identification, and short-term memory’, *J Exp Child Psychol*. 1997/04/01, 65(1), pp. 1–24. doi: 10.1006/jecp.1996.2358.

Bull, R. and Scerif, G. (2001) ‘Executive functionin as a predictor of children’s mathematics ability: Inhibition, switching, and working memory’, *Developmental Neuropsychology*, 19(3), pp. 273–293. doi: 10.1207/S15326942DN1903.

Burgess, P. (1997) ‘Theory and methodology in executive function research’, (October), p. Psychology Press: East Sussex, UK.

Burnett, A. C., Scratch, S. E., Lee, K. J., Cheong, J., Searle, K., Hutchinson, E., De Luca, C., Davey, M.-A., Roberts, G., Doyle, L. W. and Anderson, P. J. (2015) ‘Executive Function in Adolescents Born <1000 g or <28 Weeks: A Prospective Cohort Study.’, *Pediatrics*, 135(4), pp. e826-834. doi: 10.1542/peds.2014-3188.

Butcher, P. R., Kalverboer, A. F. and Geuze, R. . (2000) ‘Infants’ shifts of gaze from a central to a peripheral stimulus: a longitudinal study of development between 6 and 26 weeks’, *Infant Behavior and Development*, 23(1), pp. 3–21. doi: 10.1016/S0163-6383(00)00031-X.

Butcher, P. R., Kalverboer, A. F., Geuze, R. H. and Stremmelaar, E. F. (2002) ‘A longitudinal study of the development of shifts of gaze to a peripheral stimulus in preterm infants with transient periventricular echogenicity’, *Journal of Experimental Child Psychology*, 82(2), pp. 116–140. doi: 10.1016/S0022-0965(02)00006-1.

Carlson, S. M. (2005) ‘Developmentally sensitive measures of executive function in preschool children’, *Developmental Neuropsychology*, 28(2), pp. 595–616. doi: 10.1207/s15326942dn2802.

Carlson, S. M. (2013) ‘Executive Function Scale for Early Childhood: Manual and Sample Materials’. Institute of Child Development, University of Minnesota.

Carlson, S. M., Mandell, D. J. and Williams, L. (2004) ‘Executive function and theory of mind: Stability and prediction from ages 2 to 3’, *Dev Psychol*, 40(6), pp. 1105–1122.

Chi, M. T. (1977) ‘Age differences in the speed of processing: A critique.’, *Developmental Psychology*, 13(5), pp. 543–544. doi: 10.1037/0012-1649.13.5.543.

Choudhury, N. and Benasich, A. A. (2011) ‘Maturation of auditory evoked potentials from 6 to 48 months: Prediction to 3 and 4 year language and cognitive abilities’, *Clinical Neurophysiology*, 122(2), pp. 320–338. doi: 10.1016/j.clinph.2010.05.035.

Christian, K., Morrison, F. J. and Bryant, F. B. (1998) ‘Predicting kindergarten academic skills: Interactions among child care, maternal education, and family literacy environments’, *Early Childhood Research Quarterly*. JAI, 13(3), pp. 501–521. doi: 10.1016/S0885-2006(99)80054-4.

Chugani, H. T., Phelps, M. E. and Mazziotta, J. C. (1987) ‘Positron emission tomography study of human brain functional development.’, *Annals of neurology*, 22(4), pp. 487–497. doi: 10.1002/ana.410220408.

Clark, V. P., Fan, S. and Hillyard, S. A. (1994) ‘Identification of early visual evoked potential generators by retinotopic and topographic analyses’, *Human Brain Mapping*, 2(3), pp. 170–187. doi: 10.1002/hbm.460020306.

Colombo, J. (1993) *Infant cognition: Predicting later intellectual functioning.*, *Infant cognition: Predicting later intellectual functioning.*

Colombo, J. (2001) ‘The development of visual attention in infancy’, *Annu Rev Psychol*, 52, pp. 337–367. doi: 10.1146/annurev.psych.52.1.337.

Colombo, J. and Mitchell, D. W. (2009) ‘Infant visual habituation’, *Neurobiol Learn Mem*. 2008/07/16, 92(2), pp. 225–234. doi: 10.1016/j.nlm.2008.06.002.

Counsell, S. J., Edwards, A. D., Chew, A. T., Anjari, M., Dyet, L. E., Srinivasan, L., Boardman, J. P., Allsop, J. M., Hajnal, J. V, Rutherford, M. A. and Cowan, F. M. (2008) ‘Specific relations between neurodevelopmental abilities and white matter microstructure in children born preterm’, *Brain*. 2008/10/28, 131(Pt 12), pp. 3201–3208. doi: 10.1093/brain/awn268.

Cowan, N. (2011) ‘The focus of attention as observed in visual working memory tasks: Making sense of competing claims’, *Neuropsychologia*, 49(6), pp. 1401–1406. doi: 10.1016/j.neuropsychologia.2011.01.035.

Cuskelly, M., Einam, M. and Jobling, A. (2001) ‘Delay of gratification in young adults with Down syndrome’, *Down Syndrome Research and Practice*. Down Syndrome Education International, 7(2), pp. 60–67. doi: 10.3104/reports.115.

Demetriou, A., Christou, C., Spanoudis, G., Platsidou, M., Fischer, K. W. and Dawson, T. L. (2002) ‘The Development of Mental Processing: Efficiency, Working Memory, and Thinking’, *Source: Monographs of the Society for Research in Child Development*, 67(1), pp. 1–167. doi: 10.2307/3181583.

Diamond, A. (1985) ‘Development of Ability to use Recall to Guide Action, as Indicated by Infants’ Performance on AB’, *Child Development*, 56, pp. 868–883.

Diamond, A. (1990) ‘The development and neural bases of higher cognitive functions. Introduction’, *Ann N Y Acad Sci*. 1990/01/01, 608, pp. xiii–lvi.

Diamond, A. (1991) ‘Neuropsychological insights into the meaning of object concept development.’, *The epigenesis of mind: Essays on biology and cognition.*, pp. 67–110.

Diamond, A. (2006) ‘The Early Development of Executive Functions’, in *Lifespan Cognition, Mechanisms of change*. Oxford University press, pp. 70–95.

Diamond, A. (2012) 'Activities and Programs That Improve Children's Executive Functions', *Current Directions in Psychological Science*, 21(5), pp. 335–341. doi: 10.1177/0963721412453722.

Diamond, A. (2013) 'Executive Functions', *Annual Review of Psychology*, 64(1), pp. 135–168. doi: 10.1146/annurev-psych-113011-143750.

Diamond, A., Carlson, S. M. and Beck, D. M. (2005) 'Preschool children's performance in task switching on the dimensional change card sort task: separating the dimensions aids the ability to switch', *Dev Neuropsychol*. 2005/09/08. Lawrence Erlbaum Associates, Inc., 28(2), pp. 689–729. doi: 10.1207/s15326942dn2802_7.

Diamond, A. and Doar, B. (1989) 'The Performance of Human Infants on a Measure of Frontal Cortex Function, the Delayed Response Task', *Developmental psychobiology*, 22(3), pp. 271–294.

Diamond, A. and Goldman-Rakic, P. S. (1989) 'Comparison of human infants and rhesus monkeys on Piaget's AB task: evidence for dependence on dorsolateral prefrontal cortex', *Experimental Brain Research*, 74(1), pp. 24–40. doi: 10.1007/BF00248277.

van Dinteren, R., Arns, M., Jongsma, M. L. and Kessels, R. P. (2014) 'P300 development across the lifespan: a systematic review and meta-analysis', *PLoS One*. 2014/02/20, 9(2), p. e87347. doi: 10.1371/journal.pone.0087347.

Duffau, H. (2014) 'Diffusion Tensor Imaging Is a Research and Educational Tool, but Not Yet a Clinical Tool', *World Neurosurgery*, 82(1–2), pp. e43–e45. doi: 10.1016/j.wneu.2013.08.054.

Duvall, S. W., Erickson, S. J., MacLean, P. and Lowe, J. R. (2015) 'Perinatal Medical Variables Predict Executive Function Within a Sample of Preschoolers Born Very Low Birth Weight', *Journal of Child Neurology*. SAGE PublicationsSage CA: Los Angeles, CA, 30(6), pp. 735–740. doi: 10.1177/0883073814542945.

Ebaid, D., Crewther, S. G., MacCalman, K., Brown, A. and Crewther, D. P. (2017) 'Cognitive processing speed across the lifespan: Beyond the influence of motor speed', *Frontiers in Aging Neuroscience*. Frontiers, 9(MAR), p. 62. doi: 10.3389/fnagi.2017.00062.

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., Bolton, P., Charman, T., Baird, G. and Johnson, M. H. (2009) 'Visual orienting in the early broader autism phenotype: Disengagement and facilitation', *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50(5), pp. 637–642.

Elsmen, E., Pupp, I. H. and Hellstrom-Westas, L. (2004) 'Preterm male infants need more initial respiratory and circulatory support than female infants', *Acta Paediatrica*, 93(4), pp. 529–533. doi: 10.1080/08035250410024998.

Eryigit-Madzwamuse, S. and Wolke, D. (2015) 'Attention problems in relation to gestational age at birth and smallness for gestational age', *Early Human Development*, 91(2), pp. 131–138. doi: 10.1016/j.earlhumdev.2015.01.004.

Eryigit Madzwamuse, S., Baumann, N., Jaekel, J., Bartmann, P. and Wolke, D. (2015) 'Neuro-cognitive performance of very preterm or very low birth weight adults at 26 years', *Journal of Child Psychology and Psychiatry*, 56(8), pp. 857–864. doi: 10.1111/jcpp.12358.

Escera, C., Alho, K., Schröger, E. and Winkler, I. W. (2000) 'Involuntary Attention and Distractibility as Evaluated with Event-Related Brain Potentials', *Audiology and Neurotology*, 5(3–4), pp. 151–166. doi: 10.1159/000013877.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D. and Glisky, M. L. (1999) 'Executive functioning in preschool children: performance on A-not-B and other delayed response format tasks', *Brain Cogn*. 1999/12/11, 41(2), pp. 178–199. doi: 10.1006/brcg.1999.1117.

Espy, K. A., Stalets, M. M., McDiarmid, M. M., Senn, T. E., Cwik, M. F. and Hamby, A. (2002) 'Executive functions in preschool children born preterm: application of cognitive neuroscience paradigms', *Child Neuropsychol*. 2003/03/15, 8(2), pp. 83–92. doi: 10.1076/chin.8.2.83.8723.

Fagan, J. F. and Singer, L. T. (1983) 'Infant recognition memory as a measure of intelligence', *Advances in Infancy Research*, 2(January 1983), pp. 31–78.

Fellman, V., Kushnerenko, E., Mikkola, K., Ceponiene, R., Leipälä, J. and Näätänen, R. (2004) 'Atypical Auditory Event-Related Potentials in Preterm Infants during the First Year of Life: A Possible Sign of Cognitive Dysfunction?', *Pediatric Research*, 56(2), pp. 291–297. doi: 10.1203/01.pdr.0000132750.97066.b9.

Fischer, B. and Weber, H. (1993) 'Express saccades and visual attention', *Behavioral and Brain Sciences*, 16(3), p. 553. doi: 10.1017/S0140525X00031575.

Foreman, N., Fielder, A., Price, D. and Bowler, V. (1991) 'Tonic and phasic orientation in full-term and preterm infants', *Journal of Experimental Child Psychology*, 51(3), pp. 407–422. doi: 10.1016/0022-0965(91)90085-7.

Fry, A. F. and Hale, S. (2000) 'Relationships among processing speed, working memory, and fluid intelligence in children', *Biol Psychol*. 2000/10/18, 54(1–3), pp. 1–34.

Fuster, J. M. (1973) 'Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory.', *Journal of Neurophysiology*, 36(1), pp. 61–78.

Garon, N., Bryson, S. E. and Smith, I. M. (2008) 'Executive function in preschoolers: A review using an integrative framework.', *Psychological Bulletin*, 134(1), pp. 31–60. doi: 10.1037/0033-2909.134.1.31.

Gathercole, S. E., Pickering, S. J., Knight, C. and Stegmann, Z. (2004) 'Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age', *Applied Cognitive Psychology*. John Wiley & Sons, Ltd., 18(1), pp. 1–16. doi: 10.1002/acp.934.

van der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N. and van Engeland, H. (2001) 'Eye movements, visual attention, and autism: a saccadic reaction time study using the gap and overlap paradigm', *Biol Psychiatry*. 2001/11/03, 50(8), pp. 614–619.

Gilbert, S. and Burgess, P. (2008) 'Executive function', *Learning*, 18(3), pp. R110-4. doi: 10.1016/j.cub.2007.12.014.

Gilmore, C. S., Clementz, B. A. and Berg, P. (2009) 'Hemispheric differences in auditory oddball responses during monaural versus binaural stimulation', *International Journal of Psychophysiology*, 73(3), pp. 326–333. doi: http://dx.doi.org/10.1016/j.ijpsycho.2009.05.005.

Gilmore, R. O. and Johnson, M. H. (1995) 'Working memory in infancy: six-month-olds' performance on two versions of the oculomotor delayed response task', *J Exp Child Psychol*. 1995/06/01, 59(3), pp. 397–418.

Green, J., Charman, T., Pickles, A., Wan, M. W., Elsabbagh, M., Slonims, V., Taylor, C., McNally, J., Booth, R., Gliga, T., Jones, E. J. H., Harrop, C., Bedford, R. and Johnson, M. H. (2015) 'Parent-mediated intervention versus no intervention for infants at high risk of autism: A parallel, single-blind, randomised trial', *The Lancet Psychiatry*. Green et al. Open Access article distributed under the terms of CC BY, 2(2), pp. 133–140. doi: 10.1016/S2215-0366(14)00091-1.

Guralnick, M. J. (2012) 'Preventive Interventions for Preterm Children: Effectiveness and Developmental Mechanisms', *Journal of Developmental and Behavioral Pediatrics*, 33(4), pp. 352–364. doi: 10.1097/DBP.0b013e31824eaa3c.

de Haan, M., Bauer, P. J., Georgieff, M. K. and Nelson, C. A. (2000) 'Explicit memory in low-risk infants aged 19 months born between 27 and 42 weeks of gestation', *Dev Med Child Neurol*. 2000/06/16, 42(5), pp. 304–312.

Hack, M. (2006) ‘Young adult outcomes of very-low-birth-weight children’, *Seminars in Fetal and Neonatal Medicine*, pp. 127–137. doi: 10.1016/j.siny.2005.11.007.

Hack, M., Taylor, H. G., Drotar, D., Schluchter, M., Cartar, L., Wilson-Costello, D., Klein, N., Friedman, H., Mercuri-Minich, N. and Morrow, M. (2005) ‘Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age’, *Pediatrics*, 116(2), pp. 333–341. doi: 10.1542/peds.2005-0173.

Hadders-Algra, M. (2011) ‘Challenges and limitations in early intervention’, *Developmental Medicine and Child Neurology*. Blackwell Publishing Ltd, pp. 52–55. doi: 10.1111/j.1469-8749.2011.04064.x.

Haley, D. W., Grunau, R. E., Oberlander, T. F. and Weinberg, J. (2008) ‘Contingency Learning and Reactivity in Preterm and Full-Term Infants at 3 Months’, *Infancy*. 2008/11/01, 13(6), pp. 570–595. doi: 10.1080/15250000802458682.

Haley, D. W., Weinberg, J. and Grunau, R. E. (2006) ‘Cortisol, contingency learning, and memory in preterm and full-term infants’, *Psychoneuroendocrinology*, 31(1), pp. 108–117. doi: http://dx.doi.org/10.1016/j.psyneuen.2005.06.007.

Hammond, C. J., Potenza, M. N. and Mayes, L. C. (2011) *Development of Impulse Control, Inhibition, and Self-Regulatory Behaviors in Normative Populations across the Lifespan*. Edited by J. E. Grant and M. N. Potenza. Oxford University Press. doi: 10.1093/oxfordhb/9780195389715.013.0082.

Heathcock, J. C., Bhat, A. N., Lobo, M. A. and Galloway, J. C. (2004) ‘The performance of infants born preterm and full-term in the mobile paradigm: learning and memory’, *Physical Therapy*, 84(9), pp. 808–821.

Heathcock, J. C., Bhat, A. N., Lobo, M. A. and Galloway, J. C. (2005) ‘The relative kicking frequency of infants born full-term and preterm during learning and short-term and long-term memory periods of the mobile paradigm’, *Phys Ther*. 2004/12/30, 85(1), pp. 8–18.

Herrmann, C. S. and Knight, R. T. (2001) ‘Mechanisms of human attention: Event-related potentials and oscillations’, in *Neuroscience and Biobehavioral Reviews*, pp. 465–476. doi: 10.1016/S0149-7634(01)00027-6.

Hitzert, M. M., Van Braeckel, K. N. J. A., Bos, A. F., Hunnius, S. and 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*. Frontiers Media SA, 2, p. 106. doi: 10.3389/fped.2014.00106.

Hood, B. M. and Atkinson, J. (1993) ‘Disengaging visual attention in the infant and adult', *Infant Behavior and Development*, 16(4), pp. 405–422.

Hövel, H., Partanen, E., Huotilainen, M., Lindgren, M., Rosén, I. and Fellman, V. (2014) 'Auditory event-related potentials at preschool age in children born very preterm', *Clinical Neurophysiology*, 125(3), pp. 449–456. doi: http://dx.doi.org/10.1016/j.clinph.2013.07.026.

Howard, K., Anderson, P. J., & Taylor, H. G. (2008) 'Executive functioning and attention in children born preterm', in *Executive functions and the frontal lobes. A lifespan perspective*. New York, NY: Psychology Press, pp. 219–241.

Howard, S. J., Okely, A. D. and Ellis, Y. G. (2015) 'Evaluation of a differentiation model of preschoolers' executive functions.', *Frontiers in psychology*. Frontiers Media SA, 6, p. 285. doi: 10.3389/fpsyg.2015.00285.

Hughes, C. (2002) 'Executive functions and development: emerging themes', *Infant and Child Development*. John Wiley & Sons, Ltd., 11(2), pp. 201–209. doi: 10.1002/icd.297.

Hughes, C., Ensor, R., Wilson, A. and Graham, A. (2010) 'Tracking Executive Function Across the Transition to School: A Latent Variable Approach', *Developmental Neuropsychology*. Taylor & Francis Group, 35(1), pp. 20–36. doi: 10.1080/87565640903325691.

Huizinga, M., Dolan, C. V. and van der Molen, M. W. (2006) 'Age-related change in executive function: Developmental trends and a latent variable analysis', *Neuropsychologia*, 44(11), pp. 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010.

Hunnius, S., Geuze, R. H., Zweens, M. J. and Bos, A. F. (2008) 'Effects of Preterm Experience on the Developing Visual System: A Longitudinal Study of Shifts of Attention and Gaze in Early Infancy', *Developmental Neuropsychology*. Taylor & Francis Group, 33(4), pp. 521–535. doi: 10.1080/87565640802101508.

Huppi, P. S. (2010) 'Growth and development of the brain and impact on cognitive outcomes', *Nestle Nutr Workshop Ser Pediatr Program*. 2010/02/09, 65, pp. 137–151. doi: 10.1159/000281156.

Hutchinson, E. a, De Luca, C. R., Doyle, L. W., Roberts, G. and Anderson, P. J. (2013) 'School-age outcomes of extremely preterm or extremely low birth weight children.', *Pediatrics*, 131(4), pp. 1053–1061. doi: 10.1542/peds.2012-2311.

Isaacs, E. B., Edmonds, C. J., Lucas, A. and Gadian, D. G. (2001) 'Calculation difficulties in children of very low birthweight: a neural correlate.', *Brain : a journal of neurology*, 124(Pt 9), pp. 1701–1707. doi: 11522573.

Ishii, S. and Hashimoto-Torii, K. (2015) 'Impact of prenatal environmental stress on cortical development', *Frontiers in Cellular Neuroscience*. Frontiers Media S.A., 9, p. 207. doi: 10.3389/fncel.2015.00207.

Isquith, P. K., Crawford, J. S., Espy, K. A. and Gioia, G. A. (2005) 'Assessment of Executive Function in Preschool-Aged Children', *Mental retardation and developmental disabilities research reviews*, 11(3), pp. 209–215. doi: 10.1002/mrdd.20075.

Jaekel, J., Wolke, D. and Bartmann, P. (2013) 'Poor attention rather than hyperactivity/impulsivity predicts academic achievement in very preterm and full-term adolescents', *Psychological Medicine*, 43(1), pp. 183–196. doi: 10.1017/S0033291712001031.

Jansson-Verkasalo, E., Čeponien≐, R., Valkama, M., Vainionpää, L., Laitakari, K., Alku, P., Suominen, K. and Näätänen, R. (2003) 'Deficient speech-sound processing, as shown by the electrophysiologic brain mismatch negativity response, and naming ability in prematurely born children', *Neuroscience Letters*, 348(1), pp. 5–8. doi: 10.1016/s0304-3940(03)00641-4.

Jansson-Verkasalo, E., Ruusuvirta, T., Huotilainen, M., Alku, P., Kushnerenko, E., Suominen, K., Rytky, S., Luotonen, M., Kaukola, T., Tolonen, U. and Hallman, M. (2010) 'Atypical perceptual narrowing in prematurely born infants is associated with compromised language acquisition at 2 years of age', *BMC Neurosci*, 11, p. 88. doi: 10.1186/1471-2202-11-88.

Jansson-Verkasalo, E., Valkama, M., Vainionpaa, L., Paakko, E., Ilkko, E. and Lehtihalmes, M. (2004) 'Language development in very low birth weight preterm children: a follow-up study', *Folia Phoniatr Logop*. 2004/03/06, 56(2), pp. 108–119. doi: 10.1159/000076062.

Jensen, A. R. (1993) 'Why Is Reaction Time Correlated With Psychometric g?', *Current Directions in Psychological Science*, 2(2), pp. 53–56. doi: 10.1111/1467-8721.ep10770697.

Johnson, M. H. (2000) 'Functional brain development in infants: Elements of an interactive specialization framework', *Child development*, 71(1), pp. 75–81. doi: http://dx.doi.org/10.1111/1467-8624.00120.

Johnson, M. H. (2001) 'Functional brain development in humans', *Nature Reviews Neuroscience*, 2(7), pp. 475–483.

Johnson, S. (2007) 'Cognitive and behavioural outcomes following very preterm birth', *Seminars in fetal & neonatal medicine*. 2007/07/13, 12(5), pp. 363–373. doi: 10.1016/j.siny.2007.05.004.

Johnson, S., Fawke, J., Hennessy, E., Rowell, V., Thomas, S., Wolke, D. and Marlow, N. (2009) 'Neurodevelopmental disability through 11 years of age in children born before 26 weeks of gestation', *Pediatrics*. 2009/08/05, 124(2), pp. e249-57. doi: 10.1542/peds.2008-3743.

Johnson, S., Hennessy, E., Smith, R., Trikic, R., Wolke, D. and Marlow, N. (2009) 'Academic attainment and special educational needs in extremely preterm children at 11 years of age: the EPICure study', *Archives of disease in childhood. Fetal and neonatal edition*. 2009/03/14, 94(4), pp. F283-9. doi: 10.1136/adc.2008.152793.

Johnson, S. and Marlow, N. (2006) 'Developmental screen or developmental testing?', *Early Human Development*, 82(3), pp. 173–183. doi: 10.1016/j.earlhumdev.2006.01.008.

Johnson, S. and Marlow, N. (2016) 'Early and long-term outcome of infants born extremely preterm.', *Archives of disease in childhood*. BMJ Publishing Group Ltd and Royal College of Paediatrics and Child Health, 102(1), pp. 97–102. doi: 10.1136/archdischild-2015-309581.

Johnson, S., Moore, T. and Marlow, N. (2014) 'Using the Bayley-III to assess neurodevelopmental delay: which cut-off should be used?', *Pediatr Res*. 2014/02/05, 75(5), pp. 670–674. doi: 10.1038/pr.2014.10.

Johnson, S., Wolke, D. and Marlow, N. (2008) 'Outcome monitoring in preterm populations: Measures and methods', *Zeitschrift für Psychologie/Journal of Psychology*, 216(3), p. 135.

de Jong, M., Verhoeven, M., Lasham, C. A., Meijssen, C. B. and van Baar, A. L. (2015) 'Behaviour and development in 24-month-old moderately preterm toddlers', *Arch Dis Child*. 2015/01/16, 100(6), pp. 548–553. doi: 10.1136/archdischild-2014-307016.

Jongbloed-Pereboom, M., Janssen, A. J. W. M., Steenbergen, B. and Nijhuis-van der Sanden, M. W. G. (2012) 'Motor learning and working memory in children born preterm: A systematic review', *Neuroscience & Biobehavioral Reviews*, 36(4), pp. 1314–1330. doi: 10.1016/j.neubiorev.2012.02.005.

Kail, R. and Salthouse, T. A. (1994) 'Processing speed as a mental capacity', *Acta Psychol (Amst)*. 1994/08/01, 86(2–3), pp. 199–225.

Kail, R. V. and Ferrer, E. (2007) 'Processing speed in childhood and adolescence: Longitudinal models for examining developmental change', *Child Development*. Blackwell Publishing Inc, 78(6), pp. 1760–1770. doi: 10.1111/j.1467-8624.2007.01088.x.

Kallioinen, M., Eadon, H., Murphy, S. and Baird, G. (2017) 'Developmental follow-up of children and young people born preterm: summary of NICE guidance', *BMJ*, 358(j3514). doi: 10.1136/bmj.j3514.

Kalpakidou, A. K., Allin, M. P. G., Walshe, M., Giampietro, V., McGuire, P. K., Rifkin, L., Murray, R. M. and Nosarti, C. (2014) 'Functional neuroanatomy of executive function after neonatal brain injury in adults who were born very preterm', *PLoS ONE*. Edited by Y.-F. Zang, 9(12), p. e113975. doi: 10.1371/journal.pone.0113975.

Kane, M. J. and Engle, R. W. (2003) 'Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference.', *Journal of Experimental Psychology: General*, 132(1), pp. 47–70. doi: 10.1037/0096-3445.132.1.47.

Kavsek, M. and Bornstein, M. H. (2010) 'Visual habituation and dishabituation in preterm infants: a review and meta-analysis', *Res Dev Disabil*. 2010/05/22, 31(5), pp. 951–975. doi: 10.1016/j.ridd.2010.04.016.

Kemner, C., Verbaten, M. N., Cuperus, J. M., Camfferman, G. and Van Engeland, H. (1998) 'Abnormal saccadic eye movements in autistic children', *Journal of Autism & Developmental Disorders*, 28(1), pp. 61–67. doi: 10.1023/A:1026015120128.

Kerr-Wilson, C. O., Mackay, D. F., Smith, G. C. S. and Pell, J. P. (2012) 'Meta-analysis of the association between preterm delivery and intelligence', *Journal of Public Health*, 34(2), pp. 209–216. doi: 10.1093/pubmed/fdr024.

Kimura, D. (1961) 'Cerebral dominance and the perception of verbal stimuli', *Canadian Journal of Psychology/Revue canadienne de psychologie*. Canada: University of Toronto Press, 15(3), pp. 166–171. doi: 10.1037/h0083219.

Kimura, D. (1967) 'Functional asymmetry of the brain in dichotic listening', *Cortex*, (3), pp. 163–168.

Klein, E., Moeller, K., Kiechl-Kohlendorfer, U., Kremser, C., Starke, M., Cohen Kadosh, R., Pupp-Peglow, U., Schocke, M. and Kaufmann, L. (2014) 'Processing of intentional and automatic number magnitudes in children born prematurely: evidence from fMRI.', *Developmental neuropsychology*. Taylor & Francis, 39(5), pp. 342–64. doi: 10.1080/87565641.2014.939179.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L. and Vandegeest, K. A. (1996) 'Inhibitory Control in Young Children and Its Role in Emerging Internalization', *Child Development*, 67(2), pp. 490–507.

Kochanska, G., Murray, K. T. and Harlan, E. T. (2000) 'Effortful control in early childhood: continuity and change, antecedents, and implications for social development.', *Developmental psychology*, 36(2), pp. 220–232. doi: 10.1037/0012-

1649.36.2.220.

Kolb, B., Mychasiuk, R., Muhammad, A. and Gibb, R. (2013) 'Brain plasticity in the developing brain', *Prog Brain Res*. 2013/12/07, 207, pp. 35–64. doi: 10.1016/b978-0-444-63327-9.00005-9.

Krägeloh-Mann, I. (2004) 'Imaging of early brain injury and cortical plasticity', *Experimental Neurology*, 190, Suppl, pp. 84–90. doi: http://dx.doi.org/10.1016/j.expneurol.2004.05.037.

Kushnerenko, E. V, Van den Bergh, B. R. H. and Winkler, I. (2013) 'Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence', *Frontiers in Psychology*, 4. doi: 10.3389/fpsyg.2013.00595.

Kyllonen, P. and Zu, J. (2016) 'Use of Response Time for Measuring Cognitive Ability', *Journal of Intelligence*, 4(4), p. 14. doi: 10.3390/jintelligence4040014.

Landry, S. H. and Chapieski, M. L. (1988) 'Visual attention during toy exploration in preterm infants: Effects of medical risk and maternal interactions', *Infant Behavior and Development*, 11(2), pp. 187–204. doi: 10.1016/S0163-6383(88)80005-5.

Landry, S. H., Leslie, N. A., Fletcher, J. M. and Francis, D. J. (1985) 'Visual attention skills of premature infants with and without intraventricular hemorrhage', *Infant Behavior and Development*, 8(3), pp. 309–321. doi: 10.1016/0163-6383(85)90020-7.

Lawrence, E. J., Rubia, K., Murray, R. M., McGuire, P. K., Walshe, M., Allin, M., Giampietro, V., Rifkin, L., Williams, S. C. and Nosarti, C. (2009) 'The neural basis of response inhibition and attention allocation as mediated by gestational age', *Hum Brain Mapp*. 2008/04/17, 30(3), pp. 1038–1050. doi: 10.1002/hbm.20564.

Lazard, D. S., Collette, J. L. and Perrot, X. (2012) 'Speech processing: from peripheral to hemispheric asymmetry of the auditory system', *Laryngoscope*. 2011/11/19, 122(1), pp. 167–173. doi: 10.1002/lary.22370.

Lehto, J. E., Juujärvi, P., Kooistra, L. and Pulkkinen, L. (2003) 'Dimensions of executive functioning: Evidence from children', *British Journal of Developmental Psychology*. Blackwell Publishing Ltd, 21(1), pp. 59–80. doi: 10.1348/026151003321164627.

Li, K., Sun, Z., Han, Y., Gao, L., Yuan, L. and Zeng, D. (2015) 'Fractional anisotropy alterations in individuals born preterm: A diffusion tensor imaging meta-analysis', *Developmental Medicine and Child Neurology*, 57(4), pp. 328–338. doi: 10.1111/dmcn.12618.

Linsell, L., Johnson, S., O'Reilly, H., Wolke, W., Morris, J. K. and Kurinczuk, J. J. (no date) ‘Cognitive trajectories from infancy to early adulthood in infants born extremely preterm’, *In press*.

Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J. and Marlow, N. (2015) ‘Prognostic Factors for Poor Cognitive Development in Children Born Very Preterm or With Very Low Birth Weight’, *JAMA Pediatrics*. American Medical Association, 169(12), p. 1162. doi: 10.1001/jamapediatrics.2015.2175.

Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J. and Marlow, N. (2017) ‘Risk Factor Models for Neurodevelopmental Outcomes in Children Born Very Preterm or With Very Low Birth Weight: A Systematic Review of Methodology and Reporting’, *American Journal of Epidemiology*. Springer Publishing Company, New York, NY, 185(7), pp. 601–612. doi: 10.1093/aje/kww135.

Lobo, M. A. and Galloway, J. C. (2013) ‘Assessment and stability of early learning abilities in preterm and full-term infants across the first two years of life’, *Research in Developmental Disabilities*. Elsevier Ltd, 34(5), pp. 1721–1730. doi: 10.1016/j.ridd.2013.02.010.

Lovato, S. B. and Waxman, S. R. (2016) ‘Young Children Learning from Touch Screens: Taking a Wider View’, *Frontiers in Psychology*. Frontiers, 7, p. 1078. doi: 10.3389/fpsyg.2016.01078.

Luna, B. (2009) ‘Developmental changes in cognitive control through adolescence.’, *Advances in child development and behavior*. NIH Public Access, 37, pp. 233–78.

Luu, T. M., Ment, L. R., Schneider, K. C., Katz, K. H., Allan, W. C. and Vohr, B. R. (2009) ‘Lasting effects of preterm birth and neonatal brain hemorrhage at 12 years of age.’, *Pediatrics*. NIH Public Access, 123(3), pp. 1037–44. doi: 10.1542/peds.2008-1162.

Mackay, D. F., Smith, G. C. S., Dobbie, R. and Pell, J. P. (2010) ‘Gestational age at delivery and special educational need: retrospective cohot study of 407,503 schoolchildrenr’, *PLoS Med*. 2010/06/15, 7(6), p. e1000289. doi: 10.1371/journal.pmed.1000289.

Månsson, J., Fellman, V. and Stjernqvist, K. (2015) ‘Extremely preterm birth affects boys more and socio-economic and neonatal variables pose sex-specific risks’. Blackwell Publishing Ltd, 104(5), pp. 514–521. doi: 10.1111/apa.12937.

Mansson, J. and Stjernqvist, K. (2014) ‘Children born extremely preterm show significant lower cognitive, language and motor function levels compared to children born at term, as measured by the Bayley-III at 2.5 years’, *Acta Paediatr*. 2014/02/06. doi: 10.1111/apa.12585.

Marlow, N. (2004) ‘Neurocognitive outcome after very preterm birth’, *Archives of disease in childhood. Fetal and neonatal edition*. 2004/04/23, 89(3), pp. F224-8.

Marlow, N. (2013) ‘Measuring neurodevelopmental outcome in neonatal trials: a continuing and increasing challenge’, *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 98(6), pp. F554–F558. doi: 10.1136/archdischild-2012-302970.

Marlow, N., Hennessy, E. M., Bracewell, M. A. and Wolke, D. (2007) ‘Motor and executive function at 6 years of age after extremely preterm birth’, *Pediatrics*. 2007/10/03, 120(4), pp. 793–804. doi: 10.1542/peds.2007-0440.

Marlow, N., Hennessy, E. M., Bracewell, M. A., Wolke, D. and for the EPICure Study Group (2007) ‘Motor and executive function at 6 years of age after extremely preterm birth’, *Pediatrics*. 2007/10/03, 120(4), pp. 793–804. doi: 10.1542/peds.2007-0440.

Matthews, A., Ellis, A. E. and Nelson, C. A. (1996) ‘Development of preterm and full-term infant ability on AB, recall memory, transparent barrier detour, and means-end tasks’, *Child Dev*. 1996/12/01, 67(6), pp. 2658–2676.

McCall, R. B. and Carriger, M. S. (1993) ‘A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ.’, *Child development*, 64(1), pp. 57–79.

de Mello, R. R., Rodrigues Reis, A. B. and da Silva, K. S. (2017) ‘Cognitive performance of premature infants: association between bronchopulmonary dysplasia and cognitive skills. Cross-sectional study’, *Sao Paulo Medical Journal*, 135(4), pp. 383–390. doi: 10.1590/1516-3180.2017.0010190317.

Mikkola, K., Kushnerenko, E., Partanen, E., Serenius-Sirve, S., Leipälä, J., Huotilainen, M. and Fellman, V. (2007) ‘Auditory event-related potentials and cognitive function of preterm children at five years of age’, *Clinical Neurophysiology*, 118(7), pp. 1494–1502. doi: 10.1016/j.clinph.2007.04.012.

Millar, W. S. and Weir, C. (2015) ‘Baseline Response Levels are a Nuisance in Infant Contingency Learning’, *Infant and Child Development*, 24(5), pp. 506–521. doi: 10.1002/icd.1896.

Milne, S., McDonald, J. and Comino, E. J. (2012) ‘The Use of the Bayley Scales of Infant and Toddler Development III with Clinical Populations: A Preliminary Exploration’, *Physical & Occupational Therapy In Pediatrics*. Taylor & Francis, 32(1), pp. 24–33. doi: 10.3109/01942638.2011.592572.

Miyake, A. and Friedman, N. P. (2012) ‘The Nature and Organisation of Individual Differences in Executive Functions : Four General Conclusions’, *Current Directions in*

Psychological Science, 21(1), pp. 8–14. doi: 10.1177/0963721411429458.The.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. and Wager, T. D. (2000) 'The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis', *Cognitive Psychology*, 41(1), pp. 49–100. doi: 10.1006/cogp.1999.0734.

Moore, T., Hennessy, E. M., Myles, J., Johnson, S. J., Draper, E. S., Costeloe, K. L. and Marlow, N. (2012) 'Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: The EPICure studies', *BMJ (Online)*, 345(7886).

Moore, T., Johnson, S., Haider, S., Hennessy, E. and Marlow, N. (2012) 'Relationship between test scores using the second and third editions of the bayley scales in extremely preterm children', *Journal of Pediatrics*, 160(4), pp. 553–558. doi: 10.1016/j.jpeds.2011.09.047.

Mulder, H., Pitchford, N. J., Hagger, M. S. and Marlow, N. (2009) 'Development of executive function and attention in preterm children: a systematic review', *Dev Neuropsychol*. 2010/02/26, 34(4), pp. 393–421. doi: 10.1080/87565640902964524.

Mulder, H., Pitchford, N. J. and Marlow, N. (2010) 'Processing speed and working memory underlie academic attainment in very preterm children', *Archives of disease in childhood. Fetal and neonatal edition*. 2010/05/22, 95(4), pp. F267-72. doi: 10.1136/adc.2009.167965.

Mulder, H., Pitchford, N. J. and Marlow, N. (2011a) 'Inattentive behaviour is associated with poor working memory and slow processing speed in very pre-term children in middle childhood', *Br J Educ Psychol*. 2011/03/12, 81(Pt 1), pp. 147–160. doi: 10.1348/000709910x505527.

Mulder, H., Pitchford, N. J. and Marlow, N. (2011b) 'Processing speed mediates executive function difficulties in very preterm children in middle childhood', *J Int Neuropsychol Soc*. 2011/03/29, 17(3), pp. 445–454. doi: 10.1017/s1355617711000373.

Musiek, F. E. (1986) 'Neuroanatomy, neurophysiology, and central auditory assessment. Part III: Corpus callosum and efferent pathways', *Ear Hear*. 1986/12/01, 7(6), pp. 349–358.

Musiek, F. E., Kurdziel-Schwan, S., Kibbe, K. S., Gollegly, K. M., Baran, J. A. and Rintelmann, W. F. (1989) 'The dichotic rhyme task: results in split-brain patients', *Ear Hear*. 1989/02/01, 10(1), pp. 33–39.

Naatanen, R. (2001) 'The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm)', *Psychophysiology*. 2001/04/26, 38(1), pp. 1–21.

Näätänen, R. and Winkler, I. (1999) 'The concept of auditory stimulus representation in cognitive neuroscience.', *Psychological bulletin*, 125(6), pp. 826–859. doi: 10.1037/0033-2909.125.6.826.

Nacher, V., Jaen, J., Navarro, E., Catala, A. and Gonz??lez, P. (2015) 'Multi-touch gestures for pre-kindergarten children', *International Journal of Human Computer Studies*, 73, pp. 37–51. doi: 10.1016/j.ijhcs.2014.08.004.

NICE (2017) *Developmental follow-up of children and young people born preterm: Guidance and guidelines*. NICE.

Nicholls, M. E. (1996) 'Temporal processing asymmetries between the cerebral hemispheres: evidence and implications.', *Laterality*, 1(2), pp. 97–137. doi: 10.1080/713754234.

Noland, J. S., Steven Reznick, J., Stone, W. L., Walden, T. and Sheridan, E. H. (2010) 'Better working memory for non-social targets in infant siblings of children with Autism Spectrum Disorder', *Developmental science*. 2010/02/04. NIH Public Access, 13(1), pp. 244–251. doi: 10.1111/j.1467-7687.2009.00882.x.

Northam, G. B., Liegeois, F., Tournier, J. D., Croft, L. J., Johns, P. N., Chong, W. K., Wyatt, J. S. and Baldeweg, T. (2012) 'Interhemispheric temporal lobe connectivity predicts language impairment in adolescents born preterm', *Brain*. 2012/11/13, 135(Pt 12), pp. 3781–3798. doi: 10.1093/brain/aws276.

Nosarti, C., Giouroukou, E., Healy, E., Rifkin, L., Walshe, M., Reichenberg, A., Chitnis, X., Williams, S. C. R. and Murray, R. M. (2008) 'Grey and white matter distribution in very preterm adolescents mediates neurodevelopmental outcome', *Brain*, 131(1), pp. 205–217. doi: 10.1093/brain/awm282.

Nosarti, C., Giouroukou, E., Micali, N., Rifkin, L., Morris, R. G. and Murray, R. M. (2007) 'Impaired executive functioning in young adults born very preterm', *J Int Neuropsychol Soc*. 2007/05/25, 13(4), pp. 571–581. doi: 10.1017/s1355617707070725.

Nosarti, C., Nam, K. W., Walshe, M., Murray, R. M., Cuddy, M., Rifkin, L. and Allin, M. P. G. (2014) 'Preterm birth and structural brain alterations in early adulthood', *NeuroImage: Clinical*. Elsevier, 6, pp. 180–191. doi: 10.1016/j.nicl.2014.08.005.

Nosarti, C., Rushe, T. M., Woodruff, P. W. R., Stewart, A. L., Rifkin, L. and Murray, R. M. (2004) 'Corpus callosum size and very preterm birth: relationship to neuropsychological outcome', *Brain*. Oxford University Press, 127(9), pp. 2080–

2089. doi: 10.1093/brain/awh230.

Nosarti, C., Rushe, T. M., Woodruff, P. W. R., Stewart, A. L., Rifkin, L. and Murray, R. M. (2004) ‘Corpus callosum size and very preterm birth: Relationship to neuropsychological outcome’, *Brain*. Oxford University Press, 127(9), pp. 2080–2089. doi: 10.1093/brain/awh230.

Nowicka, A. and Tacikowski, P. (2011) ‘Transcallosal transfer of information and functional asymmetry of the human brain’, *Laterality*. 2009/08/07, 16(1), pp. 35–74. doi: 10.1080/13576500903154231.

NPEU (2013) *NPEU IMD Tools*, *Office of National Statistic Indices of Multiple Deprivation 2010*. Available at: https://tools.npeu.ox.ac.uk/imd/ (Accessed: 22 August 2017).

Odd, D., Evans, D. and Emond, A. (2013) ‘Preterm birth, age at school entry and educational performance’, *PLOS ONE*, 8(10). doi: 10.1371/journal.pone.0155157.

Oliver, A., Johnson, M. H., Karmiloff-Smith, A. and Pennington, B. (2000) ‘Deviations in the emergence of representations: a neuroconstructivist framework for analysing developmental disorders’, *Developmental Science*. Blackwell Publishers Ltd, 3(1), pp. 1–23. doi: 10.1111/1467-7687.00094.

Ortiz-Mantilla, S., Choudhury, N., Leevers, H. and Benasich, A. A. (2008) ‘Understanding language and cognitive deficits in very low birth weight children’, *Developmental psychobiology*. 2008/02/21, 50(2), pp. 107–126. doi: 10.1002/dev.20278.

Orton, J. L., Mcginley, J. L., Fox, L. M. and Spittle, A. J. (2015) ‘Challenges of neurodevelopmental follow-up for extremely preterm infants at two years’, *Early Human Development*, 91, pp. 689–694. doi: 10.1016/j.earlhumdev.2015.09.012.

Pannek, K., Scheck, S. M., Colditz, P. B., Boyd, R. N. and Rose, S. E. (2014) ‘Magnetic resonance diffusion tractography of the preterm infant brain: a systematic review’, *Dev Med Child Neurol*. 2013/10/10, 56(2), pp. 113–124. doi: 10.1111/dmcn.12250.

Paterson, S. J., Heim, S., Friedman, J. T., Choudhury, N. and Benasich, A. A. (2006) ‘Development of structure and function in the infant brain: implications for cognition, language and social behaviour.’, *Neuroscience and biobehavioral reviews*. NIH Public Access, 30(8), pp. 1087–105. doi: 10.1016/j.neubiorev.2006.05.001.

Peacock, J. L., Marston, L., Marlow, N., Calvert, S. A. and Greenough, A. (2012) ‘Neonatal and infant outcome in boys and girls born very prematurely.’, *Pediatric research*. Springer Nature, 71(3), pp. 305–10. doi: 10.1038/pr.2011.50.

Pelphrey, K. A. and Reznick, S. J. (2003) ‘Working Memory in Infancy’, *Advances in Child Development and Behavior*, pp. 173–227. doi: 10.1016/S0065-2407(03)31005-5.

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W. and Gore, J. C. (1999) ‘An fMRI study of stroop word-color interference: Evidence for cingulate subregions subserving multiple distributed attentional systems’, *Biological Psychiatry*, 45(10), pp. 1237–1258. doi: 10.1016/S0006-3223(99)00056-6.

Piaget, J. (1954) *The construction of reality in the child*. Basic Books: Routledge.

Pitchford, N. J. and Outhwaite, L. A. (2016) ‘Can Touch Screen Tablets be Used to Assess Cognitive and Motor Skills in Early Years Primary School Children? A Cross-Cultural Study’, *Frontiers in Psychology*. Frontiers, 7, p. 1666. doi: 10.3389/fpsyg.2016.01666.

Posit Science (2014) *Auditory Pathway | BrainHQ from Posit Science*, *Posit Science*. Available at: http://www.brainhq.com/brain-resources/image-gallery/auditory-pathway (Accessed: 26 January 2017).

Posner, M. and Petersen, S. (1990) ‘The attention system of the human brain’, *Annual review of neuroscience*, 13(1), pp. 25–42. doi: 10.1146/annurev.ne.13.030190.000325.

Potharst, E. S., Houtzager, B. A., van Sonderen, L., Tamminga, P., Kok, J. H., Last, B. F. and van Wassenaer, A. G. (2012) ‘Prediction of cognitive abilities at the age of 5 years using developmental follow-up assessments at the age of 2 and 3 years in very preterm children’, *Developmental Medicine & Child Neurology*. Blackwell Publishing Ltd, 54(3), pp. 202–203. doi: 10.1111/j.1469-8749.2011.04212.x.

Pozzetti, T., Ometto, A., Gangi, S., Picciolini, O., Presezzi, G., Gardon, L., Pisoni, S., Mosca, F. and Marzocchi, G. M. (2014) ‘Emerging executive skills in very preterm children at 2 years corrected age: A composite assessment’, *Child Neuropsychology*. 2013/01/31, 20(2), pp. 145–161. doi: 10.1080/09297049.2012.762759.

Putkinen, V., Niinikuru, R., Lipsanen, J., Tervaniemi, M. and Huotilainen, M. (2012) ‘Fast measurement of auditory event-related potential profiles in 2-3-year-olds’, *Developmental Neuropsychology*, 37(1), pp. 51–75.

Ramon-Casas, M., Bosch, L., Iriondo, M. and Krauel, X. (2013) ‘Word recognition and phonological representation in very low birth weight preterms’, *Early human development*, 89(1), pp. 55–63.

Rennie, D. A. C., Bull, R. and Diamond, A. (2004) ‘Executive Functioning in Preschoolers: Reducing the Inhibitory Demands of the Dimensional Change Card

Sort Task', *http://dx.doi.org.libproxy.ucl.ac.uk/10.1207/s15326942dn2601_4*. Lawrence Erlbaum Associates, Inc.

Reynolds, G. D. and Romano, A. C. (2016) 'The Development of Attention Systems and Working Memory in Infancy', *Frontiers in Systems Neuroscience*, 10(March), p. Article 15. doi: 10.3389/fnsys.2016.00015.

Reznick, J. S., Morrow, J. D., Goldman, B. D. and Snyder, J. (2004) 'The onset of working memory in infants', *Infancy*. Blackwell Publishing Ltd, 6(1), pp. 145–154. doi: 10.1207/s15327078in0601_7.

Rickards, A. L., Kitchen, W. H., Doyle, L. W. and Kelly, E. A. (1989) 'Correction of developmental and intelligence test scores for premature birth', *Journal of Paediatrics and Child Health*. Blackwell Publishing Ltd, 25(3), pp. 127–129. doi: 10.1111/j.1440-1754.1989.tb01434.x.

Roebers, C. M. (2017) 'Executive function and metacognition: Towards a unifying framework of cognitive self-regulation', *Developmental Review*, 45, pp. 31–51. doi: 10.1016/j.dr.2017.04.001.

Rose, S. A. (1983) 'Differential rates of visual information processing in full-term and preterm infants', *Child Development*, 54(5), pp. 1189–1198.

Rose, S. A. and Feldman, J. F. (1990) 'Infant cognition: Individual differences and developmental continuities', *Individual differences in infancy: Reliability, stability, prediction.*, pp. 229–245.

Rose, S. A. and Feldman, J. F. (1996) 'Memory and processing speed in preterm children at eleven years: a comparison with full-terms', *Child Dev*. 1996/10/01, 67(5), pp. 2005–2021.

Rose, S. A., Feldman, J. F. and Jankowski, J. J. (2001) 'Attention and recognition memory in the 1st year of life: A longitudinal study of preterm and full-term infants.', *Developmental Psychology*. American Psychological Association, 37(1), pp. 135–151. doi: 10.1037/0012-1649.37.1.135.

Rose, S. A., Feldman, J. F. and Jankowski, J. J. (2002) 'Processing speed in the 1st year of life: a longitudinal study of preterm and full-term infants', *Dev Psychol*. 2002/11/14, 38(6), pp. 895–902.

Rose, S. A., Feldman, J. F. and Jankowski, J. J. (2009) 'Information Processing in Toddlers: Continuity from Infancy and Persistence of Preterm Deficits', *Intelligence*. 2010/02/18, 37(3), pp. 311–320. doi: 10.1016/j.intell.2009.02.002.

Rose, S. A., Feldman, J. F. and Jankowski, J. J. (2011) 'Modeling a cascade of effects:

the role of speed and executive functioning in preterm/full-term differences in academic achievement', *Developmental science*. 2011/09/03, 14(5), pp. 1161–1175. doi: 10.1111/j.1467-7687.2011.01068.x.

Rose, S. A., Feldman, J. F. and Jankowski, J. J. (2012) 'Implications of infant cognition for executive functions at age 11', *Psychol Sci*. 2012/10/03, 23(11), pp. 1345–1355. doi: 10.1177/0956797612444902.

Rose, S. A., Feldman, J. F., Jankowski, J. J. and Van Rossem, R. (2008) 'A Cognitive Cascade in Infancy: Pathways from Prematurity to Later Mental Development', *Intelligence*, 36(4), pp. 367–378. doi: 10.1016/j.intell.2007.07.003.

Rose, S. A., Feldman, J. F., Jankowski, J. J. and Van Rossem, R. (2012) 'Information Processing from Infancy to 11 Years: Continuities and Prediction of IQ.', *Intelligence*. NIH Public Access, 40(5), pp. 445–457. doi: 10.1016/j.intell.2012.05.007.

Ross, G., Boatright, S., Auld, P. A. and Nass, R. (1996) 'Specific cognitive abilities in 2-year-old children with subependymal and mild intraventricular hemorrhage', *Brain Cogn*, 32(1), pp. 1–13. doi: 10.1006/brcg.1996.0054.

Ross, G., Tesman, J., Auld, P. A. M. and Nass, R. (1992) 'Effects of Subependymal and Mild Intraventricular Lesions on Visual Attention and Memory in Premature Infants', *Developmental psychology.* American Psychological Association, 28(6), pp. 1067–1074.

Rovee, C. K. and Rovee, D. T. (1969) ' Conjugate Reinforcement of Infant Exploratory Behavior', *Journal of Experimental Child Psychology*, 8(1), pp. 33–39. doi: http://dx.doi.org/10.1016/0022-0965(69)90025-3.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P. and Posner, M. I. (2004) 'Development of attentional networks in childhood', *Neuropsychologia*, 42(8), pp. 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012.

Rushe, T. M., Rifkin, L., Stewart, A. L., Townsend, J. P., Roth, S. C., Wyatt, J. S. and Murray, R. M. (2001) 'Neuropsychological outcome at adolescence of very preterm birth and its relation to brain structure', *Developmental Medicine and Child Neurology*, 43(4), pp. 226–233. doi: 10.1111/j.1469-8749.2001.tb00194.x.

Saavalainen, P., Luoma, L., Bowler, D., Määttä, S., Kiviniemi, V., Laukkanen, E. and Herrgård, E. (2007) 'Spatial Span in Very Prematurely Born Adolescents', *Developmental Neuropsychology*. Taylor & Francis Group, 32(3), pp. 769–785. doi: 10.1080/87565640701539535.

Saavalainen, P. M., Luoma, L., Laukkanen, E., D, M. B., Maatta, S., Kiviniemi, V. and

Herrgard, E. (2008) ‘School performance of adolescents born preterm: neuropsychological and background correlates’, *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society*. 2007/10/16, 12(3), pp. 246–252. doi: 10.1016/j.ejpn.2007.08.007.

Scerif, G. (2010) ‘Attention trajectories, mechanisms and outcomes: at the interface between developing cognition and environment’, *Developmental Science*. Blackwell Publishing Ltd, 13(6), pp. 805–812. doi: 10.1111/j.1467-7687.2010.01013.x.

Scerif, G., Cornish, K., Wilding, J., Driver, J. and Karmiloff-Smith, A. (2004) ‘Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome’, *Developmental Science*, 7(1), pp. 116–130. doi: 10.1111/j.1467-7687.2004.00327.x.

Scharoun, S. M. and Bryden, P. J. (2014) ‘Hand preference, performance abilities, and hand selection in children.’, *Frontiers in psychology*. Frontiers Media SA, 5, p. 82. doi: 10.3389/fpsyg.2014.00082.

Schwartz, B. B. and Reznick, J. S. (1999) ‘Measuring infant spatial working memory using a modified delayed-response procedure’, *Memory*. 2000/01/25, 7(1), pp. 1–17. doi: 10.1080/741943714.

Scott, S. K. and Wise, R. J. S. (2004) ‘The functional neuroanatomy of prelexical processing in speech perception’, *Cognition*, 92(1–2), pp. 13–45.

Sigman, M., Cohen, S. E., Beckwith, L., Asarnow, R. and Parmelee, A. H. (1991) ‘Continuity in cognitive abilities from infancy to 12 years of age’, *Cognitive Development*, 6(1), pp. 47–57. doi: 10.1016/0885-2014(91)90005-X.

Simms, V., Gilmore, C., Cragg, L., Clayton, S., Marlow, N. and Johnson, S. (2015) ‘Nature and origins of mathematics difficulties in very preterm children: a different etiology than developmental dyscalculia’, *Pediatr Res*. International Pediatric Research Foundation, Inc., 77(2), pp. 389–395. doi: 10.1038/pr.2014.184.

Simms, V., Gilmore, C., Cragg, L., Marlow, N., Wolke, D. and Johnson, S. (2013) ‘Mathematics difficulties in extremely preterm children: evidence of a specific deficit in basic mathematics processing’, *Pediatr Res*. 2012/11/21, 73(2), pp. 236–244. doi: 10.1038/pr.2012.157.

Slater, A. (1997) ‘Can measures of infant habituation predict later intellectual ability?’, *Arch Dis Child*. 1998/03/13, 77(6), pp. 474–476.

De Smedt, B., Noël, M.-P., Gilmore, C. and Ansari, D. (2013) ‘The relationship between symbolic and non-symbolic numerical magnitude processing skills and the typical and atypical development of mathematics: a review of evidence from brain and behavior', *Trends in Neuroscience & Education*, 2, pp. 48–55. doi: 10.1023/A.

Sokhadze, E. M., Casanova, M. F., Casanova, E. L., Lamina, E., Kelly, D. P. and Khachidze, I. (2017) 'Event-related Potentials (ERP) in Cognitive Neuroscience Research and Applications', *NeuroRegulation*, 4(1), pp. 14–27. doi: 10.15540/nr.4.1.14.

Spencer-Smith, M. M., Spittle, A. J., Lee, K. J., Doyle, L. W. and Anderson, P. J. (2015) 'Bayley-III Cognitive and Language Scales in Preterm Children.', *Pediatrics*, 135(5), pp. e1258-65. doi: 10.1542/peds.2014-3039.

Spittle, A. J., Orton, J., Doyle, L. W. and Boyd, R. (2007) 'Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants', *Cochrane Database Syst Rev*. 2007/04/20, (2), p. CD005495. doi: 10.1002/14651858.CD005495.pub2.

Spittle, A. J., Spencer-Smith, M. M., Eeles, A. L., Lee, K. J., Lorefice, L. E., Anderson, P. J. and Doyle, L. W. (2013) 'Does the Bayley-III Motor Scale at 2?years predict motor outcome at 4?years in very preterm children?', *Developmental Medicine & Child Neurology*, 55(5), pp. 448–452. doi: 10.1111/dmcn.12049.

Stechler, G. and Latz, E. (1966) 'SOME OBSERVATIONS ON ATTENTION AND AROUSAL IN THE HUMAN INFANT', *Journal of the American Academy of Child Psychiatry*, 5(3), pp. 517–525. doi: 10.1016/S0002-7138(09)62098-7.

Steele, A., Karmiloff-Smith, A., Cornish, K. and Scerif, G. (2012) 'The Multiple Subfunctions of Attention: Differential Developmental Gateways to Literacy and Numeracy', *Child Development*. Blackwell Publishing Ltd, 83(6), pp. 2028–2041. doi: 10.1111/j.1467-8624.2012.01809.x.

Stefanatos, G. A., Joe, W. Q., Aguirre, G. K., Detre, J. A. and Wetmore, G. (2008) 'Activation of human auditory cortex during speech perception: effects of monaural, binaural, and dichotic presentation', *Neuropsychologia*. 2007/11/21, 46(1), pp. 301–315. doi: 10.1016/j.neuropsychologia.2007.07.008.

Steinmann, S. and Mulert, C. (2012) 'Functional relevance of interhemispheric fiber tracts in speech processing', *Journal of Neurolinguistics*, 25(1), pp. 1–12.

Sugita, K., Iai, M., Inoue, T. and Ohta, R. (1990) 'Normative data and the effect of correction for prematurity on test scores in the psychomotor development of extremely low birthweight infants', *Brain and Development*, 12(3), pp. 334–338. doi: 10.1016/S0387-7604(12)80316-0.

Sun, J. and Buys, N. (2011) 'Prefrontal lobe functioning and its relationship to working memory in preterm infants.', *Child and adolescent health yearbook, 2011*, 4(1), pp. 89–108.

Sun, J. and Buys, N. (2012a) ‘Early executive function deficit in preterm children and its association with neurodevelopmental disorders in childhood: a literature review’, *International journal of adolescent medicine and health*. 2012/11/28, 24(4), pp. 291–299. doi: 10.1515/ijamh.2012.042.

Sun, J. and Buys, N. (2012b) ‘Executive function and its relationship with developmental disorders in preterm born children’, *International journal of adolescent medicine and health*, 24(4), pp. 291–299.

Sun, J., Mohay, H. and O’Callaghan, M. (2009) ‘A comparison of executive function in very preterm and term infants at 8 months corrected age’, *Early human development*. 2008/11/14, 85(4), pp. 225–230. doi: 10.1016/j.earlhumdev.2008.10.005.

Taylor, H. G. and Clark, C. A. C. (2016) ‘Executive function in children born preterm: Risk factors and implications for outcome’, *Seminars in Perinatology*, 40(8), pp. 520–529. doi: 10.1053/j.semperi.2016.09.004.

Taylor, H. G., Klein, N. and Hack, M. (2000) ‘School-age consequences of birth weight less than 750 g: a review and update.’, *Developmental neuropsychology*, 17(3), pp. 289–321. doi: 10.1207/S15326942DN1703_2.

Taylor, H. G., Minich, N., Bangert, B., Filipek, P. A. and Hack, M. (2004) ‘Long-term neuropsychological outcomes of very low birth weight: associations with early risks for periventricular brain insults.’, *Journal of the International Neuropsychological Society : JINS*, 10(7), pp. 987–1004.

Thelen, E., Corbetta, D. and Spencer, J. P. (1996) ‘Development of reaching during the first year: role of movement speed.’, *Journal of experimental psychology. Human perception and performance*, 22(5), pp. 1059–76.

Therien, J. M., Worwa, C. T., Mattia, F. R. and DeRegnier, R.-A. O. (2004) ‘Altered pathways for auditory discrimination and recognition memory in preterm infants.’, *Developmental Medicine and Child Neurology*, 46(12), pp. 816–824. doi: 10.1017/s0012162204001434.

Thompson, D. K., Inder, T. E., Faggian, N., Warfield, S. K., Anderson, P. J., Doyle, L. W. and Egan, G. F. (2012) ‘Corpus callosum alterations in very preterm infants: perinatal correlates and 2 year neurodevelopmental outcomes’, *Neuroimage*. 2011/12/14, 59(4), pp. 3571–3581. doi: 10.1016/j.neuroimage.2011.11.057.

Tideman, E. (2000) ‘Longitudinal follow-up of children born preterm: cognitive development at age 19’, *Early human development*. 2000/06/16, 58(2), pp. 81–90.

Tomé, D., Barbosa, F., Nowak, K. and Marques-Teixeira, J. (2014) 'The development of the N1 and N2 components in auditory oddball paradigms: a systematic review with narrative analysis and suggested normative values', *Journal of Neural Transmission*, pp. 375–391. doi: 10.1007/s00702-014-1258-3.

Tsujimoto, S., Kuwajima, M. and Sawaguchi, T. (2007) 'Developmental Fractionation of Working Memory and Response Inhibition During Childhood', *Experimental Psychology*. Hogrefe & Huber Publishers, 54(1), pp. 30–37. doi: 10.1027/1618-3169.54.1.30.

Turken, A., Whitfield-Gabrieli, S., Bammer, R., Baldo, J. V, Dronkers, N. F. and Gabrieli, J. D. (2008) 'Cognitive processing speed and the structure of white matter pathways: convergent evidence from normal variation and lesion studies', *Neuroimage*. 2008/07/08, 42(2), pp. 1032–1044. doi: 10.1016/j.neuroimage.2008.03.057.

Twomey, D. M., Wrigley, C., Ahearne, C., Murphy, R., de Haan, M. and Marlow, N. (no date) 'Touchscreen technology for early cognitive assessment in children', *In press*.

Vicari, S., Caravale, B., Carlesimo, G. A., Casadei, A. M. and Allemand, F. (2004) 'Spatial working memory deficits in children at ages 3-4 who were low birth weight, preterm infants.', *Neuropsychology*, 18(4), pp. 673–678. doi: 10.1037/0894-4105.18.4.673.

Vohr, B. (2014) 'Speech and language outcomes of very preterm infants', *Seminars in fetal & neonatal medicine*. 2013/11/28, 19(2), pp. 78–83. doi: 10.1016/j.siny.2013.10.007.

Vollmer, B., Lundequist, A., Mårtensson, G., Nagy, Z., Lagercrantz, H., Smedler, A.-C. and Forssberg, H. (2017) 'Correlation between white matter microstructure and executive functions suggests early developmental influence on long fibre tracts in preterm born adolescents.', *PloS one*. Public Library of Science, 12(6), p. e0178893. doi: 10.1371/journal.pone.0178893.

Volpe, J. J. (2009) 'Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances', *Lancet neurology*, 8(1), pp. 110–124. doi: 10.1016/S1474-4422(08)70294-1.

Wainwright-Sharp, J. A. and Bryson, S. E. (1993) 'Visual orienting deficits in high-functioning people with autism', *Journal of Autism and Developmental Disorders*, 23(1), pp. 1–13. doi: 10.1007/BF01066415.

Wass, S., Porayska-Pomsta, K. and Johnson, M. H. (2011) 'Training attentional control in infancy', *Current Biology*, 21(18), pp. 1543–1547.

van de Weijer-Bergsma, E., Wijnroks, L. and Jongmans, M. J. (2008) 'Attention development in infants and preschool children born preterm: a review', *Infant behavior & development*. 2008/02/26, 31(3), pp. 333–351. doi: 10.1016/j.infbeh.2007.12.003.

Wellman, H. M., Cross, D. and Bartsch, K. (1987) 'Infant search and object permanence: a meta-analysis of the A-not-B error', *Monographs of the Society for Research in Child Development*, 51(3), pp. 1–67. doi: 10.1037/026082.

WHO (2012) *Born too soon, The Global Action Report on Preterm Birth*. doi: 10.1186/1742-4755- 10-S1-S1.

Wiebe, S. A., Espy, K. A. and Charak, D. (2008) 'Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure.', *Developmental Psychology*, 44(2), pp. 575–587. doi: 10.1037/0012-1649.44.2.575.

Wilcox, T., Nadel, L. and Rosser, R. (1996) 'Location memory in healthy preterm and full-term infants', *Infant Behavior & Development*, 19(3), pp. 309–323. doi: http://dx.doi.org.ez.statsbiblioteket.dk:2048/10.1016/S0163-6383(96)90031-4.

Wilson-Ching, M., Pascoe, L., Doyle, L. W. and Anderson, P. J. (2014) 'Effects of correcting for prematurity on cognitive test scores in childhood', *Journal of Paediatrics and Child Health*, 50(3), pp. 182–188. doi: 10.1111/jpc.12475.

Wilson, S. L. and Cradock, M. M. (2004) 'Review: Accounting for Prematurity in Developmental Assessment and the Use of Age-Adjusted Scores', *Journal of Pediatric Psychology*. Oxford University Press, 29(8), pp. 641–649. doi: 10.1093/jpepsy/jsh067.

Winawer, J. and Horiguchi, H. (2015) 'Visual system architecture'.

Wolke, D. (1998) 'Psychological development of prematurely born children.', *Archives of disease in childhood*. BMJ Publishing Group, 78(6), pp. 567–70.

Woodman, G. F. (2010) 'A brief introduction to the use of event-related potentials in studies of perception and attention.', *Attention, perception & psychophysics*. NIH Public Access, 72(8), pp. 2031–46. doi: 10.3758/APP.72.8.2031.

Woodward, L. J., Moor, S., Hood, K. M., Champion, P. R., Foster-Cohen, S., Inder, T. E. and Austin, N. C. (2009) 'Very preterm children show impairments across multiple neurodevelopmental domains by age 4 years', *Archives of disease in childhood. Fetal and neonatal edition*. 2009/03/25, 94(5), pp. F339-44. doi: 10.1136/adc.2008.146282.

Zaidel, E. (1986) 'Callosal dynamics and right hemisphere language', *Two*

Hemispheres–One Brain: Functions of the corpus callosum. New York, Praeger, pp. 435–459.

Zelazo, P. D. (2006) 'The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children', *Nat Protoc*. 2007/04/05, 1(1), pp. 297–301. doi: 10.1038/nprot.2006.46.

Zelazo, P. D., Frye, D. and Rapus, T. (1996) 'An age-related dissociation between knowing rules and using them', *Cognitive Development*, 11(1), pp. 37–63. doi: 10.1016/S0885-2014(96)90027-1.

Zelazo, P. D., Reznick, J. S. and Spinazzola, J. (1998) 'Representational flexibility and response control in a multistep multilocation search task', *Dev Psychol*, 34(2), pp. 203–214.

Appendix 1. PDP ethical approval

The UCH Preterm Development Project:
Growing up after extremely preterm birth

Short title: UCH PDP

Protocol v.1.4
25 July 2015

Sponsor: **UCL**

Funding source:

1. **SPARKS**
2. **CHART Studentship**
3. **NIHR Senior Investigator Award** to Neil Marlow

SYNOPSIS

Title	**The UCH preterm Development Project: Growing up after extremely preterm birth**
Acronym	UCH PDP
Chief Investigator	Neil Marlow
Objectives	In very preterm children, compared to children born at full term: 1. How do cortical folding, thickness and connectivity differ at full term equivalent age? 2. Do MRI based indices of brain structure and growth in preterm infants relate to measures of neurological maturity and neurophysiologic functions measured in infancy? 3. How do measures of EEG maturation made in the preterm group relate to neuropsychological and neurophysiological function in infancy? 4. Do measures of MRI structures and neurophysiologic function relate to behavioural measures of executive function over the first year and developmental attainment at 2 years? 5. Do measures of socialisation/communication reflect early indicators of autistic symptoms in very preterm infants? 6. Do measures of neuropsychological function made in infancy relate similarly to MRI structural and spectroscopic findings after neonatal encephalopathy in term babies? (Appendix 2)
Study Configuration	Prospective cohort study
Setting	Neonatal unit and outpatient clinics
Sample size estimate	Prospective cohort study of 50 newborn babies born <32 weeks of gestation, 50 babies following neonatal encephalopathy, and a similar number of term born controls
Number of participants	50 index and up to 50 control babies 50 babies following neonatal encephalopathy
Eligibility criteria	All babies born in UCH <32 weeks of gestation without life threatening congenital malformations and considered likely to survive Babies delivered at term who received care in UCH with neonatal encephalopathy
Description of interventions	Babies will receive 2 MRI scans in the first weeks after birth and at term equivalent age; serial EEG recordings will be made over the preterm period and a range of neurobehavioural tests undertaken over the first year after discharge home
Duration of study	7 years from October 2010
Outcome measures	• Measures of brain growth and metabolite ratios on MRSpectroscopy on MRI

	• EEG maturity • Attention, processing speed, working memory and overall development as infants
Statistical methods	Frequentist statistics: Data will be compared between subgroups and over time using appropriate categorical or continuous statistical methods. Linear regression will be used to adjust associations for confounding variables. Two statistical software packages will be used SPSS and Stata in their most up to date version at the completion of the study.

POSTNATAL BRAIN GROWTH AND EARLY INFANCY OUTCOMES IN VERY PRETERM CHILDREN

BACKGROUND

Birth before 32 weeks of gestation (very preterm birth) is associated with a range of cognitive and learning problems that become more frequent with lower gestations at birth.(1) Despite attempts to ameliorate these impairments using a range of general developmental interventions either no or only small intercurrent effects have been demonstrated with little evidence of lasting benefit.(2-4) As a result of these impairments very preterm children frequently have special educational needs(5) with consequent high societal costs.(6-8) More focussed and targeted interventions are thus required to reduce this individual and societal burden. Recent research has shown that core impairments involve executive function deficits that are age and gestation dependent in their prevalence.(9) From current work at Nottingham (PI: NM), we have shown that verbal processing speed and working memory are independent predictors of cognitive function, behaviour, inattention and educational attainment in very preterm children, explaining much of the effect of prematurity on functional outcomes at 10 years of age (Mulder H PhD Thesis 2009 (submitted)). Working with educationalists from the University of Durham, NM and SJ are developing studies of preschool intervention strategies targeted for very preterm children to determine whether specific changes in the manner of presenting information may ameliorate later educational and behavioural problems in this population. However there is concern that these interventions may be too late. It is clear that we need better markers of function earlier in life in order to identify individuals at greatest risk and commence targeted interventions. We propose to use a dual approach to identifying such markers in infancy combining MR measures of brain development and neuro-behavioural measures of cognitive function in order to develop biomarkers for later impairment that can act as selection criteria for interventional studies and as short-term outcome measures for use in neonatal interventional studies.

MR measures: Brain growth over the period from very preterm birth to full term appears suboptimal compared to babies who deliver at full term. Studies have shown that both overall head growth (10) and brain development in terms of size and complexity(11, 12) are impaired in these children. Global scaling factors appear related to 2 year global developmental outcomes(12) and brain velocity or focussed measures of grey matter thickness and folding appear related to contemporaneous neurobehavioural measures.(13, 14) Associations between perinatal illness and cortical growth(15) and between global developmental outcomes and white matter microstructure(16) or apparent diffusion coefficients(17) have been described. Because on an individual basis global developmental scores are poorly predictive of later outcomes, identification of specific and trackable executive functions would provide a more effective manner of identifying children who are likely to require intervention, thereby increasing the power of future intervention studies.

Changes in the periventricular white matter also may be seen on MRI – termed diffuse excessive high intensity signal intensity (DEHSI)(18, 19) – which are thought to relate to underlying developmental changes in the brain, possibly secondary to impairment in oligodendroglial lineage development.(20) Recently we have used T2 relaxometry to quantify DEHSI in babies <32 weeks of gestation.(21) Among 62 preterm children qualitative DEHSI was associated with higher (i.e. more abnormal) relaxometry values compared to preterms without DEHSI or term control infants; the qT2R values were increased in the posterior white matter (where lesions are associated with worse sensorimotor outcomes) than in frontal or central white matter. Furthermore, apparent diffusion coefficient (ADC) values were also higher in preterm compared to control infants indicating less organized white matter structures (see Table) which been associated with global DQ at 2 years.(17) Two-year follow up of this cohort is under way.

Table: Mean (SD) MRI values in White Matter (WM) in three study groups

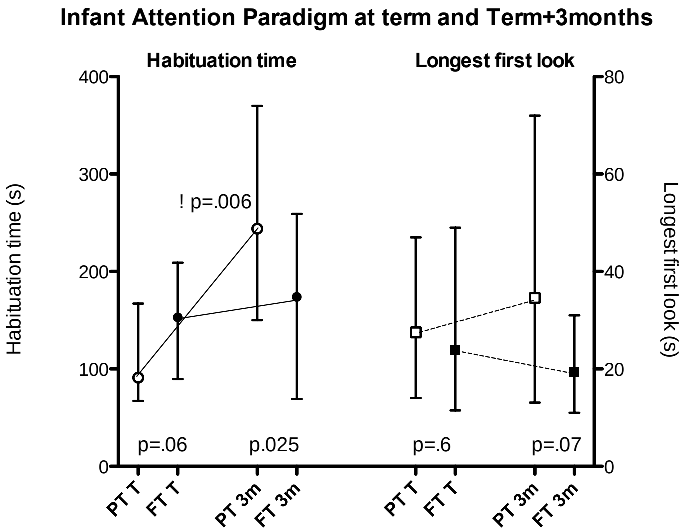
ROI	Control infants (n=7) c	Preterm infants without DEHSI (n=12) [a]	Preterm infants with DEHSI (n=41) [b]	'p' (a v b)	'p' (a v c)	'p' (b v c)
qT2 values						
Frontal WM	222 (25)	237 (34)	260 (28)	*0.03*	*0.24*	*<0.001*
Central WM	212 (26)	214 (26)	248 (38)	*0.02*	*0.84*	*0.01*
Posterior WM	221 (22)	246 (28)	294 (42)	*<0.001*	*0.05*	*<0.001*
ADC values						
Frontal WM	1.46 (0.59)	1.58 (0.14)	1.54 (0.14)	*0.40*	*0.84*	*<0.001*
Central WM	1.45 (0.13)	1.55 (0.17)	1.45 (0.18)	*0.12*	*0.98*	*0.085*
Posterior WM	1.50 (0.09)	1.72 (0.16)	1.54 (0.16)	*0.005*	*0.47*	*<0.001*

Although the specific location of abnormality is of interest, it is possible that lack of integrity anywhere in the white matter tracts will result in later deficits because of the interruption in development of effective neural connections. At the Centre for Medical Image Computing (CMIC) at UCL, we have developed image processing tools to quantify changes in brain structure and development using strategies developed to study conditions of old age where brain tissue is lost.(22-25) Through this application we intend to apply these to the development of the brain in infancy.

Electroencephalography: The EEG of the preterm baby is unique and patterns of electrical activity are seen that mirror the rapid maturational changes taking place in the brain (Boylan & Murray, in Rennie Hagmann and Robertson). The preterm EEG shows a characteristic discontinuous pattern alternating between periods of quiescence and mixed frequency bursting activity. The duration of quiescent periods termed 'interburst intervals' (IBI) are related to the degree of prematurity, with longer IBI seen in babies that are extremely premature. These periods of quiescence decrease in duration as the baby matures and prolonged IBI for gestational age have been associated with abnormal development at 3 years (Holmes et al J Clin Neurophysiol 1993). In addition, patterns such as delta brush activity, temporal saw-tooth activity and frontal sharp transients appear at specific gestational ages and then wane (Vecchierini-Blineau et al Clin Neurophysiol 2007, Brain & Development 2003). State cycling is identifiable in the preterm EEG from around 27 weeks of gestation. and Positive temporal sharp waves have also been found to persist to term in ex-preterm babies (Scher et al EEG Clin Neurophysiol 1994) (Biagioni et al 1999). Disorganised EEG patterns have been described in babies who developed white matter injury of prematurity, as have changes in spectral edge frequency (Hayakawa et al Neuropediatrics 1997; Inder et al Pediatrics 2003) (Kidokoro Pediatrics 2009) (Vermeulen et al Dev Med 2003).

By recording serial multichannel video-EEGs between birth and term, it will be possible to examine the progression of interburst interval duration, emergence and persistence (or not) of gestation-specific patterns such as delta brushes, frontal and temporal sharp transients, and the presence of any injury potentials such as positive Rolandic sharp waves . These data will provide objective measures of the speed of maturation of an individual baby's functional activity and if they prove to be an accurate predictor of later neuropsychological and neurophysiological outcome, they will enable serial EEG monitoring to be used as a basis for future intervention studies.

Behavioural and Neurophysiologic measures: Functional and behavioural correlates of MRI-based indices of abnormality using conventional or novel techniques are required to identify those infants who are at risk of developing specific deficits in the various critical domains of function prior to school entry. We have experience of behavioural (NM, SJ), neurophysiologic (MdH, NM) and psychological testing during infancy (MdH) both in normal babies and in clinical groups including preterms (26-29). Using a visual attention paradigm we have demonstrated differences between very preterm and term infants in habituation time and in the maturation of habituation between full term and 3 months post term (Figure; NM unpublished observations). Such measures can be reliably measured at 3-4 months of age at which time shorter fixation durations and higher shift rates are purported to reflect greater processing speed and more efficient disengagement or shifting of



attention and are thus linked to the development of neural attentional systems (30-32) (33, 34)). Such measures have been shown to underpin the reported relationship between infant visual recognition memory (VRM) and later IQ (35-38) and are thus purported to be stronger predictors of later intellectual outcomes than traditional developmental tests (39-41)).

Later in infancy visual event related potentials (ERP) may be used to evaluate processing of information. One ERP component reliably elicited in visual cognition tasks is believed to reflect attention allocation and is generated in prefrontal and anterior cingulate cortex (reviewed in (26)). The same task can also provide a measure of the functional integrity of other neurocognitive systems: the N290 component is elicited in response to faces and generated in occipito-temporal cortex (42) and the positive slow wave is generated in response to familiar stimuli and believed to reflect memory functions of the temporal lobe (43). In addition, latencies of ERP components provide robust measures of speed of processing.

Impaired executive functions are emergent and already evident in infancy in VP children. Piagetian A-not-B tasks and other delayed-response type tasks can be used in infancy to assess early executive functions such as working memory and inhibitory control.(44) Studies using such tasks have shown that VP children have impaired performance compared with their term counterparts in the first year of life (45) and during the preschool years, particularly for indices of working memory.(44) Furthermore differences in attention and inhibition may persist through to young adulthood.(46) Studies with non-human primates indicate that successful performance on A-not-B and delayed-response tasks involves dorsolateral prefrontal cortex (47) and in normal human infants has been related to frontal cortex EEG signals (48, 49).

Further evidence suggests that very preterm infants show impaired memory on deferred imitation tasks compared to their full-term counterparts, with the degree of impairment related to gestational age at birth.(50) Such tasks rely on the integrity of the medial temporal lobes.(51, 52) Comparison of performance at immediate and delayed recall will give an indication of whether memory impairments in VP infants are isolated to short-term working memory or also encompass long-term memory.

Thus a range of approaches may be taken to determine executive functioning during infancy. The correlation of these emerging executive deficits with MRI-derived variables will help us to identify brain-behaviour relationships so that we can determine the relative value of robust MRI, neurophysiologic and behavioural measures as potential biomarkers. Later follow-up of the assessed population using our routine service arrangements will facilitate the confirmation of association with developmental outcomes using the Bayley Scales (3^{rd} Edition) and, in later funding applications, will facilitate the identification of longitudinal stability in executive functioning in this population and provide a method of refining interventions in infancy. We believe that these investigations will lead to studies of targeted interventions within 3-5 years.

Broad hypotheses to be tested:

In very preterm children, compared to children born at full term:

1. How do cortical folding, thickness and connectivity differ at full term equivalent age?
2. Do MRI based indices of brain structure and growth in preterm infants relate to measures of neurological maturity and neurophysiologic functions measured in infancy?
3. How do measures of EEG maturation made in the preterm group relate to neuropsychological

and neurophysiological function in infancy?

4. Do measures of MRI structures and neurophysiologic function relate to behavioural measures of executive function over the first year and developmental attainment at 2 years?
5. Do measures of socialisation/communication reflect early indicators of autistic symptoms in very preterm infants?

Children recruited to this study will then be entered into our longitudinal outcome evaluation programme alongside children recruited into other perinatal studies, in order to determine the relationship of these infancy measures via research-based studies with later cognitive and executive functional measures.

PLAN OF INVESTIGATION:

Population: Babies born at <32 completed weeks of gestation with comparison children born at full term at University College Hospital London. In the first instance, preterm parents will be approached at the end of the first week after birth and permission sought to include them in the study. Exclusion criteria will be low likelihood of survival and severe congenital abnormality. Term comparison children will be recruited from the postnatal wards. Inclusion criteria for term group: gestation 37-42 weeks, birthweight between 10^{th} and 90^{th} percentile for gestation, no perinatal complications and Apgar score at 5min >7. Additional methods of recruitment will include a flyer briefly detailing the study and project providing contact details. In the case of preterm infants, the flyer will be included in an information pack sent to parents when organising routine medical follow-ups. For full-term infants, the flyer will be distributed via local parent and infant groups and their venues. Participants may also be recruited via email notices. Children will be recruited from parents resident within the North Central London Neonatal Network on the understanding that follow up evaluations and a MRI scan at full term are carried out at UCH. This is our current practice for Network-based follow up. No alteration to clinical care will be necessary as part of the project other than the two imaging procedures and EEG recordings.

MRI Scanning: Two MRI scans will be carried out using the 1.5T Siemens MR scanner at UCH. The first will be at a point where the child is physiologically stable and the equivalent of 30 weeks of gestation or two weeks after birth which ever is the later. Infants will be transported to the MRI (200y from the Neonatal Unit) using a MR-compatible incubator, with integral monitoring, ventilation/CPAP facility and head coil in situ (Lammers GmbH). Infants will be stabilised in the Neonatal Unit before transport and not disturbed during the scan process away from the unit, unless necessary. Scans are supervised by one of two "imaging" fellows with specific responsibility (trainee neonatologists) and a senior neonatal nurse (usually an ANNP). The scanning protocol will comprise:

- High-resolution T1w anatomical scan (MP-RAGE) - T1 values of T1(GM, 3T)=±1900ms, and T2(WM, 3T)=±2100ms for grey and white matter, respectively will be initially considered in order to optimize contrast.
- High resolution quantitative T1 and T2 measurements using the multiple saturation technique and a high-resolution 3D-CPMG-based methods respectively, the latter also providing T2 weighted anatomical scans.(53)
- High-resolution 3D-DTI will be used to assess the early maturation process in white matter fibres (54) with a second DTI scan using a high angular resolution diffusion to allow neurite orientation dispersion and density imaging (NODDI) (56)
- 3D-MTR scan will be acquired using a standard 3D-FLASH sequence in combination with an off-resonance pre-pulse to estimate the degree of myelination and correlate these measures with cortical development.
- Single voxel point-resolved spectroscopy (PRESS) to derive proton and phosphorus spectroscopy

The combination of multiple scans included in a modern clustering algorithms will allow clear distinction of the cortical layers in these children, as conventional segmentation algorithms might fail since they are not optimized for the type of image contrast apparent in this patient population.(55) The complete MRI protocol should not take more than 60min using modern parallel imaging techniques.

The analyses that we propose comprise three major components: cortical thickness estimation methods; connectivity matrix estimation; and correlation of the cortical thickness and the connectivity over time. In our previous work, we have demonstrated that we could estimate accurately the cortical thickness of controls and aging population. CMIC is uniquely placed to develop these methods further using its unique set of analysis tools in non-rigid registration,(24) diffusion image analysis,(25) tissue classification(22)] and cortical thickness estimation.(23) Some of these tools have been provided to the community through open-source projects and have been widely used, such as CAMINO (www.cs.ucl.ac.uk/research/medic/camino) and our newly released NiftyReg (http://sourceforge.net/projects/niftyreg).

EEG studies: A 8-channel EEG and video will be collected at weekly intervals over the first three weeks starting as soon as is feasible within the first week with one recording on a day adjacent to the first MRI scan. The neonatal EEG set up is complete in about 30 minutes and space is left for cerebral ultrasound scanning through the anterior fontanelle. The adapted hat has been in use for some time now and acts as a robust mount for our endotracheal tube holder system or as a CPAP mount system. Each EEG will be recorded for two hours to ensure state cycling in the EEG is captured but tracing will where possible carry on between nursing “cares” to minimise the disruption to the baby. A final EEG will be recorded for 2 hours on the same day as the Term MRI scan is done. Control children will have a single 2h EEG recorded.

The EEG trace is digitised and is automatically uploaded onto a central server with anonymised identifiers for confidentiality. Current work (REC Application:) is developing automatic seizure detection software and video-EEG is routinely in use on the neonatal unit. Working in collaboration with colleagues from the University College Cork (Dr Geraldine Boylan) and a second group at UCL/University of Sussex (Drs Simon Farmer and Luc Berthouse, respectively) we wish to develop new markers of EEG maturation that can be used alongside other experimental measures to build up a unique picture of preterm brain development. These colleagues will only work from anonymised EEG records with simple clinical data – sex of the baby, gestation at birth, and postnatal age for the purpose of software development. In turn their summary data will be merged with an anonymised study database for later formal analysis alongside other study measures.

We will commence this additional investigation after we have established routine recruitment into the first part of the study at a later date.

Feedback to parents: All MRI scans and EEG traces are reported by a clinical radiologist/neurophysiologist respectively. Feedback of relevant clinical information to parents occurs as part of the routine clinical care and is led by the attending neonatologist; this is recorded in the clinical case-record. Feedback of the term scan findings, where the baby will usually be an outpatient, is done by the Imaging fellow or consultant at the next available outpatient appointment as appropriate.

Clinical Assessment: Routine cerebral ultrasound scanning will be carried out as per protocol and perinatal/neonatal data abstracted from the clinical record. The imaging fellow will be trained and responsible for carrying out a NAPI neuro-behavioural assessment at 36w postmenstrual age (pma) and a formal neurological assessment (Amiel-Tison) at 36w and 40w pma. The postdoctoral psychologist (and SJ) will assess function over the first year based in our new baby lab (dedicated space within the Clinical Research Facility (ground floor EGA wing, UCH)). Because it is unclear as to which tests are likely to be the most predictive in this group we will take a range of age-appropriate approaches to function and behaviour integrated through the first year. The full test battery for preterm children is shown in the table below. At 3m general movements (routine) and Visual attention tests (experimental) are video recorded and checked/scored off line. We will also video the delayed response task at 6m and all 3 tests at 12 months. All video records will be destroyed 3 years after the conclusion of the study following write up and publication unless the parent gives explicit permission for their use for further research or teaching.

Outcome	Test	Age (Corrected)		Examiner	Time	Location
Neuro-behaviour	NAPI	36w pma		Clinician	40min	NICU
Neurological integrity	Amiel-Tison	36w/40w pma	+	Clinician	15min	NICU
	General Movements	3m		Clinician	15min	Clinic*
	Flash VEP	40w pma & 6m	+	Psychologist	15min	Baby lab
Visual function	Clinical test of vision	40w pma	+	Psychologist	5min	Baby lab
Neuropsychology	Visual attention test	3m	+	Psychologist	20minζ	Baby Lab
(Memory, attention, inhibition, planning and/or processing speed)	ERP	6m	+	Psychologist	15 min	Baby Lab
	Delayed-response task	6m	+	Psychologist	10 min	Baby Lab
	A-not-B task	12m	+	Psychologist	5min	Baby Lab
	Means-End task	12m	+	Psychologist	5min	Baby Lab
	Deferred Imitation task	12m	+	Psychologist	10min	Baby Lab
Developmental status	Bayley III	6 & 12m	+	Clinician	25/45m	Clinic* or Home-visit
Language/motor skills	Questionnaire	12m	+	Parent	20min	Baby Lab
Perinatal data & SES	Questionnaire	Discharge	+	Parent	10min	NICU
	Clinical Record Form	Discharge		Clinician	20min	NICU

*Routine follow up information; + assessments for term children; ζ Test itself takes only ~5 min but time quoted includes that for settling the baby and waiting for a suitable state to perform the test.

In addition measures of socialisation/communication will be made as detailed in Appendix 1 as part of a nested study to evaluate the appearance of abnormal social interaction reported frequently in this population at 2 years.

Parents will be approached for consent during the first week after birth. A second recruitment point into the clinical psychological evaluation will also be possible to boost numbers for the psychological and socialisation studies. Parents will be approached in outpatients to collect initial pilot data at 3m age, and also at term equivalent age, following a single MRI scan, which is routine practice for clinical prognostic purposes. For this a slightly modified PIL and consent for will be used.

Feasibility/facilities: We have reserved time for research and clinical MRI scans and are purchasing the MR-compatible incubator currently. Space is available in the CLRN-sponsored EGA wing Clinical Research Facility and we have appropriate neurophysiology equipment installed there. We have experience in making reliable and repeatable neurophysiologic measures (MdH) and behavioural assessments (SJ, NM) in young infants and have performed longitudinal cohort studies with good family retention. Indeed the UCH very preterm longitudinal cohort from the 1980s is still under investigation by the Institute of Psychiatry, with excellent retention over many years. We perform clinical and research based MRI regularly, having two clinical fellows and nurse practitioners dedicated to staffing our MRI sessions.

Outputs:

These data will be presented at national and international paediatric, radiological and psychological research meetings and written up for publication in peer-reviewed journals.

Study written and electronic records

- Each participant will be assigned a study identity code number, for use on all paper records, other study documents and the electronic database. The documents and database will also use the date of birth as a second identifier.
- All written records will be treated as confidential documents and held securely in accordance with regulations. The investigator will make a separate confidential record of the participant's name, date of birth, local hospital number or NHS number, and Participant Study Number, to permit identification of all participants enrolled in the study, for the purposes of later follow-up.
- All paper forms will be completed using black ballpoint pen. Errors shall be lined out but not obliterated by using correction fluid and the correction inserted, initialled and dated.
- The person completing each paper form shall sign and date each form.

Quality assurance & audit

Insurance and indemnity

- Insurance and indemnity for clinical study participants and study staff is covered within the NHS Indemnity Arrangements for clinical negligence claims in the NHS, issued under cover of HSG (96)48. There are no special compensation arrangements, but study participants may have recourse through the NHS complaints procedures.
- University College London has taken out an insurance policy to provide indemnity in the event of a successful litigious claim for proven non-negligent harm.

Conduct of the Study

- Study conduct will be subject to systems audit for inclusion of essential documents; permissions to conduct the study; CVs of study staff and training received; local document control procedures; consent procedures and recruitment logs; adherence to procedures defined in the protocol (e.g. inclusion / exclusion criteria, timeliness of visits); and accountability of study materials.
- The Study Manger, or where required, a nominated designee of the Sponsor, shall carry out a site systems audit at least yearly and an audit report shall be made.
- Monitoring of study data shall include confirmation of informed consent; source data verification; data storage and data transfer procedures; local quality control checks and procedures, back-up and disaster recovery of any local databases and validation of data manipulation. The Study manager, or where required, a nominated designee of the Sponsor, shall carry out monitoring of study data as an ongoing activity.
- Data will be effectively double entered using SPSS Data Entry continuous comparison techniques to ensure accurate electronic data records
- Study data and evidence of monitoring and systems audits will be made available for inspection by the REC as required.

Data Management and analysis

- Access to all study documents is limited to the study personnel (see below) excepting that the CRF and all source documents will be made be available at all times for review by the Chief Investigator, Sponsor's designee and inspection by relevant regulatory authorities.
- All study staff and investigators will endeavour to protect the rights of the study's participants to privacy and informed consent, and will adhere to the Data Protection Act, 1998. The Assessment procedures will only collect the minimum required information for the purposes of the trial. All records will be held securely, in a locked room, and a locked cabinet. Access to the information will be limited to the trial staff and investigators and any relevant regulatory authorities. Computer held data including the study database will be held securely and password protected. All data will be stored on a secure dedicated web server. Access will be restricted by user

identifiers and passwords (encrypted using a one way encryption method).

- Information about the study in the participant's medical records / hospital notes will be treated confidentially in the same way as all other confidential medical information.
- Two postdoctoral fellows (physics and computing and psychology, respectively) will carry out data analysis under supervision from the co-investigators. Where off-line analysis of data is undertaken these data are managed anonymously using a unique study identifier. Summary data are then returned to the main study database, which is again anonymised to maintain confidentiality.
- Electronic data will be backed up every 24 hours to both local and remote media in encrypted format.
- Data will be entered onto study forms and posted back to the study centre where they will be encoded for computer analysis using SPSS Data Entry and SPSS for Windows. Data will be checked in real time using the facility built into the Data Entry Module. Once clean the database will be combined with data from the 1995 cohort study for analysis.
- Data will be analysed using appropriate categorical and continuous comparisons using SPSS and STATA statistical software packages. Major outcomes will be examined for known explanatory variables (sex of the child, gestational age, multiple birth) and regression analyses will be performed to look for antecedents and associates of good and poor outcome.
- Most very preterm babies are successfully scanned at full term for clinical purposes within our unit. We anticipate being able to recruit 50 infants <32w gestational age over a 12-month period at UCH from a population of 100 babies and have funding to do this. For two-tailed comparisons of preterm and term born children on continuous measures with normal distributions differences of 0.6sd will be detected with 90% power at 5% significance or a 4 fold increase in the proportion of children with abnormal developmental scores in the preterm compared to the term group (16% versus 5%), which is comparable to other published studies. For within the preterm group correlations of scan findings with neuropsychological outcomes the power is more difficult to predict but given paired data we anticipate that the power will be sufficient.
- Electronic data will be archived at the completion of the study for a period of 25 years to allow for later follow up of this unique population. Paper records will be kept for 10 years after the last assessment is carried out on the population.

References:

1. Marlow N. Neurocognitive outcome after very preterm birth. Arch Dis Child Fetal Neonatal Ed. 2004 May;89(3):F224-8.
2. Spittle AJ, Orton J, Doyle LW, Boyd R. Cochrane Database Syst Rev. 2007(2):CD005495.
3. Avon Premature Infant Project. Arch Dis Child Fetal Neonatal Ed. 1998 Jul;79(1):F4-11.
4. Johnson S, Fawke J, Hennessy E, Rowell V, Thomas S, Wolke D, et al. Pediatrics. 2009 Aug;124(2):e249-57.
5. Johnson S, Hennessy E, Smith R, Trikic R, Wolke D, Marlow N. Arch Dis Child Fetal Neonatal Ed. 2009 Jul;94(4):F283-9.
6. Mangham LJ, Petrou S, Doyle LW, Draper ES, Marlow N. Pediatrics. 2009 Feb;123(2):e312-27.
7. Petrou S, Abangma G, Johnson S, Wolke D, Marlow N. Value Health. 2009 Jul 29.
8. Petrou S, Henderson J, Bracewell M, Hockley C, Wolke D, Marlow N. Early Hum Dev. 2006 Feb;82(2):77-84.
9. Mulder H, Pitchford NP, Marlow N. Developmental Neuropsychology. 2009;34(4):393-421
10. Wood NS, Costeloe K, Gibson AT, Hennessy E, Marlow N, Wilkinson AR. Archives of Disease in Childhood. 2003
11. Inder TE, Warfield SK, Wang H, Huppi PS, Volpe JJ. Pediatrics. 2005 Feb;115(2):286-94.
12. Kapellou O, Counsell SJ, Kennea N, Dyet L, Saeed N, et al. PLoS Med 2006;3(8):e265.
13. Tolsa CB, Zimine S, Warfield SK, Freschi M, Sancho Rossignol A, Lazeyras F, et al. Pediatr Res. 2004 Jul;56(1):132-8.

14. Dubois J, Benders M, Borradori-Tolsa C, Cachia A, Lazeyras F, Ha-Vinh Leuchter R, et al. Brain. 2008;131:2028-41.
15. Kaukola T, Kapellou O, Laroche S, Counsell SJ, Dyet LE, Allsop JM, et al. Acta Paediatr. 2009 Jun;98(6):990-5.
16. Counsell SJ, Edwards AD, Chew AT, Anjari M, Dyet LE, Srinivasan L, et al. Brain. 2008 Dec;131(Pt 12):3201-8.
17. Krishnan ML, Dyet LE, Boardman JP, Kapellou O, Allsop JM, Cowan F, et al. Pediatrics. 2007 Sep;120(3):e604-9.
18. Counsell SJ, Allsop JM, Harrison MC, Larkman DJ, Kennea NL, Kapellou O, et al. Pediatrics. 2003 Jul;112(1 Pt 1):1-7.
19. Counsell SJ, Shen Y, Boardman JP, Larkman DJ, Kapellou O, Ward P, et al. Pediatrics. 2006 Feb;117(2):376-86.
20. Riddle A, Luo NL, Manese M, Beardsley DJ, Green L, Rorvik DA, et al. J Neurosci. 2006 Mar 15;26(11):3045-55.
21. Hagmann CF, De Vita E, Bainbridge A, Gunny R, Kapetanakis AB, Chong WK, et al. Radiology. 2009 Jul;252(1):209-17.
22. Acosta O, Bourgeat P, Fripp J, Bonner E, Ourselin S, Salvado O. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv. 2008;11(Pt 1):253-61.
23. Acosta O, Bourgeat P, Zuluaga MA, Fripp J, Salvado O, Ourselin S. Medical Image Analysis. 2009 October;13(5):730-43.
24. Modat M, Ridgway GR, Taylor ZA, Lehmann M, Barnes J, Hawkes DJ, et al. Comput Methods Programs Biomed. 2009 Oct 7.
25. Rohrer JD, Warren JD, Modat M, Ridgway GR, Douiri A, Rossor MN, et al. Neurology. 2009 May 5;72(18):1562-9.
26. de Haan M. Visual attention and recognition memory in infancy. New York, NY: Psychology Press; 2007.
27. Pike AA, Marlow N. Early Hum Dev. 2000 Feb;57(2):123-35.
28. Johnson S, Whitelaw A, Glazebrook C, Israel C, Turner R, White IR, et al. J Pediatr. 2009 Oct;155(4):488-94.
29. Glazebrook C, Marlow N, Israel C, Croudace T, Johnson S, White IR, et al. Arch Dis Child Fetal Neonatal Ed. 2007 Feb 15.
30. Colombo J, Shaddy DJ, Richman WA, Maikranz JM, Blaga OM. Infancy. 2004;5(1):1 - 38.
31. Colombo J, Mitchell DW. In: Colombo J, Fagen J (ed) Individual differences in infancy. Hillsdale: Lawrence Erlbaum 1990.
32. Colombo J, Mitchell DW. Infant visual habituation. Neurobiology of Learning and Memory. 2009;92:225-34.
33. Frick JE, Colombo J, Saxon TF. Child Development. 1999 May-Jun;70(3):537-48.
34. Rose SA, Feldman JF, Jankowski JJ. Developmental Psychology. 2001 Jul;37(4):539-49.
35. Kavsek M. Journal of Applied Developmental Psychology. 2004 Aug;25(3):369-93.
36. Rose SA, Feldman JF, Jankowski JJ. A Developmental Psychology 2001 Jan;37(1):135-51.
37. Slater A, Cooper R, Rose D, Morison V. Human Development. 1989 May-Jun;32(3-4):137-47.
38. van de Weijer-Bergsma E, Wijnroks L, Jongmans MJ. Infant Behaviour and Development. 2008;31:333-51.
39. Rose SA, Feldman JF, Jankowski JJ. Intelligence. 2009;37:311-20.
40. Rose SA, Jankowski JJ, Feldman JF, Rose SA, Jankowski JJ, Feldman JF. Infant Behavior & Development.2008;31:470-80.
41. Slater A. Individual differences in infancy and later IQ. J Child Psychol Psychiatry. 1995 Jan;36(1):69-112.
42. Johnson MH, Griffin R, Csibra G, Halit H, Farroni T, de Haan M, et al. Dev Psychopathol. 2005 Summer;17(3):599-619.

43. Reynolds GD, Richards JE. Developmental Psychology. 2005;41:598-615.
44. Espy KA, Stalets MM, McDiarmid MM, Senn TE, Cwik MF, Hamby A. Child Neuropsychology. 2002 Jun;8(2):83-92.
45. Sun J, Mohay H, O'Callaghan M. Early Hum Dev. 2009 Apr;85(4):225-30.
46. Lawrence EJ, Rubia K, Murray RM, McGuire PK, Walshe M, Allin M, et al. Hum Brain Mapp. 2009 Mar;30(3):1038-50.
47. Diamond A. Ann N Y Acad Sci. 1990;608:267-309; discussion -17.
48. Fox NA, Bell MA. Ann N Y Acad Sci. 1990;608:677-98; discussion 98-704.
49. Bell MA, Fox NA. Child Dev. 1992 Oct;63(5):1142-63.
50. de Haan M, Bauer PJ, Georgieff MK, Nelson CA. Dev Med Child Neurol. 2000 May;42(5):304-12.
51. McDonough L, Mandler JM, McKee RD, Squire LR. Proc Natl Acad Sci U S A. 1995 Aug 1;92(16):7580-4.
52. Adlam A-LR et al J Cogn Neurosci. 2005; 17: 240-8.
53. Georgiades CS, Itoh R, Golay X, van Zijl PC, Melhem ER. AJNR Am J Neuroradiol. 2001 Oct;22(9):1732-7.
54. Golay X, Jiang H, van Zijl PC, Mori S. Magn Reson Med. 2002 May;47(5):837-43.
55. E.R. Melhem and X. Golay in Clinical MR Neuroimaging: Diffusion, Perfusion and Spectroscopy, J Gillard, A Waldman, PB Barker (eds), Cambridge University Press, Cambridge, UK. 2005: 647-673
56. Jiang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 2012 (e-pub ahead of print doi: 10.1016/j.neuroimage.2012.03.072)

Protocol appendix 1: Additional Measures for Social Communication Development

Many children born very preterm experience difficulties with social communication and peer relationships throughout childhood.[1, 2] Very preterm infants are also at an increased risk of developing an autism spectrum disorder and related symptoms.[3-5] These are disorders characterised by abnormalities of social interaction and communication, presenting alongside repetitive behaviours and restricted interests (APA, 2001). Children with these difficulties require highly targeted interventions, which are most successful when commenced early in life.[6] Recent work with other high-risk populations has indicated various possible early behavioural and brain markers for abnormalities in social development. These markers have not been examined in preterm populations, but could help us to identify infants at greatest risk and commence targeted interventions as early as possible.

EEG Measures: For both VP and term controls (at 6m), we wish to measure the neural correlates of eye-gaze processing using a visual event-related potential (ERP) paradigm (32-electrode Geodesic Sensor Net). In the task, which will be completed in the same ERP session as the information processing paradigm (see main Protocol), infants will be presented with images of female faces, with either direct or averted gaze. Infants will continue with trials as long as they show interest in looking (max. 150 trials, approx. 10min).

Prior research using this paradigm has shown that direct gaze elicits a larger negativity (N170) than averted gaze in infants as young as 4 months,[7, 8] particularly over mid-line scalp/occipital regions. This early sensitivity to direct eye contact is thought to be crucial to infants' adaptive social and communication development. Research has demonstrated abnormalities in the ERP responses to direct and averted gaze in certain clinical and at-risk-groups (e.g. prolonged latency of the occipital P400 in children with autism and in their infant-siblings). This task will therefore allow us to identify atypical features of eye-gaze processing within the first year of life in VP infants.

Eye-tracking Measures: We wish to measure looking behaviour of VP and term controls (at 6 and 12 months) whilst viewing a range of visual displays of social and non-social stimuli. Looking behaviour (i.e. location of first "looks" and looking time to elements of displays) will be recorded using a remote, non-intrusive Tobii corneal reflection eye-tracker (T60/T120), whilst seated on their parent's lap. Measures will include looking behaviour whilst viewing faces and eye-gaze cues, visual preference for social over non-social stimuli, and ability to disengage and shift visual attention between objects. In total, this series of tasks will last no longer than 15 minutes (including the eye-tracker calibration).

Abnormal looking-responses to faces and eye-gaze cues, and impaired disengagement of attention, have been observed in infants at-risk of developing disorders of social communication. These paradigms will allow us to examine visual scanning behaviours and preferences in this population – and with longitudinal follow-up, allow us to relate these findings to longer-term neurodevelopmental outcomes.

Observational Assessments: For both VP and term controls, we wish to administer the Autism Observation Scale for Infants (AOSI) at 6 and 12 months. The AOSI is an 18-item direct observational measure designed to detect and monitor putative signs of autism in infants aged 6–18 months. It is a semi-structured play-based assessment, in which systematic presses are used to elicit target behaviours. Target behaviours include visual tracking and attentional disengagement, coordination of eye gaze and action, imitation, early social-affective and communicative behaviours, behavioural reactivity, and various sensory-motor behaviours. The AOSI is conducted at a small table, with the infant seated opposite the examiner, on his/her parent's lap. The assessment will be videoed for rating and takes approximately 15-20 min to administer.

Following the AOSI, a 10-minute free-play session of parent and infant will be video-recorded (for both VP and term controls). Play interactions will be coded blind to group (FT/PT), according to a validated global rating scale.

Additional Questionnaire Measures: Parents will be asked to complete the Infant Behaviour Questionnaire (IBQ-R)[9] to obtain a measure of infant temperament at both time points.

References

1. Hoy EA, Sykes DH, Bill JM, Halliday HL, McClure BG, Reid MM. The social competence of very-low-birthweight children: teacher, peer, and self-perceptions. Journal of abnormal child psychology. 1992; **20**(2): 123-50.
2. Johnson S, Marlow N. Positive screening results on the modified checklist for autism in toddlers: implications for very preterm populations. J Pediatr. 2009; **154**(4): 478-80.
3. Johnson S, Hennessy E, Hollis C, Kochhar P, Wolke D, Marlow N. Screening for Autism Spectrum Disorders in Extremely Preterm Children. Acta Paediatrica. 2009; **98**: 163-.
4. Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Autism spectrum disorders in extremely preterm children. J Pediatr. 2010; **156**(4): 525-31 e2.
5. Moore T, Johnson S, Hennessy E, Marlow N. Screening for autism in extremely preterm infants: problems in interpretation. Dev Med Child Neurol. 2011; **Resubmission under reveiw**.
6. Yirmiya N, Charman T. The prodrome of autism: early behavioral and biological signs, regression, peri- and post-natal development and genetics. Journal of child psychology and psychiatry, and allied disciplines. 2010; **51**(4): 432-58.
7. Farroni T, Csibra G, Simion F, Johnson MH. Eye contact detection in humans from birth. Proceedings of the National Academy of Sciences of the United States of America. 2002; **99**(14): 9602-5.
8. Farroni T, Johnson MH, Csibra G. Mechanisms of eye gaze perception during infancy. Journal of cognitive neuroscience. 2004; **16**(8): 1320-6.
9. Gartstein MA, Rothbart MK. Studying infant temperament via the revised infant behavior questionnaire. [References]. Infant Behavior & Development. 2003; **26**(1): 64-86.

Protocol appendix 2a: Inclusion of babies recruited to longitudinal studies following Neonatal Encephalopathy entered into the Baby Brain Study to PDP neuropsychology protocol only.

Long term outcome following perinatal hypoxic-ischaemic encephalopathy (HIE) is increasingly recognised to be complex and associated with a range of only recently identified and subtle defects of working memory and executive function, which combine to produce a significant rate of school failure (figures 1 & 2) [1-3]. Because of the relatively infrequent occurrence of intrapartum asphyxia leading to encephalopathy (approximately 1.5-2.0 per 1000 live births) little attention has been directed at early detection and intervention in this group, aside from early recognition of signs of cerebral palsy using gestalt recognition of abnormal general movements.[4] Among older children acute hypoxia is recognised to cause a reduction in hippocampal size and subsequent functional deficits in working memory.[5]

Fig 1: Mean differences in Standardised Scores between Children with Moderate NE - with (a) less severe and (b) more complex clinical courses and classmates

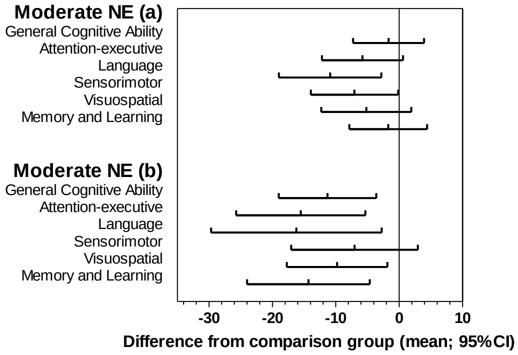
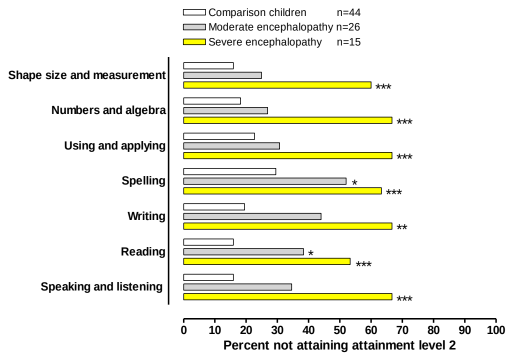


Fig2: Proportions of children with moderate NE failing to attain Key Stage 2, compared to classmates



Based on our current experience we hypothesise that among infants who have had neonatal encephalopathy, and do not go onto to develop severe cerebral palsy:

1. Measures of processing speed and working memory obtained in early infancy (3-12m) will predict poor performance on developmental tests at 2 years of age
2. Measures of socialisation made in early infancy will predict behavioural outcomes at 2 years
3. Neuropsychological measures will relate to structural brain injury evaluated on MRI, metabolic changes assessed using MR spectroscopy and background EEG in the neonatal period.
4. Commonalities in infancy assessments are present with preterm development in infancy reflecting core neuropsychological processes in infancy.

We are currently recruiting babies with encephalopathy to one randomised trial of Xenon neuroprotection (TOBY-Xenon – www.npeu.ox.ac.uk/tobyxenon; PI Robertson) and one longitudinal cohort study, the Babies Brains Study (PI Robertson). All studies include long-term follow up using the UCH-based follow up team.

We propose to recruit families of babies who have been entered into the longitudinal studies above to our BabyLab protocol, used in the UCH Preterm Development Project. Key to this project is the determination of infant measures of cognitive function, memory and socialisation at 3, 6 and 12 months after birth. We will invite parents to enter this study at discharge from hospital so that it will not compromise recruitment to these ongoing studies and we will carry out the testing alongside the routine follow up provided at UCH for these children, which also doubles as outcome evaluation at 2 years of age using a neurological assessment and the Bayley 3 Scales of Infant and Toddler Development. The advantages for this project will be correlation with high resolution MRI imaging, spectroscopy and automated EEG analysis carried out as part of these studies and determination of outcomes at 2 years. We believe that this is a unique study in the area and will provide novel data to enhance the other ongoing studies.

Population: We provide a regional service for therapeutic hypothermia as treatment for HIE and cool approximately 50 babies per year. Some of these will be referred from remote units but we anticipate approximately 20 babies will be available for this study each year. We propose to run this protocol alongside our current control group, who will act as a reference group. Recruiting over two years will result in approximately the same sample size as for the main PDP study and thus has similar power to detect differences as in the main protocol. These are exploratory studies and formal power calculation is thus not as yet possible.

Additional references for protocol appendix 2

1. Marlow N, Rose AS, Rands CE, et al. Neuropsychological and educational problems at school age associated with neonatal encephalopathy. Archives of disease in childhood. Fetal and neonatal edition 2005;90(5):F380-7.
2. Robertson CM, Finer NN, Grace MG. School performance of survivors of neonatal encephalopathy associated with birth asphyxia at term. The Journal of pediatrics 1989;114(5):753-60.
3. de Vries LS, Jongmans MJ. Long-term outcome after neonatal hypoxic-ischaemic encephalopathy. Archives of disease in childhood. Fetal and neonatal edition 2010;95(3):F220-4.
4. Prechtl HF, Ferrari F, Cioni G. Predictive value of general movements in asphyxiated fullterm infants. Early human development 1993;35(2):91-120.
5. Isaacs EB, Vargha-Khadem F, Watkins KE, et al. Developmental amnesia and its relationship to degree of hippocampal atrophy. Proceedings of the National Academy of Sciences of the United States of America 2003;100(22):13060-3.

Protocol appendix 3 Details of the 2-3 year assessments for all babies

The brain goes through notable large and dramatic changes during the first two years of life, making this a crucial age to investigate and monitor how these changes are effecting cognitive performances at each stage (1). Very few studies have performed thorough investigations of the multiple domains that are potentially affected by preterm birth at 2 years of age, and those that have touched the surface of these problems, express the need for more in depth investigations (2), such as those proposed here.

All clinical subjects are followed to 2 years (corrected for preterm birth if appropriate) by the clinical service and receive a formal outpatient neurological and developmental assessment. In addition we will invite each family to attend on another day for additional testing (control subjects will have all assessments at one visit including a Bayley-III assessment).

EEG Measures: Preterm infants are known to have difficulties with their speech development. Up to now there have been a number of studies investigating the neural processing involved in the language deficits in older children (4-5 years), once speech problems are more apparent (3, 4). The common consensus from these studies is the deficits are likely to be related to auditory processing abilities, with the opinion that assessments of the neural correlates are required at a much younger age to investigate this further (3). What we aim to do with this ERP task is to investigate auditory processing at 2 years of age to see if there are any apparent delays at this earlier stage of development. Recent MRI findings have suggested that the detection of impaired maturation of the brain microstructure after birth may have the potential to be strong predictors of language and cognitive impairment in these children. By taking a measure of their language development, both behaviourally and in terms of auditory processing, we determine any correlation between that and the structural data obtained at birth (5, 6).

For all groups (at 24-30m), we will use an auditory event-related potential (ERP) paradigm (32-electrode Geodesic Sensor Net). In one EEG session, we will first run the same resting state paradigm as in the main Protocol; this will be followed by a novel assessment: the toddlers will watch an unrelated visual video presentation whilst listening to auditory speech or non-speech-related sounds in trial sets. Subjects will continue with alternating trials as long as they show interest in the paradigm (maximum time estimation 20 min).

Eye-tracking Measures: We will measure looking behaviour and in both the VP and term infants in order to both assess their ability to successfully switch visual attention and investigate their social preferences. For visual attention, we will administer the same task as previously performed at 6 and 12m of age in order to obtain longitudinal measures of switching abilities. The GAP task displays visual cues in the center of the screen and after a period of time, in the baseline condition, the central stimulus disappears and a peripheral stimulus is shown. In the test trials, the peripheral stimulus appears but the central stimulus remains. A non-intrusive Tobii corneal reflection eye-tracker (T60/T120) is used to record the time taken for the child to disengage from the central stimulus, whilst seated on their parent's lap. We will also use the same equipment to investigate social preferences as previously determined. In total, this series of tasks should last no longer than 15 min.

Neuropsychology paradigms: A number of behavioural paradigms have been devised to assess the developing executive functions as this age. In order to try and tease apart possible specific executive function deficits, as shown in previous investigations (7), we will use a composite assessment of inhibition, working memory and cognitive flexibility skills (8). The selection of tasks should take no longer than 30 minutes to administer. These are important as end points for our infancy assessments

Observational Assessments: Following a break, for all subjects we wish to administer the Autism Diagnostic Observation Scale-2 (ADOS-2). This 30-45 min assessment is designed to highlight key behaviours and language abilities that lend to the diagnosis of autism, from very mild cases to more broad general developmental disorders. The assessment involves a series of social interactions between the toddler and the examiner in a series of structured and semi-structured tasks; the total score has defined cutoffs for autism and autism spectrum disorders. The assessment is modular and the appropriate module is determined by the trained examiner according to the developmental and language level of the child. Should a child score highly the PI (NM) will discuss with the parents if they would accept referral to community services for further management. At most this will include no more that 4 children.

Lastly, we will video record 10 minute free-play observation between the parent and the toddler. Interactions with the mother and infant will be coded blind to the study group, using to a validated global rating scale.

In total this will mean 70 minutes of assessment followed by the ADOS-2 and 10 minutes free play after a break.

Additional Questionnaire Measures: Parents will be asked to complete prior to the visit the *Toddler Behaviour Assessment Questionnaire* (9), continuing on from the IBQ recorded in the previous assessment ages; the *Family Environment Scale (FES)*, to provide an idea of the family home environment, as has been shown to have an impact on learning; lastly, the *MacArthur Communicative Development Inventory (MCDI)*, as a measure of the child's language abilities.

Additional references for Appendix 3

1. Sun J, Buys N. Executive function and its relationship with developmental disorders in preterm born children. Int J Adolesc Med Health. 2012;24(4):291-9.
2. Mansson J, Stjernqvist K. Children born extremely preterm show significant lower cognitive, language and motor function levels compared to children born at term, as measured by the Bayley-III at 2.5 years. Acta Paediatr. 2014.
3. ansson-Verkasalo E, Valkama M, Vainionpaa L, Paakko E, Ilkko E, Lehtihalmes M. Language development in very low birth weight preterm children: a follow-up study. Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP). 2004;56(2):108-19.
4. Jansson-Verkasalo E, Čeponien≐ R, Valkama M, Vainionpää L, Laitakari K, Alku P, et al. Deficient speech-sound processing, as shown by the electrophysiologic brain mismatch negativity response, and naming ability in prematurely born children. Neuroscience Letters. 2003;348(1):5-8.
5. Chau V, Synnes A, Grunau RE, Poskitt KJ, Brant R, Miller SP. Abnormal brain maturation in preterm neonates associated with adverse developmental outcomes. Neurology. 2013;81(24):2082-9.
6. Northam GB, Liegeois F, Chong WK, Baker K, Tournier JD, Wyatt JS, et al. Speech and oromotor outcome in adolescents born preterm: relationship to motor tract integrity. The Journal of pediatrics. 2012;160(3):402-8 e1.
7. Espy KA, Stalets MM, McDiarmid MM, Senn TE, Cwik MF, Hamby A. Executive functions in preschool children born preterm: application of cognitive neuroscience paradigms. Child

neuropsychology : a journal on normal and abnormal development in childhood and adolescence. 2002;8(2):83-92.
8. Pozzetti T, Ometto A, Gangi S, Picciolini O, Presezzi G, Gardon L, et al. Emerging executive skills in very preterm children at 2 years corrected age: A composite assessment. Child Neuropsychology. 2014;20(2):145-61.
9. Goldsmith HH. Studying Temperament via Construction of the Toddler Behavior Assessment Questionnaire. Child Development. 1996;67(1):218-35.

Appendix 4: Cross-sectional recruitment of typically developing 2-3 year olds for the 4th stage of the Preterm Development Cohort.

The Preterm Development Project has been designed to track a group of preterm infants and a comparison group of term born infants through their first few years of life. The aim for this study is to have a clear understanding of the executive function, information processing speeds, and social skill abilities of infants at four different points through the first three years of life. As with many longitudinal studies, there are always a number of infants that fail to attend one or more of the follow-up visits and/or withdraw before completion. In order to ensure an acceptable sized cohort comparison group (i.e. sufficiently powered), an additional cross-sectional group of children born at term are required for the two to three year time point. Below highlights the methods for the recruitment of this cohort, the inclusion/exclusion criteria for inclusion, the information to be requested and what will be expected from infants during the assessment.

Population: Term born children are to be recruited by advertisement in local nurseries. Nurseries to be approached include the ULCH nursery, GOSH nursery, and the UCL Nursery. Inclusion criteria: birth at 37-42weeks gestation, birthweight between 10th and 90th percentile for gestation, no perinatal complications and Apgar score at 5min >7. The study will be advertised, with the nurseries agreement, with fliers, and posters in the reception areas, and the parents will be asked to approach the team if they are interested in participating in the study. Once the parent has expressed interest in taking part, a member of the team will be in touch to clarify the above inclusion information before they are able to take part. If the child is eligible to take part, a visit to the CRF facility in the EGA wing will be organised. The parent will be sent an information leaflet and map to the EGA wing via email along with a confirmation of the study visit date and time. On the day of the assessment the parent will be asked to complete a consent form to provide their consent for participation in this follow-up visit only. However, the consent form will also ask parents whether they are happy to be contacted by the team about future follow-ups, should they be arranged. Parents will provide their consent for this on an opt-in basis only.

Assessment: The assessment will be identical to that run with the term-born infants currently participating in the full longitudinal Preterm Development Project. The study uses four key assessment methodologies: EEG, eye-tracking and neuropsychological assessment, and play-based standardised assessments (Bayley Scales of Infant and Toddler Development®, Third Edition; Autism Diagnostic Observation Schedule, ADOS). The preterm infants complete the Bayley assessment as part of their routine clinical follow-up within the trust, and therefore do not complete the assessment with us. These scores are obtained from clinical records, with parent permission. Where we have completed standardised assessments, parents are given feedback on their child's performance. If concerns are indicated by a child's score profile on the Bayley or ADOS (i.e. within a clinical range), the principal investigator (Prof Neil Marlow) will contact the parents, and offer to make a referral to local services for further investigation.

EEG Measures: Preterm infants are known to have difficulties with their speech development. Up to now there have been a number of studies investigating the neural processing involved in the language deficits in older children (4-5 years), once speech problems are more apparent (1, 2). The common consensus from these studies is the deficits are likely to be related to auditory processing abilities, with the opinion that assessments of the neural correlates are required at a much younger age to investigate this further (2). What we aim to do with this ERP task is to investigate auditory processing at 2 years of age to see if there are any apparent delays at this earlier stage of development. Recent MRI findings have suggested that the detection of impaired maturation of the brain microstructure after birth may have the potential to be strong predictors of language and cognitive impairment in these children. By taking a measure of their language development, both behaviourally and in terms of auditory processing, we determine any correlation between that and the structural data obtained at birth (3, 4).

For all groups (at 24-30m), we will use an auditory event-related potential (ERP) paradigm (32-electrode Geodesic Sensor Net). In one EEG session, we will first run the same resting state paradigm as in the main Protocol; this will be followed by a novel assessment: the toddlers will watch an unrelated visual video presentation whilst listening to auditory speech or non-speech-related sounds in trial sets. Subjects will continue with alternating trials as long as they show interest in the paradigm (maximum time estimation 20 min).

Eye-tracking Measures: Eye-tracking assessments will use remote eye-tracking technology (Tobii T60 corneal reflection eye-tracker) to record infants' eye-movements in response to visual stimuli whilst they sit on a parent's lap. This permits inferences to be made about infants' visual attention skills. Looking behaviour will be recorded in response to two developmentally sensitive tasks, identical to those used at earlier follow-ups at 6 and 12 months. The first is a non-social task, measuring infants' ability to successfully disengage visual attention (the Gap-Overlap Task). In this task, infants are presented with animated visual cues at the centre of a screen, which is then replaced by a peripheral stimulus. In the baseline condition, the central target disappears prior to the appearance of the peripheral target, which facilitates disengagement in order to shift attention. In the test trials, however, the central stimulus remains on-screen when the peripheral target appears, meaning that the infant has to actively disengage their attention in order to shift. The second task measures infants' responsiveness to adults' social looking cues (the Gaze-Following task). Infants are presented with short video recordings of a lady shifting her gaze from direct towards the infant, to objects positioned to her left and right. Infants' looking behaviour in response to these gaze cues is measured. In total, completion of these two tasks should take no longer than 15 min.

Neuropsychology paradigms: A number of behavioural paradigms have been devised to assess the developing executive functions as this age. In order to try and tease apart possible specific executive function deficits, as shown in previous investigations (5), we will use a composite assessment of inhibition, working memory and cognitive flexibility skills (6). The selection of tasks should take no longer than 30 minutes to administer. These are important as end points for our infancy assessments

Observational Assessments: Following a break, for all subjects we wish to administer the Autism Diagnostic Observation Scale-2 (ADOS-2). This 30-45 min assessment is designed to highlight key behaviours and social-communication difficulties associated with a diagnosis of autism spectrum disorder (ASD) according to DSM-V criteria. Researchers trained in the administration and reliable coding of the instrument will complete the assessment. The ADOS consists of both structured and semi-structured tasks, and involves a series of presses and social interactions between the toddler and the examiner. Defined cutoffs for autism and autism spectrum disorders are provided, based on total scores. The assessment is modular and the appropriate module is determined by the trained examiner according to the developmental and language level of the child. Should a child score within the clinical-range, the PI (NM) will discuss with the parents if they would accept referral to community services for further investigation and management. Based on known prevalence rates, it is expected that this will involve no more than four children.

Lastly, we will video record 10 minute free-play observation between the parent and the toddler. Interactions with the mother and infant will be coded blind to the study group, using to a validated global rating scale.

In total this will mean 70 minutes of assessment, followed by the ADOS-2 and 10 minutes free play after a break.

Additional Questionnaire Measures: Parents will be asked to complete prior to the visit the *Toddler Behaviour Assessment Questionnaire* (7), continuing on from the IBQ recorded in the previous assessment ages; the *Family Environment Scale (FES)*, to provide an idea of the family home environment, as has been shown to have an impact on learning; lastly, the *MacArthur*

Communicative Development Inventory (MCDI), as a measure of the child's language abilities.

1. Jansson-Verkasalo E, Čeponien R, Valkama M, Vainionpää L, Laitakari K, Alku P, et al. Deficient speech-sound processing, as shown by the electrophysiologic brain mismatch negativity response, and naming ability in prematurely born children. Neuroscience Letters. 2003;348(1):5-8.
2. Jansson-Verkasalo E, Valkama M, Vainionpaa L, Paakko E, Ilkko E, Lehtihalmes M. Language development in very low birth weight preterm children: a follow-up study. Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP). 2004;56(2):108-19.
3. Chau V, Synnes A, Grunau RE, Poskitt KJ, Brant R, Miller SP. Abnormal brain maturation in preterm neonates associated with adverse developmental outcomes. Neurology. 2013;81(24):2082-9.
4. Northam GB, Liegeois F, Chong WK, Baker K, Tournier JD, Wyatt JS, et al. Speech and oromotor outcome in adolescents born preterm: relationship to motor tract integrity. The Journal of pediatrics. 2012;160(3):402-8 e1.
5. Espy KA, Stalets MM, McDiarmid MM, Senn TE, Cwik MF, Hamby A. Executive functions in preschool children born preterm: application of cognitive neuroscience paradigms. Child neuropsychology : a journal on normal and abnormal development in childhood and adolescence. 2002;8(2):83-92.
6. Pozzetti T, Ometto A, Gangi S, Picciolini O, Presezzi G, Gardon L, et al. Emerging executive skills in very preterm children at 2 years corrected age: A composite assessment. Child Neuropsychology. 2014;20(2):145-61.
7. Goldsmith HH. Studying Temperament via Construction of the Toddler Behavior Assessment Questionnaire. Child Development. 1996;67(1):218-35.

Appendix 2. Ethical Approval

Health Research Authority

Research Ethics Service London - Hampstead Research Ethics Committee

23 December 2015

Study title:	**The UCH Preterm Development Project: growing up after very preterm birth**
Amendment number:	**Substantial Amendment 4**
Amendment date:	**27 July 2015**

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:

Document	*Version*	*Date*
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

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/

Appendix 3. PDP demographics questionnaire

PDP
UCH Preterm Development Project

Study Number : ________________
Infant ID: ________________
Infant Gender: ________________
Date complete ________________
Age of Infant: ________________

Brain development after very preterm birth

Questionnaire for parents - About your family

Thank you ever so much for coming back to see us in the Baby lab. As it has been a little while since we last saw you, we would like to make sure our records are still up to date. Please could you complete this form and either bring it along to your next visit with us in the baby lab or complete and send it back to a member of our team in a reply to our confirmation email that you will have received for your visit.

We realise that some of these questions are similar to the ones we have asked in the past, but we would be extremely grateful if you could complete the whole booklet as some of the information we are requesting are essential to the tasks we will ask your child to perform at the next visit.

We realise these are personal data.
All the information will be treated in the strictest confidence and will not be seen by anyone outside the study.
The family information will be coded and will be used anonymously in all our analysis
The questionnaire will also be destroyed when we have finished with it.

If you have any questions, or would like any help in completing this questionnaire, please speak to the staff member who gave you the form or you can telephone the Study office

Thank you very much for your help

To begin, please can you provide us with the following information:

i) Name of Baby Taking Part: ______________________________

ii) This form was completed by (name): ______________________ Date ____________

iii) Please state your relationship to child:

Mother ☐ 1

Father ☐ 2

Other (please specify below)* ☐ 3

If "Other", please specify (e.g. Grandmother): ______________________

Section A: Your address and contact details

Please tell us your Home Address	
Postcode	
Your telephone number at home	
Your mobile or cell phone number	
Your email address	

Section B: Your family Doctor (GP) and contact details

Your Doctor's Name	
Your Doctor's Practice Address	
Postcode	
The practice telephone number	

Section C: Another family member's address and contact details

Although this seems odd it has been very helpful for us to have the contact details of another family member (e.g. one of your baby's grandparents) to check we have the correct contacts for you should you move.

Their name	
Relationship to your child	
Their Home Address	
Postcode	
Their telephone number at home	

Their mobile or cell phone number

Their email address

Section D: About your family

1 **Who lives with you (adults)?**

No other adults ☐ 1
Husband or wife or partner ☐ 2
Other adults (aged 18 or more) ☐ 3

If other adults live with you, who are they?
(E.g. Your child's maternal grandmother etc.)

2 **How many children (aged up to 18 years) are there in the household (*including* the child taking part in the study)?**

☐ Children

Please list the dates of birth of the other children.

NAME AND DATE OF BIRTH e.g. Joe - 24/06/1994

3 **If any of the other children have received a diagnosis of learning difficulties or a developmental disorder (e.g. Autism Spectrum Disorder; ADHD; Dyslexia), please specify.**

NAME, DIAGNOSIS AND AGE AT DIAGNOSIS e.g. Hannah – ADHD, diagnosed at 5 years

4 **Are you:**

Married? ☐ 1
Single? ☐ 2
Living together? ☐ 3
Widowed? ☐ 4
Separated / Divorced? ☐ 5

5 **Are you:**

- Living with the father or mother of the study child? ☐ 1
- Living with other partner? ☐ 2
- Previously living with partner, now alone? ☐ 3
- Never living together? ☐ 4
- Other situation, e.g. family or friends? ☐ 5

6 **Is your current partner the biological father/mother of this child?**

- Yes ☐ 1
- No ☐ 2

7 **What is your current age?** ☐ Years

8 **What is your partner's current age?** ☐ Years

Section E: About your home

1 **Do you rent or own your accommodation?**

- Owner (mortgage) ☐ 1
- Council rented ☐ 2
- Private rented (furnished) ☐ 3
- Private rented (unfurnished) ☐ 4
- Housing society or co-operative ☐ 5
- Tied to occupation ☐ 6
- Other (please describe below) ☐ 7

2 **How long have you lived at this present address?** ☐ Years

3 **If less than 6 years how many moves in the last 5y?** ☐ Moves

4 **What language is spoken at home?**

- English only ☐ 1
- Other language(s) ☐ 2

Please tell us which is the other language(s)

Which language is the most predominately spoken in the home?

Please give us a percentage of the time spoken in the language (rough estimate)

Section F: About your education

1 What is your highest qualification from school or college?

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

Section H: About your work

EMPLOYMENT

1 **Are you currently in paid employment? Please tick for yourself and your partner (if applicable) as appropriate:**

	You	Your Partner
Employed	☐ 1	☐ 1
Self-employed	☐ 2	☐ 2
Unemployed	☐ 3	☐ 3
Retired	☐ 4	☐ 4
Other (please describe)	☐ 5	☐ 5

If you are currently in paid employment please complete the following questions for your current job. If you are currently unemployed, please complete the following questions for your last job.

2 **What is your current/last job? Please describe below:**

	You	Your Partner
Job title:	______________	______________
Company/organisation:	______________	______________
Type of Industry:	______________	______________
	______________	______________

3 **Please describe what you mainly do/did in this job.**

Job description: ______________ ______________

4 **How many hours a week do/did you work?** ______ hours ______ hours

5 **How many people are/were employed at the place where you work/worked?**

	You		Your Partner	
1 to 24	☐	1	☐	1
25 to 499	☐	2	☐	2
500 or more	☐	3	☐	3

6 **Are you a manager?**

	You		Your Partner	
No	☐	1	☐	1
Yes	☐	2	☐	2

7 **If you are not a manager, do you supervise other members of staff?** (not including supervision of children, patients etc).

	You		Your Partner	
No	☐	1	☐	1
Yes	☐	2	☐	2

8 **If you are self employed, do you employ other people?**

	You		Your Partner	
I work on my own/with a partner, but have no employees	☐	1	☐	1
I have 1 to 24 employees	☐	2	☐	2
I have 25 to 499 employees	☐	3	☐	3
I have 500 or more employees	☐	4	☐	4
Not applicable - not self-employed	☐	5	☐	5

Section G: About your baby

Part 1: Questions about your baby's birth.

1 **What is the date of birth of the child taking part in this study?**

dd/mm/yy ___/___/___

2 **What was your expected due date of delivery?**

dd/mm/yy ___/___/___

3 **Did you need any assistance with the birth?**

Yes ☐ 1

No ☐ 2

4 **If yes, was the birth a caesarean section/ forceps delivery/ ventouse delivery/other (please specify)?**

5 **If the birth was by caesarean section was this -**

Emergency? ☐ 1

Elective? ☐ 2

6 **What was your baby's birthweight?** ☐

7 **Did your baby need to be admitted to the neonatal unit for any reason?**

Yes ☐ 1

No ☐ 2

If yes, please tell us why

Appendix 4. Auditory processing

Upon hearing a sound, the ear transforms the raw sensory information into mechanical energy via the cochlea to the auditory nerve. From here, the processing pathway through the brain is complex and not completely understood. The speed in which one processes a sound can have a significant impact on various aspects of life, from conversing to spatial awareness (Scott and Wise, 2004). Upon reaching the auditory nerve, the sound information projects to the contralateral superior olivary nucleus via the ascending auditory pathway, reaching via the inferior colliculus it is bilaterally represented to the medial geniculate nucleus. From there it is projected to the auditory cortex bilaterally but asymmetrically with connectivity between auditory cortices via the corpus callosum (see Figure 8-2 for one proposed model). A certain amount of decussation occurs at each level of the pathway (Lazard, Collette and Perrot, 2012). However, common consensus appears to suggest the dominant fibres are those that transmit information contralaterally (Scott and Wise, 2004).

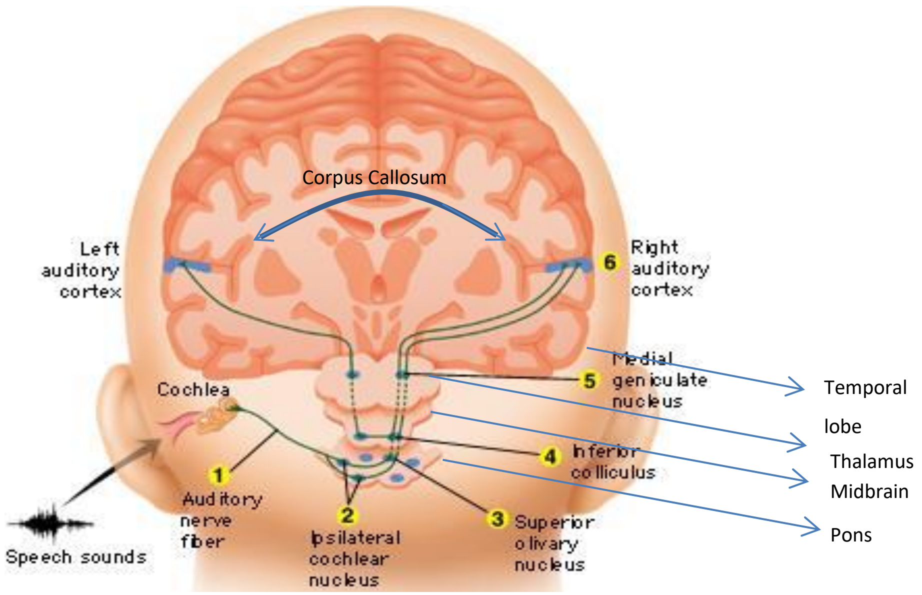


Figure 8-2. Auditory pathways modified from Posit Science, (2014).

The medial geniculate nucleus is the final relay before the primary auditory cortex, found in the thalamus. Once the auditory signal hits the Primary Auditory Cortex

(PAC), there are multiple projections to surrounding temporal areas (Scott and Wise, 2004; Lazard, Collette and Perrot, 2012). The PAC contains two structures; Heschl's gyrus, the area considered as the primary auditory cortex, and the Planum Temporale (PT). The PT is considered part of Wernicke's area and has been seen to respond to a variety of acoustic non-speech stimuli and differences in articulation. Projections from surrounding areas form the posterior and anterior stream.

The anterior stream is thought to correlate sounds to meaning, projecting to the higher cortical regions in the PFC and medial temporal lobes (Brocas and wernicka's area respectively). The posterior stream is stipulated to correlate 'speech sounds to motor representations of articulation', and associates sensory and motor stimuli, therefore thought as a step beyond initial auditory processing (Scott and Wise, 2004).

Broca and Wernicke's areas are the two best known higher cortical areas associated to speech processing. Broca's area, discovered back in 1861, is associated to speech production, predominantly on the left hand side of the brain; Wernicka's area is linked to the understanding of speech and is considered the secondary auditory cortex (Steinmann and Mulert, 2012). The main intrahemisphere connection between these two areas is the Arcuate Fasciculus (AF) and the interhemispheric connection linking the two auditory cortices runs through the posterior of the Corpus Callosum, the splenium (Steinmann and Mulert, 2012).

The Corpus Callosum (CC) is responsible for the majority of communication between the two hemispheres; for auditory, motor, cognitive and voluntary information. The contribution of the CC in auditory processing and speech perception is a matter of debate (Bamiou *et al.*, 2007); however, it is clear from numerous patient populations studies that both hemispheres are involved in typical speech interpretation (Musiek *et al.*, 1989). The two large fibre tracts involved in interhemispheric communications are the Corpus Callosum and the Anterior Commissure (AC). It is in the AC and the posterior section of the CC that Northam *et* *al.*, saw a reduced volume in preterm participants correlating to language impairments.

How these connections come into play during the processing of sound remains uncertain, however, it is clear that the two hemispheres have to communicate via the corpus callosum (CC) (Bamiou *et al.*, 2007) upon receiving auditory information (Steinmann and Mulert, 2012). Numerous auditory experiments have led to the proposal of multiple different models regarding how the CC is involved in the processing of auditory information (Bamiou *et al.*, 2007).

The first of the models, by Kimura (Kimura, 1961, 1967), utilised dichotic listening tasks where two different sounds are presented to the two ears simultaneously. The sounds presented to the right ear appear to be heard over those presented to left in typical subjects, suggesting that the dominant pathway crosses over within the brain if taking the left hemisphere as dominant for language. Kimura also suggests the ipsilateral pathways are inhibited by the contralateral pathways in this competing situation (Bamiou *et al.*, 2007). In split brain patients, where the Corpus Callosum (CC) has been severed, a left ear disadvantage is seen, suggesting that the sounds presented to the left is ear is not being processed directly on the left hand side and a decussation of hemispheres across the CC is necessary for the processing of sounds (Musiek *et al.*, 1989).

Zaidel (Zaidel, 1986) proposed two alternative models; the 'Callosal Relay'; and 'direct access' model. Within the callosal relay model, speech stimuli require callosal transfer from the right to the left hemisphere for complete comprehension of the sounds. Supporting this, results from dichotic listening tasks find a longer reaction time to verbal information presented to the left ear, suggesting the information was not processed as quickly due to the interhemispheric transfer from the contralateral hemisphere to speech centres on the left (Bamiou *et al.*, 2007).

The alternative 'direct access model' (Zaidel, 1986), suggests sounds are processed within the contralateral hemisphere to the side of presentation without the need to cross the brain. Less efficient processing occurs in the right hemisphere to verbal stimuli presented to the left ear according to this theory (Bamiou *et al.*, 2007).

When presenting sounds monaurally, the literature is mixed as to how the brain deals with this information, with some reporting the contralateral pathway remains dominant even in the absence of stimulation from the other ear (Stefanatos *et al.*, 2008); yet others suggest the interhemispheric transfer of information is no longer required as the ipsilateral fibres are no longer being inhibited by simultaneous ear stimulation (Musiek, 1986).

Further investigations are required to determine the involvement of interhemispheric connections during a monaural stimulation tasks. In terms of the more dominant auditory pathways, excluding results from split brain patients, it seems the literature would support contralateral dominance. Although large inconsistencies are apparent in the literature regarding the proportional involvement of the interhemispheric connections in speech perception it is clear that the two hemispheres have to communicate upon hearing auditory information (Steinmann and Mulert, 2012), via the corpus callosum (CC) (Bamiou *et al.*, 2007). If there are reduced volume or delays in interhemispheric transfer, evidence would suggest this would be observable in the ERP components in response to auditory stimulation, either in latency or in hemispheric dominance of response.

The speculation in relation to the preterm literature, however, is that the reduced volume of the CC and associated connections leads to a reduction in speed of processing across these interhemispheric connections, and ultimately leading to delays in language.

Auditory and language impairments in children born preterm

A common problem reported in the preterm population is delays in language production. The literature in this area is unclear, however, it is postulated that it is the speed in which infants process sounds that may determine the production of speech (Ramon-Casas 2013; Jansson-Verkasalo 2004). Axons pass though the CC to produce interhemispheric connections. Differences in size relate to the volume and number of axons that traverse this area. White matter injury may cause the various areas of the CC to have a reduced volume, but likewise poorly myelinated fibres (fewer oligodendroglia) are smaller and will occupy less of the cross sectional area. Reduction in cortex cell counts or connectivity through altered developmental trajectories will produce similar findings. Poor splenium development therefore may reflect the end stage of any of these processes. Much of the white matter damage we see on MRI scans is not easily identifiable in the neonatal period. The posterior portion of the CC is last to develop with the fastest growth observed after birth. It is this section of the CC, the Splenium, that has most widely been found to be effected by preterm birth; the area responsible hosting connections between the two auditory centres of the brain (Chiara Nosarti *et al.*, 2004).

Language deficits are often reported in preterm populations with the speculation these differences are due to the encoding of speech sounds. Most studies investigate the neural correlates associated with auditory processing in preterm children though language, but do not look at the simple processing of a sound. Auditory ERPs are a measure of activity following auditory stimulation and frequently explored with the use of Mismatch Negative (MMN) paradigms (Hövel *et al.*, 2014). A MMN paradigm investigates the auditory discrimination response; a component elicited when the brain automatically computes a change in a physical stimulus (Naatanen, 2001). The correlation of this component to the specific mechanisms within auditory perception is not clear and is therefore commonly used to investigate simple sensory stimulation (Jansson-Verkasalo *et al.*, 2003). For example, an MMN paradigm used by Jansson-Verkasalo et al., (2003) investigated this response to 3 standard syllables across term and preterm children. The amplitude of the MMN response was considerably reduced in the preterm participants and correlated to deficits within the object naming task. The authors concluded that deficits observed were reflective of preattentive auditory processing difficulties that could later contribute to language delays (Jansson-Verkasalo *et al.*, 2003).

Jansson-Verkasalo et al., continued their investigations the following year (2004) in a cohort 2 years olds. Children born preterm achieved significantly lower scores on the language comprehension tests than their term born peers which later correlated to the performance in the auditory discrimination task at 4 years. It is therefore possible that auditory processing could be responsible for aspects of the language difficulties seen in this population (Jansson-Verkasalo et al., 2004). This supports the link between the CC size and its involvement in the transmission of auditory information between the hemispheres. What remains to be determined is whether it is the speed of transmission that is to blame for this delay or whether is it due to the poor connectivity of the structure that leads to insufficient transmission of information that is the cause of the processing delays.

What is perhaps surprising, given the volumetric differences observed in the CC within ex-preterm populations, is the lack of investigations exploring the hemispheric difference in response to sounds. Although no MR data is available within this current thesis, the processing of auditory information will be explored looking at possible delays in interhemispheric transmission give the evidence discussed above. This will be the focus of the neural information processing section of this thesis.