Number Processing in Infants and Children Born Very Preterm

Mérari Jizar Lavander-Ferreira

UCL Great Ormond Street Institute of Child Health

Thesis submitted for the degree of

Doctor of Philosophy

Supervisors: Professor Michelle de Haan &

Professor Neil Marlow

April 2020

Declaration

I, Mérari Jizar Lavander-Ferreira, 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.

Mérari Jizar Lavander-Ferreira

03/04/20

Date

Abstract

Individuals born very preterm (<32 weeks; VP) have notably poorer attainment in mathematics than their term-born peers. Only a handful of studies have investigated basic numerical skills in VP children and the underlying mechanisms associated with problems with mathematics in this population are still not fully comprehended. Basic processes underlying numerical cognition can go awry very early in development and there is a lack of knowledge of early trajectories of acquisition of numerical skills in infants born prematurely. This thesis reports on a series of studies investigating number processing in very preterm infants and children. These make use of a combination of tools, such as neurodevelopmental assessments, eye-tracking, event-related-potentials, neuropsychological evaluations and experimental tasks. Specifically, cross-sectional studies investigated numerical sensitivity in VP infants aged six and twelve months. Behavioural and electrophysiological measures assessing a range of domain-general and domain-specific skills associated with mathematics performance were also investigated in VP school-aged children. The results showed that, during the first year of post-natal life, VP infants do not exhibit differential developmental trajectories in the basic ability to discriminate numerosities compared to infants born at full term, although they required a longer time to discriminate the new number of elements. Later in development, school-aged VP children demonstrated difficulties in processing basic numerical information. Electrophysiological data demonstrated that this might be associated with deficits in sensory and attention resources and not necessarily in how VP children encode number-related information. Difficulties in processing numerical information, however, have only a marginal impact on their performance in mathematics. We tentatively conclude that difficulties in mathematics in individuals born very prematurely are largely associated with domain-general skills.

Impact Statement

This research aimed to investigate a range of domains associated with mathematics performance in infants and children born very preterm. The thesis contributes to the field in three key ways. First, it is one of the first studies to assess basic numerical skills in infants born very prematurely. As such, it contributes to the early identification of cognitive processes disrupting typical trajectories of the development of numerical skills among individuals born prematurely. Second, this is the first study to employ electrophysiological measures to explore numerical representations in VP children. Electrophysiological measures help to elucidate the underlying neural mechanisms of number processing in VP children, an area that has been strikingly unexplored. Understanding the neural resources allocated to the cognitive process related to number processing helps to elucidate the cognitive profile of very preterm children struggling with maths. This knowledge will ultimately inform the design of better interventions to help them. Thirdly, using a similar approached employed by previous studies (e.g., Simms et al, 2015), we examined domain-general and domain-specific skills and their association with mathematical performance in VP children. The findings of this work replicated previous work. Given the fact that we currently face a replication crisis in different subjects of science (Munafò et al., 2017), this provides greater validity, making the results more likely to generalisable to the larger population.

Ultimately, this line of research contributes to enhance knowledge of the mechanisms of number processing, and their interplay with executive functions, in infants and children born very prematurely. This will help both in the early identification of those individuals born prematurely who are at particular risk of experiencing difficulties with mathematics, and in targeting better interventions for those who struggle at school.

Finally, the findings presented in this thesis have been disseminated within the scientific community through presentations at several academic meetings.

Acknowledgments

First and foremost, I would like to thank my supervisors Professor Michelle de Haan and Professor Neil Marlow, for their support, encouragement and guidance throughout my Ph.D. I have been working with Professor Michelle de Haan since 2012 when she was also my MSc supervisor. Professor de Haan has believed in me, supported and encouraged me to look for external funding to pursue my Ph.D. and I am enormously grateful for her faith in me. Professor de Haan has introduced me to the fantastic world of infant studies and I consider myself enormously privileged to have learned from one of the world's leading researchers in infant cognition. I also consider myself incredibly fortunate to have gained knowledge from one of the most well-recognised scientists in the field of prematurity, Professor Marlow. He has patiently taught me a range of concepts and abilities, beyond the prematurity field. Above all, both Professor de Haan and Professor Marlow exceeded their duties as supervisors in giving me emotional support during the most difficult time of my life when I lost my mother. I will be always grateful for their intellectual and emotional support during the time of my Ph.D.

This work would not have been possible without the help of many people, whose contributions I acknowledge here. I want to thank our team at the Institute of Women's Health supporting the Preterm Developmental Project (PDP). My thanks to Jade Okell, our research nurse, who booked many assessments; Carla Logon-Bishop, Professor Marlow's personal assistant, who booked rooms for our assessments, many of which were last-minute bookings; Dr Kathy Chant, who stepped in during the last stages of the PDP study to recruit a large number of families with preterm infants, thereby increasing the size of our clinical group substantially.

I want to thank researchers from the Great Ormond Street Institute of Child Health who have joined either the infants' or the children's project. Many placement, BSc and MSc students, joined my Ph.D. projects, helping in some way, either by booking and recruiting families, or in processing data. I want to register my gratitude to Dr Kayleigh Walker (née Day) for teaching me about infant assessments and introducing me to the world of EEG. I want to thank her for patiently teaching me how to place an EEG net on an infant's head — a very scary task for beginners. A special thank you also for Mariana Pereira, who patiently shared her EEG knowledge, and Professor Torsten Baldeweg, for his intellectual guidance in the very last stages of writing the last experimental chapter.

I want to extend my thanks to many colleagues and friends who kindly took the time to revise part of this work, including Mariana Pereira, Dr Laura Kaus, Dr Sarah Buck and Dr Kathy Chant. Their feedback has significantly improved this work. Further, I want to thank all present and former colleagues that made my time at the Great Ormond Street Institute of Child Health not only stimulating but also pleasant.

Next, I want to thank my family. My parents, Maheli Jizar Lavander and Wilson Cândido Ferreira, and my siblings, Marihem Joice Lavander Ferreira and Ivan Lavander Cândido Ferreira, for always giving support, encouragement, and patience over the last years. Unfortunately, my late mother was unable to see one of the greatest achievements in my life. In September 2018, her life was taken from us too early battling metastatic breast cancer. She was proud to see me pursuing my academic career as a first-generation Ph.D. candidate in the family. My special gratitude to my late mother, who guided and supported my steps. She is greatly missed.

I want to express my gratitude to Dr Sebastian Bobadilla Suarez for his technical and emotional support. My special thanks for his help, for quickly answering questions and for his tutorials about statistical analysis. Dr Bobadilla has also contributed modelling *w* scores for Study 3 (chapter 6) and made the code publicly available (https://github.com/bobaseb/webers_fraction) and has provided several comments to improve my writing progress.

Further, I want to thank the children and families that participated in this study. This research would not have been possible without their contribution.

Finally, this work would not have been possible without the financial support given by the 'Science without Borders program'. The Science without Borders program was an unprecedented Brazilian programme that sent 100,000 Brazilian researchers to study abroad implemented during the government of former President Dilma Rousseff. Due to a coup in 2016, President Dilma Rousseff was impeached, and several programmes implemented by her government were cancelled, including 'Science without Borders program'. I was fortunate to be able to finish my Ph.D. with financial support from the programme awarded by CAPES. Future generations of scientists will be impacted by the cancellation of Science without Borders and will not have the opportunity I had to be mentored by world-leading researchers in a world-leading university. My recognition goes, therefore, to the former Brazilian government lead by former President Dilma Rousseff who gave academic opportunities to several researchers who would not otherwise have such experiences.

Dedication

This work is dedicated to my parents, Maheli Jizar Lavander and Wilson Cândido Ferreira. As teachers, my parents always taught me the power that education can have in improving people's lives. Their weapon to fight against inequality was always education. This work is dedicated to them in my deepest gratitude for what they taught me and for the unconditional support they always gave me.

For my late mother, I dedicate this work in recognition of her efforts for guiding my steps to achieve this dream. I also dedicate this work to my dearest father who was born prematurely. He and his twin sister were born at home in the countryside of Brazil. To keep my father and my aunt alive, my grandmother would involve their crib with hot bottles of water constantly and restless for several weeks until they were no longer at risk of losing their lives. My dedication to my father who was a fighter since he was born and taught me, like him, to be a fighter.

Obrigada mamãe e papai por acreditarem em mim. Este trabalho também é fruto do sonho e trabalho de vocês!

List of Contents

Table of Figures

List of Tables

List of Abbreviations

ADHD	Attention-deficit/hyperactivity Disorder
ANS	Approximate Number System
AOIs	Areas of Interests
AWMA	The Alloway Working Memory Assessment
BD	Bachelor's Degree
BPD	Bronchopulmonary dysplasia
CLD	Chronic Lung Disease
CP	Cerebral Palsy
CR	Correct Response
D-KEFS	Delis-Kaplan Executive Function System
DD	Developmental Dyscalculia
DTI	Diffusion Tensor Imaging
EDD	Expected Date of Delivery
EEG	Electroencephalogram
EF	Executive Functions
EP	Extremely preterm
ERPS	Event Related Potentials
F	Female
fMRI	Functional Magnetic Resonance Imaging
FSIQ	Full Scale IQ
FT	Full Term
GA	Gestational Age
IES	Inverse Efficiency Score
IMD	Index of Multiple Deprivation Quintile
IPS	Intraparietal Sulcus
IQ	Intelligence Quotient
IVH	Intraventricular Haemorrhage
LD	Learning Disabilities
M	Male
MD	Mathematical Difficulties
MLD	Mathematics Learning Disability

MRI	Magnetic Resonance Imaging
NDE	Numerical Distance Effect
PCA	Principal Component Analysis
PDP	Preterm Development Project
PRI	Perceptual Reasoning Index
PSI	Processing Speed Index
PVL	Periventricular Leukomalacia
RLD	Reading Learning Disability
ROP	Retinopathy of Prematurity
RT	Reaction time
SES	Socioeconomic status
SLD	Specific Learning Disorder
STEM	Science, Technology, Engineering and Mathematics
Tea-Ch	Test of Everyday Attention for Children
UCL	University College London
UCLH	University College London Hospital
VCI	Verbal Comprehension Index
VP	Very Preterm
w	Weber Fraction
WIAT-II	The Wechsler Individual Achievement Test, 2nd edition
WIAT-MS	WIAT Mathematics scales
WISC – IV	Wechsler Intelligence Scale for Children - Fourth UK Edition
WM	Working Memory
WMI	Working Memory Index

1 General Introduction

Mathematical skills are crucial for our daily lives. We use a range of numerical skills in simple daily tasks such as comparing costs of products, calculating change and measuring the amount of a certain ingredient to cook a recipe. More complex numerical skills are also required regularly, such as calculating monthly bills, estimating the amount of time to travel to a distant place, understanding a chart displayed in a newspaper showing the inflation rate in the local economy, and estimating the costs to renovate a house. In our society today, a range of mathematical abilities are essential for graduation and professional success.

We use our numerical skills is a variety of ways. For instance, we can compare two different portions of berries, choosing which one has the larger quantity without counting them. This mechanism relies on a very rudimentary cognitive system that allows us to approximate different quantities without employing symbolic numbers and is a language-independent system. Even infants and animals share this naturalist way to discriminate quantities. On the other hand, we are also able to compare quantities in a very precise way. We can easily identify an increase in the cost of a daily consumable product just by comparing £4.51 to £4.52. Our symbolic numerical system is a language-dependent system and requires instruction. Together, the approximate and exact systems used by humans to compare quantities I refer to as number processing, and this is the focus of this work.

Pre-, peri- and neonatal risk factors can affect the human ability to process number information. One population at risk of having difficulties in numerical skills is individuals born prematurely (<37 weeks of gestational age) (Kiechl-Kohlendorfer, Ralser, Pupp Peglow, Pehboeck-Walser, & Fussenegger, 2013). The lower the gestational age the higher the prevalence of the difficulties with numerical skills. Generally, very preterm (VP) individuals (<32 gestational weeks) are at significantly increased risk of having those difficulties (Tatsuoka et al., 2016). Difficulties related to numerical skills range from more rudimentary number abilities, such as imprecise approximate number representation (e.g., comparing arrays of dots) to more complex tasks, such as arithmetic (Thevenot, Chazoule, Masson, Castel, & Fayol, 2016). Even individuals without severe and moderate cognitive deficits, such as neurodevelopmental delays/ intellectual disability, have difficulties in a range of numerical abilities including basic number processing skills (Libertus, Forsman, Adén, & Hellgren, 2017).

While several studies have attempted to characterise the cognitive profile of preterm individuals struggling with numerical skills (e.g., Johnson et al., 2009; Johnson, Wolke, Hennessy, & Marlow, 2011; O'Reilly, Johnson, Ni, Wolke, & Marlow, 2020), less attention has been given to the cognitive precursor of formal mathematical skills in this population. Known as the Approximate Number System (ANS), it is our rudimental system that helps us to approximate quantities (Gallistel & Gelman, 1992; Dehaene, 2011). The ANS might be thought of as the starting point for the acquisition of maths abilities (Libertus, Feigenson, & Halberda, 2013). Consequently, deficits in the ANS might lead to difficulties in the achievement of math abilities. Understanding the ANS and its relationship to math abilities is important, since it offers the opportunity to identify early children at risk of developing difficulties in maths later on.

Mathematical proficiency is likely to emerge from a combination of domain-specific and domain-general skills (Costa, Nicholson, Donlan, & Van Herwegen, 2018). Domain-specific mathematical skills are described as exclusively relevant for learning mathematics per se. These include basic quantitative skills, such as counting, number fact knowledge and calculation skills; accurate numerical representations, such as digit recognition, and performance on magnitude comparison and number line tasks (Gilmore et al., 2018). Domain-general skills, meanwhile, are defined as skills that are relevant for all cognitive learning and, in the case of mathematical performance, include predominately working memory, executive function and visuospatial skills (Gilmore & Cragg, 2018). Thus, it is crucial to investigate both domain-general and domain-specific skills in the at-risk population, since both contribute to the development of their mathematical skills.

The number of studies investigating the approximate and exact systems in the preterm population is limited, with most focusing on school-aged children (Hellgren, Halberda, Forsman, Ådén, & Libertus, 2013; Simms, Gilmore, et al., 2013b; Guarini et al., 2014; Libertus et al., 2017). In addition, not all studies comprehensively evaluate domain-general skills that are fundamental for mathematical performance (e.g., Hellgren et al., 2013; Guarini et al., 2014). More work is needed to understand how numerical cognition develops between infancy and early childhood in this population at risk of having mathematical difficulties. In particular, more insight should be gained into the early development of numerical abilities in this population and the underlying neural and cognitive mechanisms of number processing. This thesis, therefore, aims to investigate the cognitive precursors of formal mathematical skills in a very preterm population, and the interplay with a set of domain-general skills and certain domain-specific skills through a series of cross-sectional studies that employ behavioural and neuroimaging measures to evaluate both infants and school-aged children.

Firstly, in Chapter 2, I review the literature surrounding preterm birth and the common causes that might lead to academic failure, especially in the mathematical domain. I then present an overview of the development of mathematical trajectories and the interplay of general- and specific-domains on mathematical performance in typically developing children. Finally, I review the current literature surrounding numerical skills in the preterm population. Against this background, I indicate the gaps in knowledge and my specific aims in the subsequent empirical chapters. In Chapter 3, I describe the general methodology chosen in this work. In chapters 4 and 5, I investigate the early stages of numerical sensitivity in six- and twelve-month-old VP infants, respectively. In addition, in chapter 5, I also explore the association between visual working memory, which is an important domain-general skill for later mathematical performance, and numerical sensitivity in VP infants. In Chapter 6, I investigate domain-general skills, and a set of domain-specific skills, and their association with mathematical performance in VP school-aged children. The domain-general skills that I investigate include intelligence and executive functions such a processing speed, working memory, attention, planning and inhibition. The domain-specific skills, meanwhile, focus on numerical magnitude comparison, an ability that requires participants to decide which of two numerosities is the largest, using either symbolic (Arabic digits) or non-symbolic (arrays of dots) quantities. In Chapter 7, I examine the neural mechanisms associated with symbolic and non-symbolic magnitude comparison tasks in very preterm children, employing event-related potentials (ERPs). Finally, in Chapter 8, I discuss the main findings of this work, its applications and potential future directions in the field of numerical cognition in the preterm population.

Taken together, this thesis aims to enhance knowledge of the cognitive mechanisms of number processing and their interplay with executive functions in infants and children born prematurely. A better understanding of the underlying mechanisms of number cognition and its relationship to executive functions might help in the early identification of those particular individuals born prematurely who are at risk of experiencing difficulties with mathematics, thereby helping to target interventions sooner and more closely.

2 Literature Review

This chapter aims to characterise prematurity and the impact of preterm birth on the abilities crucial to the development of numerical skills and mathematical achievement. Firstly, I present an overview of the literature surrounding outcomes following preterm birth, and on the different causes and risk factors that might affect cognitive development. Next, I review studies investigating cognitive outcomes following preterm birth. Then, I provide an overview of the education outcomes in individuals born prematurely. To finalise this section, I provide the cognitive, behaviour and educational phenotype of individuals born prematurely.

In the subsequent section, I review the development of mathematical abilities in typically developing children. This section provides a perspective of what is expected in the typical development of numerical cognition. Current cognitive models and studies investigating typically developing children using both behavioural and brain imaging are discussed.

Having provided an overview of the typical development of mathematical abilities, in the next section I discuss what can go awry in an atypical developing population with mathematical difficulties. I give an overview of studies investigating numerical skills in children with mathematical difficulties, including individuals born prematurely. This allows the difficulties related to numerical skills in the preterm population to be characterised, pointing to directions of future work, including the work of this thesis.

Taken together, these overviews provide a background to the building blocks of numerical skills in typically developing children, and what is known about the development of numerical skills in children born prematurely. Against this background, it will become apparent how mathematical achievement is the academic area most affected in the preterm population and yet little is known about the development of their numerical skills, the interplay of domain-general and domain-specific abilities during infancy, and neural underpinnings of numerical skills in this at-risk population of having mathematical difficulties. Finally, the aims addressed in the subsequent empirical chapters of this thesis are presented.

2.1 Prematurity

Preterm birth is defined as childbirth occurring at less than 37 completed weeks of pregnancy, and further classified as late, moderate, very and extremely preterm, occurring at 34-36, 32-33, 28-31, and ≤27 weeks, respectively[1] (see Figure 2.1) (Quinn et al., 2016). It is estimated that 15 million babies are born prematurely worldwide annually, representing approximately 11% of all births (Chawanpaiboon et al., 2019). Approximately 85% of premature infants are born moderate to late, 10% are **very preterm (VP)** and 5% are born **extremely preterm (EP)** (WHO, 2017). With sophisticated neonatal intensive-care facilities, more than 95% of infants born before 28 weeks of gestation survive in developed countries. In contrast, only 10%, or less, of these infants survive in developing countries (Blencowe et al., 2013).

In the UK, 80,000 infants are born prematurely with an estimated annual societal economic burden of £1 billion (Mangham, Petrou, Doyle, Draper, & Marlow, 2009). Although risk factors for preterm birth have been associated with a wide range of conditions, such as

[1] Historically, birth weight has been used to identify prematurity, with definitions of low, very low, extremely low birth weight (birth weight of 2500g, <1500g, and <1000g, respectively). With more accurate ultrasound techniques to measure size in early stages of pregnancy, calculation of gestational age (GA) is now being used more frequently and accurately in clinical practice (Johansson and Cnattigius, 2010) and is considered the gold standard for GA assessment (Vogel et al., 2018). For this reason, GA will be used in preference to birth weight in this work.

history of previous preterm delivery, low-socio-economic status, low educational attainment, infection, maternal age, multiple pregnancies and tobacco use (Frey & Klebanoff, 2016), two-thirds of preterm births occur without any evident risk (Vogel et al., 2018).

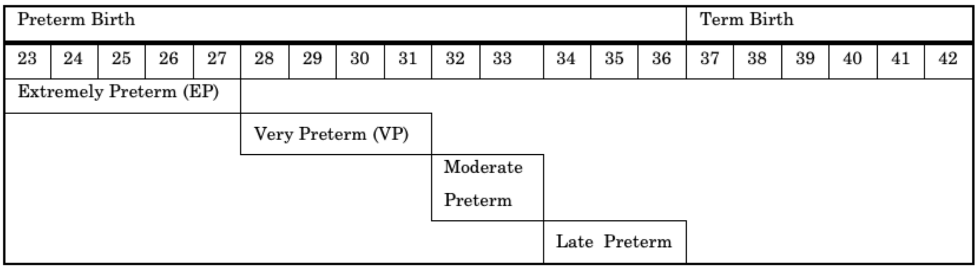


Figure 2.1: Categorisation of gestational ages by completed gestational weeks at birth.

The representation of gestation ages usually starts at 23 weeks due to the infants' viability. About 50% of babies born between 23 and 24 weeks may survive premature birth. To date, the youngest preterm infant to survive was born at only 21 weeks and 6 days of gestation.

2.1.1 Long-term effects of prematurity

The mortality rate for preterm infants has improved considerably over the last two decades (Glass et al., 2015). Nonetheless, individuals born prematurely remain vulnerable to many complications. The focus of recent research into the impact of prematurity has therefore been to understand and improve long-term outcomes of individuals born very prematurely (Wolke, Johnson, & Mendonça, 2019). Infants who do survive have higher rates of long-term morbidity, including neurologic and developmental disabilities, compared to infants born full-term (Frey & Klebanoff, 2016). Perinatal complications are inversely associated with gestational age. The shorter the gestation, the smaller the infant and the higher the risk of mortality, morbidity and disability. In particular, very and extremely preterm infants frequently experience numerous perinatal complications, whereas late and moderately preterm infants are generally spared (Glass et al., 2015). Medical problems are especially prevalent among infants born at a gestational age (GA) of 23-25 weeks (Blencowe et al., 2013). It is well established that preterm birth is associated with negative neurodevelopmental consequences, and that children born very and, in particular, extremely preterm, are most at risk (Serenius et al., 2016; Pascal et al., 2018). Thus, the degree of prematurity is a major factor in determining later development and studies often distinguish between different degrees of prematurity.

2.1.2 Neurodevelopmental disorders and prematurity

Prematurity is a leading cause of neurodevelopmental disorder (Blencowe et al., 2013). Neurodevelopmental disorders are a group of disorders in which the development of the central system is disturbed. This can include developmental brain dysfunction, which can manifest as impaired motor function, learning, language or/and non-verbal communication (Mullin et al., 2013). Neurodevelopmental disorders associated with prematurity generally include neurosensory impairments (visual and hearing loss), cerebral palsy and developmental delay in younger children, or intellectual impairment in older children (Burnett, Cheong, & Doyle, 2018). Among children born very preterm, 5 to 10% have major motor deficits, including cerebral palsy, and more than half have significant cognitive, behavioural or sensory deficits (Back & Miller, 2014). A recent meta-analysis revealed that 42% of infants born at 22 weeks GA show a moderate to severe neurodevelopmental disability; 41% at 23; 32% at 24; 23% at 25 weeks, showing a significant decrease of risk of neurodevelopmental disorders as GA increases (Ding, Lemyre, Daboval, Barrowman, & Moore, 2019).

Sequelae associated with preterm birth causing neurodevelopmental disorders suggest that the very preterm brain is susceptible to a range of injuries during the neonatal period. Intraventricular haemorrhage (IVH) and periventricular leukomalacia (PVL) are frequent brain insults that preterm infants might experience, with 25% of infants born prematurely experiencing brain injury (Gale et al., 2018). The former refers to bleeding originating in the germinal matrix, a highly vascularised foetal structure that lies at the base of the lateral cerebral ventricles. Injury to this structure may occur due to large fluctuations in blood pressure in immature blood vessels, potentially resulting in intraventricular haemorrhage (IVH) and ventricular dilatation. In severe cases, periventricular haemorrhagic infarction may also occur, which is commonly associated with the later development of cerebral palsy. The latter consists of necrosis of white matter related to hypoxia and inflammation, leading to decreased myelination, axonal degeneration and cortical volumes (Volpe, 2009). Figure 2.2 illustrates the development of cortical folding in the premature brain susceptible to brain injuries.

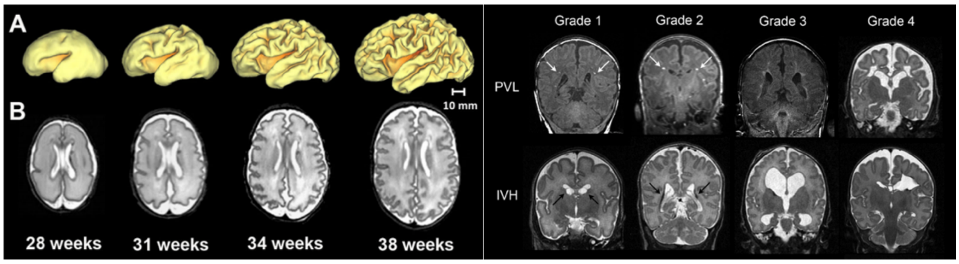


Figure 2.2: VP brain

The picture on the left-hand side illustrates the development of cortical folding in the premature human brain at different gestational ages in (a) three-dimensional surfaces and (b) axial T2-weighted images. The picture on the right-hand side exemplifies degrees of two frequent brain injuries experienced in infants born preterm: periventricular leukomalacia (PVL) and intraventricular haemorrhage (IVH) (Smyser, Kidokoro, & Inder, 2012).

2.1.3 Identifying neurodevelopmental disorders in high-risk preterm infants

2.1.3.1 Brain imaging

Due to the high incidence of neurodevelopmental impairments in the preterm population, neuroimaging techniques are employed in neonates at term equivalent age both to identify cerebral injury and its severity and to predict adverse neurodevelopmental outcomes. Cranial ultrasounds are widely used in neonatal units, although magnetic resonance imaging (MRI) is increasingly being employed. MRI offers several advantages when compared to other neuroimaging techniques, such as increased sensitivity to subtle tissue injury and detection of altered cerebral development such as global and regional reductions in cerebral growth (Inder, Wells, Mogridge, Spencer, & Volpe, 2003).

Abnormalities in the preterm brain seen on MRI include white matter abnormalities, such as cystic lesions, punctate lesions, delayed myelination, volume loss, thinning of the corpus callosum, and T2-weighted diffuse excessive high signal intensity. Delayed cortical folding, widening of the interhemispheric fissure and enlarged extracerebral space are also relatively common. Basal ganglia and thalamic lesions have been reported in neonatal MRI studies of VP infants, although relatively rarely (Anderson, Cheong, & Thompson, 2015).

Numerous studies using MRI in neonates born prematurely at term age equivalent demonstrated reasonable associations between brain abnormalities and later neurodevelopmental disorders (Nguyen The Tich et al., 2011; Brouwer et al., 2017; Balakrishnan et al., 2018). One of the first studies to examine the relationship between neonatal MRI abnormalities and early neurodevelopmental outcome demonstrated that delayed outcomes were significantly associated with white matter abnormalities and hydrocephalus in premature infants at 12-18 months of age (Miller et al., 2005). Woodward et al. (2006) investigated the associations between brain abnormalities at term equivalent age and the risks of neurodevelopmental impairments at two years of corrected age[2] in 167 very preterm children. Increased severity of white-matter abnormalities in MRI at term equivalent age was found to be associated with poorer performance on the cognitive and psychomotor scales, as well as with higher risks of severe cognitive delay, severe motor delay, cerebral palsy and neurosensory impairment (Woodward, Anderson, Austin, Howard, & Inder, 2006).

Although MRI is a potential tool to predict later neurodevelopmental outcomes in the preterm population, the technique has few limitations. For example, it has low sensitivity and high specificity for various sub-scores in predicting different outcomes, limiting the use of MRI as a screening tool. A meta-analysis has revealed that moderate and severe white-matter abnormalities seen on MRI performed around term equivalent age only predicted cerebral palsy and motor function in very preterm neonates with moderate sensitivity (72%) and specificity (62%). Other long-term outcomes such as neurocognitive and behavioural impairments had low sensitivity and high specificity for performance predicting motor, cognitive or language outcomes (van't Hooft et al., 2015). A significant proportion of preterm infants therefore experience cognitive impairments despite apparently typical brain imaging at term equivalent age (TEA), suggesting that subtler anomalies underlie these difficulties (Volpe, 2009).

---

[2]Studies investigating infants born prematurely use the term 'corrected age' or 'adjusted age' to refer to infants' age based on due date. The use of correction reflects a maturation perspective, while chronological age represents an environmentally based orientation. The consensus is that correction should occur up to two years of age (Aylward, 2020).

2.1.3.2 Developmental assessments

The British Association of Perinatal Medicine recommends developmental assessments up to two years adjusted for high-risk preterm children born under 32 weeks' gestation or with a birth weight below 1500g to monitor the long-term sequelae of preterm birth (BAPM, 2019). The National Institute for Health and Care Excellence suggested developmental assessments up to four years corrected age for children born under 28 weeks' gestation (NICE, 21017). Developmental assessments aim to monitor the long-term consequences of preterm birth, including developmental delays. Developmental delay occurs when a child does not achieve developmental milestones in comparison to peers of the same age range. The degree of developmental delay can be further classified as mild (functional age < 33% below chronological age), moderate (functional age 34%–66% of chronological age) and severe (functional age < 66% of chronological age) (Mithyantha, Kneen, McCann, & Gladstone, 2017). A significant delay is defined as performance that is two or more standard deviations below the mean on age-appropriate standardised norm-referenced testing (Bellman, Byrne, & Sege, 2013). The delay can be in a single domain (i.e. isolated developmental delay) or more than one domain. A significant delay in two or more developmental domains affecting children under the age of five years is termed global developmental delay (GDD) (Bellman, Byrne, & Sege, 2013; Choo, Agarwal, How, & Yeleswarapu, 2019). It should be noted, however, that the term developmental delay has been challenged because it suggests that the child may 'catch up', which is often not the case (Williams & Essex, 2004)

It is common clinical practice in developmental assessments to use adjusted or corrected age for preterm infants and children to account for their prematurity. Corrected or adjusted age is calculated by subtracting the number of weeks a child was born prematurely to her/his chronological age. The theoretical basis underpinning this practice is the biological and maturational perspective, which suggests that early development continues after conception either intra- or extra uterine. Correction for the degree of prematurity was therefore devised to account for the transient developmental gap that preterm infants experience (Harel-Gadassi et al., 2018).

Only the most severe impairments benefit from developmental assessments since their predictive value for an individual is relatively poor (Marlow, 2015). Wong et al. (2016) demonstrated that approximately 30% of children with mild cognitive impairment, and nearly 50% with mild communication impairment, would be classified as having no impairment in developmental assessments, showing that a substantial number of children with clinically relevant development impairments are likely to go unrecognised (Wong, Santhakumaran, Cowan, & Modi, 2016).

In the apparent absence of developmental impairments, and the presence of a typical brain imaging in early childhood, milder cognitive deficits may emerge in later childhood. A high proportion of infants born prematurely might initially perform appropriately, but difficulties may become more obvious and debilitating as the child grows older, with deficits most frequently reported when children reach school age in conditions described as 'hidden disabilities' (Whitfield, Grunau & Holsti, 1997). These are considered 'high prevalence, low severity' deficits as they affect as many as 50-70% of very preterm infants but are considered to be low severity based on the notion that the deficits are not at the level of intellectual disabilities (Aylward, 2005; Foster-Cohen, Friesen, Champion, & Woodward, 2010; Johnson, Wolke, & Marlow, 2008; Kilbride, Aylward, & Carter, 2018). These deficits include difficulties in distinctive cognitive domains contributing to later problems associated with academic performance. Although not considered severe deficits, such deficits should not be underestimated in terms of their functional impact since, despite appearing reasonably stable during development, they persist into adolescence and young adulthood (e.g., O'Reilly et al., 2020).

Next, I review cognitive, educational and behavioural outcomes following preterm birth.

2.2 Cognitive outcomes following preterm birth

Cognitive difficulties are the most common problem among preterm children (Hack et al., 2002; Hack, 2009; Jaekel, Baumann, & Wolke, 2013). A wide range of cognitive deficits have been reported, including attentional problems, working memory deficits and reduced processing speed (Johnson, Fawke, et al., 2009; Woodward et al., 2009). Difficulties can range from mild through to severe and are not consistent across preterm cohorts. Gestational age at birth correlates highly with cognitive performance, hence extremely and very preterm children have more adverse deficits in cognitive function than moderate and late preterm children (Allotey et al., 2018).

Next, I present a summary of studies reviewing different domains affected by prematurity. These comprise intelligence, executive functions, attention and processing speed. Subsequently, I provide an overview of educational outcomes affected by prematurity. Finally, based on previous reviews, I present the cognitive, behavioural and educational phenotype associated with prematurity. This will illustrate the heterogeneous profile of difficulties faced by this population, which will, in turn, be important in illustrating how these domains, especially cognitive domains, can affect individuals born prematurely, including in respect to their numerical skills, the focus of this work.

2.2.1 Intelligence

While many children born VP score in the normal range (within 1 SD of the normed mean) on standardised tests for measuring intelligence, comparisons with their full-term (FT) peers reveal that their average group scores are lower. Individuals born prematurely at various gestational ages and different stages of life consistently perform worse than their term-born peers in all intelligence domains (performance IQ, verbal IQ and Full-Scale IQ). Specifically, VP individuals perform approximately 0.8 SD below their term-born peers (Allotey et al., 2018), corresponding to approximately 12 IQ points lower (Kerr-Wilson, MacKay, Smith, & Pell, 2012; Brydges et al., 2018). Discrepancies in the cognitive performance of groups of children born very preterm may persist even when children with more severe impairments are excluded from analyses, scoring at least an average of 5 IQ points below term-born controls (Saigal, Hoult, Streiner, Stoskopf, & Rosenbaum, 2000).

Deficits in general cognitive function identified during infancy persist into early adulthood among individuals born extremely preterm, with no evidence of either substantial recovery or deterioration. A recent population-based cohort study demonstrated that scores obtained at four different time points (2.5, 6, 11 and 19 years old) were on average 25.2 points (nearly 2 SD) below their term-born peers and remained significantly lower at every assessment (Linsell et al., 2018).

2.2.2 Executive Functions

Many different definitions of executive functions (EF) are given in the literature. Most often, definitions of EF include the notion that it is an umbrella term for a set of cognitive processes that are important for active and purposeful regulation of thought, emotion and behaviour (Anderson, 2008; Diamond, 2013). There is general agreement that there are three core EFs: inhibition, working memory and cognitive flexibility. Inhibition is often described as the ability to withhold a prepotent response or thought. Inhibition skills include inhibitory control, including self-control (behavioural inhibition) and interference control (selective attention and cognitive inhibition). Working memory is the ability to hold information in short-term memory and manipulate that information at the same time (Baddeley, 1992). Cognitive flexibility, also called set shifting, or mental flexibility, is the ability to switch in a flexible manner between different tasks or different rules in the same task. From these, higher-order EFs are built such as reasoning, problem solving, and planning' (Diamond, 2013). Planning is the ability to develop strategies to reach a future goal (Shallice, 1982).

There is a growing consensus that children born very and extremely preterm are at risk of executive deficits, over and beyond the risk for lower general cognitive ability (e.g., Aarnoudse-Moens, Smidts, Oosterlaan, Duivenvoorden, & Weisglas-Kuperus, 2009; Taylor & Clark, 2016). Children with more pronounced degrees of preterm birth are at risk for reduced brain volumes in the structures associated with EF, including cerebral white matter, frontal, parietal and temporal cortices, and the basal ganglia and cerebellum. Not surprisingly, these children have a wide range of deficits in EF compared to term-born controls, with deficits proportional to the degree of prematurity (Taylor & Clark, 2016).

Aarnoudse-Moens et al. (2009) revealed that very preterm children had lower scores than their term-born peers on tasks measuring verbal fluency (-0.57 SD), working memory (−0.36 SD) and cognitive flexibility (−0.49 SD). In line with these results, Mulder et al. (2009) showed small to moderate effect sizes for semantic verbal fluency when comparing preterm children and their term-born counterparts. Differences found between preterm and term children in measures of inhibition, planning and phonemic fluency were influenced by age at assessment and/or gestational age at birth. Spatial working memory tasks used with preschool children showed that this domain is affected in preterm compared to term-born children (Mulder, Pitchford, Hagger, & Marlow, 2009).

On average, the performance of children born preterm on measures of executive control is 0.3–0.6 SD lower than that of full-term children, although effects as great as 1 SD have been reported in studies of children with extreme prematurity. Impairments in EF are found even when excluding children with low IQs or when adjusting group comparisons for IQ (Aarnoudse-Moens et al. 2009; Mulder et al. 2009).

2.2.3 Attention

Attention is another cognitive domain with sub-components, consisting of the capacity for selective focus (i.e. focusing on relevant stimuli and ignoring distracting stimuli), sustain (i.e. maintaining an alert state for an extended period), encode (i.e. holding information in temporary storage), shift (i.e. fluently transferring focus from one activity to another), and divide attention (i.e. focusing on multiple competing stimuli simultaneously) (Anderson, 2014). Numerous studies have investigated attention domains in preterm infants and pre-schoolers.

A meta-analysis exploring differences in selective, sustained and shifting attention domains between preterm and term children aged >2 years found that individuals born prematurely scored generally below their full-term-born peers in all sub-domains, with sustained attention being the domain most affected in children born extremely preterm (Mulder et al., 2009). Preterm children performed 0.4 SD below term-born peers in tasks evaluating selective attention, and deficits were greater (0.6 SD) for those born <26 weeks of gestational age. More variability was observed for sustained attention, but a moderate effect size of 0.5 SD was still revealed by the meta-analysis in favour of term children, and this increased to 0.7 SD when restricted to children born at <26 weeks of gestation. Conflicting results were found in the meta-analysis for shifting attention, with a marked deficit seen in preterm children on tasks measures by the Trails B task (0.5 SD), but no group difference was observed on sorting tasks (Mulder et al., 2009). Wilson-Ching et al. (2013) reported deficits in selective, shifting and divided attention, but performance in a measure of sustained attention was similar to matched term controls. Overall, the results imply that attention and its subdomains is another cognitive area affected by prematurity (Wilson-Ching et al., 2013).

2.2.4 Processing Speed

Processing speed refers to the time required to interpret and respond to incoming stimuli or information, and is assessed by measures of reaction time and decision time (Anderson, 2014). Specific deficits in processing speed have been previously reported for infants and children born preterm (Rose, Feldman, & Jankowski, 2002, 2011; Mulder, Pitchford, & Marlow, 2011). A meta-analysis reported poorer processing speed at different times in life in individuals born preterm. Preterm children at primary school scored 0.53 SD below their term-born peers and, at secondary school, the preterm group scored 0.35 SD lower than their term-born peers with the intermediate effect continuing after school age (Allotey et al., 2018). Even studies investigating processing speed in infants have reported slower processing speed in this population. Rose et al. (2002) investigated processing speed in preterm infants revealing that they might need as much as 30% more inspection time to perform compared to term controls (Rose, Feldman, & Jankowski, 2002). These findings were in line with a recent study from our research group revealing that infants born prematurely had slower processing speed, and were slower to shift attention and spent less time attending to or fixating on a target (Downes, Kelly, Day, Marlow, & de Haan, 2018). Together, the results suggest that processing speed is another domain affected by prematurity.

2.3 Educational outcomes following preterm birth

2.3.1 Learning Disabilities & Learning Difficulties

With a high risk of having cognitive deficits, extremely and very preterm individuals face another issue: learning disabilities and learning difficulties. In a research context, learning disabilities is frequently defined as low attainment scores <-2 SD/ or < 70 points on academic measures in comparison to the term reference group (e.g., Akshoomoff et al., 2017). LD are more frequent among EP and VP children, in comparison to term-born children, even when excluding those with neurodevelopmental disorder (Pritchard et al., 2009; Saroj Saigal et al., 2003; Litt, Taylor, Klein, & Hack, 2005; Johnson, Hennessy, et al., 2009; Pritchard et al., 2009; Hutchinson, De Luca, Doyle, Roberts, & Anderson, 2013). The estimated rates of LD in the preterm population are slightly different according to studies, mainly due to the type of study (e.g., population-based cohort and birth-cohort) and gestational age investigated (e.g., very and extremely preterm). A population-based cohort study in the UK has estimated that 2% of the extremely preterm population have specific reading learning disability (RLD), 14% specific mathematics learning disability (MLD) and 43% combined RLD and MLD, after excluding children with intellectual disabilities (Johnson et al., 2016). Slightly different rates of learning disabilities were demonstrated in a birth-cohort study of extremely preterm children, with estimated rates of 6.4% for RLD, 16.2% for

MLD and 8.3% for combined RLD and MLD (Akshoomoff et al., 2017). Nonetheless, these studies estimate higher rates of learning disabilities in this population than the rates presented by their term-born-counterparts, with mathematics learning disability being consistently more prevalent than reading learning disability. As such, extremely preterm children attend more special schools (Saigal et al., 2003; Johnson et al., 2009), are 2.85 times more likely to have special educational needs (Twilhaar, De Kieviet, Aarnoudse-Moens, Van Elburg, & Oosterlaan, 2018) and have higher chances of school repetition (Larroque et al., 2011; Hutchinson et al., 2013). Furthermore, the chances of very and extremely preterm children dropping out of school are increased with decreasing gestational age, with this only being partly explained by neurodevelopmental disorders (Mathiasen, Hansen, Nybo Andersen, Forman, & Greisen, 2010). Although the preterm population has higher rates of learning disabilities and SEN, and higher chances of repeating a grade and/or failure to complete basic school, it is reassuring that children born prematurely are typically found in mainstream education (Kelly, 2016).

Difficulties in different academic domains are also more prevalent in children born prematurely. Several studies have consistently demonstrated that extremely and very preterm children are more likely to have lower academic achievement[3] across all subjects, with mathematical achievement being the domain most affected (Saigal et al., 2000; Chyi, Lee, Hintz, Gould, & Sutcliffe, 2008; Pritchard et al., 2009; Cornelieke Sandrine Hanan Aarnoudse-Moens, Oosterlaan, Duivenvoorden, Van Goudoever, & Weisglas-Kuperus, 2011; Johnson et al., 2011; Mulder et al., 2011; Taylor et al., 2011; Akshoomoff et al., 2017). One of the first population-based cohort studies in the UK to investigate educational outcomes

[3]Low academic achievement or learning difficulties is defined here as a standard score <85 on academic measures in comparison to the term reference group.

in early adolescents born prematurely showed that, at 12 years old, VP children performed poorly in all educational assessments compared to term-born, with marked differences in mathematics beyond any other domain, even after controlling for IQ (Botting, Powls, Cooke, & Marlow, 1998). A more recent population-based cohort in the UK and Ireland supported these results, showing that extremely preterm children had significantly lower scores at 11 years old than classmates for reading (-18; CI -21 to -15) and mathematics (-27; CI -31 to -24), measured by the Wechsler Individual Achievement Test 2nd Edition (WIAT-II^{UK}) (Johnson, Hennessy, et al., 2009). Other studies supported those results, with extremely preterm children presenting 1.5 times higher risk for low math achievement than the risk for low reading achievement, with rates of 27% and 17% respectively (Akshoomoff et al., 2017). A recent meta-analysis, including 17 eligible studies comprising 2390 preterm children and 1549 controls, revealed that preterm children scored 0.71 SD below their term-born peers on arithmetic measures, showing a medium effect, in contrast to 0.44 and 0.52 SD for reading and spelling skills with small and medium-sized effects, respectively (Twilhaar et al., 2018). These results are in line with a previous meta-analysis showing a 0.60 SD deficit in mathematics scores compared with a 0.48 SD deficit in reading (Aarnoudse-Moens et al., 2009). These differences persist even after controlling for IQ (Johnson et al., 2009; Johnson et al., 2011) or excluding children who have neurosensory impairments (Pritchard et al., 2009). A review of case-control studies investigating differences between mathematical difficulties in preterm-children and their term-born counterparts shows moderate to large differences in effect sizes between VP and term-born children in standardised mathematics tests, with the greatest effect found for children who were born <26 weeks. Similar effect sizes have also been observed using curriculum-based measures and teacher reports (Simms et al., 2013b).

Already in preschool, the performance of preterm children in mathematical skills is noticeably lower than their term-born peers (Pritchard et al., 2009; Taylor et al., 2011). These difficulties appear to persist during primary school, with preterm children performing on average 0.53 SD lower on maths skills through the course of primary school (Twilhaar et al., 2018), suggesting no 'catch-up' effect. In contrast, a 'catch up' effect was found in a longitudinal study suggesting a gain of 0.10 SD in each progressive measure employed in the study. Differences in the methodology used to assess academic achievement might explain the discrepancies found between the studies. Whereas Twilhaar et al. (2018) employed typical standardised measures to assess academic achievement, Odd et al. (2013) obtained routine educational assessments mandatory in state schools in England. Nonetheless, both studies show discrepancies in academic achievement between preterm children and their term-born counterparts, although the 'catch-up' effect remains debatable (Odd, Evans, & Emond, 2013).

Furthermore, having low academic achievement has been associated not only with educational failure, but also lower adult income. A population-based cohort following 903,402 infants to adulthood revealed associations between preterm birth and adult outcomes, such as education level attained, income, receipt of Social Security benefits and the establishment of a family, even after excluding individuals with medical disabilities, such as CP (Moster, Lie, & Markestad, 2008). Another study showed specific links between preterm birth and low adult wealth, mediated by poor academic abilities and educational qualifications, in particular, mathematics attainment in middle childhood (Basten, Jaekel, Johnson, Gilmore, & Wolke, 2015). In spite of those facts, adults who were born very and extremely preterm have independent and self-supportive lives and similar rates of employment to term-born adults (Johnson & Marlow, 2017).

It should be noted, however, that approximately 2% of children born extremely prematurely were considered gifted in a large population-based cohort, in comparison to 33.5% of the same clinical population with learning difficulties. This illustrates that although prematurity markedly affects academic attainment on average, specifically math skills, a small portion of extremely preterm children have outstanding academic skills (Garfield et al., 2017).

2.4 The cognitive, behavioural and educational phenotype of preterm children

Having difficulties across various cognitive domains, the cognitive profile of those born preterm might not be apparent. Few studies have attempted to characterise the cognitive profile of preterm children.

Core deficits in working memory and visuospatial skills have been shown to mediate performance in executive function and intelligence tests (Mulder et al., 2009). These core deficits may also be implicated in attention difficulties (Nadeau, Boivin, Tessier, Lefebvre, & Robaey, 2001; Shum, Neulinger, O'Callaghan, & Mohay, 2008; Anderson et al., 2011; Mulder et al., 2011)

Preterm children are also at increased risk of behavioural and socioemotional difficulties. Children born extremely preterm may have attention, emotional or peer relationship problems that do not meet diagnostic criteria, but which may nonetheless have an impact on daily function. Some studies have suggested that the preterm behavioural phenotype consists of inattention, anxiety and social difficulties. Early attention and regulatory problems are evident in the preschool years and, by childhood, the greater specificity in outcomes points to a cluster of inattention, peer relationship problems and emotional symptoms (Johnson & Marlow, 2011, 2014; Mathewson et al., 2017).

It has been suggested that academic difficulties become more pronounced with age in VP children, perhaps due to the increasing complexity of tasks, or to the cumulative effects of early problems (Simms et al., 2013a). Cognitive difficulties observed in children born preterm are probably the result of a developmental cascade originating in early development (Oudgenoeg-Paz, Mulder, Jongmans, van der Ham, & Van der Stigchel, 2017), affecting later their academic life, with individuals born prematurely being commonly reported to lag in maths (Simms et al., 2013a). Figure 2.3 illustrates the impact of various long-term outcomes of prematurity (Wolke et al., 2019). Several domains are negatively affected by prematurity, including executive functions and academic attainment. Those difficulties seem to persist throughout adulthood.

Figure 2.3: Long-term outcomes associated with prematurity.

Approximate effect sizes for the impact of very preterm birth on long-term outcomes. Outcomes assessed in childhood and adolescence are shown in orange, and outcomes assessed in adulthood in blue. Deficits in executive functions and difficulties related to mathematics are some of the domains negatively associated with prematurity during childhood. Abbreviations: ADHD, attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; CP, cerebral palsy; DCD, developmental coordination disorder; HRQoL, health-related quality of life; IQ, intelligence quotient; SEN, special educational needs (Wolke et al., 2019).

Despite the accumulating evidence that preterm children have difficulties in formal mathematical abilities, little is known about the developmental trajectories of mathematical cognition and how these skills emerge in individuals born prematurely. Studying the early stages of typical and atypical development under conditions that can disrupt number deficits helps elucidate the mechanisms of how numerical representations are structured and change over developmental time (Ansari & Karmiloff-Smith, 2002). In populations at risk, it is imperative to understand the very basic processes underlying the sense of numerosity in its early stages, and how these can go awry. The identification of early markers of altered or delayed developmental trajectories in mathematical skills is important because of the potential for early intervention in the preterm population at increased risk of having low mathematical achievement.

Before exploring the development of mathematical abilities in individuals born preternaturally, it is important to understand the typical trajectories of numerical skills. This is the focus of the next section.

2.5 The development of numerical skills

During infancy and early childhood, humans have distinctive development stages of informal preverbal mathematical abilities. **Numerosity**, ordinality and cardinality are stages preceding the ability to perform simple arithmetic. Numerosity, or number sense, is the primitive ability to estimate the number of objects in a set (Dehaene, 1997). Two different systems are involved in the representation of numerosities: the approximate number system (ANS), representing large numerosities (>3), and the object tracking system, representing small numerosities (<3). While the ANS has its distinct proprieties, which are described in greater detail below (see section 2.5.1), the object tracking system relies on perceptual properties to discriminate two sets of small elements. By about 18 months of age, infants show an understanding of simple ordinal relationships, **ordinality** (Geary, 2000), which has been reported in infants as young as nine months (Brannon, 2002). By around 24 to 36 months, children can enumerate a small set of objects correctly in an ability known as **cardinality**. Interestingly, children learn the procedural aspects of counting first, while they still lack the conceptual understanding (Gilmore, Göbel, & Inglis, 2018). Before having a full understanding of counting, children between 2 and 4 years of age learn to count following the five counting principles: 1) **the one-to-one principle** at around 30 months of age, each object can only be counted once, each number word has to be paired with one and only one object and each object can only be paired with one number word, all objects are paired with a number word; 2) **the stable order principle**: the number orders are recited in a fixed order; 3) **the abstraction principle**: any array or collection of sets can be counted; 4) **the order irrelevance principle**: the order in which objects are counted does not matter, any order leads to the same results; 5) **cardinality principle**: the last number in the count sequence also describes how many objects there are in total in the set, so it does not only describe the order of the object, but also the quantity of the whole set (Gelman & Gallistel, 1978). Finally, **simple arithmetic** is acquired early in childhood, including sensitivity to increases (addition) and decreases (subtraction) in the quantity of sets (Geary, 2000).

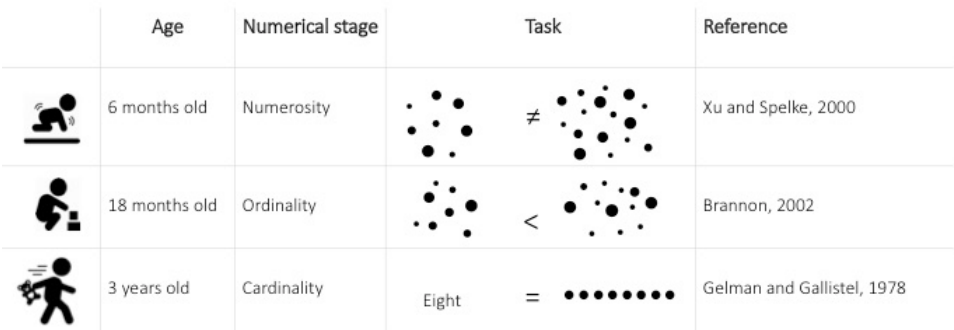


Figure 2.4: First stages of numerical development.

Even one-day-old infants show the ability to discriminate numerosities ((Izard, Sann, Spelke, & Streri, 2009). By six months, infants can discriminate quantities in a 1:2 ratio (e.g., eight *vs.* sixteen elements) (Xu & Spelke, 2000). By eighteen months of age, toddlers have the concept of 'more and less' (Brannon, 2002). By three years of age, children can make one-to-one correspondence, associating numbers of quantities with verbal symbolic numbers (Gelman & Gallistel, 1978).

The acquisition of mathematical skills is dependent on formal education (Geary, 2000). Geary termed these skills in school-taught stages as 1) **number and counting**; 2) **arithmetic**: computations; and 3) **arithmetic**: word problems. Firstly, primary school children are expected to master the counting system (e.g., learning the associated number words), gain an understanding of the base-10 system, and learn to translate, or transcode, numbers from one representation to another (e.g., verbal – 'two hundred ten' – to Arabic – '210'). Subsequently, primary school children are expected to learn the basic arithmetic facts and the computational procedures for solving complex arithmetic problems (e.g., 472+928). With sufficient practice, typically developing children will memorise most basic arithmetic facts. The ability to solve complex arithmetic problems is facilitated by the memorisation of basic facts, the memorisation of the associated procedures, and an understanding of the base-10 system. Finally, primary school children begin to solve simple word problems in kindergarten and key stage one. The primary source of difficulty in solving these problems is identifying problem type (e.g., comparing two quantities *vs.* changing the value of one quantity) and translating and integrating the verbal representations into mathematical representations (Geary, 2000).

Whilst there is some ordering in which numerical skills may emerge, development does not take place in clearly defined linear steps. Children may develop several skills in parallel, and individual children may move through skills in different orders. To reach a full understanding, however, children will need to master each of these skills. Developmental progressions can therefore be seen as approximate paths of the development of thinking, but do not necessarily entail a clear linear progression for all (Gilmore, Göbel, & Inglis, 2018).

2.5.1 The Approximate Number System

A building block for the development of mathematics includes the nonverbal ability to perceive and discriminate approximate large numerosities without counting or numerical symbols. This is supported by the **Approximate Number System** (ANS). The ANS might be thought of as the starting point for the acquisition of math abilities (e.g., Szkudlarek & Brannon, 2017). Thus, understanding the ANS and its relationship to math abilities is important, since it offers the opportunity to identify at early stages children at risk of developing difficulties in maths. Hence, many studies have focused on investigating the role of ANS, its development and neural basis and its association with mathematical performance. I review these studies in the next sections.

The origins of our symbolic number capabilities are thought to be rooted in an approximate, non-verbal representation of number known as **number sense** (Dehaene, 1997). This assumes that our innate approximate number sense guides the process of learning numerical symbols and the development of mathematics (Dehaene & Changeux, 1993; Stanislas Dehaene, 2001; Piazza et al., 2010; Feigenson, Libertus, & Halberda, 2013). The idea that symbolic and non-symbolic representations of numbers are tightly connected is sustained by evidence from correlational studies showing that individual differences in approximate numerical abilities correlate with individual differences in mathematics achievement (Halberda, Mazzocco, & Feigenson, 2008; Piazza et al., 2010; Libertus, Feigenson, & Halberda, 2011; Mussolin, Nys, Leybaert, & Content, 2011; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Bonny & Lourenco, 2013), experimental studies showing how non-symbolic approximate numerical training enhances symbolic mathematical abilities (Park & Brannon, 2013; Hyde, Khanum, & Spelke, 2014;

Park, Bermudez, Roberts, & Brannon, 2016) and longitudinal studies (Mazzocco, Feigenson, & Halberda, 2011).

The ANS is not uniquely human (Gallistel, 1990; Vallortigara, Regolin, Chiandetti, & Rugani, 2010) since various animals are capable of discriminating quantities (e.g., birds, Bogale, Kamata, Mioko, & Sugita, 2011; fish, Agrillo, Dadda, Serena, & Bisazza, 2009; Agrillo & Bisazza, 2017; Nieder, Freedman, & Miller, 2002; rodents, Meck & Church, 1983; Gallistel, 1990; lions, McComb, Packer, & Pusey, 1994; Nieder, Freedman, & Miller, 2002; primates, Cantlon & Brannon, 2006). Further evidence that humans and non-humans share the ANS comes from studies investigating single-cell recording in primates. Nieder and Miller (2004) revealed that macaques have a parieto-frontal network for numerosity homologue to the human brain (Nieder & Miller, 2004).

Human adults without formal education have also demonstrated the ability to discriminate quantities (Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004). Cross-cultural studies show that the ANS is universally shared among humans, independent of language, culture or education. For instance, Pica et al. (2004) showed that an Amazonian tribe, with a limited lexicon of number words beyond 5, was still able to perform a magnitude comparison task, comparing large sets of approximate numbers that are far beyond their naming range. Similar results were replicated in other remote tribes without formal numerical systems in their language or culture, in the Amazon (Gordon, 2004) and Australia (Butterworth, Reeve, Reynolds, & Lloyd, 2008).

Finally, a wealthy body of research has demonstrated that our innate ability to discriminate numerosities is present from infancy onwards and that the precision of the ANS develops throughout early childhood (Wynn, 1992; Xu & Spelke, 2000; Feigenson, Dehaene, & Spelke, 2004; Lipton & Spelke, 2004; Xu, Spelke, & Goddard, 2005; VanMarle & Wynn, 2006; Xu & Arriaga, 2007; Libertus & Brannon, 2010).

2.5.2 ANS during infancy

Behavioural studies with infants have systematically shown infants' sensitivity to the approximate numerical magnitudes of sets of objects. Even one-day-old infants have demonstrated the ability to discriminate large numbers of objects. Izard et al. (2009) showed that neonates looked longer at a picture of objects that matched the number of tones they heard compared to at a picture with a different number of items. Older infants are reliably able to detect changes in the number of items presented to them, even after controlling for other non-numerical aspects such as size, space and position of items (Xu & Spelke, 2000). For example, Xu and Spelke (2000) tested six-month-old infants' discrimination of the numerosities 8 *vs.* 16 dots. Infants first saw repeated presentations of either 8 or 16 dots. Controlling for non-numerical dimensions such as density and luminance ensured that infants responded to numerosity only. When tested with alternating arrays of 8 and 16 dots, infants looked longer at the numerically novel test arrays regardless of whether they had been habituated to 8 or 16, showing that they successfully responded to number (Xu & Spelke, 2000). Further experiments using a similar methodology revealed that infants' numerical discriminations are subject to a ratio limit: six-month-old infants successfully discriminate 8 *vs.* 16 and 16 *vs.* 32 dots, but fail with 8 *vs.* 12 and 24 *vs.* 32 (Xu & Spelke, 2000; Lipton & Spelke, 2003, 2004; Wood & Spelke, 2005; Xu et al., 2005; VanMarle & Wynn, 2006).

In infants, the approximate number representation is amodal, i.e., not limited to visual arrays. When tested with sequences of temporally distinct events such as sounds, six- and nine-month-old infants show the same pattern of success and failure as with visual stimuli (Lipton & Spelke, 2003, 2004; Wood & Spelke, 2005, vanMarle & Wynn, 2006). Further evidence that this ratio limit is a general, rather than stimulus-specific, limit comes from studies using sequences of actions. For instance, using jumps of a puppet and when variables of sequence duration, jump duration, jump interval, jump rate and duration and extent of motion are controlled, and rhythm is eliminated, six-month-old infants are able to discriminate 4 from 8 jump sequences, but fail to discriminate 4 from 6 sequences, whereas nine-month-old infants can discriminate between 4 and 6 sequences (Wood & Spelke, 2005).

Numerous studies have shown that numerical discrimination increases in precision as children develop: six-month-olds can discriminate numerosities with a 1:2 ratio (Xu & Spelke, 2000), ten-month-old infants succeed with a 2:3 ratio (Xu & Ariaga, 2007); three- and four-year-old children discriminate a 3:4 ratio; five- and six-year-olds discriminate a 5:6 ratio (Halberda et al., 2012) and finally adults can successfully discriminate numerosities with a ratio 9:10 (Pica et al., 2004; Halberda et al., 2008).

Overall, the ANS represents large quantities imprecisely and has its distinctive signatures. Firstly, the ANS is ratio-dependent, meaning that precision in the mental representations of number decreases as number increases, in accordance with Weber's law, known as ***w*** (e.g., two and four being easier to discriminate than six and seven). The comparison of two numbers is possible only when they differ by a sufficient ratio (Halberda et al., 2008). Secondly, numerical discrimination increases in precision as children develop (see figure 2.5). Finally, representations are successful in different modalities such as visual and auditory (Xu & Spelke, 2000; Wood & Spelke, 2005).

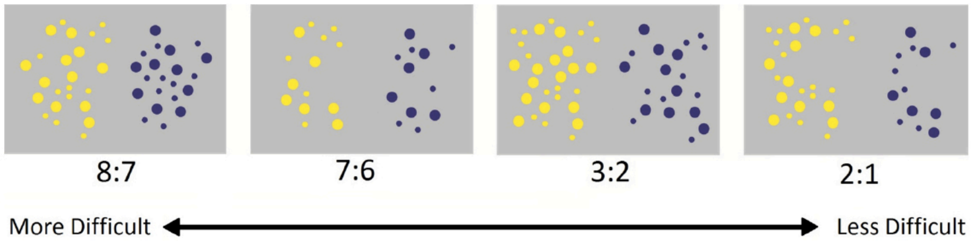


Figure 2.5: The development of our ability to discriminate non-symbolic quantities using the approximate number system (Keller & Libertus, 2015).

2.5.3 The neural bases of ANS

Measures of brain response in humans (Piazza, Pinel, Le Bihan, & Dehaene, 2007; Arsalidou & Taylor, 2011) and non-humans (Nieder, 2016) provide further support to behavioural studies showing the existence of an innate cognitive system for non-verbal number representation. Neuroimaging studies found that specific parietal areas are activated during numerical processing tasks (Dehaene, Piazza, Pinel, & Cohen, 2003; Dehaene, Molko, Cohen, & Wilson, 2004; Pinel, Piazza, Le Bihan, & Dehaene, 2004).

The intraparietal cortex is recruited during a wide range of numerical tasks and is one of the most consistently activated regions identified in a meta-analysis investigating functional Magnetic Resonance Imaging (fMRI) studies of numerical processing, both for non-arithmetic and arithmetic tasks in adults (Arsalidou & Taylor, 2011). Prefrontal areas are also found to be activated (Pesenti, Thioux, Seron, & De Volder, 2000; Gruber, Indefrey, Steinmetz, & Kleinschmidt, 2001Dehaene et al., 2003, 2004; Delazer et al., 2003; 2004; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Shuman & Kanwisher, 2004; Venkatraman, Ansari, & Chee, 2005). These results suggest that the internal representation of approximate non-verbal numerical magnitude is rooted in the **intraparietal sulcus (IPS)** (see figure 2.6 for an illustration of the IPS). Similar activations have been found in children as young as four years of age (Cantlon, Brannon, Carter, & Pelphrey, 2006). That said, children recruit parietal regions, in particular the IPS, to a lesser extent, and frontal regions to a greater extent, compared to adults (Ansari & Dhital, 2006; Cantlon et al., 2006, 2009; Kucian & Kaufmann, 2009; Holloway & Ansari, 2010) in the so-called **fronto-parietal activation shift** (Ansari, 2008). According to this view, there is a shift from an initially controlled and effortful (frontal activation) to a subsequently more automatic (parietal activation) processing of numerical magnitude (Ansari, Garcia, Lucas, Hamon, & Dhital, 2005; Rivera, Reiss, Eckert, & Menon, 2005; Kucian & Kaufmann, 2009; Holloway & Ansari, 2010).

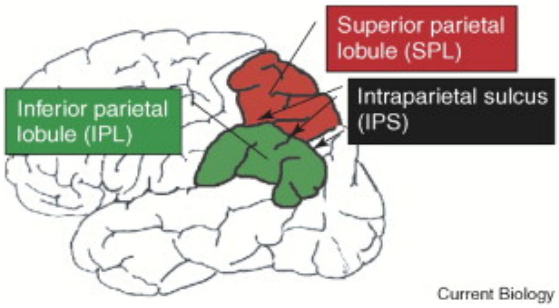


Figure 2.6: Illustration of the intraparietal sulcus (IPS) associated with numerical skills (Gobel & Rushworth, 2004).

Number-specific representations appear early in human development, demonstrating that the ANS develops prior to experience with number words or formal maths education. Even within the first year of life, the right intraparietal regions respond selectively to changes in the numerosity of visually presented sets of objects, with infants showing sensitivity to large numerosity (Izard, Dehaene-Lambertz, & Dehaene, 2008; Hyde & Spelke, 2011). Fluctuations in blood flow, as measured by functional near-infrared spectroscopy indicate that activity in the right parietal cortex of six-month-olds is modulated by changes in the number of objects within an array but not by changes in shape (Hyde, Boas, Blair, & Carey, 2010; Edwards, Wagner, Simon, & Hyde, 2016). Taken together, these results provide evidence of an innate cognitive system representing non-verbal numerical cognition rooted in parietal sites, more specifically in the IPS.

2.6 Predictors of mathematical performance

The development of mathematical abilities has been found to depend on a range of cognitive skills. Some skills are specifically associated with learning mathematics, known as **domain-specific skills**, while others are associated with learning in general, known as **domain-general skills**. Working memory, visuospatial abilities and processing speed are particularly important domain-general skills related to mathematics abilities. Domain-specific abilities that relate to mathematical abilities include counting, number fact knowledge and calculation skills, accurate numerical representations, such as digit recognition, performance on number line tasks and numerical magnitude (Gilmore et al., 2018). Among the basic mathematical abilities considered to be potential predictors of later performance, various studies have centred on the identification of numerosities through numerical magnitudes, and this is the focus of the next section.

2.6.1 Predictors of mathematical performance: numerical magnitude comparisons

The ability to process numerical magnitudes is an important domain-specific factor in the development of mathematical ability (De Smedt et al., 2013; Schneider et al., 2016). **Numerical magnitude comparisons** (see figure 2.7) refers to our basic ability to decide which of two numerosities is the largest. Numerical magnitudes are typically measured with dot (non-symbolic) and digit (symbolic) magnitude comparison tasks (Schleepen, Van Mier, & De Smedt, 2016). **Non-symbolic (dot) magnitude comparison** skills are thought to reflect the acuity of the approximate number system (ANS). **Symbolic (digit) magnitude comparison** skills are believed to index an exact symbolic representation system, which is language-dependent, develops gradually over the school years and allows for processing of discrete numbers (Ansari & Karmiloff-Smith, 2002).

The ability to process numerical magnitudes is typically measured by employing dot and digit magnitude comparison tasks that require participants to decide which of two numerosities is the largest. Non-symbolic (dot) magnitude comparison skills are thought to reflect the acuity of the approximate number system (ANS). This is a language-independent system that is present from young infancy and shared across species, enabling the estimation of quantities. Symbolic (digit) magnitude comparison skills are believed to index an exact symbolic representation system, which is language-dependent, develops gradually over the school years, and allows for processing of discrete numbers (Schleepen et al., 2016).

The performance in numerical magnitude comparison tasks (both symbolic and non-symbolic) creates the **numerical distance effect** (NDE) (Moyer & Landauer, 1967). The NDE refers to an increase in both reaction time and error rates as the numerical distance between the two comparators decreases. For example, individuals are typically faster and more accurate when comparing the numerical magnitude of 5 *vs.* 9 than when comparing 8 *vs.* 9. The NDE is indexed by the difference in reaction time and accuracy of responses to comparisons with relatively small or large distances (Price, Palmer, Battista, & Ansari, 2012).

There is a consistent association between individuals' ability to compare magnitudes (such as two arrays of dots) and their mathematical performance measure by arithmetic scores

(Schneider et al., 2017). This magnitude information may form the basis of a 'sense' of numbers that allows us to estimate the results of arithmetic operations and check whether our answers are reasonable. Many studies have therefore investigated the underlying mechanisms of numerical magnitude comparisons and the association of this ability with mathematical abilities, both in typically developing children and at-risk populations.

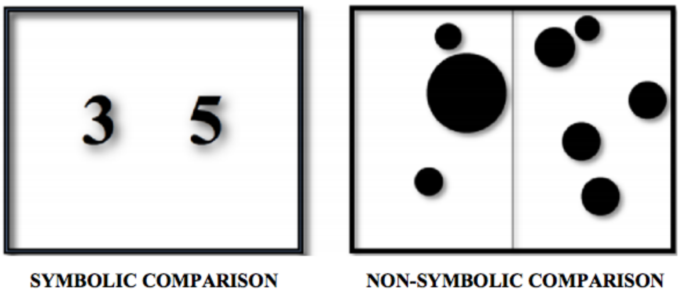


Figure 2.7: Numerical Magnitude Comparison
From Nosworthy, Bugden, Archibald, Evans, & Ansari, 2013

2.6.1.1 Symbolic and non-symbolic magnitude comparison

There is a small, but reliable, association between individuals' ability to compare non-symbolic magnitudes and their mathematical performance (Schneider et al., 2017). For example, performance in non-symbolic magnitude comparison tasks in kindergarten children predicts their maths achievement in sixth grade (Halberda et al., 2008). Moreover, longitudinal studies have shown that ANS abilities assessed at three years old predict general mathematical achievement at six years old (Mazzocco et al., 2011). In fact, it has been demonstrated that the performance of six-month-old infants in tasks evaluating their sense of numerosity could predict mathematical achievement about three years later (Starr, Libertus, & Brannon, 2013). Meta-analyses also reveal significant but modest (r=0.2) relationships between performance on ANS tasks and mathematics achievement (Chen & Li, 2014;

Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2017), although stronger relationships were found before children begin formal schooling (r=0.4). Values ranging from r= 0.58 to 0.11 were found in a recent review investigating associations between the ANS and mathematical performance. Taken together, it has been hypothesised that non-symbolic magnitude processing skills are a key cognitive factor in the development of mathematics achievement since some studies have found an association between non-symbolic processing skills and children's mathematics achievement (e.g., Halberda et al., 2008; Gilmore, McCarthy, & Spelke, 2010; Libertus et al., 2011).

On the other hand, not all studies have been able to find associations between ANS abilities, tested by non-symbolic magnitude comparison, and mathematical performance ((De Smedt et al., 2013; Vanbinst, Ansari, Ghesquiere, & De Smedt, 2016; Schneider et al., 2017). Considerable research has revealed that symbolic magnitude comparison tasks (using Arabic digits) are better predictors of mathematical achievement than non-symbolic magnitude comparison tasks (Bartelet, Vaessen, Blomert, & Ansari, 2014; Castronovo & Göbel, 2012; Desoete, Ceulemans, De Weerdt, & Pieters, 2012; Fazio et al., 2014; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Sasanguie, Defever, Maertens, & Reynvoet, 2014; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013; Vanbinst, Ceulemans, Ghesquiere, & De Smedt, 2015). A meta-analysis of 45 papers (N=17,201) supported these findings, revealing that symbolic magnitude processing was a better predictor of mathematical achievement than non-symbolic magnitude processing, with a significant correlation between symbolic number processing and mathematical competence (r=0.302) (Schneider et al., 2017). Libertus et al. (2011) demonstrated that tasks based on the ANS acuity were able to predict informal mathematical abilities (such as numbering and counting, comparing numbers to determine which is more or less, and calculating the answers to simple arithmetical problems using tokens or fingers), but not formal abilities, (such as reading and writing Arabic numerals, understanding the place value system, and the ability to recall memorised addition, subtraction and multiplication facts) (Libertus et al., 2011). In addition, ANS acuity was found to be significantly correlated with mathematical skills in pre-schoolers just beginning formal education (Libertus et al., 2011; Starr et al., 2013). These results are congruent with the studies showing that ANS acuity is significantly correlated with mathematical skills before formal education (Libertus et al., 2011; Starr et al., 2013), suggesting a refinement of the non-symbolic skills through developing symbolic number proficiency and potentially two separate trajectories for those systems.

2.6.1.2 Symbol Grounding Problem

When children first encounter symbolic representations of number, they are meaningless words or visual symbols. Therefore, children need to connect symbolic representations to their semantic meaning. How symbols, such as numerical symbols, become representations of semantic reference (such as numerosity) has been referred to as the 'symbol-grounding problem' (Harnad, 1990, 2003). In the field of numerical cognition, the most widely accepted theoretical account to resolve this problem is the notion that symbols are mapped on the preexisting ANS (e.g. Dehaene, 1997; Gallistel & Gelman, 1992), commonly known as the approximate number system (*ANS*) *mapping account* (Holloway & Ansari, 2015; Leibovich & Ansari, 2016). Alternative views suggested that early in development symbolic and nonsymbolic representations are strongly linked and this link might become weaker the more children become proficient in symbolic number processing (Lyons, Ansari and Beilock, 2012). Another approach proposes that Arabic digits are not initially mapped onto nonsymbolic representations, but that the meaning of small symbolic numbers emerges throught subitizing (the ability to discriminate small set of elements (<3) prior to acquisition of verbal counting (Starkey & Cooper, 1995)). According to this view, small numerical symbols are initially mapped on a precise representation and then, in combination with increasing knowledge of the counting list, result in an independent and exact symbolic system based on order relations between symbols. Further research in this area is necessary to provide additional evidence and resolve this debate (Reynvoet & Sasangue, 2016).

2.6.2 Predictors of mathematical performance: domain-general skills

One criticism of the studies investigating the relationship between numerical magnitude comparison and mathematical performance is the lack of consideration of the interplay with domain-general skills, such as executive functions. For instance, Gilmore et al. (2013) found that individual differences in inhibitory control, rather than ANS acuity, correlated with mathematical achievement in four- to eleven-year-olds. Once inhibitory skills were controlled for, for example, the performance in non-symbolic tasks was no longer related to mathematical abilities (Gilmore et al., 2013). Investigating the association between exact and approximate systems, working memory and maths skills in kindergarten children, Xenidou-Dervou et al. (2013) demonstrated that while non-symbolic tasks were only able to predict maths achievement indirectly, symbolic tasks were directly predictive of mathematical skills. Individual differences in working memory highly predicted individual differences in maths achievement beyond the effects of either symbolic or non-symbolic tasks (Xenidou-Dervou, De Smedt, van der Schoot, & van Lieshout, 2013).

The prediction value in tasks evaluating domain-specific and domain-general in mathematics performance is comprehensively influenced by the stage of development. For instance, Gimbert et al. (2019) showed that ANS acuity was a significant specific predictor of mathematics achievement in five-year-olds, but not in older children, and WM was a significant general predictor in seven-year-olds, but not younger children (Gimbert, Camos, Gentaz, & Mazens, 2019). This suggests that a general cognitive ability, especially WM, becomes a stronger predictor of mathematics achievement after entrance into formal schooling, at which point ANS acuity, a specific cognitive ability, loses predictive power.

Executive skills (EF) skills have been found particularly important for success in maths (Blair & Razza, 2007). The idea that the role of executive functions is particularly important for mathematical performance is sustained by evidence from correlational studies (e.g., Navarro et al., 2011; Lee et al., 2012; Wei, Guo, Georgiou, Tavouktsoglou, & Deng, 2018), experimental studies exploring participants abilities to solve arithmetic problems under different conditions that investigate the contribution of executive functions to different contexts (e.g., Holmes & Adams, 2006) and training studies (e.g., Zhang, Chang, Chen, Ma, & Zhou, 2018). Next, I review studies investigating subcomponents of EFs particularly important for mathematical performance.

2.6.2.1 Mathematical abilities and working memory

Working memory (WM) is a limited capacity system responsible for the manipulation and storage of information during the performance of cognitive tasks (Baddeley, 1986). The most widely used model of working memory comprises four subcomponents: the central executive, phonological loop, visuospatial sketchpad and episodic buffer (Baddeley, 2003). The central executive[4] is an attentional control system involved in several processes such as the selection and execution of strategies, retrieval of information from long-term memory, monitoring of input, the simultaneous storage and processing of information, and the coordination of the other components of the WM system. Two 'slave systems' lie under the central executive functions: the phonological loop and the visuospatial sketchpad. The visuospatial sketchpad involves temporary storage and rehearsal of visual and spatial information, while the phonological loop involves storage and rehearsal of phonological and auditory information. The episodic buffer, meanwhile, is considered to be responsible for the integration of information from the subcomponents of WM and long-term memory (Baddeley, 2000).

A robust body of research has provided evidence that WM is crucial for mathematical performance (Raghubar, Barnes, & Hecht, 2010). Across many studies, working memory capacity has consistently been found to be a strong predictor of arithmetic outcomes (e.g., Peng & Fuchs, 2016). Positive associations of arithmetic performance and working memory capacity have been found across all ages (Espy et al., 2004; Cragg, Keeble, Richardson, Roome, & Gilmore, 2017). Furthermore, all components of the WM system have been shown to be associated with mathematics performance and learning in children (Holmes & Adams, 2006; Swanson, 2006; Bull, Espy, & Wiebe, 2008; De Smedt et al., 2009; Alloway & Alloway, 2010; Geary, 2011; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011; Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2015). The results of meta-analysis have suggested that verbal working memory has a stronger relationship with arithmetic than

[4] There are various understandings of the central executive function. This work will follow Baddeley's definition (1986) in dividing it into updating, inhibition and shifting.

with visuospatial working memory (Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013), although it is important to note that other studies have found stronger relationships in visuospatial working memory (Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013) or no differences between the two slave systems (Peng et al., 2016; Cragg et al., 2017). While the extent to which specific components of WM explain individual differences in mathematical performance remains unclear, therefore, all studies point in the direction that maths abilities are strongly associated with WM performance.

The results from studies during preschool, primary school and adolescence suggest that younger children rely more on visuospatial working memory when learning and applying new mathematical skills, whereas older children rely more on verbal working memory after skills have been learned. For example, both Holmes & Adams (2006) and McKenzie, Bull & Gray (2003) reported strong associations between young children's mathematics attainment and visuospatial working memory, whereas older children's mathematics presented strong associations with verbal working memory. According to the authors, the results suggested that these differences might reflect a shift from the use of early visuospatial solution strategies to mature verbal solution strategies. In addition, working memory capacity appears to be linked to differences in strategy choice and strategy efficiency across children with and without mathematical difficulties. For example, children with mathematical difficulties (MD) are less likely to use direct memory retrieval to solve arithmetic questions (Geary, Brown, & Samaranayake, 1991; Bull & Johnston, 1997), as such, the phonological loop is thought to be required to store and access information in long-term memory, and therefore may be involved in the retrieval of number facts (Cragg et al., 2017).

Generally, it is assumed that the phonological loop contributes to encode and process number words and numerals and to processes that involve them, such as counting procedures to solve arithmetic problems or retrieving arithmetic facts from long-term memory. It also helps in solving mathematical word problems (written problems that have to be translated into mathematical problems to be solved). The visuospatial sketchpad contributes to solving problems recruiting mechanisms of visualisation and representing quantities on the number line. The central executive contributes in both verbal and visual information when tasks require, for example, simultaneous, sequential or active processing (Geary & Moore, 2016).

2.6.2.2 Mathematical abilities and inhibition and shifting

Inhibition is the ability to ignore information or responses that are irrelevant to certain contexts. While some studies have found that inhibition is associated with mathematical performance (Espy et al., 2004; St Clair-Thompson & Gathercole, 2006; Blair & Razza, 2007; Clark, Pritchard, & Woodward, 2010; Merkley, Thompson, & Scerif, 2016), others failed to find associations (Monette, Bigras, & Guay, 2011; Lee et al., 2012; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). In addition, the associations found were related to the more basic skill of counting rather than more cognitively demanding tasks involving calculation (Lan, Legare, Ponitz, Li, & Morrison, 2011).

Shifting is the ability to stitch attention flexibility from one task to another, or consider multiple perspectives of a situation at the same time. Some studies have documented a link between shifting and mathematical performance. For example, Lan et al. (2011) found that shifting was related to both early counting and calculation skills. A meta-analysis revealed that performance in tests of shifting is correlated with mathematics performance across a wide age range, from preschool to adolescence. It has been suggested that shifting may be involved in arithmetic ability due to the need to shift between strategies during problem-solving (Yeniad, Malda, Mesman, van IJzendoorn, & Pieper, 2013).

Thus, it has been proposed that inhibition is likely to support the acquisition of more basic mathematics skills, such as counting, whereas shifting appears to be related to the cognitive flexibility needed to apply different strategies during problem-solving. It is clear that more research is needed to understand the role of those cognitive abilities in mathematical performance. Most of the studies investigating the role of EF in mathematics provide the associations between different factors, but do not go on to develop a comprehensive understanding about the underlying causal relationships.

2.6.2.3 Mathematical abilities and processing speed

Several studies suggest that mathematical skills rely heavily on executive function domain-general processes, such as working memory. Another domain-general skill identified as an efficient predictor of arithmetic in children is processing speed (Bull & Johnston, 1997; Geary, Hamson, & Hoard, 2000; Swanson & Beebe-Frankenberger, 2004; Fuchs et al., 2006; Berg, 2008). Processing speed can be defined as the speed with which a subject executes a simple and relatively automated cognitive task (Sheppard & Vernon, 2008). Processing speed has been included as an EF sub-skill in an alternative model proposed by Anderson (Anderson, 2002). In addition to cognitive flexibility and goal setting, Anderson included information processing. Together this sub-set of cognitive skills orchestrates overall executive control (Anderson, 2002).

While some have suggested that processing speed's influence on mathematical performance is explained by the availability of resources in working memory (Case, Kurland, & Goldberg, 1982), others have found that its effect is independent of working memory capacity (Bull & Johnston, 1997; Fuchs et al., 2006, 2012). According to this view, processing speed may affect the consolidation of mathematical conceptual information in long-term memory, such as the automatisation of the counting sequence and arithmetic facts.

2.7 Current cognitive models for mathematical cognition

2.7.1 Domain-specific theories: the triple-code

The adult literature on numerical cognition has proposed the engagement of three different neural networks for number representations (Dehaene & Cohen, 1995; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). Known as the triple-code, it is the most accepted model[5] in the field of numerical cognition. It comprises an analogue magnitude representation of the mediating semantic number representation (i.e. numerosity), a verbal-phonological number representation supporting verbal counting and number fact retrieval, and a visual-Arabic number representation that comes into play upon solving written arithmetical problems (see figure 2.8). Numerical magnitude representation recruits inferior parietal areas bilaterally, whereas tasks with digit codes (Arabic numbers) activate bilateral activity in the fusiform regions, and

[5] The abstract code model and the encoding complex model are examples of other models in the numerical cognition field (Kadosh & Dowker, 2015)

finally semantic numbers, represented by verbal information, elicit activity in the left perisylvian areas and left angular gyrus (see figure 2.9).

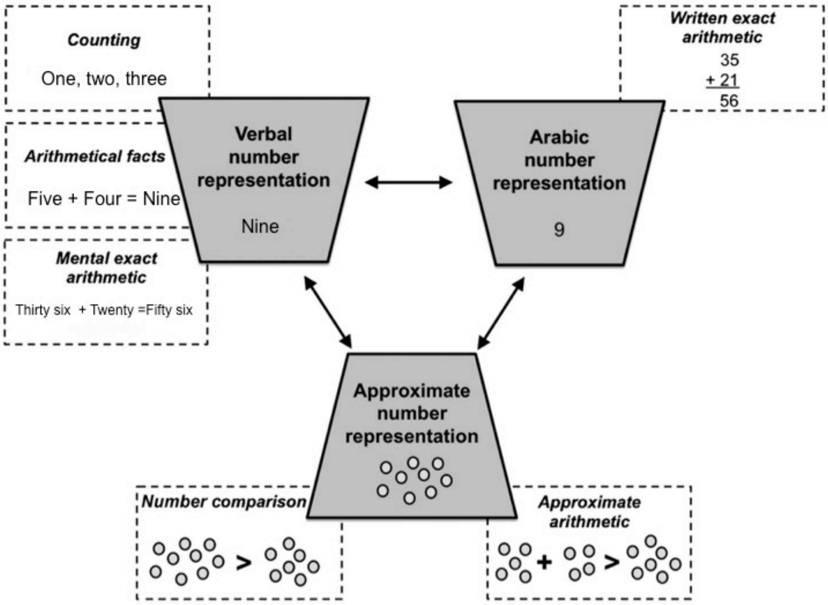


Figure 2.8: The triple-code model.

It comprises an analogue magnitude representation (i.e. numerosity), a verbal-phonological number representation (e.g., "Nine") and a visual-Arabic number representation (e.g. '9'). From Julie, Alain & Jacquelien, 2013.

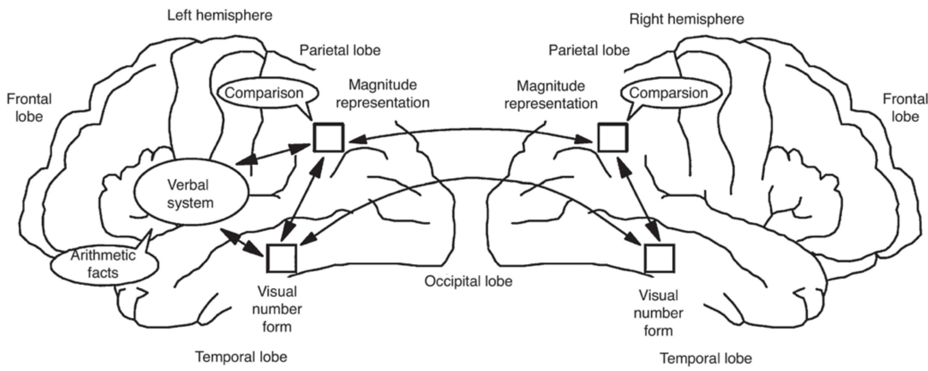


Figure 2.9: Anatomical representation of the triple-code model.

The localisation of the main areas thought to be involved in the three numerical codes is represented on a lateral view of the left and right hemispheres. Note that arrows indicate a functional transmission of information across numerical codes,and are not meant as a realistic depiction of existing neural fibre pathways, whose organisation in humans is not fully understood (Dehaene & Cohen,1995).

2.7.2 Problems with the triple-code model

Although this model assumes that number processing demands the intimate interplay of domain-specific number-related parietal as well as domain-general (pre) frontal processes involving executive control, the role of the prefrontal areas and the involvement of executive functions are not clearly specified in this model (Moeller, Willmes, & Klein, 2015). Recently, neuroimaging studies have contributed to this understanding. The results of a meta-analysis of neuroimaging studies in adults have provided robust evidence of the engagement of prefrontal functions in solving arithmetic tasks. For example, frontal brain regions are likely to be engaged in formulating and following goals (superior frontal regions) and in navigating eye movements (required in tasks evaluating number magnitudes recruiting the precentral gyrus (Arsalidou & Taylor, 2011). Meta-analyses exploring typically developed children and adults have shown age-dependent activations in non-symbolic tasks. In general, children show activations in the frontal and parietal areas, whereas adults seem to recruit generally more parietal areas (Kauffman et al., 2011). Those results are in line with the view that, during development, activations in the fronto-parietal areas are triggered by both basic and advanced numerical tasks (Ansari, 2008). The observed increases in the recruitment of parietal regions in numerical processing are taken as evidence for a specialisation of parietal functioning during ontogeny (Ansari and Dhital, 2006; Holloway and Ansari, 2010). Functional specialisation in the parietal cortex for mental arithmetic increases with age and is accompanied by a corresponding decrease of activity in the prefrontal regions. The decreasing involvement of the prefrontal areas, on the other hand, is assumed to reflect a developmental disengagement of domain-general processes related to executive control and working memory (Ansari et al., 2005; Rivera et al., 2005).

Moreover, the IPS is recruited with a rather complex functionality beyond the domain of numerical cognition. This includes, for example, spatial abilities (Culham & Valyear, 2006). Superior parts of the intraparietal cortex further play a crucial role in cognitive functions such as attention, working memory, episodic retrieval and mental imagery (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008). Lateral parts of the IPS have been implicated in cognitive functions such as spatial and working memory (Fedorenko, Duncan, & Kanwisher, 2013). Thus, the triple-code model has become insufficient to explain some of the more recent findings and further extensions of the model to include other general-domain functions are necessary. Supported by the results of a meta-analysis, Arsalidou and Taylor (2011) proposed the recruitment of several networks involved in the processing of numeral skills (see figure 2.10).

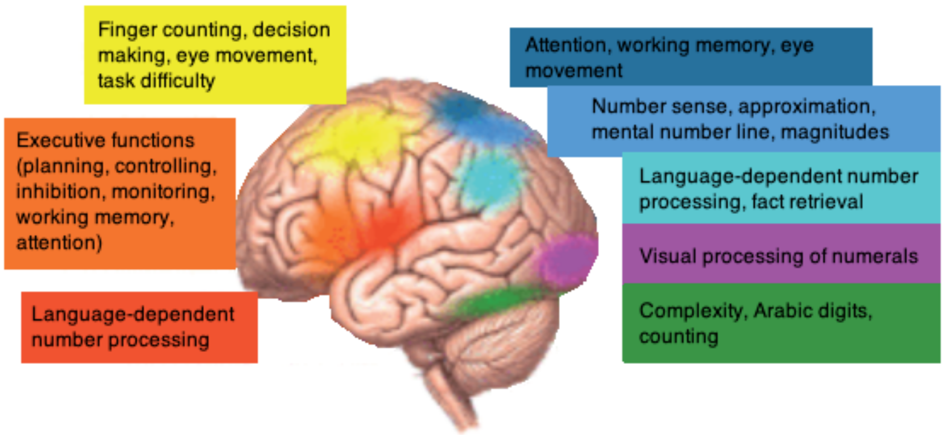


Figure 2.10: Schematic representation of neural networks supporting the acquisition of number skills. Adapted from Kadosh & Dowker (2015), pg 491.

2.7.3 Domain-general theories

Based on robust evidence that executive function skills play a critical role in the development of mathematics competence (Bull & Scerif, 2001; Kroesbergen, Van der Ven, Kolkman, Luit, & Leseman, 2009; Raghubar et al., 2010; St Clair-Thompson &

Gathercole, 2006; Yeniad et al., 2013; Gilmore & Cragg, 2018) some studies have incorporated EF processes into their models for numerical cognition (von Aster & Shalev, 2007; LeFevre et al., 2010; Cragg & Gilmore, 2014). LeFevre et al. (2010) postulate *the Pathways to Mathematics model* (Figure 2.11), focusing on the relationships between children's mathematical skills and cognitive precursors, early numeracy skills and mathematical outcomes. This model suggests three separate pathways: quantitative, linguistic and spatial attentional. Each of these pathways contributes individually to the acquisition of early numeracy abilities. Furthermore, the model proposes that the linguistic, quantitative and spatial attentional pathways vary in their contribution to mathematical performance depending on the demands of the arithmetic problem. According to this model, linguistic skills are linked to children's symbolic number system knowledge. The second skill pathway comprises quantitative abilities and processing numerical magnitudes. Spatial attention forms a third pathway with connections across a variety of numerical and mathematical skills. Sowinski et al. (2014) further expanded the quantitative pathway to include not only magnitude comparison but also counting and subitising (the ability to enumerate small quantities quickly and exactly; Clements, 1999). They also introduced a working memory pathway (the original model only focused on visuospatial attention).

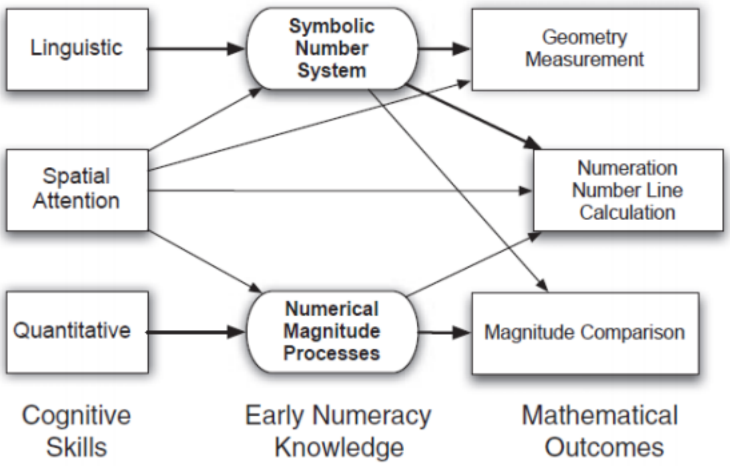


Figure 2.11: The Pathways to Mathematics model proposed by LeFevre et al. (2010).

LeFevre's model, however, lacks an explanation of the specific role of the different aspects of cognitive domains that are important in relation to numerical skills, such as working memory capacity. *The Four-Step Developmental Model of Numerical Cognition* (von Aster and Shalev, 2007) describes four stages of numerical cognition with children moving through the stages as they progress in arithmetic competency through exposure and formal schooling and an increase in working memory capacity (Figure 2.12). The first stage consists of an inherited core-system of magnitude representation, similar to Dehaene's number sense, which entails subitising and approximation abilities. This basic meaning of number is a prerequisite for the acquisition of more complex mathematical skills. Pre-school children move on to the linguistic stage of numeracy (step 2) whereby they acquire the verbal number codes. In step 3, children learn the Arabic number system and the symbolic representations of magnitudes in school. Typical mathematical skills developing at this stage are written calculations and odd even decisions. The final stage, the mental number line, develops during school years as children acquire the concept of ordinality, a second core principal of number. Von Aster and Shalev (2007) further propose that failure to establish a stage appropriately may lead to developmental delays in acquiring the follow-on stages or dyscalculia.

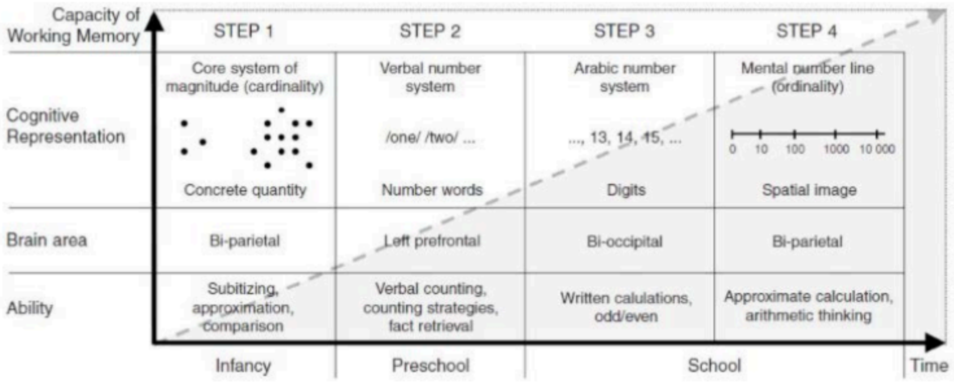


Figure 2.12: Four-step-developmental model of numerical cognition by Von Aster and Shalev, 2007.

Both the Four-Step Developmental Model of Numerical Cognition (von Aster and Shalev, 2007) and the Pathways to Mathematics model (LeFevre et al., 2010) fail to incorporate different aspects of EF through the course of a child's development, however. Addressing this, Cragg and Gilmore (2014) proposed a model suggesting specific relationships between aspects of EF and different components of mathematical proficiency. The authors proposed a framework to identify the interplay among the subcomponents of executive functions (working memory, inhibition and shifting) and specific mathematical skills (facts, procedures and concepts) (see figure 2.13). The authors also considered which conjunctions of subcomponents are more likely to be recruited at various ages. For instance, it is likely that working memory skills are important for all ages when dealing with procedural knowledge, whereas inhibition is more likely to be recruited at younger ages.

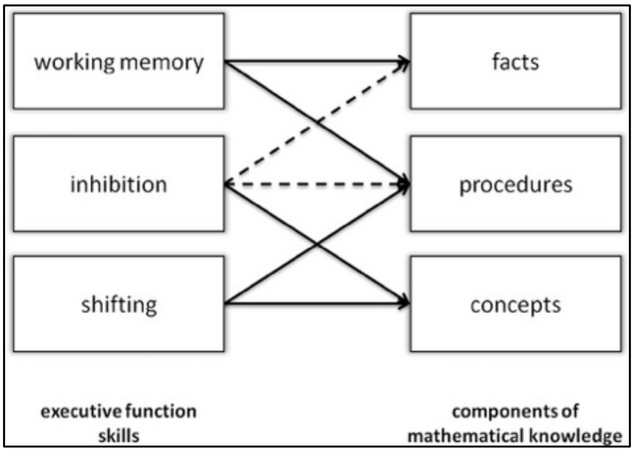


Figure 2.13: Potential subtypes of components and the relationships between executive function skills and mathematical knowledge.

Dashed lines represent relationships that change over the course of development. Model proposed by Gilmore and Cragg (2014).

More recently Cragg et al. (2017) investigated the extent to which executive function skills contribute to these three components of mathematical knowledge. Two hundred and ninety-three participants aged between eight and twenty-five years completed a large battery of mathematics and executive function tests. According to their results, inhibition skills were associated with factual knowledge and procedural skill, and working memory contributed to mathematics achievement indirectly through factual knowledge, procedural skill and, to a lesser extent, conceptual understanding. Working memory was directly associated with mathematics achievement, reflecting the role of working memory in identifying and constructing problem representations. These relationships were remarkably stable from eight years through to young adulthood. Based on the results, the authors proposed a modified version of a hierarchical framework for mathematics (Geary, 2004; Geary & Hoard, 2005) in which domain-general executive function skills, in particular working memory, support domain-specific mathematical processes, which in turn underpin overall mathematics achievement (see figure 2.14).

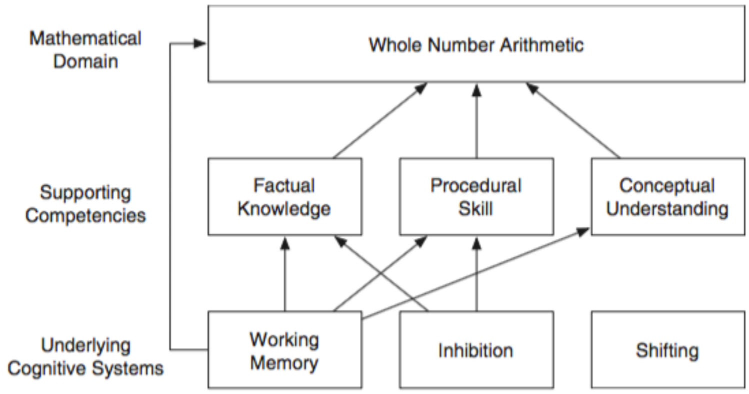


Figure 2.14: Refined hierarchical framework of the executive functions underpinning mathematics

This extends previous models by demonstrating that working memory also contributes directly to mathematical achievement (Cragg et al., 2017).

In spite of the increased focus on studies investigating the development of number processing and its relationship with domain-general skills in typically developing children, less is known in this area in the preterm population, a population at increased risk of presenting with difficulties in mathematical achievement. One way to identify the mechanisms underlying VP children's difficulties with mathematics is to examine the characteristics of term-born children who have mathematical learning disabilities.

Next, I present an overview of mathematical learning difficulties in typically developing children, followed by studies that have examined the nature of mathematical difficulties in the preterm population.

2.8 Mathematical Learning Disabilities or Developmental Dyscalculia?

Specific difficulties with number in children with otherwise normal intelligence are commonly referred to as developmental dyscalculia (DD) or mathematical learning disability, with some studies considering these as separate disorders and others considering them as different terminologies for the same condition (Szucs, 2016). Commonly, the term mathematical learning disability (MLD) has been used for children scoring within the bottom 25% in a standardised mathematical achievement test, whereas the term developmental dyscalculia has been used to refer to severe mathematical difficulties represented by performance in the lowest 5-10% in standardised mathematical achievement tests (Gilmore, Göbel, & Inglis, 2018). There is neither general agreement about what terminology should be used nor about the criterion for defining MLD, however.

For the purposes of this work, I will use the term mathematical learning disability, for two reasons. Firstly, because most papers employ this term. Secondly, in future sections, when discussing the mathematical academic profile in the preterm population, the specific cognitive difficulties that affect their performance in mathematical achievement will become apparent. Constraining the term to either mathematical learning disability or developmental dyscalculia would not clarify the cause and/or nature of their difficulties. Rather, I use the term as a way to refer to severe and/or moderate mathematical difficulties. For historical reasons, however, the term developmental dyscalculia will be used when referring to studies employing that terminology. In this case, I will be referring to individuals with severe mathematical difficulties, performing in the lowest 5-10% in standardised mathematical achievement tests, in line with the definition given by Gilmore and colleagues (Gilmore, Göbel, & Inglis, 2018).

2.8.1 Achievement-based or discrepancy-based criteria

One aspect to be considered in studies investigating mathematical learning difficulties is the different criteria employed. A variety of criteria have been used, but generally, there is no agreement about what criterion should be used in studies, and this has an impact on the heterogeneity of findings in respect to the prevalence of MLD and its comorbidities.

The most popular criteria are achievement-based and discrepancy-based. The discrepancy-based criterion is based on the marked difference between IQ and academic attainment (Simms et al., 2013a). The accepted criteria to identify an individual as having a learning disability based on IQ-achievement discrepancy is a difference of at least two standard deviations. The discrepancy-based criterion was employed for clinical purposes for decades, according to the Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-4). The discrepancy-based criterion, however, is now considered to be inadequate in that intelligence tests are sensitive to socio-economic status, meaning that learning difficulties tend to be over-diagnosed in those with lower socio-economic status (Cangöz, Olkun, Altun, & Salman, 2018). Another major problem with discrepancy-based definitions is their lack of sensitivity. For example, children with a clear discrepancy between IQ and mathematical performance could be identified as having problems, whereas children with low IQ and low mathematic performance would not be identified (Mazzocco & Myers, 2003).

In contrast, the achievement-based criterion considers the performance in standardised achievement tests, regardless of IQ, and is the method currently employed by the DSM-5. According to the DSM-5, the assessment of cognitive processing skills for a diagnosis of Specific Learning disabilities (mathematics) is no longer required, except when Intellectual Disabilities are suspected. For education purposes, the elimination of the IQ-achievement discrepancy criterion indicates that individuals with specific learning disorder (SLD) and lower IQ (e.g., an IQ score above 70 ± 5), but who do not have an Intellectual Disability, would be able to benefit from special education services.

Given these considerations, next I present an overview about mathematical learning difficulties in typically developing children.

2.8.2 Mathematical Learning Disabilities in typically developing children

The lack of criteria for identifying mathematical learning disabilities complicates attempts to estimate the prevalence of MLD. As a result, estimates for MLD have ranged from 4% to 14% (Geary, 2015). Various studies have investigated the nature and causes of mathematical learning disability, but there is still no consensus about the underlying core cognitive deficit in mathematical learning disability (Gilmore, Göbel, & Inglis, 2018). Originally, developmental dyscalculia was described as a core deficit in understanding and manipulating the quantity of sets and their numerosities (Butterworth, 1999, 2005, 2010). Other accounts, however, have proposed MLD to be the result of more general cognitive impairments, particularly in working memory, visuospatial skills, attention and executive functions (McLean & Hitch, 1999; Geary & Hoard, 2005; Donlan, 2007; Geary, 2007; Le Corre & Carey, 2007; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2014). Generally speaking, there are broadly two groups of theories addressing the core deficits underlying MLD: domain-specific and domain-general theories. The former proposes that the underlying core deficit causing MLD is a numerical one. The latter proposes that non-numerical mechanisms that are particularly important for number processing are impaired (Gilmore, Göbel, & Inglis, 2018).

2.8.3 Mathematical Learning Disabilities and domain-specific theories

One of the major domain-specific driven theories of MLD postulates that difficulties in acquiring appropriate arithmetic skills might be rooted within a specific deficit in an innate ability to represent approximate numerosities (Butterworth, 1999, 2005, 2010), known as number sense (Dehaene, 1997, 2001). This is supported by data showing that MLD children are not only poor in school arithmetic and standardised tests of arithmetic, but they are slower and less efficient at very basic numerical tasks (Iuculano, 2016). For example, Piazza et al. (2010) investigated the links between number sense and MLD, employing a non-symbolic magnitude comparison task. Preschool, school-age children and adults were tested, as well as children with MLD, identified by standardised mathematical achievement tests and matched in age and IQ with the typically developing school-age children. Results showed that typically developing children, at all ages, showed improved number acuity with age. Children with MLD showed severe impairments in numerical acuity, with ten-year-old children with MLD scoring at the level of five-year-old typically developing children (Piazza et al., 2010). Another study employing a similar methodology investigated the association between number sense and MLD. Using standardised mathematical achievement tests, 71 children were clustered in four groups: MLD, low achieving, typically achieving and high achieving. Students with MLD had significantly poorer acuity than students in all other mathematics achievement groups. This relationship persisted even when controlling for domain-general abilities, in this case, intelligence (Mazzocco, Feigenson, & Halberda, 2011).

According to the domain-specific theory, MLD is the result of an inability to form approximate representations of numerical magnitude (Halberda et al., 2008; Piazza et al., 2010; Feigenson et al., 2013). Other studies, however, have failed to find deficits in number sense in children with MLD (Rousselle & Noël, 2007; Iuculano, Tang, Hall, & Butterworth, 2008; Landerl & Kölle, 2009; De Smedt & Gilmore, 2011). For instance, Rousselle and Noel (2007) found imprecise representation in children with MLD in tasks assessing symbolic magnitude comparison, but not non-symbolic magnitude comparison. Another study also found that children with MLD are generally less efficient than children with typical development in tasks assessing number sense, but do not have imprecise representations (Landerl & Kolle, 2009). These results suggest that MLD might be linked with deficits in mapping number symbols and not necessarily with imprecise representation in number sense. Within this framework, Dehaene & Wilson (2007) proposed five subtypes of developmental dyscalculia: 1) impairment of number sense; 2) mapping impairment; 3) deficit of verbal-symbolic representations; 4) problems with executive functions; 5) deficit in spatial attention. The authors identified two possible causes of a core deficit: 1) a deficit in number sense (e.g., Piazza et al., 2010; Mazzoco et al., 2011); 2) a deficit in symbolic numerical representation (e.g., Rousselle and Noel, 2007; Landerl and Kolle, 2009). Those deficits were related to impairment in the horizontal intra-parietal sulcus (hIPS) area, and a failure to build adequate connections between non-symbolic and symbolic representations of number. Other possible causes of different subtypes of dyscalculia were also identified, namely deficits in verbal symbolic representation, executive dysfunction and spatial attention.

2.8.4 Mathematical Learning Disabilities and domain-general theories

Due to the heterogeneity of difficulties associated with MLD, often related to more general cognitive domains, various studies have tried to characterise subgroups of MLD by linking deficits between domain-specific and domain-general skills. For instance, Geary (2004) proposed three subtypes of MLD: 1) procedural subtype; 2) semantic memory subtype; and 3) visuospatial subtype. In the first group, children frequently use immature procedures, have a poor understanding of concepts underlying procedural use, frequently present errors in the execution of procedures and have difficulties in sequencing multiple steps in complex procedures. Geary proposes that those difficulties may be a result of verbal working memory deficits associated with deficits in conceptual knowledge. The second group is described as having difficulties retrieving mathematical facts, and Geary proposes that this is associated with long-term memory deficits. Finally, for the third group, children show deficits in the spatial representation of number. Henik, Rubinsten, & Ashkenazi (2015) meanwhile, proposed four types of Developmental Dyscalculia described according to cognitive deficits in conjunction with the brain mechanisms contributing to different manifestations of MLD. Firstly, what they refer to as pure DD presents with abnormalities in the IPS leading to a pure deficit in numerical processing only. Their second type is Combined developmental dyscalculia and dyslexia, which presents with abnormalities in the angular gyrus leading to deficits in associating symbols with the events they symbolise, thus resulting in comorbidities between developmental dyscalculia and dyslexia. Their third type presents with abnormalities in the IPS that affect both numerical processing and attention, creating a deficit in arithmetic processing. Finally, their fourth type presents with abnormalities in frontal areas leading to deficient executive functions, causing both deficits in arithmetic and attention, such as ADHD.

Numerous studies have attempted to describe subcategories of MLD based on arbitrary subgroups and/or focus on a single factor. Highly powered studies might show an alternative approach considering a multidimensional structure of cognitive functions and their relationships to mathematical performance (Szűcs, 2016). An elegantly-designed study by Bartelet et al. (2014) employed this type of data-driven approach to cluster 226 children with MLD into different subgroups. Results revealed six groups of MLD children: 1) mental number line difficulties group; 2) number sense difficulties group: children in this group had difficulties in tasks related to number line and dot comparison; 3) spatial difficulties group: children in this group had difficulties in tasks related to dot comparison visuospatial short-term memory; 4) access deficit group: children in this group had difficulties in tasks related Arabic knowledge and counting skills; 5) no numerical cognitive deficit group: children in this group had no difficulties in number-related tasks, but had deficits in verbal short-term working memory skills; 6) garden-variety group: children within this group had difficulties in nonverbal IQ (Bartelet, Ansari, Vaessen, & Blomert, 2014).

A number of scholars have challenged the core deficit hypothesis that postulates that developmental dyscalculia originates from the impairment of the magnitude representation of the human brain, residing in the intraparietal sulcus. For example, Szucs & Goswami (2013) found robust evidence that children with developmental dyscalculia present with major dysfunction in visuospatial short-term memory and working memory, with additional impairment in inhibitory function. The authors noted that both of these functions have been linked to the IPS. Indeed, the IPS is an area that plays an essential role not only in quantity representations but also in maintaining quantity-related information in short-term WM (Menon, 2016). Taking a systematic approach to investigate the role of verbal and visual short-term memory and working memory in developmental dyscalculia, Szűcs (2016) conducted a meta-analysis including 36 studies with 663 MLD and 1049 control participants. According to the results, two subtypes of individuals with MLD were identified. The first group, characterised with MLD, had linked reading problems and difficulties in verbal short-term memory and working memory, whereas the second group presented with difficulties with visuospatial short-term and working memory, but no problems associated with reading skills.

Taken together, these findings illustrate that MLD is a heterogeneous disorder, resulting from individual deficits in basic numerical processing and/or arithmetic functioning and highly associated with cognitive deficits. This means that arithmetic difficulties reflect individual differences in both numerical and non-numerical functions. Children with MLD have problems mastering a wide range of numerical skills such as counting skills, magnitude processing, arithmetic, transcoding between number words, digits and quantities, spatial number representation, and more domain-general skills like working memory or attentional processes, in particular, visuospatial working memory (Kucian & von Aster, 2015). Although there has been a growth in the literature investigating the causes and origins of MLD in recent years, there is still no consensus about subgroups of MLD and the interplay between domain-specific and domain-general skills in this condition. Several studies have attempted to distinguish different profiles in MLD, and it is clear that one or several deficits in domain-specific skills are implicated in association with difficulties in domain-general skills, such as working memory and/or attentional processes. It remains unclear, however, whether the high rate of visuospatial working memory deficits and executive deficits in children with developmental dyscalculia is due to high comorbidity of developmental dyscalculia with attention and/or working memory problems, or whether those deficits are core features of developmental dyscalculia itself (Gilmore, Göbel, & Inglis, 2018).

2.8.5 Brain correlates of mathematical learning disabilities

MLD was initially conceptualised as a disorder of a single brain region localised in the IPS. Indeed, neuroimaging studies suggested reduced grey matter in the intraparietal sulcus and adjacent regions, including the superior parietal lobe (Cipolotti & van Harskamp, 2001; Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Price, Holloway, Räsänen, Vesterinen, & Ansari, 2007). Other aberrant areas have been implicated in MLD, however. Recently, MLD has been characterised as a disorder of brain plasticity in multiple functional systems (Iuculano, 2016). These include not only the posterior parietal cortex, but also prefrontal areas, necessary for task/rule switching and error monitoring, ventral temporal-occipital regions, important for maintaining and manipulating information in WM, as well lateral and medial temporal cortices, implicated in the retrieval of maths facts and anchored in the medial temporal lobe. Together, these areas serve multiple cognitive functions necessary for successful arithmetic performance (Cho, Ryali, Geary, & Menon, 2011; Fias, Menon, & Szucs, 2013; Iuculano & Kadosh, 2014; Iuculano et al., 2015). For example, reduced grey matter has been reported in the anterior cingulate cortex, the left inferior frontal gyrus, and the dorsolateral prefrontal cortex (Rotzer et al., 2008). In addition, the ventral visual stream seems to be affected. Reduced grey matter volumes have also been found in the fusiform gyrus, parahippocampal gyrus, and the right anterior temporal cortex, which might obstruct the development of semantic memory representations crucial for numerical fact retrieval (Rykhlevskaia, Uddin, Kondos, & Menon, 2009).

Studies investigating white matter in children with MLD show inconsistent results. White matter deficits have been reported in the left frontal lobe and in the right parahippocampal gyrus (Rotzer et al., 2008), but reduced white matter volumes have also been found in the right temporoparietal region and the splenium of the corpus callosum (Rykhlevskaia et al., 2009). These areas are important for fact retrieval and spatial memory processing, as well as in visuospatial processing during the acquisition of mathematical skills. In addition, white matter projection fibres linking the right fusiform gyrus with the temporoparietal white matter present reduced white matter in children with MLD (Rykhlevskaia et al., 2009).

Overall, MLD can be described as a heterogeneous learning disorder that is the result of multifaceted disturbances in one or multiple neurocognitive systems, such as dorsal and/or ventral stream, frontoparietal networks, that are engaged in the performance of arithmetic operations.

2.8.6 ANS and mathematical learning disabilities

Whether children with mathematical learning disabilities are impaired in non-symbolic number comparison is unclear. Several studies have found significantly lower precision for non-symbolic number comparison in children with MLD (Price et al., 2007; Mussolin, Mejias, & Noël, 2010; Piazza et al., 2010; Mazzocco et al., 2011; Moll, Gobel, & Snowling, 2015). For example, Piazza et al. (2010) established for the first time an association between an impaired number sense and developmental dyscalculia. Using a dot magnitude comparison task, children aged between eight- and twelve-years-old with developmental dyscalculia (DD) scored at the level of five-year-old normally developing children and those scores were able to predict performance on tasks involving the manipulation of symbolic numbers. These results were interpreted as evidence for delayed development of the ANS in children with DD.

Other studies, however, have not found significant differences in non-symbolic number comparison between children with MLD and typically developing children (Rouselle and Noel, 2007; Landerl and Kolle, 2009; De Smedt and Gilmore, 2011). Disagreements in the results might reflect the possibility that deficits in non-symbolic numerical representations are not sufficient to account for the complex and heterogeneous clinical picture of MLD. Difficulties in non-symbolic numerical representations (Butterworth, 2010), in symbolic numerical representations (i.e. Mussolin, De Volder, Grandin, Schlogel, Nassogne and Noelal, 2009), or in the ability to link symbolic and non-symbolic representation (i.e. Rubinsten and Henik, 2005) have all been associated with MLD, including inhibition (Bull, Johnston, & Roy, 1999; Bull & Scerif, 2001; Espy et al., 2004; Passolunghi & Siegel, 2004), spatial processing (Rourke & Conway, 1997), and working memory (Geary, 2004; Swanson, 2006; Bull, Espy & Wiebe, 2008). Szucs et al. (2013) found inhibition impairments in children with MLD, lower visuospatial short-term memory and working memory performance, but no impairments in tasks tapping the ANS. Together, these results illustrate that MLD is a complex and heterogeneous clinical picture.

Another aspect to be considered in studies investigating the association between deficits in non-symbolic numerical representation and MLD is the diverse methodology employed to characterise them. Bugden and Ansari (2015) found that children with MLD demonstrated greater ANS deficits when visual perceptual cues were incongruent with numerical magnitude but did not show any deficits when visual perceptual cues were congruent with numerical magnitude. The authors suggest that ANS deficits in children with MLD were driven by their inability to inhibit the visual perceptual cues of the dot stimuli to choose the numerically larger quantity, and thus that individual differences in visuospatial working memory predict performance during incongruent trials in children with MLD (Bugden & Ansari, 2016). There results demonstrate that accuracy of non-symbolic magnitude judgements can be influenced by the visual characteristics of stimuli and the different methodology employed might measure different cognitive constructs contributing to different results.

Overall, a proportion of children with MLD, but not all, show impairments in non-symbolic number comparison tasks, tapping the ANS. Impairments in domain-general abilities, such as attention and working memory, are also found in children with MLD, who may or may not have impairments in the ANS, with the underlying causes of MLD still unclear. It is reasonable to infer that there are different profiles of children with MLD and not necessarily all of them present with a number sense impaired.

2.8.7 Brain correlates of ANS and mathematical learning disabilities: fMRI

The neural substrate of basic numerical processing has been evaluated using fMRI in children with MLD. Price et al. (2007) demonstrated atypical activation in the right intraparietal sulcus during a non-symbolic numerical magnitude processing task, suggesting either a weakened parietal representation of numerical magnitude in MLD and/or a reduced ability to access and manipulate numerical quantities. Kauffman et al. (2009) also found atypical activation in intraparietal areas in children with MLD, with stronger activations in inferior parietal cortices bilaterally (intraparietal sulcus, supramarginal gyrus, extending to the left angular gyrus), suggesting a compensatory neural activity in left (intra)parietal regions in children with MLD (Kaufmann, Wood, Rubinsten, & Henik, 2011). Kucian et al. (2011) demonstrated that children with MLD showed stronger activation of frontal brain regions related to modulation of numerical distance when compared to typically achieving age-matched controls, yet with comparable behavioural performance. These results suggest that performing numerical distance tasks might be more difficult for children with MLD, reflecting higher engagement of visual working memory areas (Kucian, Loenneker, Martin, & von Aster, 2011). In contrast, Kucian et al. (2006) found no significant differences between children with MLD and their typically developing peers. The results revealed similar parietal and prefrontal activation patterns in MLD children compared to controls in non-symbolic magnitude comparison tasks (Kucian et al., 2006). In line with these results, Kovas (2009) found no significant increased or decreased activation related to non-symbolic numerical estimation in inferior parietal areas of the brain (Kovas et al., 2009). A recent longitudinal study investigated neural trajectories of numerical abilities in 28 children with MLD and matched-controls. Over a period of four years, behavioural and fMRI evaluations were carried out semesterly. The results revealed that, over time, typically developing children improved in numerical abilities and showed a consistent and well-developed fronto-parietal network. In contrast, MLD children revealed persistent deficits in number processing, with brain imaging results showing an age-related activation increase in parietal regions (intraparietal sulcus), pointing to a delayed development of number processing areas. In addition, an activation increase in frontal areas was observed over time, indicating the use of compensatory (Rosenberg-Lee et al., 2015).

Overall, neuroimaging studies investigating the neural underpinnings of basic number processing in children with MLD show inconsistent results. Nonetheless, they do in general seem to point to a relative increase in activity, with bilateral activation of areas related to number-related tasks in children with MLD when compared to their matched-controls. This might be indicative of compensatory mechanisms.

2.8.8 Brain correlates of ANS and mathematical learning disabilities: ERPs

The literature investigating the neural correlates of basic numerical capacities employing event related potentials (ERPs) in children suffering from MLDs is rather sparse. For instance, investigating the neural correlates of non-symbolic magnitude processing in children with MLD, Heine et al. (2013) showed clear differences between children with MLD and their matched controls. No late parietal numerical distance effects were found for the group of children with MLD. Gomez-Velazquez et al. (2015) investigated numerical magnitude comparison in a group of children with low mathematical achievement. Lower amplitudes in components modulated by number-tasks were observed in the group of children with lower mathematical achievement (Gómez-Velázquez, Berumen, & González-Garrido, 2015). Soltész et al. (2007) reported similar results for a group of adolescents with MLD compared to age-matched controls and adults. The authors demonstrated that while ERPs indicating early, i.e., more automatic, processing steps were similar for all groups, correlates of later, i.e., more controlled stages of numerical information processing, were less homogeneous (Soltész, Szucs, Dékány, Márkus, & Csépe, 2007). Overall, with just a handful of studies, it is difficult to draw firm conclusions. However, the literature seems to point to the same directions; it can be suggested that children with MLD do show atypical modulation during non-symbolic processing. More research will likely clarify whether magnitude processing is impaired in symbolic, non-symbolic, or both formats.

2.9 Mathematical difficulties in preterm children

Although VP children have poor performance across all school subjects, specific difficulties have been found in mathematics, with preterm children's performance being 0.60 – 0.70 SD below their term-born peers. These differences persist after controlling for IQ or excluding children who have neurosensory impairments (Aarnoudse-Moens et al., 2009; Twilhaar et al., 2018). Commonly, only very preterm children may be at significantly increased risk of scoring below 1 SD in mathematic tests (Poulsen et al., 2013). The chances of having MLD are remarkably higher in children who are born very preterm (39.4%) compared to their term-born peers (14.9%) (Jaekel & Wolke, 2014), although Johnson et al. (2016) found that MLD was much higher among EP children than controls (12% *vs.* 1%).

The prevalence of MLD and mathematical difficulties may differ according to the operationalisation of the MLD definition. For example, comparing discrepancy-based and achievement-based criteria, Jaekel et al. (2014) illustrated that children across the whole GA range were at increased risk of a discrepancy-based MLD diagnosis but not of a fixed cut-off score diagnosis. Studies using the discrepancy-based criterion examining academic achievement in the preterm population have generally shown higher rates of academic difficulties amongst children born very preterm (e.g., (Grunau, Whitfield, & Davis, 2002; Litt, Taylor, Klein, & Hack, 2005). For example, Litt et al. (2005) employed both criteria when investigating academic attainment in extremely preterm children, excluding those with ID, neurosensory impairments or neurodevelopmental impairments. Using the achievement-based definition, 15% of children in the EP group, 12% of children in the VP and 4% of children in the full-term group met the criterion for an RLD. The rates for MLD were 14%, 5% and 3% for the EP, VP and control group, respectively. Combined RLD and MLD were

25%, 10% and 7% for the EP, VP and control group, respectively. When using a discrepancy-based definition, differences were found in the rate of MLD in the EP group, representing 40% of the sample, relative to the full-term group (20%) (Litt, Taylor, Klein, & Hack, 2005). Although MLD rates were higher in the preterm group using both methods, greater differences were found between the term and preterm group for MLD when using the discrepancy-based method.

On the other hand, Simms et al. (2013a) suggested that either the achievement-based criterion or consistently poor mathematical achievement over a period of two school years could be a more appropriate measure for the preterm population. Since children born prematurely are generally impacted in their general cognitive abilities, employing a discrepancy-based criterion would not identify potential mathematical difficulties.

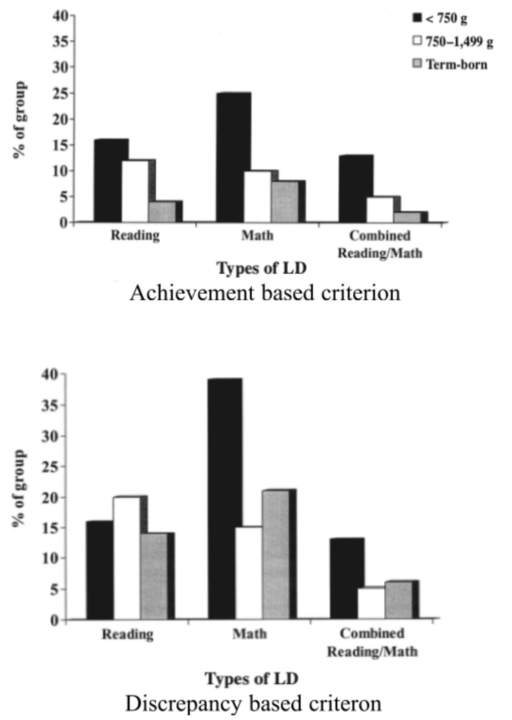


Figure 2.15: Learning disabilities in preterm children.

Regardless of the criteria, academic-achievement (on top) or discrepancy-achievement (on bottom), maths is the academic domain most affected in children born prematurely (Litt et al., 2005).

In order to understand the underlying causes of mathematical difficulties in this population, it is also important to disentangle the subcomponents of mathematical performance. Most existing studies rely merely on standardised tests. These are very general tests, however, with composite measures of attainment in mathematics and thus do not allow exploration of specific areas of difficulty. Given the wide variation in mathematics tests, comparisons between studies become problematic. When identical measures are used, such as the Woodcock–Johnson-III, a similar pattern of difficulties is observed across studies, with VP children displaying larger deficits in the Applied Problems subscale compared with the Math Fluency subscale (Taylor et al., 2011). This indicates greater difficulty with the application of mathematical concepts, rather than with knowledge of basic mathematics facts, indicating that problems in mathematics in preterm children may be related to the application of domain-specific skills in more complex mathematical problem-solving scenarios, rather than to perform in low-level mathematical tasks (Simms et al., 2013a). Investigating both mathematical achievement and specific maths skills in the very preterm children, Simms et al. (2015), found significantly poorer mathematical achievement, counting proficiency and use of less sophisticated strategies when solving simple arithmetic problems when compared to their term-born peers. Importantly, group differences in strategy use and counting were negated after controlling for working memory and visuospatial skills. Even just after school entry, preterm children already show difficulties in mathematics strongly associated with EF (Simms et al., 2015).

In fact, research has suggested that preterm birth negatively affects mathematical abilities mediated by general-domain skills, such as EF. For example, Espy et al. (2004) investigated the relationship between pre-schoolers' mathematical skills and executive functions in preterm children. Working memory and inhibition were related to a mathematical composite score after controlling for maternal education and child vocabulary (Espy et al., 2004). Longitudinal studies have described difficulties in general processing speed in preterm infants from the first year of life, with cascading effects on cognitive abilities at preschool (Rose, Feldman, Jankowski, & Van Rossem, 2008). It is reasonable to infer that early difficulties in processing speed and working memory would affect not just mathematical abilities, but also other academic skills, such as reading and writing (Anderson et al., 2003; Sansavini, Guarini, & Caselli, 2011; Kovachy, Adams, Tamaresis, & Feldman, 2015). Nonetheless, mathematical attainment is the most affected academic domain. Mathematical difficulties, in particular, show an area of specific difficulty, even when taking general cognition into account (Taylor et al., 2009). Recent studies have stressed the fact that visuospatial processing (Geldof, Wassenaer-Leemhuis, de Kieviet, Kok, & Oosterlaan, 2011; Johnson et al., 2011), perceptual-motor abilities (De Rose et al., 2013), and executive functions, such as processing speed and working memory (Mulder, Pitchford & Marlow, 2010; Rose, Feldman & Jankowski, 2011) are important factors underlying academic attainment, and those might mediate the effects of preterm birth on maths achievement at school age (Mulder et al., 2010; Rose et al., 2011).

Altogether, these research studies indicate that a high proportion of children born very preterm, even without intellectual disabilities and neurosensory impairments, show difficulties in mathematical abilities. These deficits appear even prior to school entry and are intimately associated with domain-general skills, such as visuospatial processing and executive functions.

2.9.1 Brain correlates of mathematical difficulties in preterm children

Studies using neuroimaging techniques elucidating why preterm children present a specific susceptibility to mathematical failure are rather scarce. Isaacs et al. (2001) conducted a voxel-based morphometry study in a group of adolescents who had been born very and extremely preterm and who presented with deficits in calculation abilities together with normal range intelligence and reading scores (Isaacs, Edmonds, Lucas, & Gadian, 2001). The authors demonstrated that the preterm group had less grey matter in the left intraparietal sulcus (IPS), an area particularly important for processing basic numerical processing information. The authors concluded that impairments in these types of low-level skills were responsible for poor achievement in mathematics. Since the IPS also plays an important role in recruiting areas involved in visuospatial working memory, this could explain why visual working memory difficulties were predictive of mathematical failure in very preterm children, although these domains were not assessed in this study. Furthermore, a longitudinal study by Ulman et al. (2015) used neonatal magnetic resonance imaging at term equivalent to predict skills important for mathematical achievement, such as working memory, in those children when aged five and seven. The results identified localised regions around the insula and putamen positively associated with early mathematics at five and seven years, even after covarying for important perinatal clinical factors. This study highlights the importance of working memory with mathematical skills in this population early in life (Ullman et al., 2015).

2.9.2 The ANS in preterm children

Deficits in formal mathematical skills have been extensively recognised in preterm children (Aarnoudse-Moens et al. 2009; Johnson et al., 2011), but only a handful of studies have investigated their basic numerical skills (Hellgren et al., 2013; Simms, Gilmore, et al., 2013b; Guarini et al., 2014; Tinelli et al., 2015; Libertus et al., 2017). Even though it is debatable whether this is the best predictor, it has been claimed that the ANS is a potential tool to elucidate primary mechanisms in numerical cognition for identifying groups at risk and for targeting early, such as the preterm population. Due to the different methodologies applied in these studies, the underlying mechanism of mathematical performance in the preterm population is unclear. On the one hand, research shows that difficulties in mathematical performance are linked to deficits in numerical representations independent of domain-general skills (Hellgren et al., 2013; Libertus et al., 2017). This implies that individuals born prematurely with mathematical problems might have the classic dyscalculic profile, with a core deficit in understanding and manipulating the quantity of sets and their numerosities (Butterworth, 1999, 2005, 2010). On the other hand, studies have been showing that difficulties in mathematics performance in the preterm population are driven by executive functions affecting, albeit not exclusively, numerical representations. The decrease of gestational age plays a crucial role in identifying those with difficulties in numerical representations, with individuals with lower gestational ages (e.g., EP) being more susceptible to have numerous cognitive difficulties, including deficits in numerical representations. On the other hand, studies have suggested that differences in performance in tasks evaluating the ANS between preterm children and their term-born peers are mediated by difficulties in executive functions (Simms et al., 2013b; Guarini et al., 2014; Simms et al., 2015, Tinelli et al., 2015). Next, I review studies investigating numerical representations in individuals born prematurely according to gestational age (extremely preterm and very preterm).

Studies investigating numerical representation in extremely preterm children are sparse, but generally indicate difficulties in numerical representation. The mechanisms underlying those difficulties, as mentioned previously, are inconclusive. For example, Hellgren et al. (2013) investigated the approximate number system in EP children alongside general cognitive abilities (working memory, processing speed, and visual attention). The authors claimed that extremely preterm children had specific impairments in the approximate number system, and that those deficits were not a consequence of a general cognitive deficit, poor working memory, poor attention, or slow processing speed. Surprisingly, the authors concluded that deficits in numerical representations were not associated with difficulties in domain-general skills, although EP children performed significantly poorer than their full-term peers in the domain-general tasks. The authors also claimed that difficulties in mathematical performance could be associated with deficits in numerical presentation, but this was not investigated in their study. Another study investigated numerical representations in EP children employing a number estimation task (Simms et al., 2013b). Children were shown a set of dots on a single page that varied in quantity and were asked to give a verbal response with alternatives (e.g.: "Do you think there are 20, 40, 60, or 80 dots?"). The results demonstrated that preterm children perform significantly worse than their term-born peers. Significant correlations were observed between numerical representation and mathematical performance, even after controlling for domain-general skills, but only for preterm children. The authors concluded that the results indicated that EP children's attainment in mathematics was associated with their underlying accuracy of numerical representations and was not simply a component of their general cognitive ability. Libertus et al. (2017) suggested that domain-general skills were not associated with numerical representation in extremely preterm children. According to their results, strong associations were found between ANS acuity and maths performance (r= 0.40). Preterm children showed significantly lower ANS acuity than their term-born peers, even when controlling for verbal IQ, perceptual reasoning skills, working memory and attention. The authors argued that the ANS acuity task was a unique predictor of term-born children's math ability, even when controlling for processing speed, concluding that extremely preterm children do have deficits in the ANS. The authors' claims remain debatable, however, since another study systematically investigating the tasks employed in this study (Panamath), suggested that perceptual skills highly influence the numerical domain and that discrimination is not based uniquely in numeric representation, but rather perceptual cues (Bugden and Ansari, 2016).

Similar results were observed in studies investigating numerical representations in VP children. Generally speaking, VP children face difficulties in numerical representations. Arguably, these difficulties seem to be associated with deficits in executive functions affecting basic and advanced numerical abilities. For example, Guarini et al. (2014) investigated very preterm children testing magnitude comparison (symbolic and non-symbolic), number knowledge (e.g., counting, and reading and writing Arabic numerals), and intelligence in VP children aged between six and eight years old. The results showed that preterm children were as accurate as their term-born peers, but were significantly slower in non-symbolic magnitude comparisons, suggesting deficits in processing rather than numerical representations. Preterm children aged six years old were faster but less accurate than their full-term peers when comparing Arabic digits. At eight years old, preterm children achieved a similar numerical symbolic level to typically developing children, but were slower than their term-born controls. This could be explained by the fact that, at six years, preterm children made several errors, while full-term children were more accurate, spending more time searching for the correct response. This implies the involvement of more elaborate monitoring skills at eight-years-old and better strategies, showing a maturation of EF and catch-up effects. Tinelli et al. (2015), meanwhile, investigated non-symbolic magnitude comparison in very preterm children and did not find significant differences between children born prematurely and controls, implying that the ANS was not compromised in this population.

Perhaps the most well-designed study investigating the impact of prematurity in both domain-specific and domain-general skills related to mathematical performance is that of Simms et al. (2015). The authors carefully investigated different domains of mathematical performance, not just by standardised testing, but also by designing different tasks to assess math-specific domains. A comprehensive assessment was undertaken of general cognitive abilities (IQ, working memory, processing speed, visuospatial skills and inhibition) and mathematical achievement (measured by a WIAT-II) using standardised measures, whereas specific mathematics skills (symbolic and non-symbolic magnitude comparison, number line estimation, digit recognition, counting skills, number fact knowledge, arithmetic strategy and arithmetic concepts) were assessed with experimental tests. Very preterm children were found to have significantly poorer mathematical achievement than term-born children, as well as poorer working memory and visuospatial skills. Differences between inhibition and processing speed were not found between groups, however. Although preterm children had poorer performance in specific mathematics skills, there was no evidence of imprecise numerical representations. Difficulties in mathematics were rather associated with deficits in visuospatial processing and working memory.

To sum up, a small number of studies have investigated numerical representations in prematurity with inconsistent results. Some studies claim that this population has an impaired representation of numerosities (Hellgren et al., 2013; Libertus et al., 2017) unrelated to general-domain skills, supporting the core-deficit theory. In contrast, other studies claim that impaired representation in this population is caused by multiple cognitive components (Simms et al., 2013b; Guarini et al., 2014) such as in processing speed and visuospatial skills. Only two studies have investigated the association of numerical representations and mathematical performance (Simms et al., 2015; Libertus et al., 2017). Simms et al. (2015) found that very preterm children do not have imprecise representation and their difficulties related to maths are associated with domain-general skills, such as visuospatial skills. In contrast, Libertus et al. (2017) suggested that difficulties with maths are associated with imprecise numerical representations. No difficulties with symbolic representations were found in any study.

A summary of the studies reviewed here is provided in table 2.1. A few considerations are necessary before drawing any conclusions, however. Firstly, the gestational age seems to play an important role in the development of numerical representation. The lower the gestational age the poorer the numerical representations. Thus, gestational age seems to affect both basic and advanced numerical skills. Secondly, studies investigating extremely preterm children are likely to be influenced by IQ, which could explain the finding of imprecise number representation. For example, in the study conducted by Simms et al. (2013b), general cognitive factors accounted for a substantially larger proportion of the variance (70%) than numerical representations. Some studies, however, claim that, even after controlling for general-cognitive abilities, difficulties in numerical representations persist (Libertus et al., 2017). Thirdly, the variety of measures, both for standardised assessments and experimental tasks, makes comparison between studies difficult. For instance, while Simms et al. (2015) employed the WIAT-II to assess mathematical abilities in their study, Libertus et al. (2017) employed the arithmetic subtest from the WISC-IV. Similar problems are faced when investigating numerical domain-specific skills, with different experiment designs found across studies.

Given these considerations, taken together, the studies suggest that brain maturation at birth affects the development of numerical abilities. Extremely preterm children show difficulties in basic and advanced numerical skills and it is debatable whether these deficiencies are associated with domain-general skills. In contrast, it seems evident from the literature that difficulties with mathematical performance in very preterm children is driven by domain-general skills, most distinctly by executive functions. Problems with EF seem to affect both basic and advanced numerical skills, including our sense of numerosity.

Table 2.1: Summary of studies investigating numerical representation in the preterm population

Study	Age group	GA	N preterm	N control	ANS task	Mathematics difficulties*	ANS difficulties	Domain-general abilities tested	Academic abilities tested
Hellgren et al., 2013	6	EP	65	47	NS comp (acc, RT, w)	Not tested	Yes	IQ, processing speed, verbal working memory and visual attention	Not tested
Simms et al., 2013	11	EP	219	153	Number estimation	Yes, associated to general cognitive difficulties	Yes	IQ, sensorimotor, visuospatial processing, attention and executive function.	Mathematics and reading
Libertus et al., 2017	6.5	EP	82	89	NS comp (acc, RT, w)	Yes, associated to ANS	Yes	IQ, working memory, attention	Mathematics
Guarini et al., 2014	6 and 8	VP	140	60	NS and S comp (acc and RT)	Not tested	No, slower processing speed	IQ	Experimental measures of numerical knowledge
Tinelli et al., 2015	8	VP	28	26	NS comp (w)	Not tested	No	IQ	Not tested
Simms et al., 2015	8-10	VP	115	77	NS and S comp (acc and RT)	Yes, associated to visuospatial skills	No	IQ, working memory, processing speed, visuospatial skills and inhibition	Mathematics and subcomponents of numerical knowledge

GA=gestational age; NS = non-symbolic; S = symbolic; comp = comparison task; acc = accuracy; RT = response time; w = Weber fraction.
*Mathematical difficulties according to achievement-based criterion on standardised assessments.

2.9.3 Brain correlates of number processing in preterm children

Studies investigating the neural basis of numerical representations in the preterm population are even scarcer than behavioural studies. In fact, only one study has investigated non-symbolic and symbolic numerical representations in adults born prematurely, testing very, late and moderate preterm individuals. In total ten individuals born prematurely (M=32,2 GA weeks SD=2.94) and ten term-born matched adult controls underwent an fMRI scan while performing symbolic and non-symbolic numerical magnitude comparison tasks. Calculation fluency scores were also obtained. Overall, no differences were found between the preterm and the control in terms of their behavioural performance in the numerical magnitude comparison tasks. The fMRI results, however, showed an increased activation response in inferior frontal and parietal regions when comparing non-symbolic magnitudes in the preterm group (figure 2.17). Individuals with lower calculation fluency exhibited greater signal change for the non-symbolic dot conditions in the frontal and parietal regions. The results signal that although the behaviour data did not indicate differences between groups, the preterm group might have developed compensatory strategies to discriminate numerical representations, explaining the combination of greater activations and similar behavioural performance (Clark, Liu, Wright, Bedrick, & Edgin, 2017).

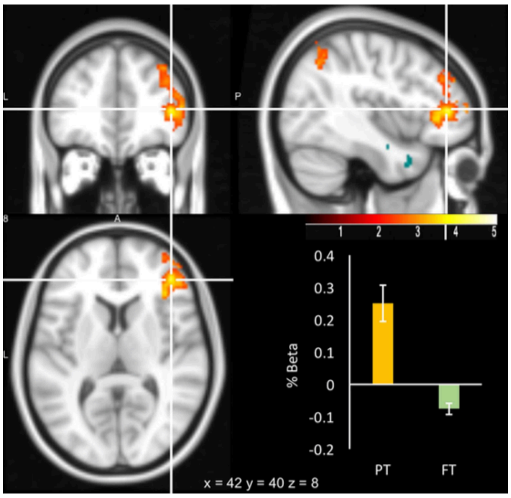


Figure 2.16: Schematic representation of activations found in the study by Clark et al. (2017).

Red colour illustrates greater activations in preterm and green colour shows the greater activations exhibited by full-term in a non-symbolic magnitude comparison task. The preterm group showed increased activity in the middle frontal gyrus and right inferior parietal lobule during the non-symbolic task. Individuals with lower calculation fluency exhibited greater signal change for the non-symbolic comparison tasks in the frontal and parietal regions.

The number of studies exploring symbolic representations in the preterm population is slightly more representative, although still discrete and solely based on magnetic resonance imaging. For example, a voxel-based morphometry study tested sixteen very preterm children aged between six and seven years old alongside intelligence, maths proficiency and symbolic numerical representations (Starke et al., 2013). The results revealed increased grey matter volumes in their left IPS and decreased white matter density in their fronto-parietal brain structures, which are relevant to number skills. The extent of grey and white matter in the right anterior IPS was the only region significantly correlated with behavioural performance in the symbolic comparison tasks. Klein et al. (2014) tested both symbolic magnitude comparison and physical Stroop task (when participants are asked to compare the physical magnitudes[6] of the

[6] Physical magnitudes are also known as intentional and automatic number magnitudes, respectively.

presented number). The results indicated that the lower the general cognitive ability and maths proficiency the more widespread were the number-relevant fronto-parietal activations associated with number magnitude processing. Additionally, decreasing maths proficiency was associated with increased activation in the left inferior frontal cortex. While increasing gestational age was associated with more (intra)parietal activation, decreasing gestational age was related to more activation in frontal cortex areas. In line with these results, a recent study explored the association of gestational age with the neural correlates of number processing in six- and seven-year-old children born prematurely (n=16), as measured by intentional and automatic number magnitude processing. The results indicated that gestational age reliably predicts the frontal-to-parietal shift of activation observed for the symbolic magnitude comparison (Klein et al., 2018).

Taken together, those results indicate, firstly, that preterm might have a compensatory strategy to discriminate numerical representations, eliciting wider activations in networks recruited for numerical tasks. Secondly, the gestational age is inversely associated with the number of areas elicited. Thirdly, not all studies showed correlations between behavioural performance in numerical magnitude tasks and mathematical performance, but those with lower mathematical performance show wider networks modulated by numerical tasks.

2.10 Summary

Individuals born prematurely, especially those born before 32 weeks gestational age, are particularly at risk for deficits in executive function (Taylor & Clark, 2016) and visuospatial skills (Simms et al., 2015). Difficulties in these cognitive domains are likely to have a direct effect on their academic abilities. In fact, academic difficulties are evident in this population, with mathematical performance being particularly at risk, among other academic domains (Taylor, Espy, & Anderson, 2009). Problems in formal mathematical abilities were observed in distinctive mathematical subdomains. For example, struggles in excelling in counting proficiency, and less sophisticated strategies when solving simple arithmetic problems are common difficulties associated with this population (Simms et al., 2013a). The underlying mechanisms of those difficulties are rather unclear, however. On the one hand, research has shown that difficulties with maths in this population are associated with domain-general skills, such as EF, working memory and visuospatial skills (e.g., Simms et al., 2015). On the other hand, studies have claimed that difficulties with maths are solely associated with domain-specific skills (e.g. Libertus et al, 2017).

Research coming from typically developing children has shown that the building blocks for mathematical performance are based both in domain-general and domain-specific skills. While domain-general skills, such as EF, working memory, processing speed and visuospatial skills, are particularly important for the typical development of mathematical abilities, numerical magnitude comparison is an important domain-specific skill also crucial for the development of mathematical abilities (Gilmore et al., 2018). The ability to discriminate non-symbolic quantities seems to have an important role before the acquisition of formal mathematical abilities, whereas our exact symbolic representation has a predictive value after the start of formal education (Mazzocco, Feigenson, & Halberda, 2011). Together, EFs domain-general skills and numerical magnitude comparison — a domain-specific skill — are important predictors of formal and advanced mathematic skills. Certain abilities in both domains are possible to assess even during infancy (Feigenson, Dehaene, & Spelke, 2004). For example, the ability to discriminate quantities is possible to assess even in infants, and has been shown to be linked to later mathematical abilities during childhood (Starr et al., 2013). Hence, testing those abilities seems to be a potential tool to identify those at risk of struggling at maths.

Formal and advanced numerical skills have been extensively investigated in the VP population (Jaekel & Wolke, 2014; Basten et al., 2015; Johnson et al., 2016). Few studies, however, have explored this population's basic numerical abilities, with just a handful of studies investigating the performance of numerical magnitude comparison (Hellgren et al., 2013; Guarini et al., 2014; Tinelli et al., 2015; Libertus et al., 2017). The results are ambiguous due to the mixture of methodologies applied and different gestational ages investigated. Whereas it seems that EP children do have difficulties in discriminating non-symbolic quantities (e.g., Libertus et al., 2017), the origins of those difficulties are unclear. Studies investigating VP children are more conclusive, however, in showing slower performance than their peers (e.g., Guarini et al, 2014). This suggests that the nature of their difficulties in numerical representation is associated with deficits in EFs, an umbrella of cognitive skills often impaired in the preterm population and crucial for mathematical proficiency. Although differences between VP children and their peers are clear, difficulties in numerical representation account for very little of the differences in their performance in standardised mathematical tests.

Studies investigating numerical representations in the VP population, however, are based mostly on behavioural outcomes. Only a few studies have investigated the neural correlates of basic numerical representation in the preterm population (Clark et al., 2017; Klein et al., 2014, 2018). Employing fMRI, these studies indicate that individuals born prematurely recruit wider brain areas when performing numerical talks and the greater activation is associated with low mathematical proficiency.

Remarkably, no studies have investigated the early foundations of numerical abilities in infants born prematurely. It is crucial to understand the very basic processes underlying the sense of numerosity in its early stages, and how these can go awry. The identification of early markers of altered or delayed developmental trajectories in mathematical skills is important because of the potential for early intervention in the preterm population at increased risk of presenting low mathematical achievement.

2.11 Rationale, Aims and Hypothesis

In contrast to the large body of research into cognitive processes and academic performance in school-aged children born prematurely, there is a lack of research that seeks to evaluate the early trajectories of the emerging building blocks of numerical skills in children born very prematurely, a high-risk population for presenting with mathematical difficulties. Identifying those building blocks and their delayed emergence may allow for the early identification of children at risk of later mathematical difficulties and thus help to formulate interventions targeted to assist those struggling at school. Thus, my first research question is: *Do infants born very prematurely show imprecise numerical representations?*

The foundations of mathematical abilities are built upon both domain-general and domain-specific skills. Thus, it is crucial to investigate both domains during infancy and their associations. Thus, my second question is: *Is numerical sensitivity in VP infants associated with their visual working memory capacities?*

Our knowledge regarding numerical representation in children born prematurely is based on only a handful of studies. It seems unclear whether VP children have imprecise numerical representations and if so, whether those difficulties account for difficulties in mathematical attainment. Thus, my third question is: *Are difficulties in mathematical performance in VP children associated with domain-general or domain-specific skills?*

Finally, there has been little use of neuroimaging techniques to investigate basic numerical competency in older children who had been born preterm. The application of such techniques may shed light on the neural correlates of poor performance, particularly in maths. My final question, therefore, is: *Do VP children have with atypical neural correlates of numerical representation (symbolic and non-symbolic)?*

2.11.1 Aims

In this thesis, studies examining the impact of prematurity in number processing will be presented investigating a group of very preterm infants and children (VP henceforth) with relatively uncomplicated neonatal courses, compared to a group of term children (term henceforth), combining eye-tracking technology, neuropsychological assessments, experimental measures and ERP outcomes.

The specific study objectives are:

1) To investigate numerical sensitivity in VP infants in the first year of life. Numerical sensitivity was assessed using eye-tracking technology to measure infants' looking behaviour at two-time points. A number familiarisation task was designed to address the ratio-dependency exhibited by typically developing infants at six months of age (ratio 1:2) and twelve-months of age (ratio 2:3), respectively (Study 1

and Study 2). Development measures were also incorporated to run out any developmental delay. Differences between VP and FT infants were investigated.

2) To explore the relationship between visual working memory and numerical sensitivity between the term and VP cohorts at twelve months of age (Study 2). The ability to discriminate a number sensitivity and a Piagetian A-not-B visual working memory tasks were employed. Differences between VP and FT infants were investigated.
3) To investigate the relationship between domain-general and a set of domain-specific skills (numerical magnitude comparison) and mathematical performance in VP and FT children aged between eight and ten years old (Study 3). Domain-general skills included intelligence, processing speed, working memory, attention, planning and inhibition, all assessed by standardised tests. Domain-specific abilities comprised numerical magnitude (symbolic and non-symbolic) tested by experimental paradigms.
4) To explore differences in neural correlates between the term and VP children when comparing numerical magnitudes (Study 4). Employing ERPs, children aged between 8 and 10 years old were subjected to two experimental tasks testing symbolic and non-symbolic numerical comparison. Differences between VP and FT infants were investigated.

2.11.2 Hypothesis

This thesis is broken down into four experimental studies: numerical sensitivity in six-month-old VP infants (Study 1); the association of numerical sensitivity and visual working memory in twelve-month-old VP infants (Study 2); the relationship of domain-general and domain-specific skills (numerical magnitude comparison) and mathematical performance in school-aged children (Study 3); and neural correlates of numerical magnitudes in VP children (Study 4). Accordingly, the following hypotheses were made for each aspect of these studies.

2.11.2.1 Numerical sensitivity in six-month-old VP infants

I hypothesise that if VP infants do not have imprecise numerical representations at early stages of development (six months of age), they will be able to discriminate the same age-dependent ratio that full-term infants have previously demonstrated (ratio 1:2 for six-month-old infants). Given the fact that VP infants have shown difficulties in domain-general abilities at early stages of development, such as processing speed, they might require a longer inspection time, indicating slower processing speed. This will be demonstrated if VP infants showed a statistically significantly longer inspection time in relation to the novel number of elements, compared to FT infants.

2.11.2.2 Numerical sensitivity in twelve-month-old VP infants

Similar to Study 1, I hypothesise that if individuals born very prematurely do not have imprecise numerical representations at early stages of development (twelve months of age), they will be able to discriminate the same age-dependent ratio tasks that full-term infants have previously demonstrated (ratio 2:3 for twelve-month-old infants). In addition, given the fact that both numerical sensitivity and visual working memory are critical building blocks for abilities later associated with mathematical achievement, I also hypothesise that the ability to discriminate a numerical task is associated visual working memory during infancy. Thus, VP infants who successfully discriminate a numerical task will demonstrate higher visual working memory skills.

2.11.2.3 The relationship of domain-general and domain-specific skills and mathematical performance in VP children

I hypothesise that if difficulties in mathematical performance in VP children are driven by domain-general skills, deficits in domain-general EF functions will account for a great proportion of the variance in their mathematical performance, replicating previous results found in the literature. In addition, significant differences between term and preterm groups in reaction times, but not accuracy, in non-symbolic numerical magnitude comparison, would illustrate that difficulties in numerical representations are driven by deficits in domain-general skills, such as processing speed.

2.11.2.4 Neural correlates of numerical magnitudes in VP children

I hypothesise that if VP children do not have difficulties in numerical representation, either symbolic or non-symbolic, no differences between term and preterm groups will be found in the neural resources recruited for number-related tasks. If the nature of their difficulties is related to domain-general skills, however, significant differences will be found between groups in the recruitment of neural resources associated with general cognitive domains. This will be noticeable in cognitive phases prior to encoding information, and in wider activations in topography maps.

3 General Methods

This chapter aims to describe the general methodological background of this thesis. Since this work is focused on distinct age groups (infants and school-aged children), this section will address the general methodology applied according to age cohorts. Methods of recruitment will be described with inclusion and exclusion criteria, followed by a general overview of the methods employed in the study and the approaches to data analysis. Specific methods relevant to each chapter are also included throughout.

3.1 Infants' studies

Study 1 (chapter 4) and 2 (chapter 5) investigated numerical discrimination in infants aged six and twelve months old. In addition, Study 2 examined general cognitive abilities and working memory skills. The sections to follow will outline recruitment, inclusion and exclusion criterion and measures employed to address the aims for the infant studies (see section 2.11.1 for research aims).

3.1.1 Recruitment of infants

Infants from Studies 1 and 2 were recruited from a database of families who expressed an interest in participating in the 'The University College London Hospital Preterm Development Project: growing up after very preterm birth" (PDP), which aimed to investigate the early brain and social-cognitive development of children born very preterm. PDP was initiated before the start of this study and additional measures to address these study questions were later included. The long-term aim of the PDP study was 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.

The preterm group was recruited from the Neonatal Unit at University College London Hospital (UCLH). A neonatologist or a research nurse approached very preterm (VP)[7] parents at the end of the first week after birth inviting them to participate in the study. An information leaflet was given to parents at the time of consent, 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. 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 UCLH Babylab.

3.1.2 Power Calculations of infants' studies

A power calculation was conducted assuming a Cohen's $d = 0.3$ with means and standard deviation based on a previous study investigation of differences in infant cognition between term and preterm groups. Specifically, we based our sample size calculations on novelty effects observed in a habituation paradigm investigating infants in the first year of life, with M=58.68, SD=5.24 for the term group and M=57.06, SD=5.75 for the preterm group (Rose et al, 2008). As is standard $\alpha=0.05$ was used, with desired power 80%. The calculation for group comparisons suggested a sample size of 181 infants for each group. Because of attrition rates are high with infants' population, a conservative sample size would need to recruit approximately 217 infants, considering 20% attrition. This

---

[7] The term *'very preterm (VP)'* represent the whole clinical group of participants born prematurely, including children born very and extremely preterm. The EP group is a subgroup of the VP group This is a common practice adopted in the literature in which the terminology of the gestational age reflects both subgroups of the preterm group. For example, Haster & Akshoomoff (2017) investigated math abilities and other general cognitive skills in a group of pre-schoolers with gestational ages ranging from 24 to 32 weeks and refereed to the clinical group as 'very preterm'.

sample size is not possible within our group. For practical reasons, our strategy was to recruit as many participants as possible within a certain time period (March 2015 to March 2017). Given the design of Study 1 and Study 2 are identical, apart from infants age, the same caveat applies for both studies.

3.1.3 Infant Sample

A total of 21 very preterm infants comprised the clinical group of Study 1 and 2. Term-born infants were recruited from antenatal classes and postnatal wards at UCLH. In order to increase the sample size of term participants for Study 2, additional infants were recruited via email and posters around the university. A total of 39 term-born infants comprised the control group of Study 1 and 2. Figure 3.1 illustrates the schematic representation of the recruitment of participants comprising Study 1 and 2.

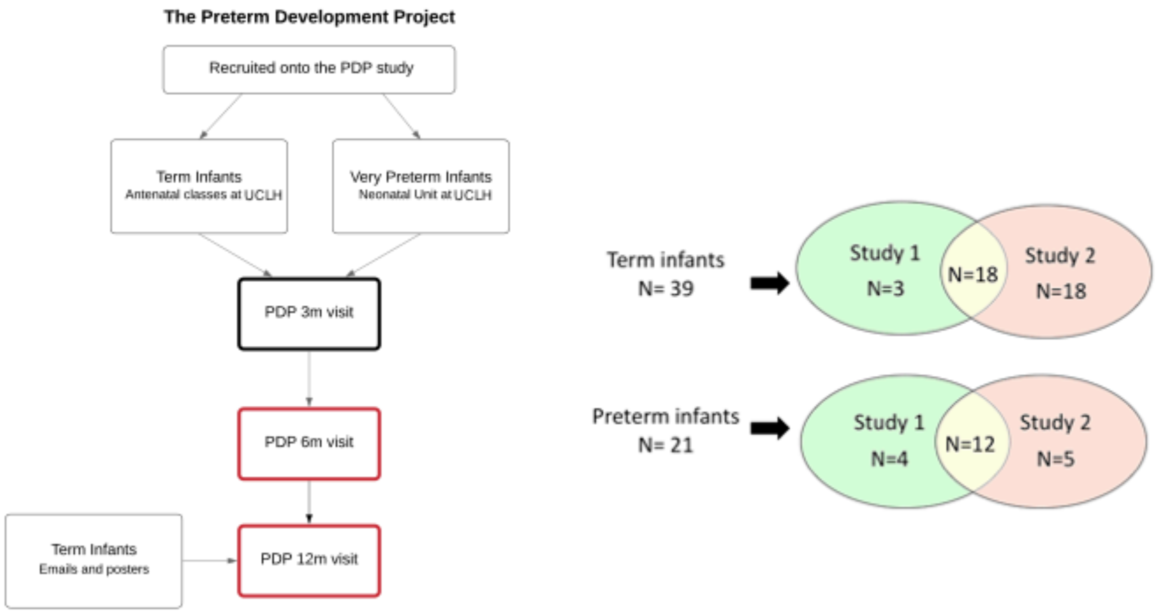


Figure 3.1: The Preterm Development Project (PDP) study structure.

Red boxes in the flow chart (left-hand side) indicate the stages of the study at which the data was collected for this thesis. The Venn diagram on the right-hand side indicates the number of participants who joined only Study 1, only Study 2, and participants who took part in both studies.

3.1.4 Inclusion/exclusion criteria for infant studies

Infants who were born at less than 31 weeks and 6 days gestational age were eligible for inclusion in the preterm group (very and extremely preterm infants). The cut off for gestational age was informed by previous literature (as discussed in section 2.1.1), whereby infants born extremely and very preterm frequently experience numerous complications, including delayed cognition and academic difficulties, whereas late and moderate preterm infants experience fewer complications. Preterm infants were excluded if there was a severe congenital abnormality or a low likelihood of survival.

Infants eligible for the term group were born between 37 and 42 weeks of gestation, had a birthweight between the 10^{th} and 90^{th} percentile for gestational age, no perinatal complications, an APGAR[8] score above 7 at five minutes and generally in good health. Infants were excluded from the term group if there was a diagnosis of a chronic condition or medical illness that could affect the developmental outcome, or if they had a hearing or visual impairment.

3.1.5 Approval for infant studies

The study was approved by the London Research Ethics Committee 2 (Reference 10/H0720/80) and was registered with the Research and Development Department of UCLH (appendix 3.1).

3.1.6 Infant's appointments

VP infants received the normal clinical outpatient follow-up assessments as part of standard preterm care, with additional

---

[8] Apgar score is an index score evaluating the condition of a newborn infant based on a rating of 0, 1, or 2 for each of the five characteristics of colour, heart rate, response to stimulation of the sole of the foot, muscle tone and respiration. Lower scores indicate a poorer condition with ten being the maximum score.

visits to the Babylab as part of the study. Term and preterm infants attended three study appointments within their first year, at three, six and twelve months of age. For the purpose of these assessments, VP infants were corrected for gestation by using their Expected Date of Delivery (EDD). The experimental paradigms included in the PDP study aimed to assess the development of executive functions, social abilities and numerical skills in a cohort of term and VP infants. For the purposes of this thesis, only experimental designs tapping into numerical skills were included, therefore only data obtained from the six- (Study 1) and twelve- (Study 2) month-old follow-up assessments were used. This age group was selected based on the extensive literature investigating infants' numerical abilities with a similar age range (Wynn, 1992; Xu & Spelke, 2000; Brannon, 2002; Xu, 2003; Xu, Spelke & Goddard, 2005; Xu & Arriaga, 2007). Additional infants were recruited for each time point to increase sample size. See appendix 3.2 for the parents' information sheet.

All infants were assessed individually at the Babylab at the Institute for Women's Health, University College London (UCL). Cognitive assessments were conducted by a research psychologist (Mérari Ferreira). Written informed consent for the participant in this study was obtained from a parent or legal guardian of the infant (appendix 3.3). Participants received a certificate and a t-shirt for taking part in the study and travel expenses were reimbursed.

Our ultimate aim was to investigate numerical discrimination in infants during the first year of life. To do this, we initially planned to carry out a study recording event-related potentials in 6-month-old infants (Study 1), and a study investigating looking time preference in 12-month-old infants (Study 2). In Study 1, we tried to replicate Hyde and Spelke's study recording ERPs. Between May 2015 and September 2017, a total of 66 infants were assessed using their paradigm. Due to difficulties faced during data collection (infants fussiness, noisy data, insufficient number of trials visited by infants, incomplete sessions), we decided to drop this study. This was replaced by a study investigating looking time preference with infants 6-month-old infants. Thus, both Study 1 and Study 2 investigated numerical sensitivity in infants during the first year of life employing looking time measures (please see section 3.1.9.2 for more details). Data collection of Study 1 was carried out between March 2017 and September 2017 and Study 2 between July 2016 and March 2018. Due to the late implementation of Study 1, data from Study 1 and 2 were collected concurrently. Consequently, we could not adapt the paradigm on Study 2 based on preliminary data from Study 1.

3.1.7 Infants' Medical History

As previously discussed in section 2.1.1, neonatal complications following premature birth are extensive. Several perinatal factors have been associated with adverse cognitive outcomes, such as infection, respiratory and neurological complications. In addition, infants born prematurely have an increased likelihood of visual deficits, such as retinopathy of prematurity (ROP). Thus, medical factors that are more likely to affect infants' cognitive performance were obtained from neonatal medical records.

3.1.8 Demographic data

In addition to perinatal complications, being born male and maternal education are important factors in predicting the developmental outcomes of preterm infants (Yaari et al., 2018). Parents were therefore asked to complete a questionnaire designed to obtain background information about the participant's family. This included address, parents' educational background, socioeconomic status and medical history. Parental education was categorised by those with qualifications of a bachelor's degree (BD) or higher, and those with qualifications below this level. Socioeconomic status (SES) was defined by the Index of Multiple Deprivation quintile (IMD), a nationally-available score related to postcode which is categorised into five groups with a score of 1 being assigned to the least deprived neighbourhood and 5 to the most deprived (NPEU, 2013) (Appendix 3.4).

3.1.9 Infant Outcomes

3.1.9.1 Developmental measures

To rule out any group differences due to more general neurodevelopmental impairment, we[9] looked at scores obtained from the Ages and Stages Questionnaire (ASQ) for Study 1 and Bayley-III scores for Study 2.

3.1.9.1.1 The Ages & Stages questionnaire

Caregivers of the participants of Study 1 completed the ASQ (ASQ-3; Squires, Bricker & Potter, 2009), specifically the section appropriate for six-month-old infants. The questionnaire is a screening tool that gives scores for five developmental domains (communication, gross motor, fine motor, problem-solving and personal/social development). Each domain consists of six questions answered by caregivers as 'yes', 'sometimes', or 'no' and coded as 10, 5 or 0, respectively. An example question for the sub-scale 'problem solving' is: 'Does your baby pass a toy back and forth from one hand to the other?' The questionnaire provides cut-offs for each domain, grouping infants into three categories: 'below the cut-off', 'close to the cut-off', and 'above the cut-off'. Scores below the cut-off for the problem-solving domain were used as an exclusion criterion for this study. The ASQ is the most commonly used parent-completed questionnaire for screening development (Hornman, Kerstjens, de Winter, Bos, & Reijneveld, 2013).

[9] Instead of using the first person in the singular, I rather used the plural, as a reflection of the collective effort of all the researchers who were part of this work.

Sensitivity and specificity are good (75% and 81%) and there is a modest agreement with the Bayley-III (r = 0.56) (Schonhaut, Armijo, Schönstedt, Alvarez, & Cordero, 2013). Recently, it has been shown that the ASQ is comparable with the Bayley-III when identifying preterm children at risk, supporting the view that the ASQ can be used as a screening tool for developmental delay (Schonhaut, Pérez, Armijo, & Maturana, 2020). We chose to employ a screening questionnaire due to the time constraints for the assessment of this age group.

3.1.9.1.2 Bayley-III

The Bayley is the most widely used tool for the assessment of early development. Currently in its third edition (Bayley, 2006), the primary objective is to identify children with developmental delay through an individually administered assessment of children aged in the range 1–42 months. The Bayley-III has three main subtests: the Cognitive Scale, which includes items such as attention to familiar and unfamiliar objects, looking for a fallen object and pretend play; the Language Scale, which explores understanding and expression of language, for example, recognition of objects and people, following directions, and naming objects and pictures; and the Motor Scale, which assesses gross and fine motor skills, such as grasping, sitting, stacking blocks and climbing stairs. Raw scores of successfully completed items are converted to scale scores and to composite scores. 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 two 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.

Infants born very preterm in the UK are followed up at four-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 sought from the parents of these infants to access medical records and the relevant Bayley-III scores were obtained. Scores from VP infants at twelve months of corrected age were used for Study 2.

Term infants were assessed on the Bayley-III during the PDP assessments at twelve months. The PDP researchers were taught to administer the Bayley-III by Ms B Hutchon (a paediatric occupational therapist and national trainer for Bayley assessments), who is responsible for the follow-up clinics within the North Central London Network, including UCLH. Consistency between assessments of the term and VP cohorts was achieved by following the same administration practices as those adopted in clinic.

A few concerns with the current version of the Bayley-III have been reported, including low sensitivity of the tool to detect mild cognitive impairments. For example, Johnson, Moore and Marlow (2014) investigated the agreement between classifications of delay made using the previous version of the Bayley (BSID-II) and Bayley-III. Their results showed that the Bayley-III produces higher scores than its predecessor. The authors recommend that Bayley-III cognitive and language scores <85 provide the best definition of moderate and severe developmental delay. Anderson and Burnett (2017) proposed that the Bayley-III overestimates development, resulting in a lower level of identification of children with developmental delay, suggesting that the Bayley-III is a poor predictor of later cognitive and motor impairments. Given this, within this thesis, the cut-off score of <85 was used to identify those at possible risk of delay.

3.1.9.2 Looking time measures

In order to investigate numerical sensitivity in infants, we employed looking time measures. Looking time measures are a primary tool for assessing mental processes in infancy (Oakes, 2010). Fantz (1964) was the first to notice that infants prefer to look at novel rather than familiar stimuli. The preference for a novel stimulus is commonly interpreted to indicate the infant's recognition of the familiar stimulus. Looking time measures have been at the forefront of behavioural research in many domains including, memory, social and numerical cognition (Reynolds, 2015).

Habituation and familiarisation paradigms are the most common looking time measures to assess infants' cognition. In habituation studies, infants are presented with a stimulus, or set of stimuli, and their looking time is recorded. Typically, infants' looking time decreases, or habituates, with repeated exposure to the stimulus, and it increases to novel items. Habituation is calculated by averaging infants' looking during blocks of trials and comparing those averages as the session progresses (Oakes, 2010). In familiarisation studies, all infants are presented with a fixed number of familiarisation trials regardless of changes in attention (e.g., Quinn, Yahr, Kuhn, Slater, & Pascalils, 2002; Kovack-Lesh, Horst, & Oakes, 2008). Preferential looking time is relatively easy to use with infants ranging from new-born to toddlers. Although habituation and familiarisation paradigms do not necessarily require a computer, with advances of new technologies such as eye-tracking technique, it has become easier to test infants' responses to a variety of stimuli.

The literature exploring infants' numerical discrimination is extensive. Using the habituation paradigm, several studies have been conducted testing discrimination of numerosity in prelinguistic infants at different ages and using different ratios. For example, at six-months-old, infants can successfully discriminate between sets of 8 dots versus 16 dots (a ratio of 1:2), but fail to discriminate 8 versus 12 dots (a ratio of 2:3) (Xu & Spelke, 2000). Using the same ratios, six-month-old infants successfully discriminate between arrays of 16 versus 32 discs, but not 16 versus 24 (Xu, Spelke & Goddard, 2005). At ten months, infants are able to discriminate a 2:3 ratio (e.g., eight from twelve elements), but not a 4:5 ratio (e.g., eight from ten elements) (Xu & Arriaga, 2007). Studies investigating atypical populations employ a familiarisation paradigm rather than habituation to ensure experiments are brief enough for atypical infants to complete (e.g., Van Herwegen, Ansari, Xu, & Karmiloff-Smith, 2008).

We employed a familiarisation paradigm in our studies, investigating the total amount of time spent looking at novel and familiar stimuli. First look preference, which is the automatic first saccadic[10] reaction to either of two numerical displays, was also investigated. Looking time was recorded by an eye-tracker. For numerical discrimination, we used the ratio-limit according to age; a 1:2 ratio for six-month-old infants (Study 1), and a 2:3 ratio for 12-month-old infants (Study 2).

---

[10] A saccade is a rapid, conjugate, eye movement that shifts the centre of gaze from one part of the visual field to another. Saccades are mainly used for orienting gaze towards an object of interest.

3.1.9.2.1 Eye-tracking principles

In order to record looking preference, we employed an eye-tracking technique. Eye-tracking was first used in the 19th century by Lamare and Hering, who observed and described the movements of the eyes during reading Since the invention of video recordings, it has become possible to re-examine eye movements retrospectively. In the last decade, eye-trackers with in-built infra-red diodes have become increasingly accurate and easy to use (Wade, 2010).

The measurement of eye movements is a powerful tool for investigating perceptual and cognitive functions in both infants and adults. In pre-verbal infants, looking time measures are a major gateway to the developing brain (Gredebäck, Johnson, & Hofsten, 2010). It is particularly advantageous in developmental research since it allows implicit non-verbal data collection. Although looking patterns do not directly reveal information about brain functioning and real time neural computations, they enable conclusions to be drawn about what a child is processing. Eye-tracking techniques have made the process of recording, assessing and analysing looking time measures much more precise. Previously, the majority of infant studies employed preferential looking, familiarisation or habituation paradigms, where looking times were manually coded and inferences made as to whether infants were able to make discriminations

Aslin (2007) therefore stated that "It is no exaggeration to say that without looking time measures, we would know very little about nearly any aspect of infant development". Moreover, they enable researchers to investigate several aspects of looking behaviour besides looking durations, such as fixation durations, patterns of fixations, number of saccades, saccade latencies, and directions of saccades (Aslin, 2007).

3.1.9.3 A-not-B paradigm

In addition to investigating numerical sensitivity, in Study 2 we also wanted to examine the association between numerical discrimination and visual working memory. The Piagetian A-not-B paradigm (Piaget, 1954; Diamond, 1985) is a long-standing assessment of executive functions, particularly working memory in children within their first year. In the standard version of this task, infants watch as a desirable object is hidden in one of two possible locations, a brief delay is imposed, and then infants are allowed to reach. The A-not-B error, also known as a perseverative error, is an error in the mental perception of objects seen in infants before the age of one year. This illustrates a child's ability to mentally represent objects also known as object permanence. The standard A-not-B paradigm challenges working memory networks following the short delays (Schwartz and Reznick, 1999; Espy et al., 2002; Reynolds and Romano, 2016). The classic 'AB error' occurs when the participant continues to reach for the object in the original hiding location after observing the reversal of the hiding location (Diamond, 1985, 2001).

We employed the standard version of A-not-B task to assess working memory abilities. While the participant watches, the object of an infant's interest is placed in one of the wells and the participant's view of the object is then occluded. Following a brief delay, the participant is allowed to retrieve the object from one of the wells. After two successful retrieval trials, the location of the hidden object is reversed (again while the participant observes). If the infant retrieves the object correctly, the procedure is repeated with a delay of five seconds before being allowed to search for the object. A similar procedure is carried out with a ten-second delay. 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. Parents were asked to contain the baby's arms while the experimenter was counting loudly during the delay. The results were coded as succeeded or failed in 0s, 5s or 10s.

It should be noted that a range of cognitive processes have been proposed to account for the AB error. Working memory and inhibitory control are considered the major contributors to A-not-B object permanence performance (Diamond, Prevor, Callender, & Druin, 1997; Diamond, Cruttenden, & Neiderman, 1994). Thus, not only do infants need to keep the current location of a hidden object in working memory throughout different manipulations of a hiding site, but they must also inhibit reaching back toward a previously rewarded hiding site when they see the object being hidden in a different spatial location.

3.2 Children studies

The aim of Study 3 (chapter 6) was to investigate domain-general and domain-specific skills and the relationship with mathematical skills in very preterm children. Study 4 (chapter 7) examined neural correlates of numerical magnitude comparisons employing electrophysiological measures in very preterm children. VP and term control children aged between eight and ten years old were recruited to take part in studies 3 and 4. This age range was selected because difficulties in mathematics start to emerge after the commencement of formal education. The following sections describe recruitment, inclusion and exclusion criterion and measures employed to address the aims for the children's studies (section 2.11.1).

3.2.1 Recruitment of VP children

Children born prematurely were recruited for studies 3 and 4 at UCLH. Very preterm children (<32 weeks gestational age) born between 2006 and 2009 at UCLH under the care of the Consultant

Neonatologist (Professor Neil Marlow) and discharged from the hospital with a 'normal outcome'[11] were identified and invited to participate in the study. Initial contact with the family was made via a letter from the Consultant Neonatologist (Professor Neil Marlow). Parents were sent a letter introducing them to the study (appendix 3.5). Interested parents completed a consent form (appendix 3.6) agreeing to be contacted by a researcher from the study team (Mérari Ferreira) and returned it to the investigators via freepost. If no response was received within two weeks, a further reminder was sent after which the family was determined to be a non-responder and one attempt was made by the researcher to contact the family by telephone. Families received study information sheet for the parent (appendix 3.7) and the child. Once the family of a preterm child had agreed to participate, the researcher contacted them to confirm the inclusion and exclusion criteria. Thirty-eight VP children were recruited for studies 3 and 4. Figure 3.2 provides a schematic representation of recruitment for the preterm group.

[11] Normal outcome was classified as infants discharged from hospital with no major cerebral damage (i.e., periventricular leukomalacia (PVL), intraventricular hemorrhage (IVH > grade II), hydrocephalus, retinopathy of prematurity (ROP > stage II) and congenital malformations based on the results of ultrasound scan.

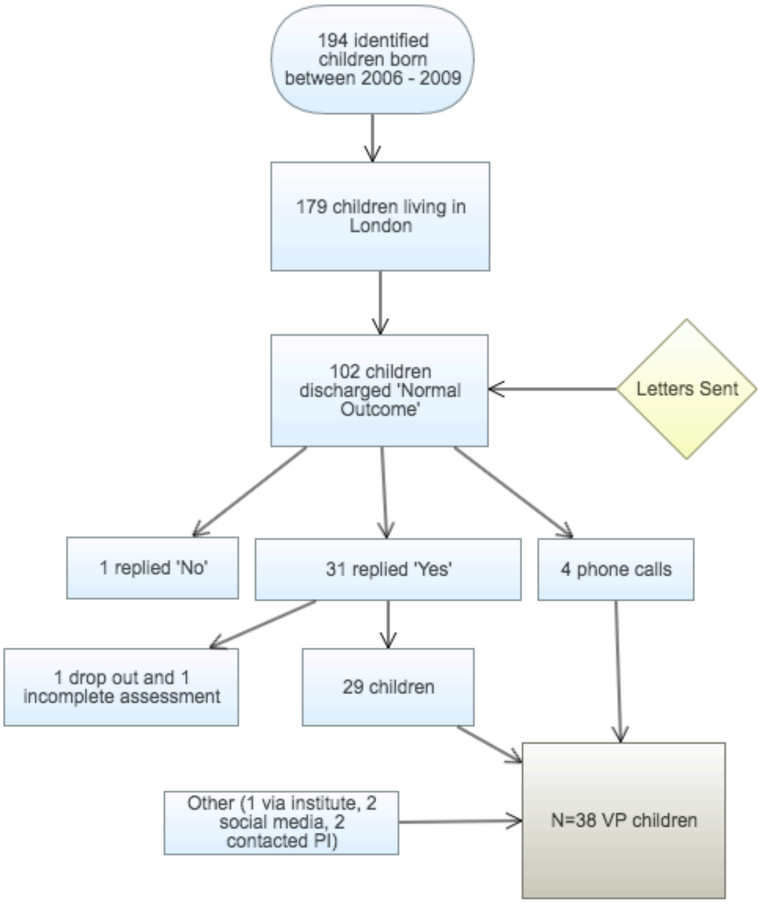


Figure 3.2: Schematic representation of recruitment to the preterm group for studies 3 and 4.

3.2.3 Recruitment of term-born children

Initially, we planned to recruit full-term children matching very preterm children by sex, school, class and ethnicity. This would have been achieved by asking head teachers to identify three children of the same class, gender and ethnicity as the VP child from their school. We would have then randomly identified which of the three the head teacher should first pass a study invitation letter to. If no response was received within two weeks we would have asked the head teacher to pass the information to one of the two remaining children. This method has been previously employed by our research group, which has extensive experience in recruiting preterm and term participants. In our previous studies, however, the assessments were carried out in schools, which considerably increased responses from participants. In our current study, assessments were carried out in our lab in order to collect electrophysiological data and thus this method of recruitment was unsuitable to recruit term children. Thus, recruitment of the term-born group was made via emails and posters around the university, and local schools were contacted in order to recruit control children. Once families had agreed to be contacted, the researcher contacted those families by telephone to confirm inclusion and exclusion criteria and a visit to the lab was arranged.

3.2.4 Power calculations of children's studies

A power calculation was conducted assuming a Cohen's $d = 0.3$ with means and standard deviation based on a previous study investigation differences in numerical representation between term and preterm groups (Libertus et al, 2017). Specifically, we based our sample size calculations on combined scores of accuracy and reaction times in a non-symbolic magnitude comparison tasks assessing school-aged children, with M=1.16, SD=1.83 for the term group and M=0.004, SD=0.74 for the preterm group. As is standard, α=0.05 was used, with desired power 80%. The calculation for group comparisons suggested a sample size of 19 children for each group. Given that not all our sample would have difficulties in maths, we decided to recruit as many participants as possible within a certain time period (June 2016 to June 2018).

3.2.5 Research approval for children's studies

The study was approved by Hampstead Research Ethics Committee 15/LO/1687 (appendix 3.8). It was registered with the Research and Development Department of UCLH. Written informed consent for the participants in this study was obtained from the parent or legal guardian of the children (appendix 3.9). Written assent was obtained directly from the participants.

3.2.6 Inclusion/ exclusion criteria for term-born children

The exclusion criteria for the control group were: 1) children born below 37 weeks gestational age, 2) children attending special schools, 3) presence of neurosensory disability, 4) any history of neurological, psychological or psychiatric disorder (such as brain injury, epilepsy, ADHD or learning disability), 5) using medication that could interfere with cognition.

3.2.7 Inclusion/ exclusion criteria for VP children

The exclusion criteria for preterm children were: 1) children not attending mainstream school, 2) severe disabilities causing a child to be unable to perform behavioural tests. Children unable to attend mainstream schools were not included as these schools do not follow a standard curriculum and the primary outcome of this study is curriculum-based attainment in mathematics. Given the UK policy for integration, it was expected that some preterm children in mainstream school would have significant functional disability. These children were included where study assessments could be completed, as the goal was to assess the full spectrum of attainment for preterm survivors in mainstream education. In addition, neurosensory disability was an exclusion criterion as it could preclude the child from participating in study assessments. Table 3.1 shows all of the inclusion and exclusion criterion for studies 3 and 4.

Table 3.1: Inclusion and exclusion criterion for studies 3 and 4

	Inclusion	Exclusion
Full-term participants	Born at >37 weeks of gestational age;	Born at <37 weeks of gestational age;
	Be fluent in English	Attending special schools;
	Age between 8-10 years old;	Presenting neurosensory disability;
		History of neurological disorder (such as brain injury, epilepsy);
		History of psychiatric disorder (such as ADHD). Taking medication that might interfere with cognition.
Preterm participants	Born at <32 weeks of gestational age;	Born at >37 weeks of gestational age;
	Be fluent in English;	Attending special schools;
	Age between 8-10 years old;	Presenting neurosensory disability.

3.2.8 Children's Appointment

All children were assessed individually by a research psychologist (Mérari Ferreira) at the Wolfson Centre, Institute of Child Health, UCL. Where possible, all assessments were carried out on the same day. Where this was not possible, assessments were completed within two months of the initial visit, either by a participant returning to the lab facilities, or a home-visit. Children were given regular breaks during assessments. All children were given a certificate for taking part in the study and a £10 Amazon voucher. Participants' families were refunded for any expenses related to the study, including travel and food.

3.2.9 Children's Outcomes

3.2.9.1 Children's Medical History

Similar to section 3.1.5, medical factors that could affect the children's performance were obtained from their medical records and a similar procedure was carried out.

3.2.9.2 Demographic data

Similar to section 3.1.6, parents were asked to complete a questionnaire to obtain information about the participant's family background (appendix 3.10). This included address, parents' educational background, socioeconomic status, medical history, ethnicity and linguistic background, as recommended by previous studies (Sansavini et al., 2011).

3.2.9.3 Children's assessments

For Study 3 (chapter 6), the main aim was to examine domain-general and domain-specific (numerical magnitude comparison) skills in association with mathematical abilities in school-age children born prematurely. The main aim for Study 4 (chapter 7) was to investigate neural correlates of numerical magnitude comparisons in the same population. The following sections will describe the procedures used to address these aims.

3.2.9.3.1 Domain-general skills

Based on previous literature (i.e. Simms et al, 2013b; 2015), a protocol was developed to focus on aspects of domain-general abilities that were indicated to be affected in VP children, potentially impacting their mathematical abilities. Thus, intelligence and executive functions including processing speed, working memory, attention, planning and inhibition composed the domain-general abilities.

3.2.9.3.1.1 Intelligence

Intelligence was assessed using the Wechsler Intelligence Scale for children (WISC-IV). This test comprises fifteen subtests, ten of which form the core battery. The ten core subtests provide four subtest indexes, namely Verbal Comprehension (VCI), Perceptual Reasoning (PRI), Working Memory (WMI) and Processing Speed (PSI). Taken together, the ten core subtests comprise the Full Scale IQ (FSIQ). The reliability coefficients for the WISC–IV composite scales range from 0.88 (Processing Speed) to 0.97 (Full Scale). The reliability coefficients of the WISC–IV composite scales are identical to or slightly better than corresponding composite scales in the WISC–III (Wechsler, 2005). Table 3.2 summarises the core subtests of the WISC-IV used in this study.

Table 3.2: Core subtests of WISC-IV

Index	Test	Description
Verbal Comprehension (VCI)	Similarities	Two words are presented, and the participant must decide how they are alike, e.g. "How are milk and water alike?"
	Vocabulary	The participant needs to define words, e.g. "What does clock mean?"
	Comprehension	The participant is asked questions about social and other situations, such as: "Why should children not be allowed to work in factories?"
Perceptual Reasoning (PRI)	Block Design	The participant is required to copy a pattern using coloured blocks in a certain amount of time.
	Picture Concepts	The participant is shown either two or three rows of pictures and has to choose one picture from each row that share a common characteristic.
	Matrix Reasoning	The participant is presented with a matrix of abstract pictures in which there is one picture missing, and needs to choose which of a number of possible options the missing picture is.
Working Memory (WMI)	Digit Span	This test has two subtests. Firstly, the participant listens to a series of numbers and is required to say them back to the examiner. Secondly, the same procedure is followed, but the participant is required to say the sequence of numbers backwards.
	Letter-Number Sequencing	The child listens to a series of letters and numbers and is required to repeat them back with the letters in alphabetical order and the numbers in numerical order.
Processing Speed (PSI)	Coding	The participant is presented with a key in which the numbers 1 to 9 are each paired with a different symbol. The participant's task is to use this key to put in the appropriate symbols for a list of numbers between 1 and 9.
	Symbol Search	The participant has to look at two target symbols and then examine a group of symbols to see if the target symbols are repeated.

3.2.9.3.1.2 Executive Functions

3.2.9.3.1.2.1 Attention

The Test of Everyday Attention for Children (TEA-Ch) assesses various dimensions of attention. It comprises nine subtests evaluating selective attention, sustained attention, or attentional control. Four core subtests were administered: Sky Search, Score!, Creature Counting and Sky Search DT, assessing selective attention, sustained attention, attention control and dual task proficiency, respectively. The reliability of the subtests ranges from 0.65 to 0.85 (Manly et al., 2001). Although the core subtests consist of four items, nine total scores are acquired. Table 3.3 describes the subtests used in this study, the corresponding dimension of attention measures and the scores generated by the subtest. Outcome scores highlighted in bold illustrate the outcomes used for the analysis.

Table 3.3: Subtests of Tea-Ch

Dimension of attention	Test	Description	Outcome scores
Selective attention	Sky Search	Participants must visually scan a large array of stimuli and identify targets as quickly as possible.	1. Number of correctly identified targets 2. Time per target 3. **Attention score**
Sustained attention	SCORE!	Participants listen to a string of audio beeps that vary in number and length of inter-stimulus intervals. Between 1-15 beeps are presented, and the participant must keep track of the number of beeps presented and verbally report the total to the examiner at the end of each trial.	**1. SCORE sustained attention**
Attentional control: set switching and inhibition	Creature counting	Participants must count the number of visually presented targets, according to the direction of counting implied by an explicit cue (arrow).	**1. Creature counting total correct** 2. Timing score
Dual Task Proficiency	Sky Search DT	Participants must visually scan for targets (as in Sky Search) whilst counting a string of beeps over a number of trials (as in SCORE). The outcome variable reflects the time and accuracy cost of performing two tasks at the same time.	**1. Dual task score**

3.2.9.3.1.2.2 Inhibition and Planning

The Delis-Kaplan Executive Function System (D-KEFS) (Delis, Kaplan, & Kramer, 2001) evaluates various aspects of executive functions, including flexibility of thinking, inhibition, problem-solving, planning, impulse control, concept formation, abstract thinking, and creativity in both verbal and spatial modalities. The D-KEFS is composed of nine subtests: Trail Making Test, Verbal Fluency Test, Design Fluency Test, Colour-Word Interference Test, Sorting Test, Twenty Questions Test, Word Context Test, Tower Test, and Proverb Test. The Colour-Word Interference Test assesses inhibition, whereas the Tower Test is used to assess spatial planning skills, the ability to establish a rule set, and the ability to follow a set of instructions. Table 3.4 describes the subtests for this study. Items highlighted in bold illustrate the outcomes used for the analysis in the chosen test. The outcome score highlighted in bold illustrates the outcome used for the analysis.

Table 3.4: Subtests of D-KEFS

Test	Item	Description
Colour-Word Interference Test	Colour naming	The participant is presented with a page containing a series of red, green and blue squares and is asked to say the names of the colours as quickly as possible without making mistakes.
	Word reading	The participant is presented with a page containing the words “red”, “green”, and “blue” printed in black ink and is asked to read the words aloud as quickly as possible without making mistakes.
	Inhibition	The participant is presented with a page containing the words “red”, “green”, and “blue” printed incongruently in red, green or blue ink and is asked to say the colour of the ink in which each word is printed as quickly as possible without making mistakes.
	Inhibition/switching	The participant is presented with a page containing the words “red,” “green,” and “blue” written in red, green, or blue ink. Half of these words are enclosed within boxes. The participant is asked to say the colour of the ink in which each word is printed (as in the third trial), and to read the word aloud (and not name the ink colour) when a word appears inside a box, as quickly as possible without making mistakes.
Tower Test	**Planning**	Using a board with three vertical pegs of equal length and five coloured disks varying in size from small to large, the participant must move from a predetermined starting position, in the fewest moves possible, to a specified ending position displayed by the examiner. The participant must follow two rules while completing the test: move only one disk at a time, and never place a larger disk over a smaller disk. The participant is timed, and the outcome is the total number of moves.

3.2.9.3.1.2.3 Working Memory

The Alloway Working Memory Assessment (AWMA) (Alloway, 2007) is a computerised assessment of working memory. This test is composed of twelve subtests assessing verbal and visuospatial short-term memory and working memory. Verbal short-term memory can be assessed by Digit Recall, Word Recall and Non-word Recall. Verbal working memory can be tested by Listening Recall, Counting Recall and Backwards Digit Recall. Visuospatial short-term memory can be tested using Dot Matrix, Mazes Memory, Block Recall. Finally, visuospatial working memory can be assessed by Odd One Out, Mister X and Spatial Recall. There are three versions of AWMA: AWMA Screener, AWMA Short Form and AWMA Long Form. AWMA Screener comprises two working memory tests and is suitable for screening working memory difficulties. AWMA Short Form comprises four tests and is recommended to screen individuals suspected of having memory difficulties, but where the specific area of their difficulty is not known. AWMA Long Form comprises all twelve subtests and is recommended for confirmation of significant working memory problems for individuals identified as having working memory problems in the classroom.

The AWMA screener version was used to assess working memory, composed by Listening Recall and Spatial Recall. In the Listening Recall subtest, the participant listened to a sequence of sentences (e.g., "Dogs have four legs") and at the end of each sentence had to judge whether the sentence was true or false. Subsequently, the participant must recall the final word of the sentence(s) in exactly the same order as it was presented (e.g., "legs"). The subtest began with a block of one sentence and was then increased to a block of six sentences. In the Spatial Recall subtest, the participant was presented with a pair of identical shapes, in which the right shape has a red dot above it and is rotated clockwise or counter clockwise.

During the presentation, the participant must identify whether the shape with the red dot is the same or the opposite of the other shape. The shape with the red dot may take three different orientations; thus, the red dot can be found in three different locations that need be remembered. A block of trials starts with a presentation of a pair of shapes and increases to the sequential presentation of seven pairs of shapes. At the end of each block, the participant must recall the positions of the red dots by pointing at the picture with three possible locations marked exactly in the same order as the sets were presented. It should be noted that the AWMA was discontinued in December of 2016 and it was not possible to renew the license of the software, thus not all participants completed the AWMA screener.

3.2.9.3.2 Domain-specific skills

Domain-specific skills are thought to be responsible for the processing of numerical concepts. This includes a basic mechanism that engages mental representations of numerical quantities, known as *numerical magnitude comparisons*. Numerical magnitude comparisons refer to our basic ability to decide which of two numerosities is the largest. Usually, numerical magnitudes are measured with dot (non-symbolic) and digit (symbolic) magnitude comparison tasks (Schleepen et al., 2016). Symbolic (digit) magnitude comparison skills are believed to index an exact symbolic representation system (Ansari & Karmiloff-Smith, 2002). Non-symbolic (dot) magnitude comparison skills are thought to reflect the acuity of the approximate number system (ANS). (Dehaene, 1997).

The most widely-used outcomes for numerical magnitude comparison are numerical distance and ratio effects (Moyer & Landauer, 1967). These refer to the phenomenon that performance is better (i.e., error rates are lower and reaction times are shorter) when the numerical distance is relatively large, e.g., comparing 1 *vs*. 9 is easier than comparing 5 *vs*. 6 (Moyer & Landauer, 1967). In such tasks, speed and accuracy increase with age and experience (Sekuler and Mierkiewicz, 1977). Many studies investigating numerical representations using magnitude comparison tasks have simply reported participants' accuracies: the proportion of trials they answered correctly (e.g., Gilmore, Attridge, & Inglis, 2011; Lourenco, Bonny, Fernandez, & Rao, 2012; Nys et al., 2013; Wei, Yuan, Chen, & Zhou, 2012; Fuhs & McNeil, 2013; Kolkman, Kroesbergen, & Leseman, 2013) or, less commonly, the number of trials they answered correctly in a given time (e.g., Nosworthy et al., 2013). Measures combining response time and accuracy data have also been proposed in the literature. For example, the inverse efficiency score was introduced as a measure of ANS acuity that controls for a potential speed-accuracy trade-off (Sasanguie, De Smedt, Defever, & Reynvoet, 2012a; Sasanguie, Van den Bussche, & Reynvoet, 2012b). This score can be calculated by dividing the mean RT of correct responses by the proportion of correct responses (Bruyer & Brysbaert, 2011). The efficiency score is given in milliseconds and can be interpreted like mean RT (i.e., the smaller the efficiency score, the higher ANS acuity).

Other measures have been used as alternative approaches to indexing individuals' acuity in magnitude comparison tasks, such as Weber fraction (Halberda et al., 2008, 2012; Piazza et al., 2010; Inglis, Attridge, Batchelor, & Gilmore, 2011; Libertus et al., 2011; Lyons & Beilock, 2011; Mazzocco et al., 2011b; Castronovo & Göbel, 2012; Price et al., 2012; Bonny & Lourenco, 2013; Sasanguie et al., 2013). Weber's law states that, as the ratio between the magnitudes of two stimuli increases, the easier it will be to differentiate between the two stimuli.

Models of the ANS utilise Gaussian curves that represent different numerosities, or mental representations of quantity. The spread of these curves is determined by the Weber fraction, hence allowing analysis and comparison of participant performance. Under this interpretation, when an individual observes an array of *n* dots, they form an internal representation. An individual's Weber fraction can be estimated by calculating the value of *w* which best fits their behavioural data (Inglis & Gilmore, 2014).

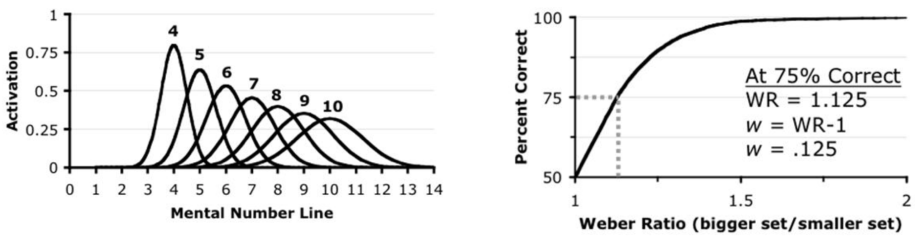


Figure 3.3: Internal mental number line measure by Weber Fraction (w).

Figure 3.3 represents the internal mental number line measured by the Weber fraction, which is calculated from data obtained from tasks assessing the ANS using individual accuracy's performance, as illustrated in the figure on the right-hand side. From panamath.org.

It has been suggested that accuracy-based measures are more informative about the underlying mechanisms of numerical representations (Dietrich, Huber, & Nuerk, 2015). In line with this view, Inglis and Gilmore (2014) revealed that accuracy and Weber fractions are strongly related, R^2 = .86, suggesting that they are measuring the same underlying construct. The authors recommend that given the superior psychometric properties of simple accuracy figures, the best way of indexing the acuity of an individual's numerical representation is simply to report their accuracy. There is no consensus in the literature about which measures should be employed to index numerical magnitude comparison, however. Given this, four main outcomes were considered in this study: **accuracy, reaction time, Weber fraction (w) and inverse efficiency scores (IES)**.

3.2.9.4 Educational performance

The Wechsler Individual Achievement Test, 2nd edition (WIAT-II) (Wechsler, 2005) was used to evaluate academic attainment. Mathematics scales (WIAT-MS) comprise numerical operations and mathematical reasoning. The test for numerical operations is a non-timed paper-and-pencil test that assesses performance in simple operations such as addition and subtraction. The mathematical reasoning test involves solving a series of real-life mathematical problems which are orally presented, for example. telling the time, and using money. Single word spelling and dictation was also assessed.

3.2.9.5 Electrophysiological measures

The aim of Study 4 was to investigate neural correlates for the comparison of numerical magnitudes in VP children employing ERPs. ERPs are measured by employing of electroencephalography (EEG). EEG is a commonly-used neuroimaging tool that offers high temporal resolution and allows direct insight into neurophysiological processes (Michel & Murray, 2012). EEG uses electrodes placed on the head to measure the electrical activity that is mostly generated by cortical neurons (Luck, 2005). The signal is amplified so that it can be digitally recorded in real time. Event-related brain potentials are small parts of the continuous EEG recording. ERPs cannot typically be seen within the raw EEG recording, because of their very small amplitude (Teplan, 2002). They, therefore, need to be singled out from the continuous recording by creating an average of recording periods, known as epochs, which are time-locked to repeated presentations of the same stimulus. This allows the spontaneous EEG fluctuations, unrelated to stimulus presentation, to be averaged out, resulting in the ERP wave, which reflects only the activity persistently related to the time-locked presentation of stimuli (Beres, 2017). See figure 3.4.

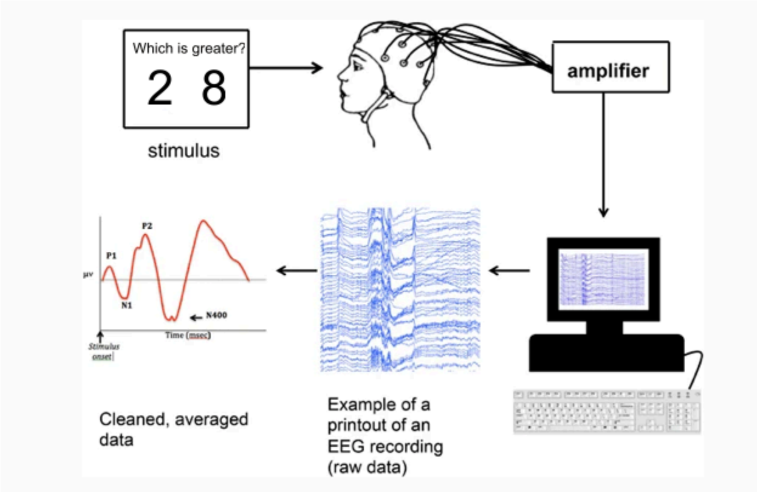


Figure 3.4: The process of EEG recording

A number of electrodes are placed on a participant's head, which enables the electrical brain activity to be measured on the surface of the scalp. EEG is recorded and amplified while an event is presented. The signal is then averaged. Adapted from Beres (2017).

Different components can be distinguished in the ERP waveform that are thought to be related to different processing stages depending on the eliciting stimulus and experimental paradigm. ERPs give a millisecond-by-millisecond account of cortical processing, but are limited in spatial resolution. ERP components are commonly characterised through measures such as mean or peak amplitude, onset latency and peak latency. Latency is the time from the onset of the stimulus to the point when the amplitude reaches its maximum peak. Note that in this study, negativity is plotted upwards. Amplitudes reflect the amount of neural activity that is used to process stimuli, and latencies reflect the velocity of the processing. The peak is usually measured where the amplitude reaches a maximum (Beres, 2017). See figure 3.5.

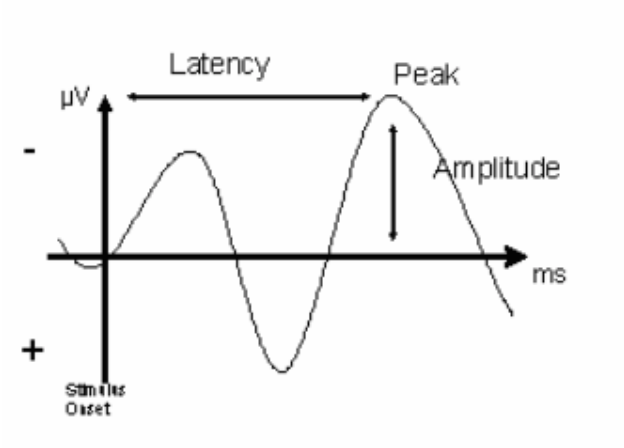


Figure 3.5: Nomenclature of the ERP components

A peak is a maximum deflection either on the negative or positive scale, the amplitude is the maximum expression of the wave (in µV), and latency (in ms) indicates the time from the stimulus onset to its maximum deflection.

Usually, early components (P100, N100, P200) are linked with basic, low-level perception and are thought to be automatic in nature. This means, that as long as a perceptual stimulus is presented, it should be elicited. Other components, which come later (usually after 250 ms) represent conscious cognitive processing and can be elicited in certain experimental conditions. This distinction of early (automatic) and late (more conscious) components illustrates the general idea of ERP components (Beres, 2017).

3.3 General statistical approach

All analyses were performed with SPSS Statistics 22.0 for Mac, with the significance level set at 0.05 and p-values were adjusted for multiple comparisons via Bonferroni correction, when applied. Data were examined for normality using histograms and the Shapiro Wilks 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 were reported; for non-normally distributed data (non-parametric) medians and ranges were reported.

Independent t-tests were used to compare the performance of the term and preterm groups on all standardised measures, and Cohen's *d* was calculated to determine standardised effect sizes across tests. Effect sizes were defined as small (0.2–0.3), medium (0.3–0.5), or large (>0.5). Analysis of covariance (ANCOVA) was carried out, controlling for the effects of other continuous variables that were not of primary interest. For example, maternal education, gender, SES and IQ are independent variables that are well-known in the literature to affect cognitive performance.

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. Sphericity assumptions were tested using Mauchly's test. In the case of violated sphericity assumptions, the Greenhouse-Geisser correction was applied (Greenhouse and Geisser, 1959). Post-hoc analyses were based on Student's t-tests.

Correlation analyses were based on bivariate correlations.

4 Number familiarisation in six-month-old very preterm infants

4.1 Abstract

Individuals born prematurely have a high prevalence of difficulties in mathematical skills, with a marked impact on their lives. A potential predictor of mathematical attainment is the ability to discriminate two sets of bigger elements (equal or more than three elements) mediated by the approximate number system (ANS). Infants demonstrate the ability to discriminate between large versus smaller quantities, and therefore evaluating the ANS might be a potential tool for early identification of atypical neurocognitive mechanisms in this at-risk population. Here, we investigated the ANS in thirteen very preterm and fourteen full-term infants aged six months using a number familiarisation task. Overall, infants born prematurely did not differ from controls in numerosity discrimination. Looking time to the novel number of elements was equal to chance for both groups. Odd ratios revealed a similar number of infants looking longer to the novel number in both groups. We tentatively conclude that number processing may not be generally impaired in preterm infants at this stage of development, but that further research must be conducted expanding sample size and evaluating other periods of numerical development.

4.2 Introduction

The improvement of medical technologies in recent years has permitted a considerable increase in the survival rates of premature babies (Kyser, Morriss, Bell, Klein, & Dagle, 2012), yet this has at the same time raised other issues. Importantly, while more babies survive, they may still not be developing optimally, increasing the risks of cognitive dysfunction and academic failure (Joseph et al., 2016). With regards to academic achievement, the most substantial deficit that preterm children have is in mathematics, markedly affecting their academic attainment. Deficits in maths are found in children born very preterm (less than 32 weeks of gestational age) (Taylor, Espy & Anderson, 2009; Johnson et al., 2011), even when excluding children with intellectual disabilities and when matching for IQ (Taylor et al., 2009). These deficits persist over time (Aarnoudse-Moens et al., 2011), illustrating that mathematics is a particular area of difficulty in this population. Despite the extensive body of research investigating mathematical achievement in the preterm population, it is unclear when difficulties in numerical skills start to emerge. It has been proposed that a system that represents large numerosities (>3), known as the approximate number system (ANS), might be a predictor of mathematical attainment (Mazzocco, Feigenson & Halberda, 2011; Bonny & Lourenco, 2013; Libertus, Feigenson & Halberda, 2013) and that it is possible to evaluate this system in infants as young as six months old (Xu & Spelke, 2000). Very basic processes underlying numerical cognition can go awry very early in development, thus the evaluation of numerical skills in early stages of development might also be a useful tool to identify children at risk for later maths difficulties. This study aims to investigate whether difficulties in numerical sensitivity, a precursor to later mathematical skills, can already be observed in infancy in individuals born very prematurely.

Prior to language acquisition, human infants are able to discriminate different quantities of small sets of items or events without counting, an ability known as numerosity (Geary, 2000). Even one-day-old babies demonstrate the ability to discriminate between different numerosities (Izard et al., 2009). Two systems are claimed to be involved in representing numerosities in infants: one representing large numerosities (>3), the ANS, and another representing small numerosities (<3), the object file system (OFS).

The ANS is thought to operate like a mental number line, with increasing overlap between number magnitudes as the numerosities increase (Mussolin, Nys, Leybaert, & Content, 2012). As a consequence, infants can discriminate numbers according to two values upon their ratio, but not their absolute difference, known as Weber's Law (Cordes & Brannon, 2009). The acuity of the ANS seems to improve as children develop (Feigenson, Dehaene & Spelke, 2004). Several studies have been conducted to test discrimination of numerosity in prelinguistic infants at different ages and using different ratios. For example, six-month-old infants successfully discriminate twofold changes (8 *vs.* 16, 16 *vs.* 32), but fail to discriminate 1.5 fold changes (8 *vs.* 12, 16 *vs.* 24) (Xu & Spelke, 2000; Lipton & Spelke, 2003; Xu, Spelke & Goddard, 2005). At ten months, infants can discriminate 1.5 fold changes (8 *vs.* 12), but fail to discriminate 1.3 fold changes (8 *vs.* 10 elements) (Xu & Arriaga, 2007).

Testing numerosities in populations at risk of having difficulties in numerical skills can help to identify early atypical trajectories in number processing. For instance, an early impairment in the ANS was found in infants with William syndrome (WS), a genetic syndrome with a high prevalence of numeracy deficits (Van Herwegen et al., 2008). Employing looking behaviour measures, Van Herwegen et al. (2008) assessed small and large numerosities in nine toddlers with WS (mean chronological age was 35 months and mean mental age was 22 months). Their results revealed that while toddlers with WS were able to discriminate small numerosities, large number discrimination was impaired from an early age onwards. Difficulties in basic numerical skills have also been reported in toddlers with fragile X syndrome, a condition caused by a single-gene mutation on the X chromosome. Toddlers with fragile X syndrome have demonstrated early atypical trajectories in numerical cognition, showing impairments in ordinal numerical sequences, an ability otherwise displayed by eleven-month-old infants (Owen, Baumgartner, & Rivera, 2013).

To date, no study has investigated early numerical skills during infancy in the preterm population, a population at risk of mathematical failure. Thus, this study aimed to investigate numerical sensitivity in infants born very preterm (less than 32 weeks of gestational age) when they were six months corrected age, in comparison to full-term infants. Using a number familiarisation task, we tested number discrimination between 8 *vs*. 16 dots (ratio 1:2) (Xu & Spelke, 2000). We hypothesised that very preterm infants would succeed in discriminating the twofold change (1:2 ratio), in line with results previously reported in the literature showing that school-aged children do not have domain-specific deficits in numerical representation, but rather domain-general deficits that affect numerical discrimination (Guarini et al., 2015, Simms et al., 2015). If numerical abilities were already impaired during infancy in individuals born prematurely, however, preterm infants would not be able to discriminate the twofold change (1:2 ratio), showing domain-specific deficits in early development, in line with reports suggesting domain-specific deficits in school-aged children (Hellgren et al., 2013; Libertus et al., 2017). To rule out any group differences being simply due to a more general neurodevelopmental impairment, we looked at scores obtained from the Ages & Stages questionnaire, a screening tool questionnaire with a modest agreement with the Bayley-III (r= 0.56) and good sensitivity and specificity (75% and 81%) (Schonhaut et al., 2013).

4.3 Methods

4.3.1 Participants

A total of 37 infants were recruited for Study 1. Mean corrected age for term-born infants (n= 21) was 6 months and 16 days (range: 6 months and five days to 7 months and 0 days) and preterm infants (n=16) 6 months and 16 days (range: 6 months and 0 days to 7 months and 0 days). Mean chronological age for preterm infants was 8 months and 15 days (range: 8 months and 16 days to 10 months and 10 days). Mean gestational age for the preterm group was 27 weeks and 2 days of mean gestation (24 weeks and 1 day to 31 and 4 days) and for the full-term group was 39 weeks and 3 days of mean gestational age (range: 38 weeks and 2 days to 41 weeks and 6 days). Data from a total of 10 infants (3 preterm and 7 term infants) was excluded due to fussiness or calibration issues (n=5), quality of eye tracking recording[12] inferior than 50% (n=2) and not enough data available during test phase[13] (n=3). The final sample included data from 14 full-term infants (6 boys and 8 girls) and 13 preterm infants (8 boys and 5 girls).

4.3.2 Demographics

Table 4.1 summarises the general characteristics of the study participants. Independent t-test for corrected age at assessment and maternal age, chi-square analysis of maternal education and socio-economic status revealed that preterm and full-term samples

---

[12] The quality of the recording was automatically calculated by *Tobii Studio*, dividing the number of eye tracking samples that were correctly identified, by the numbers of attempts throughout the recording. A record with 100% of quality means that both eyes were found during the recording, whereas 50% of quality means that just one eye was found for the full recording or both eyes for only half the time.

[13] Infants needed to spend at least 2500 ms fixating the stimuli during the test phase and contribute data on both familiar and novel stimuli in order to be included in the analysis. This criterion is consistent with previous studies investigating looking time measures in infants (e.g., Richmond & Nelson, 2009; Richmond, Zhao, & Burns, 2015).

did not differ significantly in the mentioned variables. Table 4.2 shows the neonatal morbidities of the VP children.

Table 4.1: General characteristics of study participants.

	Term n=14	**Preterm** n= 13
Gestational age (weeks), mean (SD)	40.36 (1.25)	28.38 (2.84)
Range	(38^{+1} -41^{+6})	(24^{+1} – 31^{+2})
Birthweight (g), mean (SD)	3435 (0.51)	1119 (0.40)
Range	(2630 – 4173)	(570 – 1560)
Sex (M:F)	6:8	5:8
Chronological age (days), mean (SD)	197.23 (11.12)	284.80 (16.91)
Range	(181 – 216)	(269 – 309)
Corrected age (days), mean (SD)	199.15 (7.77)	205.20 (4.54)
Range	(185 – 210)	(199-210)
Maternal age (years), mean (SD)	36.78 (5.83)	36.53 (5.14)
Range	(21 – 46)	(27 - 46)
Maternal Education		
<BD (%)	14.30	7.70
> BD (%)	85.70	92.30
SES (IMD quintile)		
First (%)	14.30	9.10
Second (%)	0.00	36.40
Third (%)	21.40	18.20
Fourth (%)	28.60	27.30
Fifth (%)	21.40	9.10

BD: Bachelor's degree

Table 4.2: Neonatal morbidities among preterm children.

	N	Very Preterm (n=7)	N	Extremely Preterm (n=6)
Multiple births				
IVH / PVL: I – II				
Grade I	1	13.18%	1	16.60%
Grade II	0	-	1	16.60%
ROP				
Grade I	0	-	3	50.00%
Grade II	1	13.18%	2	33.33%
Grade III	0	-	1	16.60%
CLD/ BPD				
Mild	1	13.18%	2	33.30%
Moderate	0	-	3	50.00%
Severe	1	13.18%	0	-

IVH: intraventricular haemorrhage; PVL: periventricular leukomalacia; ROP: retinopathy of prematurity; CLD: chronic lung disease; BPD: bronchopulmonary dysplasia. Mild: O2 at 28 days not 36 weeks, moderate: O2 36 weeks < 30% low flow; Severe: O2 36w >30% +/or ventilation/CPAP/high flow.

4.3.3 Measures

4.3.3.1 The Ages & Stages questionnaire

Caregivers of the participants completed the Ages & Stages Questionnaire (ASQ), specifically the section appropriate for six-month-old infants. From the total sample tested in the number familiarisation task, twenty-three families completed the Ages & Stages questionnaire (13 out of 14 full-term infants and 10 out of 13 preterm infants). Sessions ran longer than expected explaining attrition rate. Few families were unable to return questionnaires by free post, although researchers tentatively contacted families kindly requesting to return completed questionnaires.

4.3.3.1 Number Familiarisation Task

To investigate number sensitivity in six-month-old infants born prematurely, we employed a visual number familiarisation paradigm. We used the ratio-limit (1:2) that six-month-old infants have been shown to be able to discriminate successfully (Xu & Spelke, 2000; Lipton &7 Spelke, 2003; Xu, Spelke & Goddard, 2005). The paradigm comprised paired displays of 8 *vs.* 16 dots. A similar paradigm has been previously used in a longitudinal study investigating sleep patterns and development cognition in infants during their first year (Pish, 2015).

4.3.3.3 Apparatus

Infants' looking behaviour was recorded using a Tobii TX060 eye-tracker (1024 x 768 pixels monitor). The eye-tracker comprised an infrared camera that automatically captured corneal refraction and, once calibrated, tracked where the infant was looking on the screen. Tobii Studio 3.0.2 was employed for infants' eye calibration, displaying of the task, and recording the infants' looking behaviour.

4.3.3.4 Stimuli

Stimuli consisted of pairs of black dot arrays in a white circle background (1152 x 495 pixels) containing either 8 or 16 dots in the ratio of 1:2 for 6-month-old infants (Xu & Spelke, 2000; Lipton & Spelke, 2003; Xu, Spelke & Goddard, 2005). Stimuli were created using a script developed by Piazza et al. (2004), which generated dot arrays to vary in item size during familiarisation trials and to control for total luminance during test trials. Dot radii and dot position varied randomly between sets during familiarisation trials, but dot size was constant during test trials. All stimuli were black and white to avoid any bias because of colour preference. This method made it possible to control continuous perceptual variables that could otherwise be confounders.

4.3.3.5 Design

Infants were familiarised to paired displays of either 8 or 16 elements for six consecutive trials presented for seven seconds. Following familiarisation, infants were presented with two paired test trials for seven seconds, in which the number of dots changed within one of the sides (e.g., 8 dots on one side and 16 dots on the other). Figure 4.1 illustrates the design of the experiment. Half of the infants were familiarised with 8 dots and half of the infants with 16 dots. Additionally, the side of the novelty was counterbalanced, given a total of four test designs. This means that two tests involved infants being familiarised with 8 dots and tested with 16 dots, counterbalancing the side of the novelty during the test phase. Similarly, two tests were designed where infants were familiarised with 16 dots and tested with 8 dots, counterbalancing the side of the novelty during the test phase.

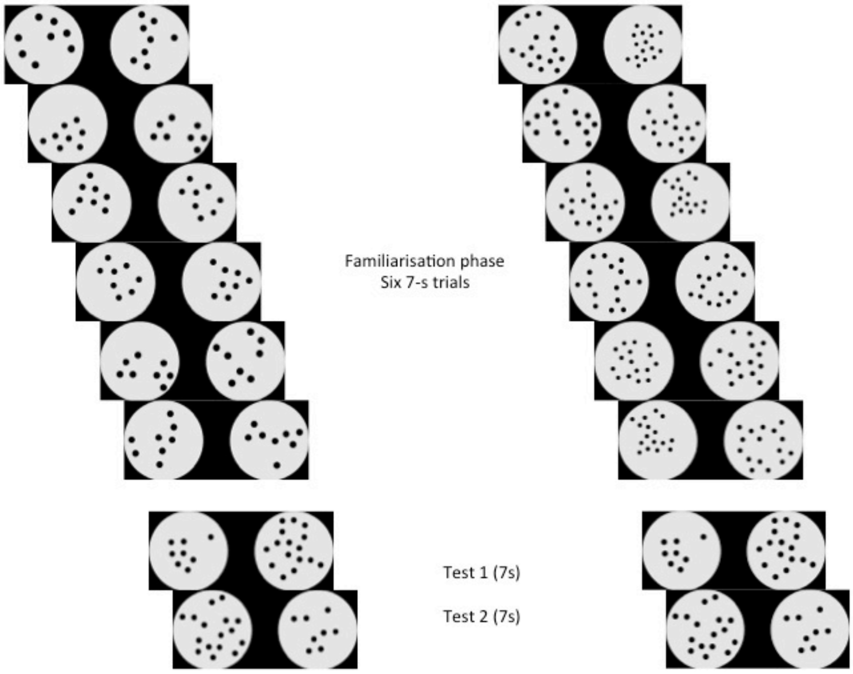


Figure 4.1: Schematic representation of the familiarisation and test trials displays employed in Study 1.

The side of novelty and familiar stimuli were counterbalanced during the test phase.

4.3.3.6 Procedure

Infants sat in a car seat approximately 70 cm from the TV screen attached with the Tobii TX060 eye-tracker. The position of the infant was adjusted until the eye-tracker detected the infant's eyes. Nine-point calibration sequences were employed using a bouncing-shaker stimulus combined with a squeaky sound to draw the infant's attention to each calibration point. The accuracy of the calibration was checked and repeated if necessary.

An attention-grabber was presented before each trial to ensure that the infant was fixating the middle of the screen and that a good track could be obtained. Once the infant was looking at the screen, the trial started. After six familiarisation trials, the infant was presented with two test trials.

4.4 Data analysis

4.4.1 The Ages & Stages questionnaire

Mann-Whitney tests were used to compare performance between preterm and term children on the Ages & Stages Questionnaire, and effect size was calculated across subscales.

4.4.2 Number Familiarisation Task

4.4.2.1 Defining areas of interest

For the number familiarisation task, first, areas of interests (AOIs) were manually determined using Tobii Studio 3.0.2, by selecting the white circle containing the dots in each stimulus (see figure 4.2). The raw eye-tracking data were subjected to a fixation filter that determined the length of each fixation and the corresponding area of interest. The filter defined a fixation as a period of time in which the eye position does not move more than 35 pixels for at least 140 ms (8–11^{o} degrees visual angle at the infant's viewing distance). Individual total looking times of fixations in AOI were analysed during the familiarisation and test phases.

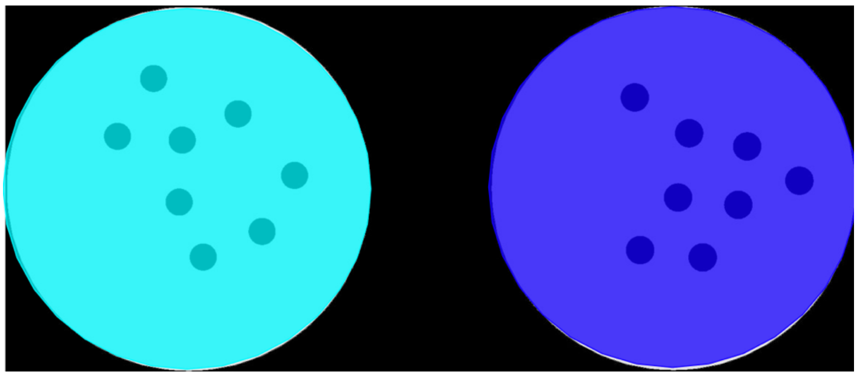


Figure 4.2: Example of area of interest (AOI) selected for a familiarisation trial.

4.4.2.2 Familiarisation phase

We wanted to investigate whether infants looking time decreased during the familiarisation phase. We hypothesized that both groups would decrease their looking time when comparing the first and last familiarisation trials. As such, we employed paired t-test, contrasting the first and last familiarisation trials. Additionally, we also wanted to investigate whether infants regained interest during test phase, when novelty was introduced during test phase. We hypothesised that infants would demonstrate regained interest by increasing their looking time during the first trial of test phase. Thus, we employed paired t-tests, contrasting total looking time of the last familiarisation trial and the first test trial.

4.4.2.3 Test phase

A mixed-design 2 x 2 x 2 ANOVA was used to explore differences between *Novelty* (novelty and familiarity) and *Trial* (first and second) as the within-subjects factors, and *Group* (term and preterm) as the between-subject factor. The dependent variable was z-scored looking time (seconds). In order to investigate whether infants would show differences in looking times towards novelty according to the number of new elements displayed (16 *vs.* 8), a second 2 x 2 x 2 x 2 mixed-design ANOVA was carried out, with *Novelty* (novelty and familiarity) and *Trial* (first and second) as the within-subjects factors, and *Group* (term and preterm) and *Novelty Bias* (16 and 8) as the between-subjects factors.

We also explored proportional looking times between novelty and familiarity. If infants' attention was directed to novelty, we would expect infants to fixate in the new stimuli more than 50% of the time on average, looking disproportionately at the novelty, fixating significantly longer than would be predicted by chance. Thus, one-sample t-test was carried out to explore the proportion of looking time at the novelty against 50% for each group. A similar procedure was employed to verify looking time by chance according to the test trial. Thus, proportional looking time was investigated between novelty in the first and second test trials against 25%.

Subsequently, we individually identified participants who successfully looked longer at the novelty. Participants were identified when they looked more than 50% of total looking time at novelty during test phase. A chi-square test was carried out to test differences in odds ratios between groups. We then individually inspected the first look, coding this as successful or unsuccessful for individuals who, respectively, did or did not present the first fixation oriented to the AOI towards novelty. If infants gazed first towards novelty, the trial was coded as successful and if their first saccade was towards the familiarised number, it was coded as failed. All trials with looks that did not clearly go in one direction, or that did not start in the middle of the screen were excluded from analyses. Odds ratios were examined between individuals who successfully gazed first towards novelty, according to groups for each test trial.

Finally, independent t-tests were carried out between all trials (familiarisation and test phases) to compare differences between groups (term *vs.* preterm).

4.5 Results

4.5.1 The Ages & Stages questionnaire

Data from the Ages & Stages questionnaire were not normally distributed. Thus, Mann-Whitney U tests were carried out to examine differences between preterm and term groups on communication, gross motor, fine motor, problem solving and personal social subscales of the Ages & Stages questionnaires. This revealed that none of the subtests showed statistically significant differences in ranked distributions between groups, with small effect sizes for all subscales. No infants were excluded from the sample due to scores below or close to cut-offs in the problem-solving domain. Table 4.3 summarises the scores obtained from the Ages & Stages. Median and individual scores for each domain of The Ages & Stages questionnaire are illustrated in figure 4.3. Dotted lines represent scores below cut-offs given by the Ages & Stages questionnaire.

Table 4.3: Results of each subscale of The Ages & Stages questionnaire

	Term			Preterm			
Subscales	N	Median	IQR	N	Median	IQR	*p*
Communication	13	45	(42,52)	10	45	(32.5, 51,25)	0.522
Gross Motor	13	45	(35,52.5)	10	40	(33.7, 51,25)	0.483
Fine Motor	13	50	(40, 55)	10	45	(40, 60)	0.879
Problem Solving	13	50	(45, 55)	10	50	(40, 56.25)	0.976
Personal Social	13	50	(45,55)	10	45	(33.75, 55)	0.522

Data not normally distributed. Comparison between groups based on Mann Whitney U test.

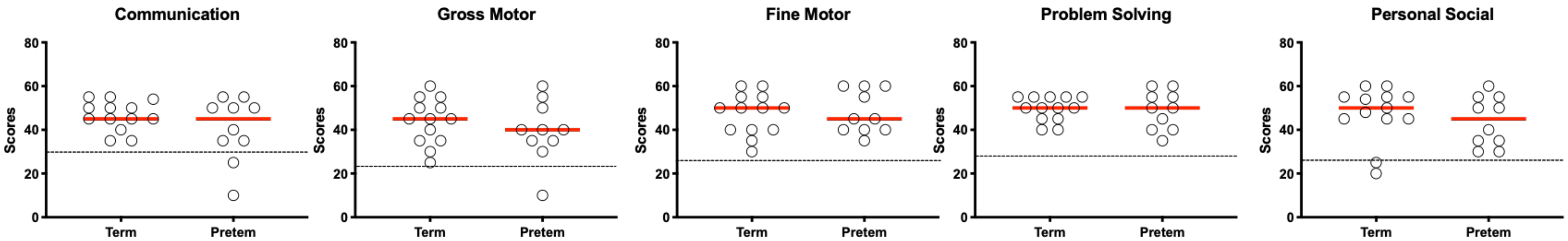


Figure 4.3: Medians and individual scores for each subscale of The Ages & Stages questionnaire.

Dotted lines represent cut-offs given by the questionnaire. Overall, no significant differences were found between groups for any of the subscales.

4.5.2 Number Familiarisation Task

4.5.2.1 Familiarisation phase

As data from the familiarisation task were not normally distributed, we standardised the data using individual z-scores for the familiarisation and test phase separately, and subsequent analyses were carried out.

To investigate differences between familiarisation trials and groups, we carried out paired t-tests between the *First* and *Last* familiarisation trials. As expected, this revealed a significant difference between the *First* (M= 5.70 seconds, SE = 1.01) and *Last* (M= 4.72 seconds, SE = 1.75) familiarisation trials ($t(26)=2.317$, $p=0.029$), suggesting that infants' looking time significantly decreased during the familiarisation phase. To investigate whether infants decreased their looking time during the familiarisation phase and regained interest in the test phase, we also contrasted looking time between the *last familiarisation* and the *first test* trials. Here, the paired t-test revealed no differences between trials ($t(26)=0.974$, $p=0.339$). Figure 4.4 shows proportional looking time during familiarization trials (graph A) and proportional looking time to the last and familiarisation trial and the first test trial (graph B).

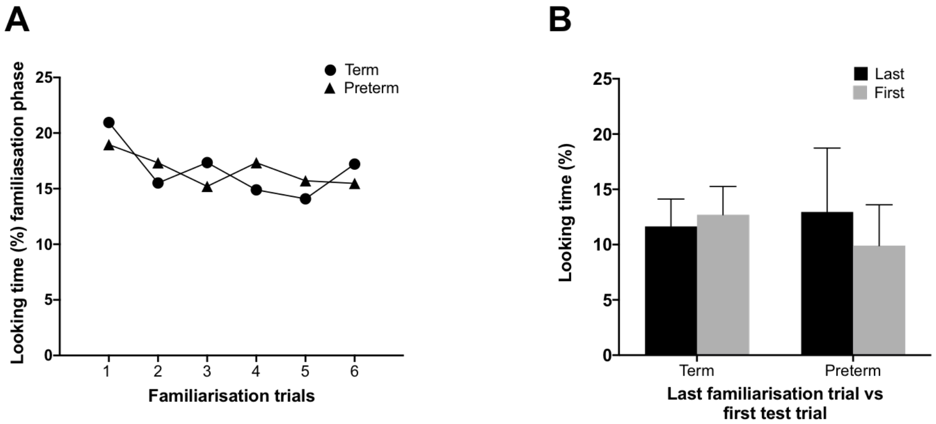


Figure 4.4: Looking time during familiarisation phase at six months of age.

A) Proportional looking time infants spent during familiarisation trials. Infants, regardless of group, significantly decreased their attention when comparing first and last familiarisation trials. B) Proportional looking time between the last familiarisation trial and the first test trial. No significant differences were observed between the last familiarisation trial and the first test trial either in the term or the preterm group ($p>0.05$). Error bars represent standard deviation of the data.

4.5.2.2 Test phase

A mixed-design 2 x 2 x 2 ANOVA was carried out to investigate differences between *Novelty* (new *vs* old), *Test Trial* (first *vs* second) and *Group* (term *vs* preterm). An interaction between *Group* and *Trial* was found ($F(1,25) = 4.597$, $p=0.042$). The post hoc paired t-test revealed no significant difference between the first and second new trial for the preterm group ($t(12)=-2.099$, $p=0.058$). Similarly, for the term group, the post hoc paired t-test revealed no significant difference between the first and second new trial ($t(13)=0.065$, $p=0.945$). Marginal differences in looking time between novelty trials for the preterm group, but not the term group, explained this interaction. No other main effects or interactions were observed. We carried out a second 2 x 2 x 2 x 2 mixed ANOVA including *Novelty Bias* (16 *vs* 8) as a between-subject factor investigating whether infants would show differences in looking times towards novelty according to the number of elements. No other interaction or main effect was observed on *Novelty Bias*. This suggests that differences observed were not related to the type of number of quantities displayed on task.

We investigated whether looking time towards novelty was above chance. One sample t-test was employed to test whether novelty preference was different than 50% according to group. This revealed no significant differences for either groups (term: t(13)=1.811, *p*=0.093; preterm: t(12)=0.793, *p*=0.443), suggesting that, overall, infants were looking to novelty equally, or just above, chance. We then explored looking time towards novelty for each test trial (looking time different than 25%) according to groups. One sample t-test revealed no significant differences for the first test trial for either the term t(13)=0.883 *p*=0.394) or the preterm group t(12)=-1.061, *p*=0.309). Similarly, no significant differences were found in the second test trial in the term group (t(13)=-1.061, *p*=0.309), or the preterm group (t(13)=2.017, *p*=0.067). This is illustrated in Figure 4.5.

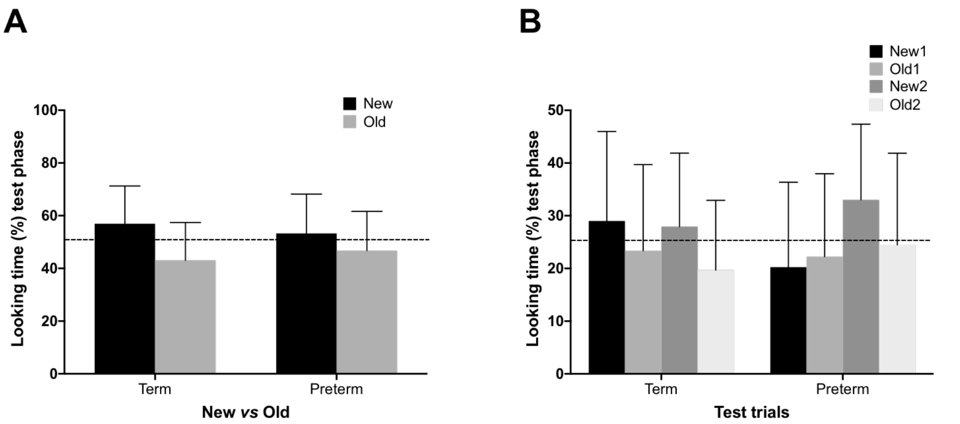


Figure 4.5: Proportional looking time during test phase at six months of age.

A) Proportional looking time spent during test phase. Both groups of infants showed no preference above chance. B) Proportional looking time spent during the test phase according to the test trial. Both groups of infants showed no preference above chance according to test trial. Dotted line present percentage of looking time by chance. Error bars represent standard deviation of the data.

Subsequently, we individually identified infants that displayed longer looking times towards novelty. We used proportional looking time to identify participants who looked more than 50% of time to novelty A chi-square of independence was performed to examine the relationship between groups and the ability to discriminate novelty. The relationship between these variables was not significant (x^2 (1, N=27) = 0.054, p=0.816). Both groups of infants were equally likely to display longer looking times to novelty. Half of the term-born and 45% of the preterm infants looked longer to the novel stimuli.

Next, we explored first looking orientation towards novelty to each test trial separately. This revealed that 41% of the term-born infants displayed the first fixation towards novelty in the first trial, in comparison to 46% of the preterm infants. A chi-square of independence revealed no difference between groups (x^2 (1, N=26) = 0.181, p=0.671). Approximately 69% of the term-born infants displayed the first fixation to novelty in the second test trial in comparison to 28% of the preterm infants (x^2 (1, N=27) = 4.464, p=0.035). Figure 4.6 illustrates an example of an infant displaying the first fixation towards the new number of elements during a test trial. No differences between groups were found for condition bias when examining infants who displayed the first fixation look towards novelty for a bigger number of elements, in this case, 16 elements (x^2 (1, N=13) = 0.737, p=0.135), nor novelty for a smaller number of elements (x^2 (1, N=13) = 0.737, p=0.391), in this case, 8 elements.

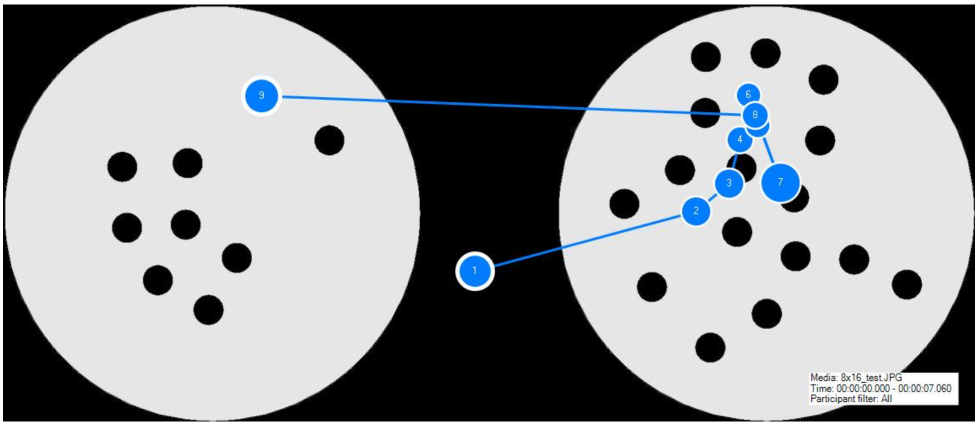
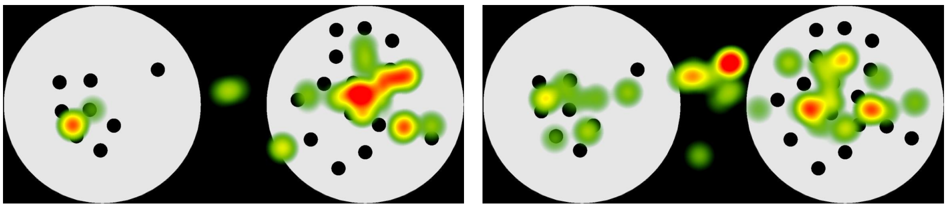


Figure 4.6: Example of a gaze plot during test phase

Gaze plots were individually inspected to verify whether infants displayed the first fixation in the AOI towards novelty.

Table 4.4 summarises the results of the number familiarisation task, comparing the term and the preterm group at six months of age. Figure 4.7 represents a heat map of a test trial of the looking time of the eye-tracking data. Visual inspections of eye-tracking heat maps suggest differences between groups in search looking patterns, although those observations were not systematically investigated in this study.



Term group **Preterm group**

Figure 4.7: Heat map of the eye-tracking data illustrating a trial during test phase.

Red areas represent a longer time for fixations whereas green areas represent a shorter time. In this example, infants were familiarised with eight numbers of elements. The figure on the left-hand side represents looking time displayed by the term group and the figure on the right-hand side represents the preterm group. Appendix 4.1 illustrates heat maps of all test trials according to the group and the number of elements during the test phase

Table 4.4: Looking time performance (seconds) in the number familiarisation task by full-term and preterm infants at 6 months old.

	Term			**Preterm**			**Differences between full-term and VP infants**
Measures	N	Mean	SD	N	Mean	SD	Mean differences (95% CI)
Familiarisation phase							
Trial 1	14	5.85	1.71	13	5.55	1.40	0.29 (-0.95 to 1.54)
Trial 2	14	5.26	1.25	13	4.26	1.92	1.00 (-2.76 to 2.28)
Trial 3	14	4.83	1.93	13	4.82	1.75	0.01 (-1.46 to 1.47)
Trial 4	14	5.18	1.28	13	4.17	1.77	1.00 (-0.21 to 2.23)
Trial 5	14	4.88	1.60	13	3.94	2.08	0.93 (-0.53 to 2.40)
Trial 6	14	4.77	1.61	13	4.66	1.96	0.11 (-1.30 to 1.53)
Test phase							
New 1	14	2.75	1.73	10	2.36	1.28	0.38 (-0.89 to 1.66)
Old 1	13	2.60	1.82	12	1.91	1.19	0.69 (-0.58 to 1.96)
New 2	14	3.13	1.79	13	2.57	1.20	0.55 (-0.69 to 1.80)
Old 2	13	1.89	0.94	12	2.29	1.31	-0.40 (-1.36 to 0.55)

All differences between groups $p > 0.05$, with effect size < 0.01.

4.6 Discussion

To date, this is the first study to have investigated early numerical sensitivity in VP infants. Using the ratio limits where prior work found successful discrimination in six-month-old full-term infants (Xu & Spelke, 2000), our results revealed no significant differences in discrimination within the 1:2 ratio-limit, both in the control group and in the preterm group. Here, we failed to replicate the findings established in previous studies investigating term-born infants (Xu & Spelke, 2000; Lipton & Spelke, 2003; Xu, Spelke & Goddard, 2005). Our sample of VP and full-term infants displayed no preference in their looking behaviour for either familiar or novel number of elements. The preference for number novelty was similar to, or just above chance for both groups, and rates of infants who looked significantly longer to novelty were similar between groups.

A few hypotheses could explain the null results in our study. Firstly, when contrasting looking time between the last familiarisation trial and the first test trial, infants, regardless of the group, did not increase their attention, indicating that they were not fully familiarised. Typically, when infants are fully familiarised, preference to novel stimuli is observed, but when not fully familiarised, infants prefer familiar stimuli (Rose, Gottfried, Melloy-Carminar, & Bridger, 1982; Hunter, Ames, & Koopman, 1983; Hunter & Ames, 1988; Roder, Bushnell, & Sasseville, 2000). These findings suggest that infants were not fully familiarised and results from the test phase support that infants were unable to notice changes in stimuli. Preferences from novelty were equal or just above chance.

Factors that affect novelty and familiarity responses in infants have been extensively studied and two factors have emerged as the main determinants of the type of response elicited from infants. The first factor contributing to novelty preference is stimulus complexity. Very simple stimuli may lead to a novelty preference, whereas more complex stimuli, which necessitate more elaborate processing, may lead to a familiarity preference (DePaolis, Keren-Portnoy, & Vihman, 2016). In our study, we controlled for continuous perceptual variables following the same procedure employed in previous studies (Piazza et al., 2004; Pish, 2015). When investigating whether infants showed a preference for more complex stimuli, i.e. larger quantities, no significant differences were observed, either for total looking time or first gaze, indicating that stimuli complexity might not have played a role during novelty discrimination in our study. The second factor that affects infant responses to familiar or novel stimuli is familiarisation time. In a typical experiment, an infant is familiarised to a stimulus and the infant's attention is measured relative to a similar but novel stimulus (see Rose et al., 1982, for an example). With a brief familiarisation time, infants show a preference for the familiarised stimuli, but as exposure time in the familiarisation phase increases, the preference shifts to the novel stimulus. Our results seem to converge to the fact that not all infants were given enough time to be familiarised and therefore did not display novelty preference.

Based on the existing literature regarding infant looking behaviour, we reasoned that our experimental design did not allow enough stimulus exposure for infants to familiarise themselves fully to the familiar number of elements, indicating that infants need additional time to process new information. In fact, deficits in processing speed are already observed in the first year of life of infants born prematurely, who have been shown to require as much as 30% more inspection time to perform as well as term controls (Rose et al., 2002). A meta-analysis revealed that preterm infants' habituation and dishabituation are significantly poorer compared to term infants. Preterm infants' performance in looking behaviour measures is moderated by risk factors, such as gestational age and brain abnormalities (Kavšek & Bornstein, 2010). However, given our null results, we could not draw any conclusions whether preterm infants display numerical sensitivity at six months of age, or whether they might require additional time to display novelty preference due to slower processing speed.

One main difference between our study and the previous studies investigating numerical sensitivity in infants is the design of the task chosen. We initially wanted to ensure complete similarity with Xu and Spelke's (2000) original study, replicating their habituation study. Typically, in habituation studies, infants are presented with a stimulus, or set of stimuli, until the average looking time in some block of trials decreases to a pre-specified criterion (e.g., 50% of what it was in the first block). Habituation is calculated by averaging infants' looking in blocks of trials and comparing those averages as the session progresses (Oakes, 2010). In our study, however, we decided to use a familiarisation task. In familiarisation studies, all infants are presented with a fixed number of familiarisation trials regardless of changes in attention (e.g., Quinn et al., 2002; Kovack-Lesh, Horst and Oakes, 2008). The fixed time displaying the repeated stimuli, however, might be insufficient for the infant to decrease their attention. In our pilot study, replicating Xu and Spelke's (2000) number habituation task, infants did not tolerate the time exposed to the habituation task, exhibiting high rates of fussiness. This is why we decided to use familiarisation rather than habituation in the present study. To achieve that, we replicated a number familiarisation paradigm previously employed by Pish (2015). In her study, six familiarisation trials were presented for five seconds. We tentatively tried to replicate her study, but longer times were necessary for familiarisation. We piloted data with six familiarisation trials lasting either seven seconds or nine seconds. Only six trials with seven seconds were tolerated by infants in our study. Numbers of trials in habituation and familiarisation studies are diverse. A study investigating ordinality in infants revealed that they required an average of eight trials to meet the habituation threshold (Brannon, 2002). A study investigating number sense in toddlers with Williams syndrome employed nine familiarisation trials (Herwergen et al., 2008). Generally, there is no consensus about either habituation or familiarisation criteria across infant studies, and thus studies are likely to differ in how many infants included in the final analyses did not actually habituate. In fact, infants' studies focusing on memory suggested that only participants with a minimum of 14000 ms total looking time during the familiarisation phase and a minimum fixation of 2500 ms looking time during the test phase should be included (Richmond et al., 2015). It might be the case that preterm infants, in addition to requiring additional time to process novel information, also show high rates of inattentiveness. In fact, it has been reported that very preterm children present significantly more inattention symptoms than their term-born peers but there was no excess of hyperactivity/ impulsivity (Johnson & Marlow, 2014).

Future studies should investigate differences in visual exploration patterns between term and preterm groups. Eye-tracking techniques permit systematically exploration of the area of the screen that infants look at when scanning a numerical array. For example, eye-tracking data revealed that infants with Downs syndrome tend to scan an overall array and did better on large number discrimination (Karmiloff-Smith et al., 2012), whereas those with William syndrome, whose tracking was confined to focussed areas, performed better on small number discrimination (Van Herwegen et al., 2008). In our study, only visual inspections of heat maps of test trials (appendix 4.1) were possible due to the program employed. However, future studies could systematic explore differences on visual strategies according to groups.

Additional limitations of this study should be addressed. One of the main limitations is the small sample size, suggesting that this study lacks in statistical power. If we were to conduct the experiment with a larger sample size, the power calculation suggests that we would need at least 181 infants per group to be able to detect differences between groups when discriminating numerical novelty (α=0.05, 1-β=80%, two sample, two-sided t-test, assumed effect size Cohen's d = 0.3). This suggested sample size is unfeasible when working with infants, especially an atypical population such as infants born prematurely. In addition, it is not even clear that such a small effect would be clinically relevant. Developmental studies frequently face difficulties regarding the sample size especially in studies into clinical populations. Recently, a large-scale infancy project has been in place aiming to replicate studies in different cognitive areas comprising an effort of several researchers from more than 30 countries, the 'ManyBabies' project. The project goals are to assess key findings in infancy in consensus paradigms and increasing diversity and numbers of participants (https://manybabies.github.io/projects/). Large-scale infant studies are crucial to better understand infants' development, increasing representative samples and providing normative standards for studies investigating infants with clinical conditions. Future studies should make efforts to employ paradigms tested in large and representative samples, facilitating direct comparisons. In addition, given the high prevalence of retinopathy of prematurity (ROP)[14] in VP infants, we would recommend that future studies

[14] Retinopathy of prematurity (ROP) is caused by abnormal development of retinal blood vessels in premature infants. It causes abnormal blood vessels to grow in the retina, and can lead to blindness.

testing visual performance should include a visual acuity test in order to eliminate the possibility of low performance due to visual impairment. A study investigating visual acuity in six-month-old infants showed that both monocular and binocular visual acuities are worse in premature infants than in full-term infants at the same chronological age (Spierer, Royzman, & Kuint, 2004).

Overall, we failed to replicate results previously reported in studies investigating numerical sensitivity in six-month-old infants (Xu & Spelke, 2000; Lipton & Spelke, 2003; Xu, Spelke & Goddard, 2005). With caution against reasoning from a null result, we tentatively conclude that number processing may not be generally impaired in preterm infants at this stage of development, but difficulties in domain-general abilities, such as poorer processing speed and inattentiveness, might affect their ability to discriminate novelty quantities successfully. Speculatively, our results appear to be in line with previous studies suggesting that difficulties in basic numerical skills in preterm school-children are mediated by difficulties in domain-general abilities (Guarini et al., 2014; Simms et al., 2015, Tinelli et al., 2015). Further research must be conducted, increasing sample size and evaluating other periods of numerical development

5 Number familiarisation and visual working memory abilities in 12-month-old very preterm infants

5.1 Abstract

Infants are able to discriminate quantities with increasing precision during their first year of life. By the age of ten months, infants are able to reliably discriminate on a 2:3 ratio (8 from 12 elements). The early ability to discriminate numerosities has been associated with later mathematical performance. At the same time, general cognitive abilities, such as visual working memory, are also linked with later mathematical performance. Infants born very prematurely (less than 32 weeks of gestational age) are at an increased risk of having problems with working memory capacities and mathematical difficulties. Yet, numerical discrimination and its association with visual working memory abilities in infants born very prematurely has been largely unexplored. We investigated whether infants born very prematurely displayed similar numerical sensitivity to full-term infants by the age of twelve months. Additionally, we evaluated general cognitive abilities according to composite scores from the Bayley-III, and visual working memory using the A-not-B task, and the association of each of these with numerical sensitivity. Thirteen very preterm and 24 full-term infants aged twelve months were included in this study. Overall, very preterm infants were able to discriminate the 2:3 ratio, but took longer to do so successfully. General cognitive abilities and working memory performance were similar between groups and not associated with performance in numerical discrimination. We tentatively conclude that numerical discrimination may not be generally impaired in preterm infants at this stage of development.

5.2 Introduction

Children born very prematurely (<32 weeks of gestational age) have more frequent difficulties in mathematics compared to typically developing children (Simms, Cragg, et al., 2013a). Identification of early predictors of attainment would be useful to allow targeted interventions (Gersten, Jordan, & Flojo, 2005). Previous studies have suggested that performance in numerical discrimination (Libertus, Feigenson & Halberda, 2013; Mazzocco, Feigenson & Halberda, 2011) and working memory (Bonny & Lourenco, 2013) are potential predictors of mathematical abilities. Even in infants, it is possible to assess both numerical sensitivity (e.g., Xu & Spelke, 2000) and working memory abilities (Gilmore & Johnson, 1995). Thus, the aim of this study was to investigate whether differences in numerical discrimination, a precursor to later mathematical skills, could already be observed between preterm and full-term participants during infancy; and whether numerical sensitivity was related to visual working memory abilities.

Developmental studies investigating numerical trajectories have systematically demonstrated that as infants get older they are able to discriminate between displays that differ in smaller ratios (for more detailed discussion, please see section 2.5.2). The ability to discriminate numerosities in early stages of numerical development has been associated with later mathematical performance. For example, the numerical discrimination performance in six-month-old infants predicted standardised math scores in the same children three years later (Starr et al., 2013). In fact, a meta-analysis revealed significant but modest ($r=0.2$) relations between performance in numerical discriminations and mathematics abilities (Fazio et al., 2014; Chen & Li, 2017; Schneider et al., 2017).

Studies investigating numerical discrimination in infants as predictive of later maths abilities, however, have failed to consider the possibility that the relationship between numerical sensitivity and mathematics performance is mediated by other domain-general abilities such as visual working memory (WM). Moreover, working memory has been found to be an important domain-general skill for predicting mathematical performance (e.g., Fanari, Meloni, & Massidda, 2019). Investigating the association between numerical discrimination, working memory and maths skills in kindergarten children, Xenidou-Dervou et al. (2013) demonstrated that individual differences in WM highly predicted mathematical performance beyond any other domain-specific tasks, such as numerical discrimination. In fact, von Aster and Shalev (2007) proposed a developmental model for numerical cognition (The Four-Step Developmental Model of Numerical Cognition), giving a central role to working memory in support of a range of numerical abilities, from basic numerical discrimination to the child's ability to solve complex arithmetic. This model suggests both an increase in working memory capacity during development and the recruitment of working memory capacity even during infancy to support numerical sensitivity, although no empirical evidence is provided.

Studies examining early working memory have shown that children born preterm have deficiencies in working memory throughout childhood (Taylor & Clark, 2016). Working memory difficulties are significantly associated with later risks of academic difficulties at school age in children born preterm, especially maths (Simms et al., 2015). A task that permits the investigation of early mechanisms of working memory in infants is the A-not-B task, wherein infants' ability to remember the spatial location of hidden objects is tested (Diamond, 1985, 2001). In fact, studies have shown an association between working memory (as measured by tasks such as the A-not-B task) and the cognitive and academic outcomes of children born preterm (Woodward, Edgin, Thompson, & Inder, 2005).

An extensive body of research indicates that infants born very prematurely are at risk of displaying deficits in visual working memory (Lowe, MacLean, Shaffer, & Watterberg, 2009) and mathematical difficulties (Taylor et al., 2009). To date, however, the association of early ability to discriminate numerosities and visual working memory, both building blocks skills for later mathematical achievement, has not been explored in the very preterm infants, a population of high risk at mathematical performance. Thus, the aim of this study was first to investigate whether atypical trajectories in numerosity sensitivity can already be observed in preterm infants during the first year of postnatal life. Secondly, we examined whether numerical sensitivity, a precursor to later mathematical skills, is related to working memory, a predictor to mathematical performance. We hypothesised that full-term infants would be able to discriminate the ratio-limit of 2:3 in a numerical sensitivity task, in agreement with results reported in the literature (Xu & Arriaga, 2007). We expected, however, that if their numerical abilities are already compromised in infancy, preterm infants would not be able to discriminate twofold change (2:3 ratio). Furthermore, we hypothesised that preterm infants' ability to discriminate numerosities would be associated with visual working memory performance, tested by the A-not-B task, since this is a domain-general skill that is often impaired in the preterm population and an important domain associated with later mathematical performance. To rule out that any group differences might be due to more general neurodevelopmental impairment, we looked at scores obtained from the Bayley-III.

5.3 Methods

The methods were the same as in Study 1, except for the following.

5.3.1 Participants

A total of 53 infants were recruited for Study 2. Data collection happened between March 2016 and March 2018. Mean corrected age for term-born infants (n= 36) was 12 months and 23 days (range: 11 months and 13 days to 13 months and 15 days) and preterm infants (n=17) 12 months and 24 days (range: 12 months and 5 days to 13 months and 8 days). Mean chronological age for preterm infants was 15 months and 10 days (range: 14 months and 22 days to 16 months and 29 days). Mean gestational age for the preterm group was 27 weeks and 6 days of mean gestation (24 weeks and 1 day to 31 and 2 days) and for the full-term group was 40 weeks and 6 days of mean gestational age (range: 38 weeks and 3 days to 42 weeks and 1 day). Data from a total of 16 infants (5 preterm and 10 term infants) was excluded due to fussiness or calibration issues (n=6), quality of eye tracking recording inferior than 50% (n=5) and not enough data available during test phase (n=5). The final sample included data from 24 full-born infants (15 boys and 9 girls) and 13 preterm infants (6 boys and 7 girls).

5.3.2 Demographics

Table 5.1 summarises the general characteristics of the study participants. Independent t-test for corrected age at assessment and maternal age, and chi-square analysis of maternal education and socio-economic status revealed that preterm and full-term samples did not differ significantly in respect to the mentioned variables. Sex is reported as ratio data. Table 5.2 shows the neonatal morbidities of the VP children.

Table 5.1: General characteristics of study participants.

	Term n=24	**Preterm** n= 13
Gestational age (weeks), mean (SD)	40^{+6} (2.52)	27^{+6} (1.06)
Range	24^{+1} - 31^{+2}	38^{+3} - 42^{+1}
Birthweight (g), mean (SD)	3501.25 (400.65)	958.46 (357.34)
Range	2700 – 4300	535 – 1560
Sex (M:F)	15:9	6:7
Chronological age (days), mean (SD)	383 (14.84)	475 (23.60)
Range	356 – 417	454 – 514
Corrected age (days), mean (SD)	388 (14.86)	389 (9.40)
Range	358 – 425	370 – 393
Maternal age (years), mean (SD)	35.5 (3.86)	35.53 (5.96)
Range	28-44	27-46
Maternal Education		
<BD (%)	4.20%	23.00%
> BD (%)	95.80%	77.00%
SES (IMD quintile)		
First (%)	17.40%	5.90%
Second (%)	8.70%	35.30%
Third (%)	21.70%	17.60%
Fourth (%)	21.70%	29.40%
Fifth (%)	30.40%	5.90%

BD: Bachelor's degree

Table 5.2: Neonatal morbidities of the preterm children.

	N	Very Preterm (n=6)	N	Extremely Preterm (n=7)
Multiple births				
IVH / PVL: I – II				
Grade I	1	16.60%	1	14.20%
Grade II	1	16.60%	2	28.50%
5.9ROP				
Grade I	-	-	2	28.50%
Grade II	1	16.60%	3	42.80%
Grade III	-	-	-	-
CLB/ BPD				
Mild	1	16.60%	2	28.50%
Moderate	1	16.60%	4	57.10%
Severe	1	16.60%	-	-

IVH: intraventricular haemorrhage; PVL: periventricular leukomalacia; ROP: retinopathy of prematurity; CLD: chronic lung disease; BPD: bronchopulmonary dysplasia. Mild: O2 at 28 days not 36 weeks, moderate: O2 36 weeks < 30% low flow; Severe: O2 36w >30% +/or ventilation/CPAP/high flow.

5.3.3 Measures

Working memory abilities were assessed using the A-not-B paradigm (see section 3.1.9.3). Bayley-III (see section 3.1.9.1.2) was employed to rule out neurodevelopmental delays.

5.3.3.1 Number Familiarisation Task

To investigate number sensitivity in twelve-month-old infants born prematurely, we employed a visual number familiarisation paradigm. We used the ratio-limit (2:3) that ten-month-old infants have successfully been able to discriminate (Xu & Arriaga, 2007). The paradigm comprised paired displays of 8 *vs* 12 dots.

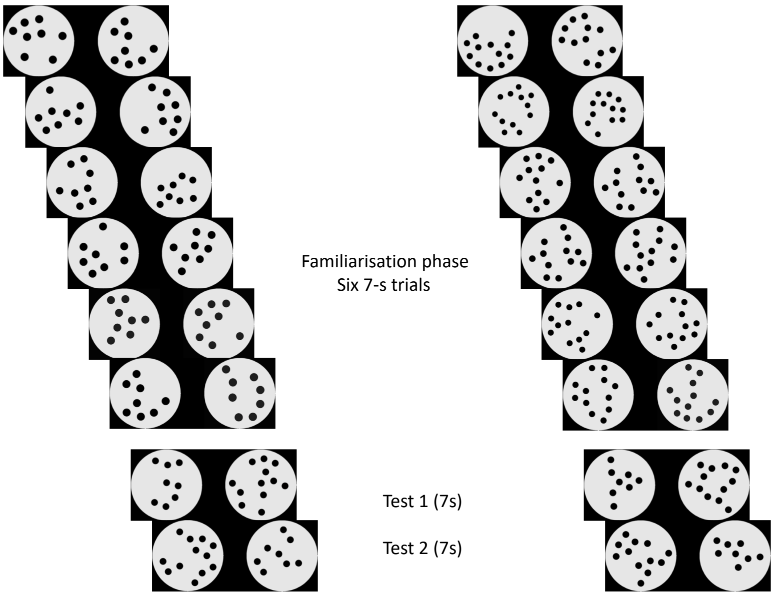


Figure 5.1: Schematic representation of the familiarisation and test trials displays employed in Study 2.

5.4 Data analysis

5.4.1 Bayley-III

Independent t-tests were used to compare performance on a standardised measure (Bayley-III) between preterm and control infants and effect size. Cohen's d was calculated across subscales. Effect sizes were defined as small (0.2–0.3), medium (0.3–0.5), or large (>0.5).

5.4.2 Number familiarisation task

For the number familiarisation task, we ran the same analyses employed in Study 1. In contrast to Study 1, the data were normally distributed in this study. Therefore, raw data were used to run all analyses (for more details, please see section 4.4.2).

5.4.3 A-not-B paradigm

For the A-not-B paradigm, participants' performance was identified according to the 'time to AB error' and four categories were created: 1) participants who failed to retrieve an object with no delay ('Fail'); 2) participants who succeeded to retrieve an object with no delay, but failed at 5 seconds ('0s'); 3) participants who succeeded to retrieve the object both at 0 seconds and 5 seconds, but failed at 10 seconds ('5s'); 4) participants who successfully retrieved the object in all time points of the task, including at a 10 seconds delay ('10s') (see section 3.1.9.3). Time-to-event analysis was carried out to investigate differences between groups.

5.4.4 Associations between number familiarisation and general cognitive abilities

We wanted to explore the associations between numerical discrimination and cognitive abilities within each group separately. Thus, Pearson's correlations were carried out investigating associations between proportional looking time towards novelty and general cognitive abilities from Bayley-III (Cognition, Expressive and Receptive Language, and Fine and Gross Motor Skills) for term and preterm group separately. Correlations above 0.4 were considered strong; correlations between 0.2 and 0.4 were considered moderate, and those below 0.2 were considered weak (Akoglu, 2018).

5.5 Results

5.5.1 Bayley-III

Independent t-tests were carried out to examine differences between groups in Cognition, Language and Motor skills from the Bayley-III composite scores. After Bonferroni correction, significant differences between groups were only observed in the Language scale (term group: M=102.17, SD=11.69; preterm group: M=89.90, SD=9.22; $t(32)= 3.047$, $p=0.005$, $d=1.1$), with big effect sizes. Table 5.3 and figure 5.2 summarise the results of the Bayley-III.

Table 5.3: Results of each subscale of Bayley-III

	Term			Preterm			Differences between full-term and VP infants		
Subscales	N	Mean	SD	N	Mean	SD	Mean differences (95% CI)	*p*	Effect size (Cohen's d)
Cognition	24	111.25	11.25	12	102.50	12.33	8.75 (0.40 to 17.09)	0.040	0.7
Language	23	102.17	11.69	11	89.90	9.22	12.26 (4.06 to 20.46)	0.005*	1.1
Motor	23	104.08	11.27	11	97.36	7.87	6.72 (-0.96 to 14.40)	0.084	0.6

*Remains significant after Bonferroni correction.

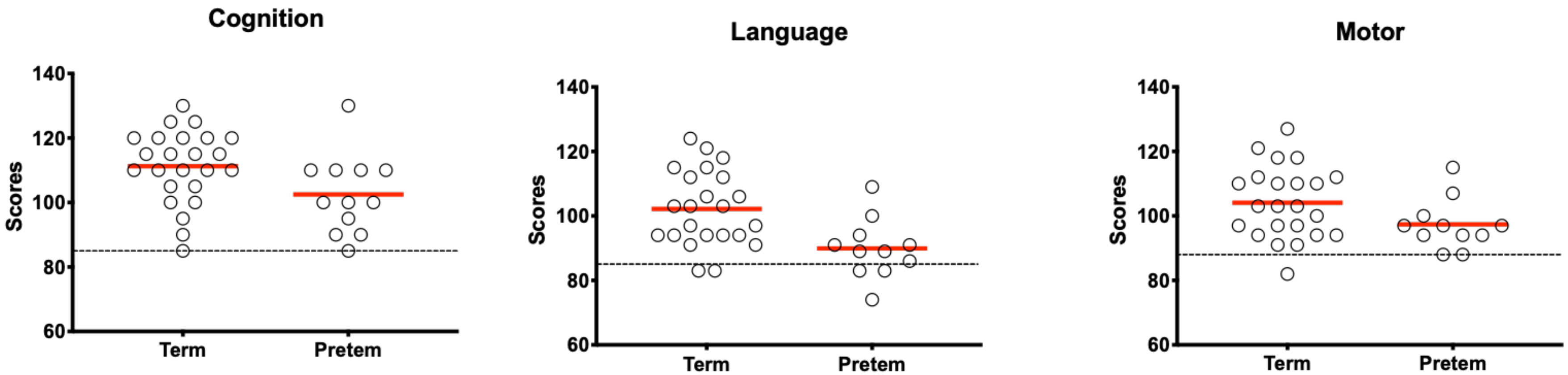


Figure 5.2: Means and individual scores for each subscale of Bayley-III.
Dotted line indicates one standard deviation of Bayley-III.

5.5.2 Number Familiarisation Task

5.5.2.1 Familiarisation phase

A paired t-test was carried out between the *First* and *Last* familiarisation trials. As expected, this revealed a significant difference between the *First* (M= 5.49 seconds, SE = 1.17) and *Last* (M= 4.16 seconds, SE = 1.72) familiarisation trials (t(36)=3.822, p=0.001), confirming that infants' looking time decreased during familiarisation phase. To confirm whether infants decreased their looking time during the familiarisation phase and regained interest in the test phase, we also contrasted looking time between the last familiarisation trial and the first test trial. The paired t-test revealed no significant differences (t(36)=0.640, p=0.526) between the last familiarisation trial (M=4.16 seconds, SE= 1.72) and the first test trial (M=3.90 seconds, SE=0.68). See figure 5.3.

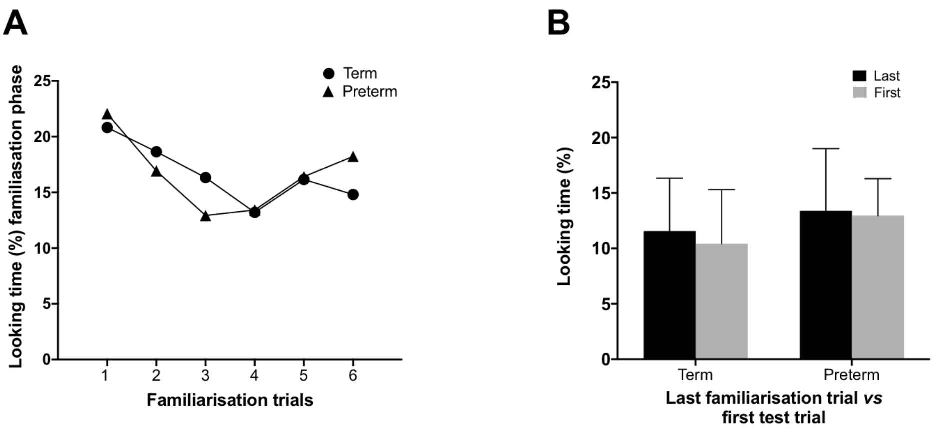


Figure 5.3: Looking time during familiarisation phase at 12 months of age.

A) Proportional looking time expended by infants during familiarisation trials. Infants, regardless of group, significantly decreased their attention comparing between the first and last familiarisation trials. B) Proportional looking time between the last familiarisation trial and the first test trial. Infants did not increase their attention during the first test trial when compared to the last familiarisation trial. Error bars represent the standard deviation of the data.

5.5.2.2 Test phase

A 2 x 2 x 2 mixed ANOVA was carried out to investigate *Group* (term *vs* preterm, between-subjects), *Novelty* (new *vs* old, within-subjects) and *Test Trial* type (first *vs* second, within-subjects). The dependent variable was looking time (seconds). A main effect of *Group* ($F(1,23) = 7.470$, $p=0.046$) was found, with preterm infants displaying significantly longer looking times (M=2.42 seconds, SE=0.21) than term infants (M= 1.85 seconds, SE=0.16) during test phase. As expected, a main effect of *Novelty* ($F(1,23) = 8.508$, $p=0.008$) was found, with infants displaying longer looking times towards novelty (M=2.44 seconds, SE=0.16) than familiarity (M= 1.83 seconds, SE=0.17). No main effects of *Test Trial* were observed. An interaction between *Novelty* and *Group* was also revealed ($F(1,23) = 5.233$, $p=0.032$). Post-hoc pairwise comparisons revealed no significant differences between novelty and familiarity for the term group ($t(23)=0.996$, $p=0.329$). Conversely, a significant difference was observed in the preterm group between novelty and familiarity ($t(12)=-1.799$, $p=0.045$), explaining the interaction.

Similar to the previous ANOVA, we carried out a second 2 x 2 x 2 x 2 mixed ANOVA, but here we included *Novelty Bias* (12 *vs* 8) as a between-subject factor to investigate whether infants would show differences in looking times towards novelty according to the number of elements. No other interaction or main effect was observed regarding *Novelty Bias.*

We investigated whether proportional looking time towards novelty was above chance. One sample t-test was employed to test novelty preference (proportional looking time against 50%) in each group, separately. This revealed no significant differences for either group (term: $t(23)=0.894$, $p=0.380$; preterm: $t(12)=1.287$, $p=0.222$), suggesting that, overall, infants were looking to novelty equally to chance. We also explored proportional looking time towards novelty for each test trial (proportional looking time against 25%) in each group, separately. The sample t-test revealed no significant differences for the first test trial for either the term t(23)=-0.5083 p=0.616) or the preterm group t(12)=1.524, p=0.153). Similarly, no significant differences were found in the second test trial in either the term group (t(23)=1.234, p=0.230) or the preterm group t(12)=0.130, p=0.889). See figure 5.4.

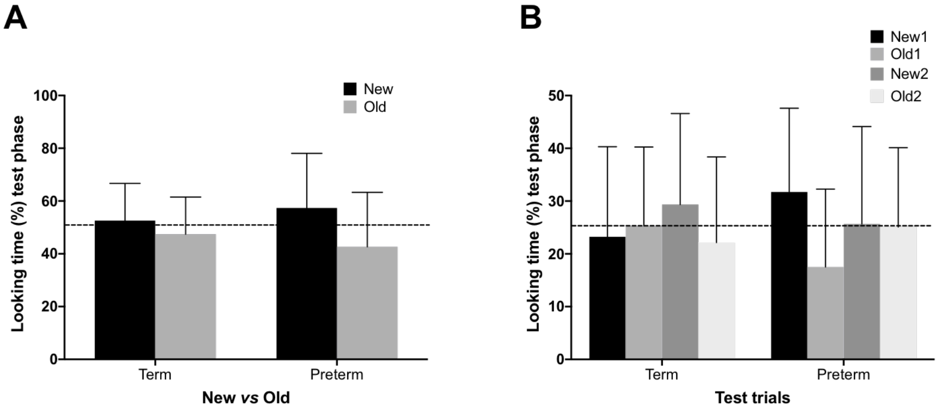


Figure 5.4: Proportional looking time during test phase at twelve months of age.

A) Proportional looking time spent during test phase. Both groups spent longer times looking towards novelty, but preferences were equal to chance. B) Proportional looking time spent during test phase compared to the test trial. Dotted line present percentage of looking time by chance. Error bars represent standard deviation of the data.

Subsequently, we identified individual infants that displayed longer looking times towards novelty, finding that this applied to approximately half of the full-term infants but to two thirds of the preterm group. The relationship between groups and the ability to discriminate novelty was not significant (χ^2 (1, N=37) = 1.857, p=0.172). Therefore, both groups of infants were equally likely to display longer looking times towards novelty.

Next, we explored orientation of first look towards novelty within each test trial separately. This revealed that only one fourth of the term-born infants displayed the first fixation towards novelty in the first trial, in comparison to 46.2% of the preterm infants. A chi-square of independence revealed no difference between groups (χ^2

(1, N=37) = 1.675, *p*=0.189). Both groups had approximately 50% of infants displaying the first fixation towards novelty in the second test trial (χ^2 (1, N=36) = 0.009, *p*=0.599). Table 5.4 summarises performance in the number familiarisation task in Study 2. Appendix 5.1 illustrates heat maps of test trials according to group and number of elements displayed during test phase.

Table 5.4: Looking time performance (seconds) in the number familiarisation by full-term and preterm infants at twelve months old.

	Term			Preterm			Differences between full-term and VP infants		
Measures	N	Mean	SD	N	Mean	SD	Mean differences (95% CI)	*p*	Effect size (Cohen's d)
Familiarisation phase									
Trial 1	24	5.45	0.99	13	5.55	1.49	-0.93 (-0.97 to 0.74)	0.310	<0.1
Trial 2	24	5.01	1.51	13	4.39	1.71	0.61 (-0.48 to 1.72)	0.596	0.3
Trial 3	24	4.41	1.52	13	3.45	1.83	0.96 (-0.18 to 2.10)	0.020	0.5
Trial 4	24	3.57	1.62	13	3.47	1.51	0.99 (-1.01 to 1.21)	0.996	<0.1
Trial 5	24	4.34	1.48	13	4.15	1.65	0.18 (-0.89 to 1.26)	0.740	0.1
Trial 6	24	3.94	1.71	13	4.57	1.73	-0.62 (-1.82 to 0.57)	0.097	0.3
Test phase									
New 1	23	1.75	1.34	13	2.87	1.52	-1.12 (-2.11 to -0.12)	0.215	0.7
Old 1	22	2.03	1.41	12	1.82	1.55	0.21 (-0.85 to 1.29)	0.273	0.1
New 2	21	2.41	1.39	11	2.93	1.68	0.52 (-1.66 to 0.62)	0.685	0.3
Old 2	20	1.85	0.97	12	2.33	1.03	-0.47 (-1.21 to 0.26)	0.535	0.4

5.5.3 A-not-B paradigm

Table 5.5 summarises AB performance with the number of participants according to time to AB error. Time-to-event analysis revealed no significant differences between groups (χ^2=0.051, p=0.820). Both term and preterm infants therefore had similar rates of time to AB error. See figure 5.5.

Table 5.5: AB performance

AB error	Term	Preterm	Total
Fail	3	4	7
0s	9	5	14
5s	11	2	13
10s	1	2	3
Total	24	13	17

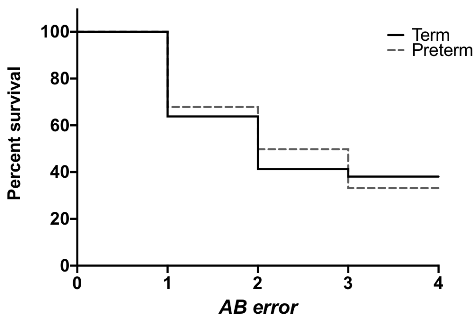


Figure 5.5: AB error according to groups

5.5.4 Associations between number familiarisation and general cognitive abilities

Next, we investigated associations between proportional looking towards novelty and general cognitive abilities from the Bayley-III (cognition, expressive and receptive language, and fine and gross motor skills) for each group separately. Only the term group had a significant association between expressive language and proportional looking time towards novelty (r=0.490, p=0.018). Surprisingly, gross motor skills were also significantly associated to proportional looking time towards novelty (r=0.468, p=0.021). As expected, a significant correlation was observed between expressive and language skills (r=0.445, p=0.033), but only in the term group. No other significant correlations were observed in the term group. However, after Bonferroni corrections, no significant correlations survived. Unexpectedly, no significant correlations were observed in any measures employed in the preterm group. Composites scores from cognition, language and motor skills were not significantly associated with proportional looking time towards novelty in either groups, apart from motor skills (r=0.468, p=0.024) in the term group. After Bonferroni corrections, no significant correlations survived.

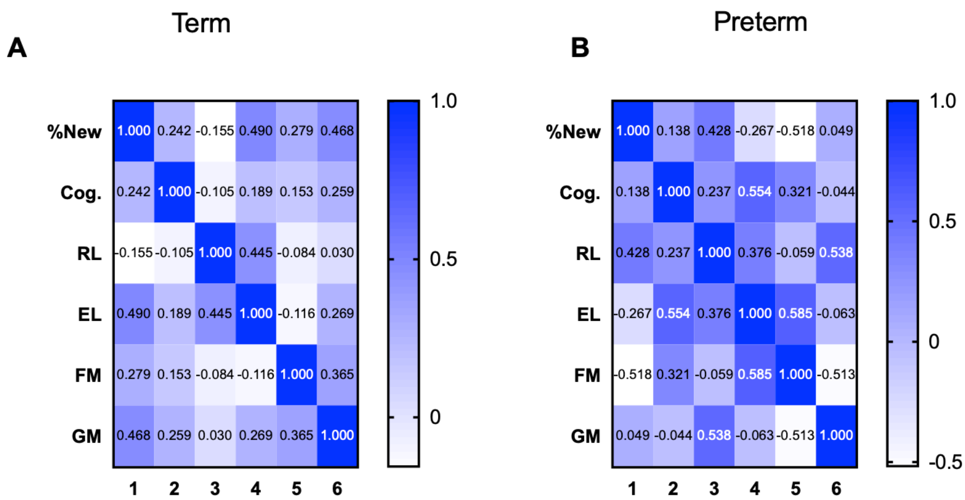


Figure 5.6: Pearson's Correlations between proportional looking time towards novelty and subdomains of the Bayley-III.
Correlational matrix illustrating first order correlations between proportional looking time towards novelty and subdomains of the Bayley-III for the term group (A) and the preterm group (B). New% (1) represents proportional looking time towards novelty; Cog. Stands for Cognitive (2), EL stands for Receptive Language (3), EL stands for Expressive Language (4), FM stands for Fine Motor skills (5); GM stands from Gross Motor Skills (6). Light colours represent either low or negative correlations, whereas bright blue represent higher correlations.

5.6 Discussion

To date, this is the first study to investigate the development of numerical sensitivity and visual working memory in a preterm population in the first year of life. Previous studies have suggested that school-aged children born extremely preterm show difficulties in discriminating numerical quantities (Hellgren et al., 2013; Libertus et al., 2017), but it has been unclear when those difficulties start to emerge and whether they are associated with domain-general deficits in early stages of development. Our study contributes to a better understanding of early stages of the development of numerical skills in the preterm population, as well as understanding the early performance of visual working memory, a domain-general skill associated with later maths skills. The first aim of this study was, therefore, to explore early numerical sensitivity in twelve-month old preterm infants employing a numerical familiarisation paradigm. Using the ratio limit where prior work has found successful discrimination in ten-month-old full-term infants (Xu & Arriaga, 2007), our results revealed that preterm infants were able to discriminate within the 2:3 ratio-limit, but displayed significantly longer looking time to process the novel numerical information than full-term-born infants.

Our results should be interpreted cautiously as differences between groups were observed in the raw data. When using proportional looking times, those differences were no longer observed, in which changes the interpretation of the results. Recently, Csibra and collaborators (2016) recommended a few steps for analysing looking time measures in infancy research. Interpreting looking time proportionally were one of them. According to the authors, it might be more intuitive to employ proportional looking time as an infant who is fast in processing a stimulus may display 10% (rather than, say, 1 s) shorter looking times than others, and novel stimuli may increase looking by 40% rather than by a fixed amount of time (say, 4 s). Thus, it is plausible that the factors influencing the duration of looking measured from a specific zero point are not additive but multiplicative in nature (Csibra et al, 2016). As such, our results should be carefully interpreted as significant differences were only observed when using raw data. Here, when considering proportional looking times, the results indicated that infants, regardless of group, were not given enough time to discriminate novelty, looking to familiarity or novelty at chance. In turn, the validity of our results should be interpreted carefully given that any significant differences between novelty and familiarity might have happen at chance.

Taking these considerations, our results revealed at twelve-month-old corrected age, preterm infants were able to discriminate the ratio-limit of 2:3 in a visual numerical sensitivity task where previous reports showed that ten-month-old term infants successfully discriminated the same ratio-limit (Xu & Arriaga, 2007). In contrast to the full-term infants, however, the preterm babies looked significantly longer at the numerical novelty. This suggest that preterm infants are able to discriminate numerical novelty and that difficulties in numerical cognition in the VP population do not emerge at this stage of development. This in turn implies, however, that, different from what was expected, the full-term infants did not display longer look times towards novelty.

There are two possible explanations for these differences in looking times between groups. Firstly, the age group of our sample was slightly older than previous studies. Whereas we investigated numerical sensitivity in twelve-month-old infants, Xu and Arriaga (2007) tested a sample of ten-month-old infants. This raises the possibility that our numerical familiarisation paradigm might not be age-appropriate for the term group, but appropriate for the preterm group. Here, perhaps, the preterm group was old enough to engage with the task, displaying longer looking times towards novelty than the term group. The term group, however, was much less engaged with the task and might not have felt as challenged by it. In fact, at eleven months old, full-term infants are expected to excel at discriminating sequences of numerosities, in a posterior developmental stage of numerical skills known as ordinality (Brannon, 2002). Thus, it is reasonable to conclude that our task was appropriate for the preterm group, who engaged with the number discrimination tasks, but not for the term group. Based on this assumption, our results indicate that infants born prematurely displayed a similar pattern of performance with advancing age as the controls, in line with the delay hypothesis. The delay hypothesis postulates that VP children can catch-up in performance of tasks with increasing age. This contrasts with the deficit hypothesis, whereby VP children display poorer performance than controls (Ritter, Nelle, Perrig, Steinlin, & Everts, 2013). Longitudinal studies assessing different stages of numerical development of the preterm population should provide evidence to fully support the delay hypothesis.

A second explanation is that preterm infants might need additional time to process new information. The literature is clear that individuals born prematurely present with difficulties in processing speed (Rose et al., 2002; Mulder, Pitchford, & Marlow, 2010; Allotey et al., 2018). From this perspective, the fact that preterm infants displayed longer looking times suggests that they need additional time to process novel information compared to term infants. Additionally, we observed that term infants significantly decreased their looking time during the familiarisation phase, when comparing the last with the first familiarisation trial. This was not the case for VP infants, however. This finding suggests that VP infants might need additional time to process new information, in accordance with our results from Study 1. As a matter of fact, deficits in processing speed are already present in the first year of life among infants born prematurely, with evidence suggesting that they require as much as 30% more inspection time to perform as well as term controls (Rose et al., 2002). For more a detailed discussion see section 4.6.

The second aim of our study was to investigate visual working memory performance. With the use of the A-not-B paradigm, we were unable to demonstrate differences in performance between the term and VP infants at twelve months of age. In the A-not-B task, the infant has to wait to respond, making it possible to assess the emergence of visual working memory abilities, coupled with inhibitory control. Our results show no impact of prematurity in these domain-general abilities: approximately 70% of the VP infants and 87% of term infants displayed some level of working memory and inhibitory abilities at twelve months of age. In older cohorts, working memory has been identified as a specific area of difficulties in preterm cohorts (Mulder, Pitchford & Marlow, 2010), but the literature surrounding the performance of preterm infants in working memory paradigms is not consistent. For instance, preterm infants outperformed term-born infants on an AB paradigm when exploring the maturation of performance from seven to fifteen months of age (Matthews et al., 1996). Ross et al. (1992) found preterm infants to be significantly less successful in the A-not-B paradigm at ten months of age (Ross, Tesman, Auld, & Nass, 1992). Sun and Buys (2011) found that preterm infants displayed a poorer performance on the A-not-B task compared to term-born infants. Finally, individual differences between infants born prematurely in the AB paradigm were observed due to gestational age and gender (Van De Weijer-Bergsma et al., 2010).

The role of executive function skills in mathematical achievement is well established. Executive function skills, including monitoring and manipulating information in mind (working memory), and suppressing distracting information and unwanted responses (inhibition) play a critical role in the development of mathematics proficiency (Cragg & Gilmore, 2014). The prediction value in tasks evaluating domain-specific and domain-general skills in mathematics performance is comprehensively influenced by the stage of development. For instance, Gimbert et al. (2019) showed that numerical discrimination was a significant specific predictor of mathematics achievement only in five-year olds and WM was a significant general predictor only in seven-year olds, suggesting that a general cognitive ability, especially WM, becomes a stronger predictor of mathematics achievement after entrance into formal schooling, whereas ANS acuity, a specific cognitive ability, loses predictive power. Gilmore et al. (2015) showed that inhibition skills were related to conceptual knowledge in older participants, whereas procedural skills related to inhibition in younger participants. It is well-known that children born prematurely present with difficulties in mathematics associated with cognitive problems, especially executive functions (Simms et al., 2013a for a review). Yet, it remains unclear when atypical trajectories start to emerge and the interplay of numerical skills with cognitive abilities. We tentatively explored early cognitive mechanisms later associated with mathematical performance. At twelve months old, neither working memory skills nor general cognitive abilities were linked to numerical discrimination in VP infants. It is plausible to conclude that, at this stage of development, difficulties associated with domain-specific and domain-general are not noticeable in infants born prematurely. Alternatively, it might also be plausible that at early stages of development, these domains are not directly associated, but are later in the process of development. The second approach is consistent with the literature on the cascade effect, whereby deficits in different subdomains of executive functioning have been shown to be correlated with mathematical abilities, but to be distinct from one another. Negative effects of prematurity in mathematical achievement were completely mediated by the three executive functions and speed in a cascade of effects: prematurity, slower processing speed, poorer executive functioning (working memory) and lower achievement in math (Rose et al., 2011). We could speculate that, at this stage of development, difficulties in processing speed are evident in infants born prematurely, but that there are no direct associations with other EFs and numerical discrimination.

General cognitive abilities were also assessed to take account of any neurodevelopmental delays identified by the Bayley-III. Our results revealed no significant differences between groups in the cognitive and motor abilities from the Bayley-III. Additionally, no significant associations between cognitive abilities and numerical skills, in either the term or the preterm group were observed. We conclude that, at this stage of development, VP infants without neurodevelopmental delays do not have difficulties in numerical discrimination and that there is no association between numerical sensitivity and visual working and general cognitive abilities (i.e. cognitive skills from Bayley-III).

Our results showed significant differences between groups in language skills. The preterm group performed on average 12 points below than the full-term group on the language scale. Our results are in line with previous studies. For example, Brósch-Fohraheim et al. (2019) has shown that preterm children had significantly lower scores in language development and the expressive communication subscale from the Bayley-III (Brósch-Fohraheim, Fuiko, Marschik, Resch, & Liu, 2019). In fact, recently, more focus has been given to the role of language in early numerical skills. LeFevre et al. (2010) found that the linguistic pathway (including vocabulary and phonological awareness) was a more consistent predictor of a range of early numeracy skills than domain-general abilities, including executive functions. A recent longitudinal study investigating the role of language and executive functions in early numerical skills in preschool children showed that verbal skills was a significant predictor of counting and number transcoding, beyond any other domain-general skills (Simmons et al, 2020). We explored the role of language and numerical discrimination investigating the associations of proportional looking time towards novelty and language skills (expressive and receptive) from the Bayley-III. Only the term group showed significant associations between the expressive language and numerical discrimination. This is not surprising given the fact that our results suggest that infants in our study looked at novelty at chance.

A few considerations should be addressed regarding the number familiarisation task. Firstly, as expected, looking time during the familiarisation phase gradually decreased for both groups. Intriguingly, looking time increased between the fourth and fifth familiarisation trial, with both groups displaying similar traits. This is an unexpected finding and it remains unclear why infants displayed longer looking times in the last trials. The increased looking time during the last familiarisation trials, however, could explain null differences between the last familiarisation trial and the first test trial (as seen on graph B figure 5.3). Secondly, the fact that both term and preterm infants displayed proportional looking time towards novelty equal to chance (as seen on graph A and B in figure 5.4) might imply that infants generally did not discriminate novelty in our task. Considerations of task design should also be emphasised, as previously discussed with respect to Study 1 (please see section 4.6).

In summary, this study provided evidence that VP infants at twelve months corrected age are sensitive to numerical discrimination using the 2:3 ratio, but needed longer looking times than term infants. At this stage of development, general cognitive abilities and working memory performance were not associated with performance in numerical discrimination. We conclude that numerical discrimination may not be generally impaired in preterm infants at this stage of development. These results should be interpreted carefully, however, due to concerns with the tasks employed.

6 Domain-general and domain-specific abilities and their relationships to mathematical performance in very preterm children

6.1 Abstract

The prevalence of poor mathematical achievement among very preterm children (VP; <32 weeks) is high in comparison to their term-born peers. Yet, the underlying cause of VP children struggling with maths is unclear. On the one hand, research has shown that VP children might have imprecise numerical representations leading to difficulties in maths; on the other hand, evidence shown that those difficulties are driven by deficits in executive functions. Thus, we aimed to investigate the performance of VP children and their full-term-born peers in tasks evaluating executive function domain-general and domain-specific abilities, and their associations with mathematical performance. Domain-general skills were assessed by standardised tests evaluating intelligence, processing speed, working memory, attention, planning, inhibition. Domain-specific skills were assessed by experimental tasks testing symbolic and non-symbolic magnitude comparison. Mathematical performance was tested by a standardised assessment. Thirty-eight very preterm children were compared with 30 full-term-born children aged between eight and ten years old. As expected, VP children had significantly poorer scores in mathematical performance, even after excluding children with low IQ and controlling for socio-economic status. In addition, VP children have significantly lower scores compared to their term-born peers in intelligence, processing speed, working memory, inhibition and planning. The results from the domain-specific

measures revealed significant differences between groups in non-symbolic numerical representations, even after controlling for intelligence, correcting for multiple comparisons. Associations between mathematical performance and domain-general and domain-specific skills revealed different patterns between groups. Interestingly, only the VP group showed significant positive associations between numerical magnitude comparisons and maths performance. Domain-general skills (intelligence, working memory, processing speed and attention) were also closely linked to mathematical scores in the preterm group. Our results replicate previous finds showing that VP infants have difficulties with non-symbolic numerical representation, but that their mathematical performance is linked to their executive function skills.

6.2 Introduction

In spite of the high rates of difficulties in mathematical attainment faced by children born prematurely (Aylward, 2005; Hack et al., 2005; Simms et al., 2013) it remains unclear whether those children have low mathematical achievement due to difficulties in domain-general or domain-specific abilities, or both. Domain-general skills are defined as skills that are relevant for all cognitive learning and, in the case of mathematical performance, include predominately working memory, executive function and visuospatial skills. Domain-specific mathematical skills are described as exclusively relevant for learning mathematics per se. These include basic quantitative skills, such as counting, number fact knowledge and calculation skills; accurate numerical representations, such as digit recognition, and performance in magnitude comparison and number line tasks (Passolunghi & Lanfranchi, 2012). Thus, it is crucial to investigate both domain-general and domain-specific skills in the at-risk population since both contribute to building mathematical proficiency.

An extensive body of research has shown that VP children have deficits in executive function domain-general abilities (e.g., Bayless & Stevenson, 2007; Mulder et al., 2009; Taylor & Clark, 2016) and that these have a negative effect on their maths attainment (e.g., Cornelieke Sandrine Hanan Aarnoudse-Moens, Weisglas-Kuperus, Duivenvoorden, van Goudoever, & Oosterlaan, 2013; Adrian, Haist, & Akshoomoff, 2019; Costa et al., 2018). On the other hand, only a handful of studies have investigated numerical representation in those born preterm, including numerical magnitude comparisons (Hellgren et al., 2013; Simms et al., 2013; Guarini et al., 2014; Simms et al., 2015; Tinelli et al., 2015; Libertus et al., 2017).

Numerical magnitude comparisons refer to our basic ability to decide which of two numerosities is the largest. Sensitivity to numerical magnitudes is a basic domain-specific skill thought to be an especially important mathematical competency, particularly in the first years of schooling (Bugden & Ansari, 2011; De Smedt, Noël, Gilmore, & Ansari, 2013; Lyons et al., 2014). While symbolic magnitude comparison tasks ask which of two Arabic numbers is larger, non-symbolic magnitude comparison tasks request a participant to estimate which set of elements (usually an array of dots), has more quantities, without counting. There is a reliable association between the performance in numerical magnitude comparison and mathematical scores (Schneider et al., 2017), making this a potential tool to identify those at high risk of showing difficulties with maths (Orrantia et al., 2018).

Deficits in non-symbolic magnitude comparison have been shown in children born prematurely (e.g., Helgren et al., 2013; Libertus et al, 2017). It is unclear, however, whether difficulties in numerical representation, measured by non-symbolic magnitude comparison, account for difficulties in mathematical performance. For example, Libertus et al. (2017) claimed that extremely preterm children have imprecise numerical representations leading to difficulties in mathematical performance that cannot be explained only by deficits in domain-general abilities. On the other hand, other studies have suggested that difficulties in mathematical achievement in preterm children are mediated by difficulties in domain-general abilities (Simms et al., 2013; Simms et al., 2015). For instance, Simms et al. (2015) showed that very preterm children did not have poorer performance in tasks assessing non-symbolic numerical magnitude, and, rather, difficulties in mathematical performance were associated with deficits in visuospatial processing and working memory, both domain-general abilities. Thus, the role of numerical magnitude comparison in the development of mathematical performance in individuals born prematurely remains inconclusive.

Differences in methodology employed in previous studies, however, make it difficult to directly compare the results. For example, when testing numerical representations, Hellgren et al. (2013) and Libertus et al. (2017) employed spatially intermixed trials. Guarini et al. (2014) and Simms et al. (2015) chose less visually demanding stimuli, since individuals born prematurely have visuospatial difficulties. It has been suggested, therefore, that the mixed results are due to methodological issues in assessing these basic numerical representations, including the possibility that these tasks may tap into perceptual processing or inhibitory control rather than number processing (De Smedt et al., 2013; Dietrich et al., 2015; Gebuis, Cohen Kadosh, & Gevers, 2016; Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016). Moreover, domain-general abilities have not been comprehensively tested in all studies. Intelligence and working memory are often contemplated, but processing speed, attention, planning and inhibition are less investigated in this population, and all of these are important domains for mathematical attainment. For instance, Libertus et al. (2017) only employed standardised measures of intelligence and working memory. Furthermore, incomplete measures of mathematical attainment in part of their sample, and different number of trials for experimental measures compromised their findings. Likewise, intelligence was the only domain-general assessed by Guarini et al. (2014), although other domain-specific abilities such as number knowledge, were assessed in their study. Simms et al. (2013) investigated several domain-general abilities in a sample of extremely preterm children, although the only measure for domain-specific ability was an estimation task comprising of few items. Simms et al. (2015) carried out the most comprehensive study investigating domain-general and domain-specific skills and their relationships with mathematical attainment in a sample composed by 115 children born prematurely. Both domain-general and domain-specific abilities were cautiously and comprehensively investigated in their study, revealing that very preterm children performed at a similar level as their term-born peers in non-symbolic magnitude comparison tasks, and that difficulties in mathematics were associated with deficits in visuospatial processing and working memory.

Overall, it is unclear whether very preterm children have difficulties processing numerical magnitude comparisons, especially non-symbolic representations. Like most research of this kind, there are obstacles in isolating particular skills or cognitive functions. Choosing the correct stimuli, task design, and methodological procedures are of utmost importance.

Thus, we investigated the relationship of domain-general and a set of domain-specific abilities (numerical magnitude comparison) and mathematical performance in children born very prematurely. We hypothesised that those born preterm would have lower scores compared to their term-born peers in all domain-general skills, in line with previous results. In addition, we hypothesised that if difficulties in mathematical performance are driven solely by deficits in numerical representation, assessed by the performance in numerical magnitude comparison, would be closely linked to mathematical performance, in line with previous results demonstrated by Libertus et al. (2017). Alternately, if poor mathematical proficiency is driven by domain-general skills, we anticipated that a great variability of executive function measures would account for mathematical performance, replicating previous results shown by Simms et al. (2015). We investigated different outcomes for numerical magnitude comparison (accuracy, RT, IES and, *w*. Besides, different criteria for mathematical learning disabilities were explored by using both achievement-based and discrepancy-based criteria.

Thus, the specific aims of this study were to:

1. Compare performance between the VP group and the term group in domain-general, domain-specific skills and educational outcomes, specifically mathematical performance;
2. Explore different criteria for mathematical learning difficulties (discrepancy-based and achievement-based criteria);
3. Investigate different outcomes in numerical magnitude comparisons (accuracy, RT, IES and *w*);
4. Explore associations between domain-general and domain-specific skills and maths performance;
5. Investigate whether mathematical performance is better predicted by domain-general or domain-specific skills.

6.3 Methods

6.3.1 Participants

Data collection was carried out between October 2016 and March 2018. During this period, 38 very preterm children (12 VP, 11 males and 26 EP, 8 males) and 30 term-born children (20 males) joined the study.

Table 6.1 summarises the general characteristics of the study participants. Independent t-test for age of evaluation, chi-square analysis of gender, bilingualism, hand preference, maternal education and socio-economic status revealed that preterm and full-term samples did not differ significantly in respect to the mentioned variables. Gender and handedness were reported separately as ratio data. Table 6.2 shows the neonatal morbidities of the VP children. Approximately a fifth of the participants of the preterm groupware identified to have neonatal brain injuries.

Table 0.1: General characteristics of study participants.

	Term n=30	**Preterm** n=38
Gestational age (weeks), mean (SD)	-	26.74 (1.8)
Range	-	(23^{+3} – 29^{+6})
Birthweight (g), mean (SD)	34872 (438)	876.05 (220)
Range	(2620 – 4310)	(500 – 1400)
Sex (M:F)	20:10	19:19
Age (months), mean (SD)	112.03 (10.50)	110.13 (11.68)
Range	(97 - 129)	(96 – 129)
Bilingual (%)	16.00	39.30
Handedness (Right: Left)	23:7	32:6
Maternal age (years), mean (SD)	43.9 (4.5)	42.6(5)
Range	(32 – 51)	(32 – 55)
Maternal Education		
< BD (%)	15.80	33.00
> BD (%)	84.20	66.00
SES (IMD quintile)		
First (%)	27.20	5.40
Second (%)	13.70	16.20
Third (%)	24.10	24.30
Fourth (%)	20.60	16.20
Fifth (%)	13.70	37.80

BD: Bachelor's degree

Table 0.2: Neonatal morbidities of the preterm children.

	N	Very Preterm (n=12)	N	Extremely Preterm (N= 24)
Multiple births	5	41.00%	5	20.00%
IVH / PVL: I - II				
Grade I	3	25.00%	2	8.00%
Grade II	0	-	3	12.50%
ROP				
Grade I	1	8.30%	0	-
Grade II	3	25.00%	6	25.00%
Grade III	0	-	2	8.30%
CLD/ BPD				
Mild	3	25.00%	8	33.30%
Moderate	2	16.60%	5	20.80%
Severe	3	25.00%	7	25.90%

IVH: intraventricular haemorrhage; PVL: periventricular leukomalacia; ROP: retinopathy of prematurity; CLD: chronic lung disease; BPD: bronchopulmonary dysplasia. Mild: O2 at 28 days not 36 weeks, moderate: O2 36 weeks < 30% low flow; Severe: O2 36w >30% +/or ventilation/CPAP/high flow.

Figure 6.1 illustrates the frequencies of gestational ages composing the preterm group. The total sample was represented by 68.4% of extremely preterm children.

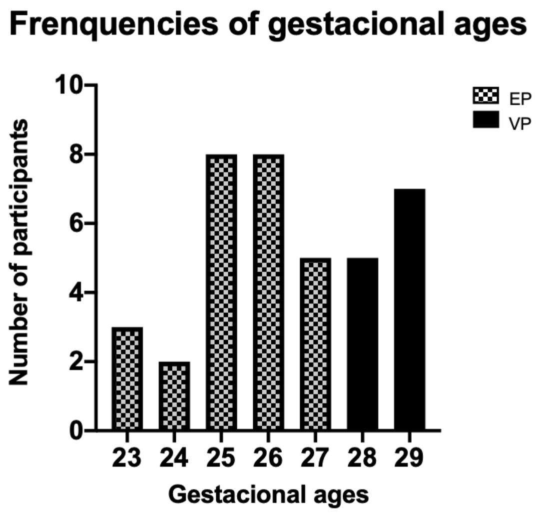


Figure 0.1: Frequency of gestational ages for the preterm group.

6.3.1.1 Neurologic, psychiatric and/or psychological conditions for the preterm group

Table 6.3 summarises formal/ suspected diagnoses in the preterm children according to parents' reports. Formal diagnoses included: two participants diagnosed with dyslexia, three participants were identified having global delay (two of them receiving full-time support at school), two were diagnosed with mild cerebral palsy and one participant identified with developmental coordination disorder. According to parental reports, 28% of the children in the preterm group received help at school (ranging from full time to one hour).

Table 0.3: Formal/ suspected diagnosis according to parent report in the preterm group.

Diagnosis or Suspected	**Number reported in screening**
Global delay/ Intellectual disability	5
Autism/ASD traits	3
Anxiety	2
Attention problems/ ADHD	2
Cerebral Palsy	2
Learning difficulties/ disabilities	5 (maths=2, reading and spelling = 3)
Developmental Coordination Disorder	4

6.3.2 Measures

Domain-general skills and education outcomes were assessed by standardised tests, whereas domain-specific abilities were investigated by experimental measures. Domain-general skills comprised intelligence and executive functions including processing speed, working memory, attention, planning and inhibition. Educational outcomes comprised mathematical performance and basic literacy skills. Domain-specific skills comprised experimental measures of symbolic and non-symbolic numerical magnitude comparison. See section 3.2.9.3.2.

6.3.2.1 Numerical magnitude comparison

To assess domain-specific numerical representation, two experimental tasks were designed: a dot magnitude comparison task and a symbolic magnitude comparison task. The designs of our experimental tasks were based on previous studies to record electrophysiological (EEG) activity during the performance of the task (Heine, Tamm, Wißmann, & Jacobs, 2011). The results of the electrophysiological data will be discussed in chapter 7.

Stimuli were presented as a pair of random-dot arrays and Arabic numerals, respectively. Each trial started with the display of a fixation cross (500-700 ms), followed by a pair of numerical values (max 3000 ms), and then a grey screen (500 ms). Participants were asked to focus on the fixation cross to prevent eye movement. The stimuli presentation and responses (response time and accuracy) were recorded using E-prime 2.0. The tasks required children to decide which of two paired stimuli was the numerically larger one (see figure 6.2). The children were instructed to respond to each trial by pressing one of two target buttons on a response device select either the quantity on the left or the right as the larger of the two numbers. For each stimulus type (symbolic and non-symbolic), pairs of stimuli were simultaneously presented to the child on a computer screen. Each task had 16 practice trials. Participants had to answer at least twelve correct answers to proceed to the test phase. 96 trials were administered, divided into three blocks. The images appeared on a 17-inch LCD monitor. Participants were placed at a distance of approximately 60 cm from the screen. The presentation of tasks was counterbalanced among participants.

6.3.2.1.1 Stimuli

6.3.2.1.1.1 Non-symbolic

A total of 96 images of paired dot arrays were used as a stimulus with 48 shown in each numerical distance (small and large). We decided to design our tasks based on previous studies using grey and white stimuli that were not visually demanding for our clinical group. The stimuli were shown on a grey background (255, 255, 255, RGB) and dots were white (128, 128, 128, RGB). The images containing dot arrays varied on individual item size, inter-item spacing, total occupied area and total luminance in order to control for visual parameters that have been shown to influence the perception of numerosity, employing the same parameters described by Piazza et al. (2004). In half of the images, the intensive parameters (individual item size and inter-item spacing) of the dot arrays were confounded with number and the extensive parameters (total occupied area and total luminance) were equated. The remainder of the images were equated on the intensive parameters (individual item size and inter-item spacing) and varied on the extensive parameters. Extensively and intensively controlled images were presented with equal frequency. The sides of paired stimulus containing bigger quantities were counterbalanced across trials, meaning half of the trials having bigger quantities were placed on the left-hand side and the other half of the trials on the right-hand side. Paired images were contained in a fixed space, and the group of dots was presented separately in the same image. Each array was located in one of the two 208 312 px^2 (64,896 px^2) rectangles separated by a monochromatic partition 34 312 px (10,608 px^2), which together formed a larger 450 312 px^2 (126,360 px^2) rectangle. For the dot magnitude comparison task, dot arrays were presented in pairs with a *small numerical distance* of 4: 11-7, 12-8 (ratio 1.5), 13-9, 14-10 (ratio 1.4), 11-15, 12-16 (ratio 1.3) and *large numerical distance*: 14-7 (ratio 2); 15-8 (ratio 1.8), 16-9 (ratio 1.7).

6.3.2.1.1.2 Symbolic

A total of 96 pairs of Arabic numerals were shown in 64-point Arial font. These images were presented in the same colour as the non-symbolic images; that is, a grey background with white colours for numerals. For the symbolic magnitude comparison task, Arabic numbers ranging from 1 to 9 were presented in pairs with a *small*

(1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8, 8-9) or *large* (1-6, 2-7, 3-8, 4-9) numerical distance.

Because numerosity can be directly interfered from the symbolic number notations but only for the numerosities 1-4 for the non-symbolic notation, using the same numerical distance for both tasks would lead to a difference in task difficulty. Thus, larger numerical distances were used in the non-symbolic task compared to with the symbolic tasks, which allowed comparable acuity rates and response time. This follows the procedure employed by Gebuis et al, (2009).

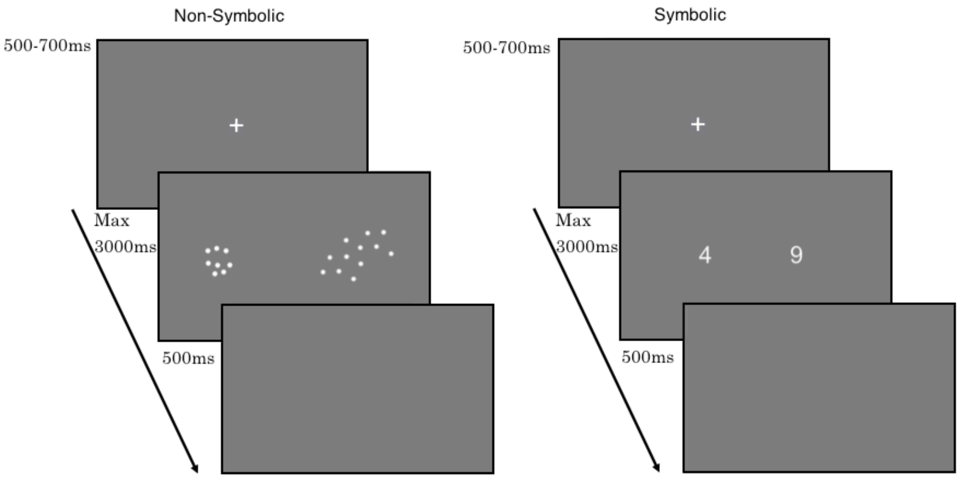


Figure 0.2: Schematic representation of non-symbolic and symbolic magnitude comparison tasks.

Table 6.4 provides a summary of the assessments used in this study, the domain evaluated in each assessment, the type of the scores and the outcomes used for each domain.

Table 0.4: Summary of the assessments used in this study for domain-general, educational attainment and domain-specific outcomes.

Domain	Test	Subdomain	Assessment	Main outcome
Domain-general				
Intelligence	WISC-IV	Verbal ability	Verbal Comprehension (VCI)	Standardised score
	WISC-IV	Non-verbal ability	Perceptual Reasoning (PRI)	Standardised score
	WISC-IV	General ability	Full Scale IQ (FSIQ)	Standardised score
Executive functions	WISC-IV	Processing speed	Processing Speed (PSI)	Standardised score
	WISC-IV	Verbal working memory	Working memory (WMI)	Standardised score
	AMWA	Verbal working memory	Verbal WM	Percentile
	AMWA	Visual working memory	Visual WM	Percentile
	TEA-Ch	Selective attention	Sky Search	Scaled score
	TEA-Ch	Sustained attention	Score!	Scaled score
	TEA-Ch	Attention control	Creature Counting	Scaled score
	TEA-Ch	Dual task	Sky Search DT	Scaled score
	D-KEFS	Inhibition	Colour-Word Interference Test (CWIT)	Scaled score
	D-KEFS	Planning	Tower Test	Scaled score
Educational Attainment	WIAT-II	Mathematics	Numerical Operations + Mathematical Reasoning	Standardised score
	WIAT-II	Mathematics	Numerical Operations (NO)	Standardised score
	WIAT-II	Mathematics	Mathematical Reasoning (MR)	Standardised score
	WIAT-II	Literacy Skill	Spelling	Standardised score
	WIAT-II	Literacy Skill	Word Reading	Standardised score
Domain-specific				
Non-symbolic		Numerical representation	Dot magnitude comparison	RT, CRs, w, IES
Symbolic		Symbolic representation	Symbolic magnitude comparison	RT, CRs, IES

RT= reaction time; CRs= correct answers; w= Weber fraction, IES = Inverse Efficiency Score

6.4 Data analysis

6.4.1 Domain-general skills and educational outcomes

Independent t-tests were used to compare performance between preterm and control children on all domain-general skills and educational outcomes for standardised measures. Cohen's d was calculated to determine standardised effect sizes across tests. Effect sizes were defined as small (0.2–0.3), medium (0.3–0.5), or large (>0.5). In addition, since the control group performed significantly above the standard population mean in many of the cognitive measures, indicating a generally better level of cognitive functioning compared to the standard population, ANCOVAs were carried out to control for the effect of full IQ on the previous analyses, with full IQ as a covariate. One sample t-test were also carried out comparing the term and the preterm group to the standard population mean, respectively (appendix 6.3 and 6.4). Given that socioeconomic status (SES) was marginally different between groups (40% of the preterm participants were in the most deprived quintile in comparison to 13% of the term), ANCOVAs controlling for SES were also carried out.

6.4.2 Mathematical performance

Discrepancy-based and achievement-based criteria for mathematical performance were also investigated in both groups. For the discrepancy-based criterion, discrepancies were calculated by subtracting the standard IQ scores from standard scores obtained from the WIAT-II Mathematics. For the achievement-based criterion, percentiles were used and individuals with percentiles equal or inferior than 25 were identified.

6.4.3 Domain-specific skills

We tested a set of domain-specific skills using symbolic and non-symbolic magnitude comparison tasks. Children's accuracy as a percentage of correct test trials, and their average reaction time based on correct responses for each condition (i.e., task and distance) were measured according to each group. Only trials in which children responded correctly within 200 and 2500 ms were analysed, in line with the same criterion used by Szucs & Soltesz, (2007). This means that 1.55% of the correct responses were excluded from the analysis, similar to the rates found by Guarini et al. (2014). No participant had a level of accuracy (correct responses) equal to or below chance (50%) across all conditions. One dataset from the symbolic magnitude comparison task in the term group and one dataset from both tasks in the preterm group were not included in the sample due to equipment failure. The final sample was composed of 37 preterm children and 30 full-term children.

An arcsine transformation was carried out to normalise the distribution of the correct answers, in line with the analysis used by Guarini et al. (2014). Thus, all analyses performed for correct answers were based on the values from the arcsine transformation but, for practicality, all original values were reported (mean and SDs).

Accuracy and reaction time were the main outcomes, but additional measures were calculated. Firstly, *Weber fractions (w)* were calculated based on accuracy of the answers for the non-symbolic magnitude task to determine each child's *w*. In line with the model proposed by Halberda et al. (2008), we employed the same model to calculate *w* in our sample. Thus, each child's responses over all 96 trials were fit in the model. The free parameter *w* (i.e., the Weber fraction) is a scaling parameter for Gaussian distributions centred on integers on the mental number line. The best-fitting *w* parameter was found via Nelder-Mead optimisation using custom python code (https://github.com/bobaseb/webers_fraction, last accessed on 21st May 2019).

Inverse efficiency scores (IES) were also calculated to control for potential speed-accuracy trade-off in both tasks. This has previously been used as a measure of ANS acuity (Sanguine et al., 2012a; Sanguine et al., 2012b; Bartelet et al., 2014). This score can be calculated by dividing the mean RT of the correct responses by the proportion of correct responses (Bruyer and Brysbaert, 2011). The efficiency score can be interpreted in the same way as *w* (i.e., the smaller the efficiency score, the higher ANS acuity).

Finally, independent t-tests were used to compare performance between preterm and control children on all experimental measures, and Cohen's d was calculated to determine standardised effect sizes across tests. Effect sizes were defined as small (0.2–0.3), medium (0.3–0.5), or large (>0.5).

6.4.4 Associations between domain-general and domain-specific skills and maths performance

We wanted to explore the associations between mathematical performance and domain-general and domain specific skills within each group separately. Thus, Pearson's correlations between standardised mathematical scores (Mathematics, Numerical Operation and Mathematical Reasoning) and domain-general and domain-specific abilities were conducted for preterm and term group separately. Correlations above 0.4 were considered strong; correlations between 0.2 and 0.4 were considered moderate, and those below 0.2 were considered weak.

6.4.5 Predictors for mathematical achievement

Finally, we wanted to investigate whether performance in maths was predicted by domain-general or domain-specific skills. Since a large number of measurements were available for both domain-general and domain-specific abilities, Principal Component Analysis (PCA) was applied to produce a small number of derived variables.

Principal components analysis is a way of doing a linear transformation on a set of variables. The minimum number of variables is two though it is normally applied on larger sets of variables. For example, Ahlheim and Love (2018) applied singular value decomposition (SVD) to fMRI data with sixteen dimensions or less. PCA is mathematically equivalent to mean-centred SVD, a common linear algebra operation. As such, PCA describes a change in coordinates of the data based on the variance of the principal components, thus the name. Specifically, it rotates and stretches the data to find the new coordinates with principal components as the new axes. When the number of variables is much higher than the number of observations, PCA can serve as a dimensionality reduction technique to improve power by focusing on only a handful of principal components as opposed to all of them (i.e., the number of components is equal to the number of variables in the dataset).

For the data discussed here, we can see that we cannot do inferential statistics on all variables due to the small sample size, so we can also use PCA here to our advantage. If we assume that there are two groups with differences in our data (term and preterm as two different clusters), and if the difference is relatively significant, then there will be a principal component describing that difference. This justifies the use of the technique to use components as predictors in a regression analysis.

In our study, PCA was an alternative approach that allowed us to carry out linear regression with only a few components extracted from PCA. However, this may also demand a cautious interpretation of the data, given that PCA works best for homogenous groups, bigger numbers of observations and larger sample size. It can obscure interpretation of the analyses and in some cases skew differences between heterogeneous groups, especially when applying linear transformations to high dimensional data (i.e., large number of independent variables) (Jolliffe & Cadima, 2016).

The varimax rotated method was employed. The derived variables were then used in place of the larger number of original variables to simplify the subsequent analysis of the data. Standardised scores were transformed to Z-scores according to group. Subsequently, two separate PCAs were carried out to generate components of domain-general variables (Verbal Comprehension, Perceptual Reasoning, Processing Speed, Working Memory, Sky Search, Score!, Sky Search DT, Colour-Word Interference Test and Tower Test) and domain-specific variables (correct answers, response times for the symbolic and non-symbolic tasks according to the distance and w). Scores from the Automatic Working Memory (verbal and visual working memory) and Creature Counting from Tea-Ch tests (attentional control) were not included due to the amount of missing values. Missing data for the variables included in the PCA were replaced by the mean of each variable according to group (see appendix 6.1). The Kaiser-Meyer-Olkin measure was used to check whether sampling adequacy was above 0.5. Bartlett's Test of Sphericity rejected the null hypothesis that the correlation matrix of the original variables is an identity matrix. Components were selected based on visual inspection of the inflexion points for each scree plot (i.e., a plot ordering components by their eigenvalue). Eigenvalues relate to the amount of variance that each component explains and selection by visual inspection is standard practice in many fields (see appendix 6.2). To foreshadow the selection, the inflexion point was observed at the second component for the domain-general PCA and the third component for the domain-specific PCA. To keep selection consistent between the domain-general and domain-specific PCAs, we selected the top three components from each PCA. Pearson's correlations were carried out between derived variables and original variables in order to explain the variance of the components. Subsequently, linear regressions were conducted to identify predictors of mathematics attainment using the components derived from the PCA (mathematics as the dependent variable and the derived components and group as independent variables).

6.5 Results

6.5.1 Domain-general skills and educational outcomes

Table 6.5 provides descriptive statistics for all domain-general abilities and educational outcomes. Independent t-tests examining differences between preterm and term groups were carried out for all standardised measures. To avoid false-positive results, since a generally increased performance was found in the term group, the term group was compared against the standard population mean. Thus, one sample t-tests against the standard population mean were conducted to verify the performance of the term group. Standard population means were considered according to the outcome measures; hence the standard population mean for standard scores was considered to be 100 (measures from the WISC-IV and WIAT-II), scale score 10 (measures from the Tea-Ch and D-KEFS) and percentile 50 (measures from AWMA). The results of the one sample t-tests for all standardised assessments can be found in appendixes 6.3 and 6.4 for the preterm and term group, respectively. Appendix 6.5 indicates descriptive statistics for all standardised assessments examining differences between preterm and term groups, excluding children with IQ equal to or below 85.

After Bonferroni corrections for multiple comparisons (α=0.002), VP children significantly had lower scores compared to full-term children in cognitive measures, including verbal comprehension (VCI), perceptual reasoning (PRI), verbal working memory (WMI), processing speed (PSI) from WISC-IV; inhibition (CWIT) and planning (Tower Test) from DKEFS, and visual working memory from AWMA. All mentioned measures had large effect sizes for between-group differences. When controlling for IQ and SES, only inhibition (CWIT) and planning (Tower Test) remained significantly different between groups. Regarding educational outcomes, after Bonferroni corrections for multiple comparisons, and controlling for IQ and SES, differences between groups were evident for all measures related to mathematical performance, including numerical operations and mathematics reasoning (from the WIAT-II), in contrast no differences were observed in literacy skills after adjustment.

A schematic representation of the results from the independent t-tests between both groups (term *vs* preterm), and one sample t-tests between the preterm group and the standard population mean is illustrated in table 6.6. A plus sign signals a better performance, whereas a minus sign signals a worse performance. When a level of significance was found (defined as p=0.05), it is signalled with ✓. Results that remained significant after Bonferroni corrections (α = 0.002) are illustrated with ✓✓. Effect sizes are reported as bands, where a small effect size is reported as 1, moderate as 2 and large as 3.

It became evident that, overall, the preterm group had lower scores compared to both the term population and the standard population mean in all cognitive and educational measures, although those differences were not always statistically significant. The verbal comprehension index was the only cognitive measure in which the VP preterm children had better performance when compared to the standard population mean, although differences were not significant. Significant differences, after Bonferroni corrections, between the preterm group and the standard population mean were observed in planning (Tower Test from DKEFS), selective attention (Sky Search from Tea-Ch) and dual task (Sky Search-DT), with large effect sizes, signalling an area of difficulty in the preterm population represented by this sample. Figure 6.3 illustrates the performance of the preterm and the term group in measures of intelligence from the WISC-IV. Since we employed two measures for testing verbal working memory (WISC-IV and AWMA), we carried out Pearson's correlations between them. Outcomes measures from WISC-IV and AWMA assessing verbal working memory were found to be moderately positively correlated ($r(48) = 0.363$, p =0.001). Figure 6.4 illustrates the performance of both groups in working memory performance from AWMA, and a scatterplot showing the correlation between the results obtained regarding verbal working memory from WISC-IV and AWMA. Figure 6.5 illustrates the performance of the preterm and the term group in measures of attention from the Tea-Ch. It also illustrates the performance from DKFES in respect to inhibition and planning. Figure 6.6 demonstrates children's performance in maths and literacy skills from the WIAIT-II.

Table 0.5: Children's performance on domain-general abilities and educational attainment (before and after Bonferroni corrections).

	Term			Preterm			Difference between control and VP children		
Test	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
WISC-IV									
VCI	30	115.77	10.19	37	104.46	15.93	11.30 (4.59 to 18.016)	0.001[a,c]	0.8
PRI	30	111.47	15.40	38	99.00	13.58	12.46 (5.43 to 19.49)	0.001[a,c]	0.8
WMI	30	104.93	11.24	38	94.85	13.20	10.11 (4.08 to 16.15)	0.001[a,c]	0.8
PSI	30	106.40	17.89	38	93.79	13.70	12.61 (4.92 to 20.29)	0.002 [a,c]	0.7
Full Scale	30	114.80	12.41	37	98.86	15.97	15.93 (8.82 to 23.047)	<0.001[a,c]	1.1
Tea-Ch									
Sky Search	30	9.60	3.08	38	7.89	3.20	1.70 (0.169 to 3.241)	0.030	0.5
Score!	30	10.17	3.31	38	8.79	3.61	1.37 (-0.321 to 3.075)	0.110	0.3
Creature Counting	28	10.36	3.18	25	9.12	2.63	1.23 (-0.387 to 2.862)	0.132	0.4
Sky Search DT	30	7.47	3.34	36	6.06	3.77	1.41 (-0.360 to 3.182)	0.116*	0.3
D-KEFS									
CWIT	29	11.34	2.22	35	8.29	3.15	3.05 (1.666 to 4.452)	<0.001[a,b,c]	1.1
Tower Test	30	11.10	2.42	38	8.71	2.16	2.38 (1.275 to 3.504)	<0.001[a,b,c]	1.0
AWMA									
Verbal WM	16	64.31	29.30	32	47.31	30.62	17.00 (-1.61 to 35.61)	0.072	0.5
Visual WM	16	78.44	21.24	32	45.83	32.06	32.61 (16.93 to 48.28)	<0.001[a,c]	1.1
WIAT-II									
Mathematics	30	124.03	17.64	38	99.24	19.37	24.79 (15.74 to 33.84)	<0.001[a,b,c]	1.3
Numerical Operations	30	124.70	15.68	38	103.76	18.06	20.93 (12.61 to 29.25)	<0.001[a,b,c]	1.2
Mathematical Reasoning	30	116.03	15.12	38	95.21	17.88	20.82 (12.66 to 28.97)	<0.001[a,b,c]	1.2
Word Reading	10	116.60	5.77	33	103.82	14.00	12.78 93.54 to 22.01)	0.008[c]	1.2
Spelling	9	117.22	11.37	31	101.19	14.06	16.02 (5.64 to 26.41)	0.003[c]	1.1

[a] Remains significant after applying Bonferroni correction ($\alpha = 0.002$). [b] Remains significant after controlling for full composite IQ (ANCOVA) and [c]Remains significant after controlling for SES (ANCOVA). *Sky Search DT was not normality distributed. Mann-Whitney revealed no differences between groups ($U = 418.500$, $p=0.116$)

Table 0.6: Schematic representation of the results from domain-general and educational attainment

Test	Preterm *vs* Term	*p*	Effect size	Preterm *vs* Standard mean	*p*
WISC-IV					
Verbal Comprehension	-	✓✓	3	+	
Perceptual Reasoning	-	✓✓	3	-	
Working Memory	-	✓✓	3	-	✓
Processing Speed	-	✓✓	3	-	✓
Full Scale	-	✓✓	3	-	
Tea-Ch					
Sky Search	-	✓	2	-	✓✓
Score!	-		2	-	✓
Creature Counting	-		2	-	
Sky Search DT	-		2	-	✓✓
D-KEFS					
Colour-Word Interference Test	-	✓✓	3	-	✓
Tower Test	-	✓✓	3	-	✓✓
AWMA					
Verbal WM	-		2	-	
Visual WM	-	✓✓	3	-	
WIAT-II					
Mathematics	-	✓✓	3	-	
Numerical Operations	-	✓✓	3	+	
Mathematical Reasoning	-	✓✓	3	-	
Word Reading	-	✓	3	+	
Spelling	-	✓	3	+	

+ = better performance; - worst performance. ✓= p=0.05 as found, ✓✓ remained significant after Bonferroni corrections (α = 0.002). Effect sizes were reported as 1,2 and 3, representing small, moderate and large, respectively.

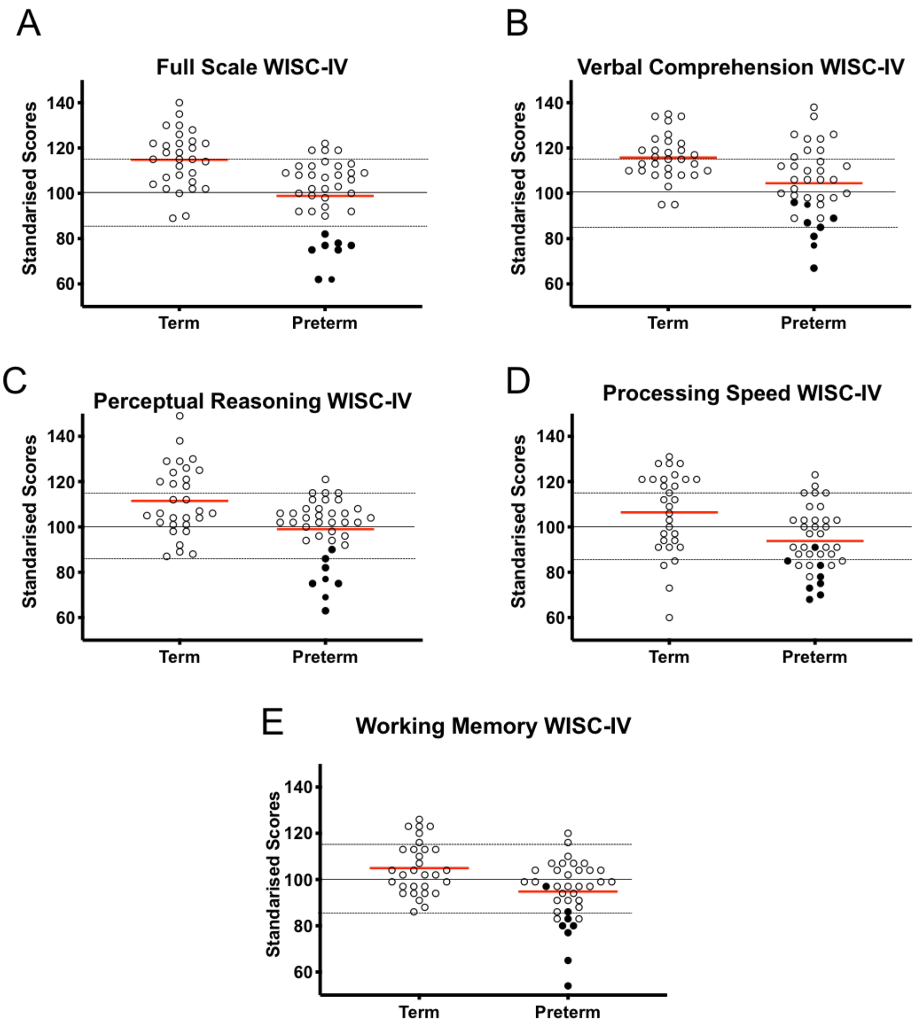


Figure 0.3: Scatter dot plots showing the performance of the preterm and full-term groups on intelligence

Graph A: full scale. Graph B: verbal comprehension (VCI). Graph C: perceptual reasoning (PRI). Graph D: processing speed. Graph E: working memory (verbal working memory from WISC-IV). VP children significantly had lower scores compared to their term-born peers in all cognitive measures. Solid lines indicate the standard population mean (100 for composite scores from the WISC-IV). Dotted lines indicate one standard deviation from the mean. Red lines represent the group mean. Participants with IQ < 85 are identified in black.

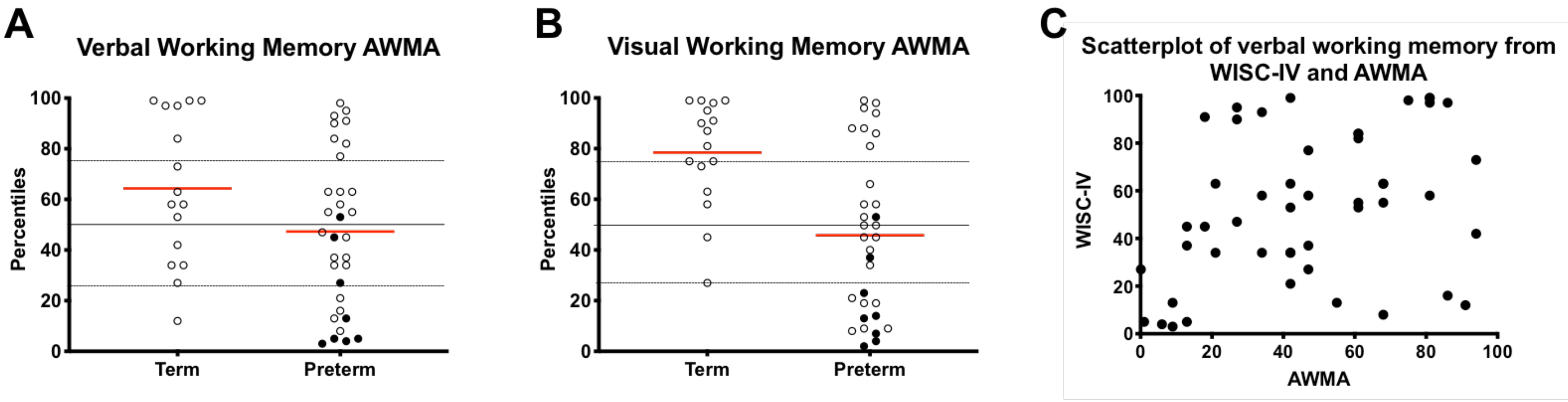


Figure 0.4: Scatter dot plots showing the performance of the preterm and full-term groups on working memory

Graph A: verbal working memory. Graph B: visual worming memory (from AWMA). Overall, VP children significantly had lower scores compared to their term-born peers in visual working memory. Solid lines indicate the standard population mean (50 for percentile scores from the AWMA). Dotted lines indicate one standard deviation from the mean. Red lines represent the group mean. Participants with IQ < 85 are identified in black. Graph C: Scatterplot showing correlation between results obtained from the WISC-IV and AWMA in respect to verbal working memory. Pearson correlation revealed a moderate correlation between the two measures.

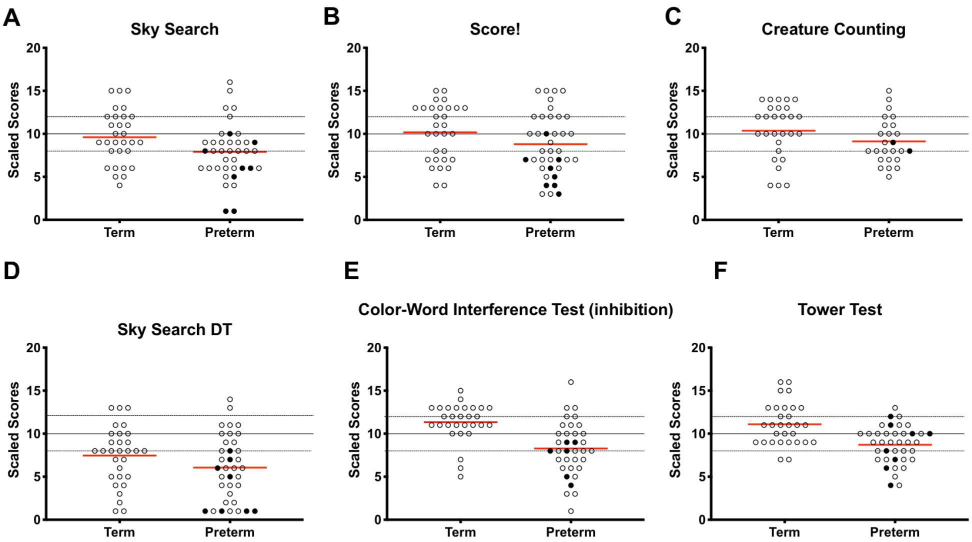


Figure 0.5: Scatter dot plots showing the performance of the preterm and full-term groups on attention from the Tea-Ch and inhibition and planning from DKFES.

Tea-Ch: Graph A: Sky Search, for selective attention. Graph B: Score!, for sustained attention. Graph C: Creature Counting, for attention control. Graph D: Sky Search DT, for dual task).

DKFES: Graph E: Colour Word Interference Test, for inhibition. Graph F: Tower Test, for planning.

Overall, VP children significantly had lower scores compared to their term-born peers in measures of inhibition and planning. Solid lines indicate the standard population mean (100 for composite scores from the WISC-IV and 50 for percentile scores from the AWMA). Dotted lines indicate one standard deviation from the mean. Red lines represent the group mean. Participants with IQ < 85 are identified in black.

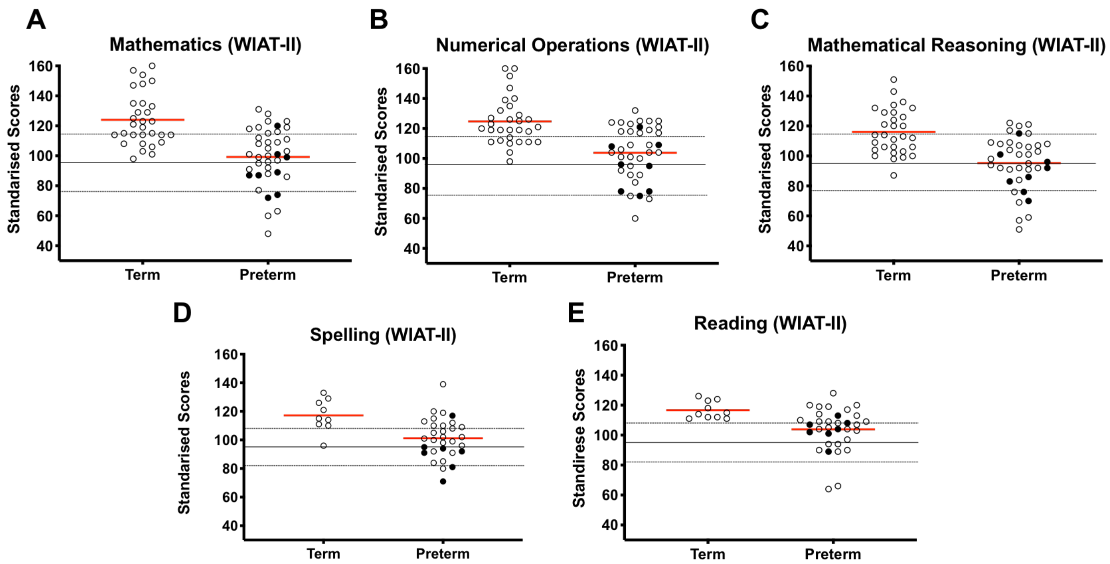


Figure 0.6: Scatter dot plots showing the performance of the preterm and full-term groups on math and literacy skills from the WIAIT-II

Graph A: Mathematics. Graph B: Numerical Operations. Graph C: Mathematical Reasoning. Graph D: Spelling. Graph E: Reading. Solid lines indicate the standard population mean (100 for composite scores from the WISC-IV and 50 for percentile scores from the AWMA. Dotted lines indicate one standard deviation from the mean. Red lines represent the group mean. Participants with IQ < 85 are identified in black. Overall, VP children significantly had lower scores compared to their term-born peers in all educational measures, but only differences in maths outcomes survived Bonferroni correction for multiple comparisons and analysis controlling for IQ and SES.

6.5.2 Mathematical performance: discrepancy-based and achievement-based criteria

Next, we wanted to investigate different criteria for identifying children with mathematical learning difficulties. We employed discrepancy-based and achievement-based criteria. The discrepancy-based criterion is based on the marked difference between IQ and academic attainment (Simms et al., 2013a). The achievement-based criterion considers the performance in standardised achievement tests, regardless of IQ (for further discussion, please see section 2.8.1). Using the discrepancy-based criterion, only three children from the preterm group were identified with mathematical learning difficulties and one child from the term group. In contrast, when applying the achievement-based criterion, ten preterm children were identified with mathematical learning difficulties. Scores from the discrepancy-based criterion significantly differed between groups ($t(65)=-2.837$, $p=0.006$), with the term group performing approximately 10 points above what would be expected from their IQ scores in mathematical achievement (Mean= -9.23, SD= 12.48), whereas the preterm group performed in accordance with expectations (Mean=-0.68, SD=12.11). See figure 6.7.

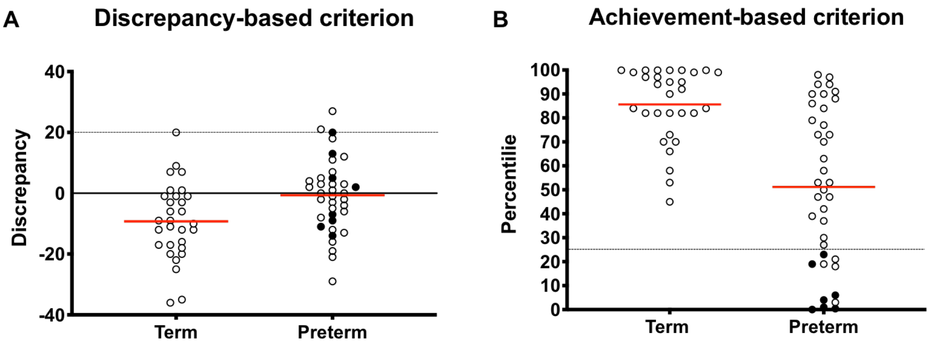


Figure 0.7: Scatter dot plots illustrating the performance for mathematical achievement of the preterm and full-term groups

A) discrepancy-based criterion and B) achievement-based criterion. Dotted lines indicate cut-offs for mathematical learning difficulties (2 SDs from the mean). Red lines represent the group mean. Participants with IQ < 85 are identified in black. Most of the preterm children with low IQ were below the cut-off for mathematical achievement when the achievement-based criterion was used, whereas the discrepancy-based criterion only identified one child, suggesting that the achievement-based criterion is more appropriated for the preterm population in our sample.

6.5.3 Domain-Specific Skills

Independent t-tests were carried out to explore differences between groups in domain-specific outcomes assessed by symbolic and non-symbolic numerical magnitude comparisons. The effect of differences in conditions (small and large distances) on accuracy, RTs, and inverse efficiency scores were explored. For the non-symbolic task, we also explored differences between groups in *w*. After Bonferroni corrections for multiple comparisons, significant differences between groups were observed for accuracy for non-symbolic representations, regardless of the condition. Significant differences were observed in small distance (t(65)=5.461, *p*=<0.001, d=1.3) and large distance (t(65)=3.751, *p*=<0.001, d=1.8). Significant differences persisted after controlling for IQ (F(1,64)=12.607, *p*=0.001) and SES F(1,64)=23.283, *p*=<0.001), with large effect sizes. The preterm group had significantly higher values for *w* (t(65)= -2.896, *p*=<0.005, d=1.05), implying a worse performance than the term group. Similarly, poorer performance was also observed in the inverse efficiency score for the non-symbolic task (t(65)=-3.021, *p*=<0.004, d=1.06), but differences did not persist after applying Bonferroni corrections. Mean RTs (SD), CRs (SD), efficiency scores (SD) and *w* (for non-symbolic task) according to group, task and distance are displayed in table 6.7. Figure 6.8 demonstrates children's performance in numerical magnitude comparisons (accuracy and reaction time). Figure 6.9 illustrates IES scores for symbolic and non-symbolic tasks. Figure 6.10 shows w scores for both groups.

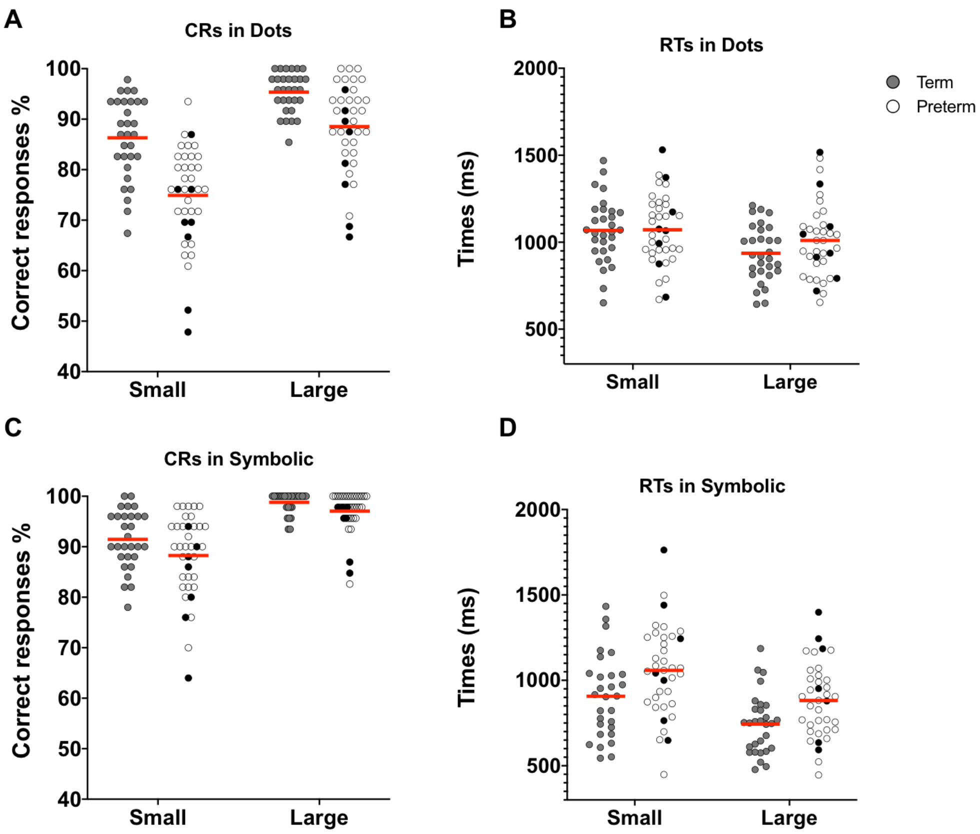


Figure 0.8: Children's performance on numerical magnitude comparisons

Graphs A and B show results from the non-symbolic task for correct responses and reaction time, respectively. Graphs C and D show results for the symbolic tests, for correct responses and reaction time, respectively. Red lines represent the group mean. Participants with IQ < 85 are identified in black. The preterm group is represented by white dots and the term group is represented by light grey dots.

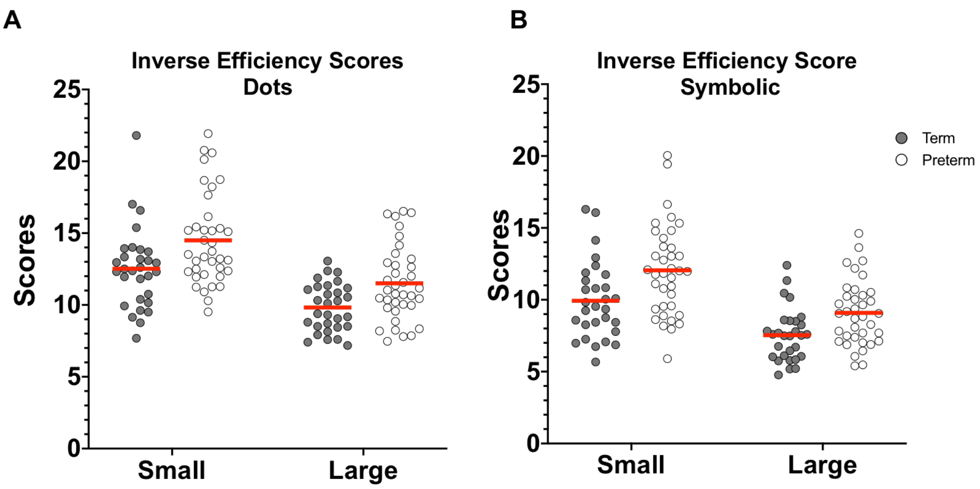


Figure 0.9: IES scores for non-symbolic (A) and symbolic (B) magnitude tasks.

Red lines represent the group mean. Participants with IQ < 85 are identified in black. The preterm group is represented by white dots and the term group is represented by light grey dots.

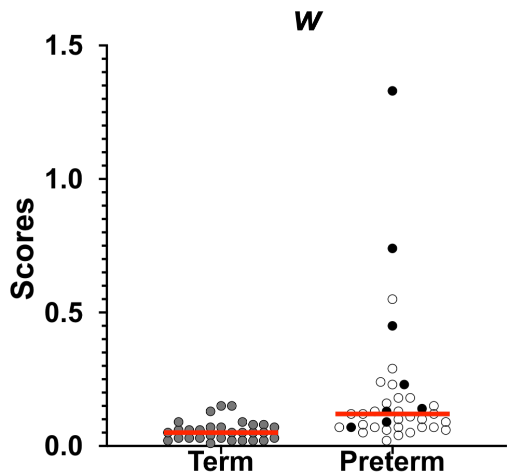


Figure 0.10: w scores.

Red lines represent the group mean. Participants with IQ < 85 are identified in black.

Table 0.7: Children's performance in domain-specific tasks in respect to RTs, CRs, w and IES.

			Term			**Preterm**		**Difference between control and VP children**		
Task	Distance	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
Dots										
CRs (%)	Large	30	95.30	3.90	37	88.51	8.73	6.83 (3.39 t 10.27)	<0.001[a,c]	0.9
RTs (ms)		30	936.42	160.03	37	1010.02	213.02	-73.59 -167.41 to 20.21)	0.122	0.3
CRs (%)	Small	30	86.30	7.97	37	74.89	9.71	11.41 (7.00 to 15.81)	<0.001[a,b,c]	1.3
RTs (ms)		30	1067.83	184.26	37	1071.44	200.47	-3.60 (-98.50 to91.29)	0.940	<0.1
w		30	0.05	0.03	37	0.18	0.242	-0.12 (-0.21 to -0.40)	*0.005[b,c]	0.7
IES	Large	30	9.82	1.65	37	11.50	2.59	-1.68 (-2.77 to -0.59)	0.003 [a,c]	0.7
	Small	30	12.52	2.80	37	14.49	3.16	-1.97 (-3.45 to -0.49)	0.010	0.6
Symbolic										
CRs (%)	Large	29	98.80	2.06	37	97.06	4.20	1.73 (0.037 to 3.43)	0.230	0.5
RTs (ms)		29	743.81	176.03	37	881.87	215.18	-138.05 (-236.65 to -39.45)	0.007	0.7
CRs (%)	Small	29	91.40	5.7	37	88.27	8.154	3.17 (-0.395 to 6.751)	0.074	0.4
RTs (ms)		29	906.81	240.98	37	1058.34	262.90	-151.53 (-277.15 to -25.91)	0.019	0.6
IES	Large	29	7.53	1.84	37	9.09	2.25	-1.55 (-2.58 to -0.51)	0.004	0.7
	Small	29	9.92	2.67	37	12.05	3.15	-2.12 (-3.58 to -0.66)	0.005	0.7

CRs= Correct responses; RTs= Reaction time; *w* = Weber fraction; IES = Inverse Efficacy Score. [a]Remains significant after applying Bonferroni correction ($\alpha = 0.003$). [b]Remains significant after controlling for full composite IQ (ANCOVA). [c]Remains significant after controlling for SES (ANCOVA). *Variable not normality distributed. Mann-Whitney confirmed differences between groups for *w* (U = 186.000, p=<0.001), and efficiency scores for dots (U = 342.000, p=0.007).

6.5.4 Associations between domain-general and domain-specific and mathematical performance

Next, we investigated associations between mathematical performance and domain-general and domain-specific skills for each group, separately. This revealed different patterns between groups. Firstly, for domain-general skills, in the VP group, Pearson's correlation revealed significant correlations between maths scores (WIAT-II), visual working memory (AWMA), dual task (Sky Search DT- Tea-Ch) and sustained attention (Score! – Tea-Ch). Overall, IQ (r=0.786, p=<0.001), verbal working memory (r=0.771, p=<0.001), processing speed (r=0.580, p=<0.001), visual working memory (r=0.550, p=0.001) attention (dual task, r=0.483, p=0.003; and sustained attention, r=0.407, p=0.011) were the domain-general skills with the highest correlations for the VP group.

The term group had significant associations between maths scores and inhibition (Colour-Word Interference Test – inhibition – D-KEFS) and attentional control (Creature Counting – Tea-Ch). IQ (r=0.699, p=<0.001), verbal working memory (r=0.583, p=0.018) and processing speed (r=0.488, p=0.006) were the domain-general skills with the strongest correlations for mathematics for the term group.

Regarding domain-specific abilities, in the VP group, correct answers (CRs) for both the non-symbolic (r=0.455, p=0.005) and symbolic tasks (r=0.332, p=0.040) and w (r=-0.382, p=0.050) were significantly associated with mathematical scores. Reaction time (r=-0.384, p=0.040), and inverse efficiency scores (r=-0.380, p=0.042) in the symbolic task were the only measures associated with mathematical scores for the term group. Figure 6.11 shows correlational matrix for the domain-general and domain-specific skills for the term and the VP groups.

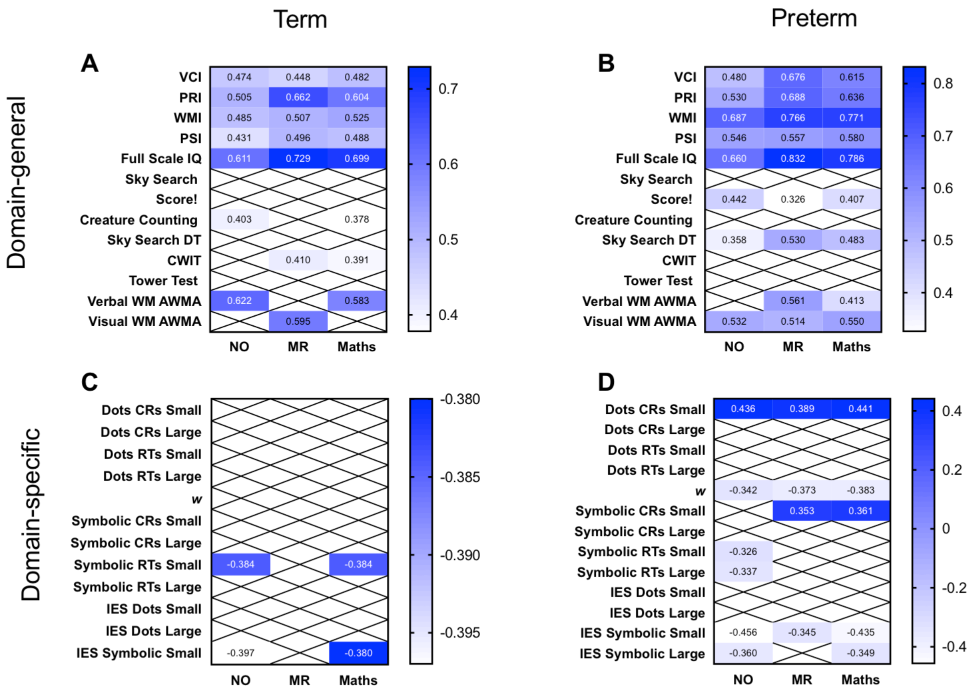


Figure 0.11: Pearson's Correlations between domain-general, domain-specific and mathematical performance

Correlational matrix illustrating first order correlations between domain-general (top row) and domain-specific skills (bottom row) and mathematical scores for the term (left-hand side) and preterm group (right-hand side). X represent correlations without significant values (p=0.05). NO stands for Numerical Operations; MR stands for Mathematical Reasoning. Light colours represent either low or negative correlations, whereas bright blue represent higher correlations.

6.5.5 Predictors for mathematical performance

6.5.5.1 Principal Component Analysis

Finally, we wanted to investigate which domain could better predict mathematical performance. Thus, Principal Component Analysis identified components that explained the variance in the measures employed in this study. PCA was performed in the matrix containing the domain-general and domain-specific measures separately. The first three components were picked based on visual inspection of the inflection point in the scree plot for each domain (see appendix 6.2). For domain-general skills, the amount of variance explained by each of the top three components was 36.9%, 14.4% and 13.1%, respectively (cumulatively explaining 64.50% of the variance of the original variables). For domain-specific skills, the amount of variance explained by each of the top three components was 35%, 30.6% and 11.7%, respectively (cumulatively explaining 77.3% of the variance of the original variables). Pearson's correlation was carried out between the derived components and the z-scores of the original measures to explain the three components selected from the PCA for each domain. Table 6.8 illustrates the correlations of the derived components and the original measures for domain-general abilities, and table 6.9 illustrates the correlations of the derived components and the original measures for domain-specific abilities.

Table 0.8: Pearson's correlations between z-scores from original measures and components derived from PCA for domain-general abilities.

Domain-General			
	Comp.1	Comp.2	Comp.3
VCI	**0.809***	-0.029	0.063
PRI	**0.665***	**0.327**	0.140
WMI	**0.764***	0.139	0.152
PSI	**0.397***	**0.563***	**0.327**
Sky Search	-0.069	**0.852***	0.225
Score!	**0.381***	**0.591***	**-0.391***
Sky search DT	**0.373***	0.069	**0.634***
CWIT	0.205	0.040	**0.392***
Tower Test	-0.068	0.095	**0.650***

Comp = Component. *Remains significant after applying Bonferroni correction (α = 0.005).

Table 0.9 Pearson's correlations between z-scores from original measures and components derived from PCA for domain-specific abilities

Domain-Specific			
	Comp. 1	Comp.2	Comp. 3
Dots CRs Small	-0.090	**0.738***	0.038
Dots CRs Large	0.007	**0.671***	**0.298**
Dots RT Small	**0.836***	0.213	0.073
Dots RT Large	**0.858***	0.096	0.063
w	-0.102	-0.670*	**-0.244**
Symbolic CRs Small	**0.260**	**0.238**	**0.737***
Symbolic CRs Large	-0.093	0.149	**0.834***
Symbolic RT Small	**0.830***	-0.146	0.040
Symbolic RT Large	**0.816***	-0.199	-0.031

Comp=Component. *Remains significant after applying Bonferroni correction (α = 0.005).

Regarding the correlations between the z-scores from the original measures and the components derived from PCA for domain-general abilities, the first component is mostly correlated with measures derived from the WISC-IV, with strong correlations with VCI (r=0.809, *p*=<0.001), WMI (r=0.764, *p*=<0.001), PRI (r=0.665, *p*=<0.001) and PSI (r=0.397, *p*=0.001). The second derived component is mostly correlated with attention measures, with the strongest correlation coming from Sky Search (r=0.852, *p*=<0.001), followed by Score! (r=0.591, *p*=<0.001) and PSI (r=0.563, *p*=<0.001). Finally, the third component is mostly correlated with the Tower Test (r=0.650, *p*=<0.001), Sky Search DT (r=0.634, *p*=<0.001) and CWIT (r=0.392, *p*=<0.001).

Regarding the correlations for domain-specific abilities, measures of RT explained most of the correlations for the first component, with strong correlations for both tasks (dots and symbolic) and both distances (small and large). Specifically, RT for dots at a large distance (r=0.858, *p*=<0.001), RT for dots at a small distance (r=0.836, *p*=<0.001), RT for symbolic small distance (r=0.830, *p*=<0.001), RT for dots at a large distance (r=0.816, *p*=<0.001). The second component was mostly correlated with accuracy for the dot magnitude comparison task, for both distances (DOTS CRs Small; r=0.738, *p*=<0.001; DOTS CRs Large; r=0.671, *p*=<0.001). Lastly, the third component was mostly correlated with accuracy for the symbolic magnitude comparison task, for both distances (Symbolic CRs Large; r=0.834, *p*=<0.001; Symbolic CRs Small; r=0.737, *p*=<0.001).

6.5.5.2 Predicting mathematical performance: linear regression

Next, linear regression analysis was carried out, with children's mathematical attainment (WIAT-II MS) as the dependent variable and the components derived from PCA for domain-general, domain-specific and group as potential predictors of their mathematical ability. As seen in table 6.10, group and components from domain-general were the only significant predictors for mathematical achievement. The model explains 67% of the variance in WIAIT-MS scores, as shown by the adjusted R-squared.

Table 0.10: Regression model predicting children's math scores

	R^2	Adjusted R^2	B	*p*
Model	0.705	0.671		
Component 1 DG			0.535	0.001
Component 2 DG			0.202	0.503
Component 3 DG			0.155	0.120
Component 1 DS			-0.031	0.683
Component 2 DS			0.066	0.404
Component 3 DS			-0.016	0.836
Group			-0.559	<0.001

DG=Domain-general;
DS= Domain-specific;
B=Standardised coefficients

Since *Group* was a significant predictor for mathematical performance, separate linear regressions were carried out in respect to *Group*. This revealed that while the model only explained 36% of the variance for the term group, it explained 59% for the preterm group (as shown by R-squared in table 6.11). The first component derived from domain-general abilities (verbal and non-verbal abilities, verbal working memory, processing speed, selective and sustained attention) was significant for both groups, whereas the second component derived from domain-general abilities (non-verbal ability, processing speed, selective and sustained attention) was only significant for the VP group. No components extracted from the measures testing domain-specific skills were significant, either for the term or the VP groups.

Table 0.11: Regression model predicting children's math scores

	Term				Preterm			
	R^2	Adjusted R^2	B	*p*	R^2	Adjusted R^2	B	*p*
Model	0.500	0.369			0.662	0.597		
Comp. 1 DG			0.679	0.001			0.682	<0.001
Comp. 2 DG			0.236	0.199			0.312	0.017
Comp. 3 DG			0.221	0.180			0.151	0.212
Comp. 1 DS			0.060	0.751			-0.111	0.337
Comp. 2 DS			-0.014	0.935			0.184	0.131
Comp. 3 DS			0.048	0.766			-0.160	0.219

Comp=Component; B=Standardised coefficients.

6.6 Discussion

Our results showed that the preterm group had lower scores compared to the term group on numerous outcomes. In particular, a 15-point average difference in measures of intelligence were found between the term and the preterm group. Since a general increased performance was observed in the control group, it could be debatable, whether comparing the preterm group to the term group would be valid. Set against that, however, when the preterm group was compared to the standard population mean (using a one sample t-test) the two groups performed similarly. Further, the absolute difference between groups in measures of intelligence is in agreement with previous studies. For example, previous studies have shown that preterm children performed approximately 12 points below their term-born peers (Kerr-Wilson et al., 2012).

Thus, despite the apparent indication that the term group had an inflated performance in several measures, our results suggest the very preterm group also had an elevated performance. As a result, differences between groups remained similar to what previous studies reported and subsequent interpretations should not be disregarded.

6.6.1 Maths difficulties and mathematical learning disabilities

The criterion employed to identify learning difficulties can greatly impact individuals if not employed appropriately. In our study, this could be further examined by comparing the discrepancy-based and achievement-based criteria, identified different individuals according to criterion employed. In fact, it has been suggested that more conventional identification methods, such as low standardised mathematics scores irrespective of IQ (e.g., <25th percentile on a standardised test) may be more appropriate in individuals born prematurely (Simms et al., 2013). Our results support this view. Correctly identifying those struggling at school is crucial in order to target interventions and improve educational outcomes.

6.6.2 Domain-general abilities following preterm birth

As expected, the VP group had lower scores compared to the term group in all domain-general skills. Working memory, processing speed, sustained, dual task, inhibition and planning were the domains where VP children had significantly poorer performances when compared either with the term group and the standard population mean. In fact, difficulties in domain-general executive functions have been extensively reported in the preterm population (e.g., for review see van Houdt, Oosterlaan, van Wassenaer-Leemhuis, van Kaam, & Aarnoudse-Moens, 2019). Deficits in EF are of special interest in assessing outcomes of preterm birth because of their value in predicting academic achievement, specifically mathematics (e.g., Mulder et al., 2009; Taylor et al., 2009; Taylor and Clark, 2016). It is well established that formal mathematics abilities are influenced by domain-general skills, such as working memory, inhibitory control, attention, task flexibility and switching (Bull et al., 2008; LeFevre et al., 2010; Gifford and Rockliffe, 2012; Friso-van den Bos et al., 2013). For example, Rose et al. (2011) reported that working memory scores accounted for the variability in mathematics achievement among children born preterm. Aarnoudse-Moens et al. (2013) demonstrated that deficits in EF were, over and above low IQ, an important predictor for poor mathematics attainment following very preterm birth. Mulder et al. (2011) demonstrated that preterm children had lower scores compared to their term-born peers in inhibition, working memory, verbal fluency, and shifting. In addition, the authors suggested that slow processing speed mediated difficulties in these cognitive domains, except response inhibition. A recent meta-analysis revealed that preterm children performed 0.5 SD lower in working memory and cognitive flexibility skills and 0.4 SD lower on inhibition. Taken together, our results are in line with previous studies showing that VP children have difficulties in several cognitive domains crucial for mathematical achievement.

6.6.3 Numerical representation following preterm birth

Those born preterm significantly had lower scores compared to the term group in terms accuracy in the non-symbolic magnitude comparison task, regardless of the distance, controlling for IQ and excluding participants with low IQ (appendix 6.6). No significant differences were observed in the symbolic magnitude comparison task, implying that (non-symbolic) numerical representation is an area of difficulty in the very preterm group. The results of this study are in line with previous studies reporting difficulties in numerical representations in the very preterm population (Hellgren et al., 2013; Simms et al., 2013; Libertus et al., 2017). For example, similar to the results found by Libertus et al. (2017), we also found that the preterm group showed significantly lower acuity in tasks assessing numerical representation than their full-born peers. The higher *w* scores supported the view that the preterm group have difficulties with numerical representations. Indeed, Libertus et al. (2017) has previously reported higher scores in *w*, in line with the outcome of this study. It has been suggested that accuracy in non-symbolic magnitude comparison and *w* are highly correlated. For example, Inglis and Gilmore (2014) showed that *w* and accuracy scores were strongly associated (R^2 =0.846). In line with this, our results showed a strong association between both measures (r(67)=0.717, p=<0.001), supporting that both measures index numerical representations. Taken together, the outcomes for numerical representation, accuracy and *w*, suggest that the preterm group had difficulties with numerical representations.

It should be noted that there is a major difference between this study and previous investigations claiming imprecise representations in the preterm population. Whereas in this study we included both very and extremely preterm children to increase our sample size, previous studies that reported imprecise numerical representations focused mostly on extremely preterm children (Hellgren et al., 2013; Simms et al., 2013; Libertus et al., 2017). This implies that gestational age might be an important variable associated to the performance of numerical representations. In fact, more than two thirds of the sample in this study comprised extremely preterm children. In order to confirm whether gestational age is a main factor affecting numerical representation, we carried out additional analysis comparing very preterm and extremely preterm to the control group, separately (appendix 6.7 and 6.8). This confirmed that extremely preterm children had poorer performance, whereas difficulties in numerical representation in the very preterm were generally absent. Thus, it is reasonable to suggest the lower the gestational age the higher the prevalence of the difficulties with basic numerical representations.

The finding that extremely preterm children have difficulties in numerical representation should be interpreted cautiously, since the results are based solely on accuracy and w scores and do not take into account reaction time. For example, Guarini et al. (2014) revealed that extremely preterm children were as accurate as the term-born children, but were significantly slower, suggesting difficulties in processing speed rather than numerical representations. Our results point in an opposite direction: preterm group was as fast as their term-born peers, but not as accurate. This might suggest that, in our study, the difficulties related to numerical representation could be explained by difficulties in inhibition rather than processing speed. In fact, significant differences between the VP and term groups, and between the VP group and the standard population mean, were found in respect to inhibition. This suggests that inhibition plays a particularly important role in non-symbolic comparison performance in the VP population.

It has been proposed that interference control skills, an inhibition ability, plays an important role in non-symbolic comparison performance as a result of the way dot stimuli are created (Szucs et al., 2013). Gilmore et al. (2013) suggested that for a participant to respond accurately to a non-symbolic comparison task trial, they must inhibit irrelevant and misleading visual information, such as dot size and convex hull, and respond solely based on numerosity estimations. This would explain why differences between groups in our study were observed in the non-symbolic, but not in the symbolic comparison task. In addition, results from IES supported this view. IES is a measure that accounts for accuracy/reaction time trade-offs. After Bonferroni corrections, no significant differences between groups were observed in IES measures. It seems reasonable to conclude that VP children had poor performance in numerical representation due to difficulties in inhibitory skills, rather than imprecise numerical representations.

6.6.4 Associations between maths performance and domain-general and domain-specific skills following preterm birth

A main goal of this chapter was to investigate the associations between domain-general, domain-specific and maths performance in the preterm group. Different patterns between the VP group and term group became evident when investigating the subdomains of domain-general and domain-specific abilities (see figure 6.8). For instance, visual-working memory and attention (sustained and dual task) were significantly associated with maths performance in the preterm group. In this regard, Xenidou-Dervou et al. (2013) demonstrated that working memory was highly associated with maths achievement beyond the effects of any other domain-specific abilities. Working memory is critical for mathematical performance, and both spatial working memory and visuospatial skills have been shown to be strong predictors of mathematical performance in children born prematurely (Litt et al., 2012; Simms et al., 2015). For instance, Simms et al. (2015) demonstrated that very preterm children had significantly poorer mathematics achievement associated with deficits in visuospatial processing and working memory. In line with previous findings, our results seem to support the contention that visual-working memory has a particular important role associated with maths attainment in the VP population.

Hitherto, associations between numerical representations and maths scores have been reported in a group of EP children. For example, Simms et al. (2013b) showed that EP children's performance in mathematics was associated with their underlying accuracy in numerical representations and was not simply a component of their general cognitive ability. Surprisingly, however, here, only the VP group showed significant associations between numerical magnitude comparison tasks and mathematical performance. Previous studies have reported associations between symbolic and non-symbolic magnitude comparison tasks and mathematical performance in typically developing children, with stronger associations found in symbolic tasks (Castronovo and Göbel 2012, Desoete, Ceulemans et al. 2012, Sasanguie, Gobel et al. 2013, Bartelet, Vaessen et al. 2014, Lyons, Price et al. 2014, Sasanguie, Defever et al. 2014; Fazio et al., 2014). However, previous studies also reported null results between correlational measures between mathematical performance and non-symbolic tasks. For example, Holloway and Ansari (2009) demonstrated that performance in non-symbolic comparison tasks was statistically unrelated to children's mathematics scores. It remains unclear why in our study the control group did not show associations between symbolic representations and maths.

Overall, our results show different patterns between the term and the VP group in respect to associations between domain-general and domain-specific skills and maths performance. Differences between groups showed that while the maths abilities of VP children were linked to attention, visual working memory and numerical magnitude comparison, planning was only linked with maths performance in the term group. This could suggest that VP children might have different cognitive strategies related to maths performance. This could be supported by predicting maths performance with both domain-general and domain-specific skills for each group. This is the focus of the discussion in the next session.

6.6.5 Predicating maths performance following preterm birth

Linear regression using components extracted from principal component analysis confirmed different patterns between groups for predicting attainment in mathematics. Intelligence, working memory, processing speed, sustained attention and dual task were measures from the first component that were significantly predictive for mathematical achievement for both groups. Selective attention was the only additional measure from the second component that was significantly predictive for mathematical achievement for the preterm group, however. Interestingly, components extracted from domain-specific measures did not have significantly predictive value for mathematical achievement for any group. The results indicate a different relationship between domain-general abilities and performance in mathematics for preterm and full-term children, but no predictive value for domain-specific abilities for both groups. In other words, the first component derived from domain-general abilities, such as intelligence, working memory and processing speed, had a high predictive value for attainment in maths for both groups, whereas for the preterm group additional measures of attention were also predictive.

Using different methods, our results replicated previous findings suggesting that domain-general abilities account for a great variance of the mathematical performance in the preterm population beyond domain-specific skills. For example, Simms et al. (2013) reported that, in a group of extremely preterm children, 70% of the variance in the performance of domain-general abilities accounted for attainment in mathematics. In our study, similarly, 60% of the variance in VP children's domain-general abilities accounted for their attainment in mathematics. Conversely, only 36% of that variance in attainment was explained by domain-general abilities in the term group. This implies that there are differences in the cognitive strategies recruited for attainment in mathematics by the VP group and the term group.

Our results also suggest that IQ could explain a great proportion of the variance in the mathematical attainment, since the first component of domain-general abilities had the strongest correlation with measures of intelligence. Given that we decided to include children with low IQ in our sample, this is plausible. A number of studies have suggested that preterm children's difficulties with mathematics originate from difficulties related to domain-general abilities, especially IQ. For instance, Jaekel et al. (2014) showed that mathematical difficulties in preterm children are related to deficits in IQ, rather than being specific learning difficulties. In contrast to this view, Simms et al. (2013) argued that learning difficulties related to mathematics in the preterm population may not arise solely as part of difficulties with domain-general abilities, but may involve additional deficits in domain-specific abilities, such as numerical representations. Although the linear regressions did not confirm the role of domain-specific abilities in the attainment in mathematics for the preterm group, differences between groups in domain-general and domain-specific measures were evident.

Working memory and processing speed were also domain-general skills that had strong correlations the first component extracted from the PCA, a component that explained much of the variance in mathematical performance in both groups. Both visuospatial and verbal working memory components are likely to be involved in learning mathematics, since solving mathematical problems may elicit visuospatial as well as verbal representations and strategies (Logie, Gilhooly, & Wynn, 1994; Imbo & LeFevre, 2010). Different components of working memory such as the central executive, visuospatial sketchpad and the phonological loop have been shown to be associated with mathematics performance (Geary, et al., 2004; Swanson and Beebe-Frankenberger, 2004; Holmes and Adams, 2006; Swanson, 2006; Imbo and Vandierendonck, 2007; Bull, Espy and Wiebe, 2008; De Smedt et al., 2009; Alloway and Alloway, 2010; Meyer et al., 2010; Raghubar et al., 2010; Geary, 2011; Toll et al., 2011; Friso-van den Bos et al., 2013; Van der Ven, et al, 2013). Processing speed has also been associated with maths development in typically developing children (Bull and Johnston, 1997; Swanson et al., 2004; Schatschneier et al. 2004; Berg, 2008). Both working memory and processing speed have previously been reported to contribute to attainment in mathematics in children born prematurely. In line with our results, Mulder et al. (2009a) reported that processing speed and working memory accounted for approximately 60% of the variance in the performance of mathematical attainment in a group of children born very preterm.

Lastly, attention was another measure with strong correlations in the first two components extracted from PCA. In fact, sustained, selective and dual task were all subdomains of attention with significant correlations with the PCA components. Mulder et al. (2009b) also showed that attention is an area of weakness in preterm children associated with poor attainment in mathematics.

Considering all together, intelligence, working memory, processing speed and attention were the domains that greatly accounted for the performance in mathematical attainment in the preterm group, in line with previous results (e.g., Jaekel et al., 2014; Mulder et al., 2009a; Mulder et al., 2009b). Domain-specific skills did not significantly account for performance in maths.

6.6.6 Clinical and educational implications

The results of this study have implications for clinical and educational settings. Our results indicate that children born prematurely do not present the profile of developmental dyscalculia. Here, we demonstrated that although the preterm population does exhibit poorer numerical representations than their term-born peers, this was due to difficulties in inhibitory skills, rather than imprecise numerical representations. Our results also demonstrated that only domain-general abilities were significant predictors of mathematical performance.

The term developmental dyscalculia has often been employed to describe a core deficit in understanding and manipulating the quantity of sets and their numerosities (Butterworth, 1999, 2005, 2010). There is no generally agreed-upon functional definition of developmental dyscalculia, however. Thus, it should be clear that when we suggest that the preterm population does not have the profile of developmental dyscalculia, we consider that the preterm population does not present a core deficit in understanding and manipulating the quantity of sets and their numerosities. Rather, VP children difficulties with attainment in mathematics seems to be associated related to domain-general abilities. This view is in line with the core deficits in domain-general framework (see section 2.7.3). According to this view, mathematical learning disabilities could be the result of more general cognitive impairments, particularly in working memory, visuospatial skills, attention and executive functions (McLean and Hitch, 1999; Geary and Hoard, 2005; Donlan et al., 2007; Geary et al., 2007; Le Corre and Carey, 2007; Szucs et al., 2014).

Indeed, previous investigations have demonstrated that the aetiology of mathematical learning difficulties in the preterm population is rather associated with domain-general abilities, such as working memory, visuospatial skills, attention, and executive functions. For instance, Simms et al. (2015) demonstrated that although preterm children have poorer performance in specific mathematics skills, difficulties in mathematics were associated with visuospatial processing and working memory. Guarini et al. (2014) demonstrated that although six-year-old VP children had imprecise numerical representations, by the age of eight preterm children were able to catch up, suggesting a delay in numerical representations, rather than a deficit. In addition, slower reaction time was associated with those imprecise representations, suggesting difficulties in processing speed, a domain-general ability.

Taken all together, and when set alongside previous findings (e.g. Simms et al., 2015) the results of our study suggest that the aetiology of mathematical learning difficulties in the preterm population is related to domain-general abilities, in line with the core deficits in domain-general framework. Our study replicates previous results in suggesting that children born prematurely do not have the profile of developmental dyscalculia according to the core deficit in domain-specific framework. Mathematical difficulties in the preterm population seem rather to be associated with domain-general abilities, with difficulties in intelligence, working memory, processing speed and attention, markedly affecting their attainment in mathematics.

Our results might help in the better targeting of interventions for VP children experiencing difficulties with mathematics. They also suggest that interventions targeting working memory, processing speed and attention abilities, alongside mathematics-specific skills, may be beneficial for children born preterm. (Mulder et al., 2010; Simms et al., 2015). Interventions focusing merely on only one of these areas of deficit may have limited results. Interventions targeting improving working memory abilities have been extensively investigated, yet evidence for transfer to academic performance is lacking (Colmar & Double, 2017; Sánchez-Pérez et al., 2018). The adaptive computerised working memory intervention 'Cogmed' has received interest in terms of transfer to performance in untrained working memory skills, attention and non-verbal IQ (e.g., Klingberg et al., 2002). Recently, there has been some success with this intervention with small groups of very preterm pre-schoolers (Klingberg, Forssberg, & Westerberg, 2002) and extremely preterm adolescents (Grunewaldt, Løhaugen, Austeng, Brubakk, & Skranes, 2013); with improvements in a variety of memory tasks and attention. For domain-specific skills, simple board games have been noted to improve the internal numerical representations of children from low-income backgrounds, with evidence of transfer to simple addition problem fact retrieval, a core skill in basic mathematics (Siegler & Ramani, 2009). The use of concrete manipulatives, such as blocks or rods, in the classroom has also had some success in improving mathematical performance (Dowker, 2004). Alternatively, the development of mathematics teaching methods that place fewer demands on both working memory and visuospatial skills may be beneficial; such as breaking complex tasks down into simple steps and the use of concrete manipulatives and structured worksheets to scaffold visuospatial information. These types of whole-class interventions may also have wider benefits for all children in the classroom (Simms et al., 2013). Further research is needed to develop and assess the efficacy of interventions for very preterm children.

6.6.7 Limitations

Despite our best efforts, our study has numerous limitations. For example, one obvious limitation is the representativeness of our sample. The relatively high cognitive performance of both groups is particularly evident. The process of recruitment of the study could potentially explain this. Recruitment of the control group was merely done via emails and posters in the university. Thus, a great number of participants had highly educated parents working in the university. This was also applicable for the preterm group, since the participants were born at UCLH, a hospital attending a population living and working close to University College London, where the study was carried out. In addition, the increased cognitive performance evident in both groups might be explained by the characteristics of the sample, such as educational and social-economic status, both variables well known to affect cognitive performance (Hackman & Farah, 2009; Duncan, Magnuson, Kalil, & Ziol-Guest, 2012). Fisher's exact test comparing groups for maternal educational and SES confirmed no differences between groups, but parental education and SES might not reflect the overall population.

The design of the study also had a few limitations. Our sample included a limited number of participants born prematurely presenting with low mathematical performance besides of the sample mixing very and extremely preterm children. We decided to recruit preterm children regardless of screening for mathematical achievement. In accordance with the literature, we expected that approximately 44% of children would present difficulties with mathematics and around 70% of the sample would have other academic cognitive difficulties (Johnson et al., 2009). In the event, however, less than 30% of our sample presented with difficulties with maths. Yet, despite this percentage being smaller than expected, we were still able to demonstrate significant differences in mathematical performance between the preterm and the control group. Ideally, the optimal design of the study would be to replicate that employed by Isaacs et al. (2001). In their study, preterm children with and without mathematical difficulties were recruited and matched with term-born children with and without mathematical difficulties. Due to constraints in time and resources, however, this design was not implemented in the current study.

Our experimental measures investigating domain-specific abilities were based on accuracy and reaction time. In addition to those outcomes, we also calculate the Weber fraction (w), for the non-symbolic magnitude comparison, and inverse efficiency score. The latter measure was an attempt to consider reaction time measures in combination with accuracy. We included this measure in an attempt to avoid misinterpretation since analysing accuracy in the absence of reaction time (or vice versa) may elicit a serious problem in assessing individual ANS acuity. For example, a participant might have a low w estimate because s/he focused on making fast decisions even if s/he had high ANS acuity. Inverse efficiency scores lack in theoretical connections to the ANS, however (Chen and Li, 2014; Fazio et al., 2014; Dietrich et al., 2015). Diffusion model (Ratcliff & McKoon, 2008) is another measure that estimates individual ANS acuity by combining a sequential sampling model and an existing theoretical model of the ANS (e.g., Park & Starns, 2015). This model explains RT distributions of both correct and incorrect trials in the numerosity comparison task in terms of the rate of information accumulation (or "drift rate"), among many other model parameters. The drift rate is determined by the quality of the stimulus information, and thus it represents the quality of internal quantity representation in the present context. One advantage of the diffusion model is that it provides separate estimates of speed-accuracy trade-off and evidence quality (e.g., Ratcliff & McKoon, 2008), in contrast to measures that are solely based on accuracy, such as w, and that are thus potentially influenced by participant-level variation in speed-accuracy settings. To date, no studies have employed this approach to investigate numerical magnitudes in the preterm population and future studies would benefit employing this model.

6.6.8 Study's strengths

Despite the limitations presented above, the study also has its strengths. First, we carried out a comprehensive evaluation comprising domain-general and domain-specific abilities in relation to mathematics performance in a preterm population. The PRISM study (Simms et al., 2013b, 2015) is the only previous study to have comprehensively investigated both domains in children born prematurely. This comprehensive assessment allowed us to have measures of different domain-general abilities such as intelligence, processing speed, working memory, attention, planning and inhibition, thus building up an overall view of the interplay of domain-general and a set of domain-specific skills in the role of mathematical attainment in the preterm population.

Second, since multiple comparisons were carried out to explore both domains, we decided to apply Bonferroni corrections to adjust for multiple comparison, minimising the likelihood of committing type I error. In addition, we employed principal component analysis – the first time this type of analysis has been employed in this type of study. It could be argued that the selection of variables for running the regressions should be theory-driven. In other words, inhibition, working memory and shifting (see section 2.6.2 for a more detailed discussion) could be selected as independent variables and mathematical attainment as the dependent variable, using a more conventional method. In contrast, PCA allows the inclusion of all measures obtained in this study. The negative side of using PCA is that after extracting derived variables the components might not be as interpretable as the original measures. Furthermore, PCA requires standardisation of the data, hence we used z-scores. Taken together, however, PCA allow more efficient data usage, especially in experiments with small sample size, as is the case here.

Another advantage of this study is the type of stimulus we decided to employ in the experimental tasks. Hellgren et al. (2013) and Libertus et al. (2017) employed an intermixed presentation of stimuli in their paradigm, adding extra visuospatial demands on participants. Here, we decide to use simple, but controlled, stimuli for intensive and extensive parameters. Several studies have employed the same type of stimuli used in the current study (e.g., Hyde & Spelke, 2009; Heine et, 2011; Hyde & Spelke, 2011). Previous studies showed substantial deficits in visuospatial processing in very preterm children related to their proficiency in mathematics (Simms et al., 2015). Thus, experimental stimuli should take into account the cognitive deficits of the population under study and carefully control for variables that could potently interfere with the performance of the individuals.

6.6.9 Future directions

This study contributes to a better understanding of the underlying mechanisms of individuals born very prematurely at particular risk of experiencing difficulties with mathematics and its relationship to domain-general and domain-specific abilities. Future studies should address the limitations of this study, such as having bigger sample sizes with subgroups of very and extremely preterm children. Furthermore, other cognitive abilities recognised to contribute to attainment in mathematics, such as visuospatial skills (Gilligan, Flouri, & Farran, 2017; Young, Levine, & Mix, 2018) should also be addressed in future assessments, as well as the investigation of other domain-specific skills. Lastly, longitudinal studies investigating the development of mathematical skills in preterm children would reveal when difficulties occur and what they are. Since the development of typical trajectories of mathematical abilities is likely to emerge from a combination of multiple foundations, both domain-specific and domain-general skills need to be systematically investigated.

7 Neural Correlates of Numerical Magnitude Comparisons in Very Preterm Children

7.1 Abstract

Numerical magnitude comparison refers to our basic ability to decide which of two numerosities is the largest. It is typically assessed by comparing two sets of quantities, whether symbolic Arabic numbers or non-symbolic (e.g., dot) elements. Research has demonstrated that this ability to process and represent numerical magnitudes correlates with mathematical achievement. It has been hypothesised that the assessment of the ability to compare numerical magnitudes is a potential tool for identifying populations at risk of struggling with maths, including very preterm children. Behavioural studies have revealed that very preterm (VP) children have poorer performance than their term-born peers in non-symbolic numerical magnitude comparison tasks. It is not clear, however, whether these deficits arise from difficulties in processing number-related information (Hellgren et al., 2013; Libertus et al., 2017) or in domain-general skills, such as processing speed (Guarini et al., 2014). A technique that could help to elucidate the neural mechanisms underlying the processing of numerical magnitudes in this at-risk population is event-related potentials (ERPs). Two ERP components over posterior scalp sites have been found to be sensitive to numerical magnitude comparison for both numerical formats: a negative component that peaks around 150 ms post stimulus (N1) and a positive component that peaks around 200 ms post stimulus (P2p). Particularly, the P2p is currently considered to be a marker of approximate magnitude representation (Hyde and Spelke, 2012). Previous work using ERPs has shown that children with mathematical learning difficulties are less sensitive to numerical changes, implying that they have

imprecise numerical representation (Heine et al., 2013). To date, however, no study has used ERPs to investigate neural correlates of numerical magnitude comparisons in VP children. Thus, our aim was to examine the underlying differences between term and preterm children in the neural mechanisms of numerical magnitude comparison. Thirty children born very preterm (<32 weeks) and 24 term-born children aged between eight and ten years old joined the study. EEG data were obtained during testing using a symbolic and non-symbolic numerical comparison paradigm. We hypothesised that if children born prematurely showed deficits in non-symbolic numerical magnitude processing, they would be less sensitive to numerical changes. This would reflect their recruitment of fewer neural resources when encoding numerical information, indexed by the P2p component; similar to what has been observed more generally in children with mathematical learning difficulties. Our results showed no significant differences between children born prematurely and their term-born counterparts in the encoding phase (P2p), either for the symbolic and the non-symbolic numerical magnitude comparisons. There were, however, significant differences between groups prior to the encoding phase (N1). Our results suggest that children born very prematurely do not show imprecise numerical representation in the same way as children with mathematical learning disabilities. On the other hand, preterm children recruit the right inferior parietal regions involved in numerical magnitude comparisons with wider activations compared to the term group, indicating potentially compensatory mechanisms.

7.2 Introduction

It has been argued that the ability to measure numerical magnitude is a potential tool for identifying those at increased risk of struggling with maths, such as VP children. See section 2.6.1 for detailed explanation about numerical magnitude comparison).

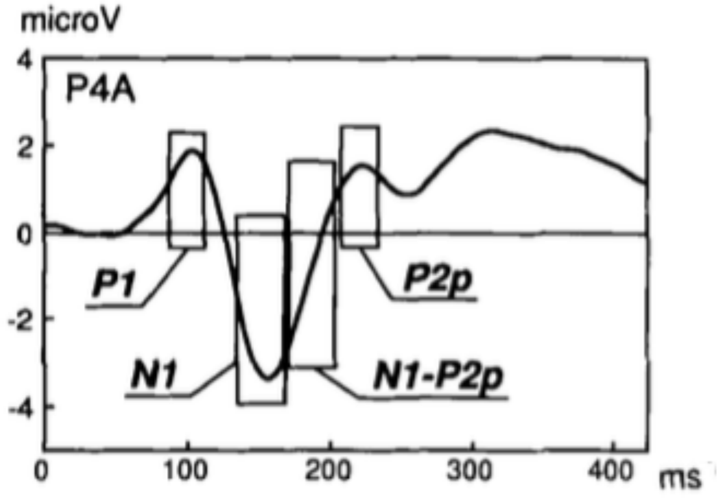
Only a small number of studies have examined numerical magnitude processing in children born prematurely (Hellgren et al., 2013; Guarini et al., 2014; Simms et al., 2015; Libertus et al., 2017) and the results are inconclusive. Whereas a few studies have argued that individuals born prematurely have deficits in non-symbolic numerical magnitude comparisons, reflecting specific imprecise numerical representations (Hellgren et al., 2013; Libertus et al., 2017), others have argued that their poor performance is associated with domain-general abilities, such as processing speed and inhibition (Guarini et al., 2014; Simms et al., 2015). For a detailed discussed, please see section 2.9.2.

A technique that could elucidate the underlying neural mechanisms of numerical magnitudes processing in VP children is event-related potentials (ERPs). Arguably, ERPs are the most effective way to investigate information processing speeds, since it has the advantage of high temporal resolution (in the range of milliseconds), meaning that it can capture neural signatures of cognitive processes long before the overt responses are made (Soltesz, Goswami, White, & Szucs, 2011). ERPs reflect changes in electrical activity in response to a stimulus or event (Riggins, Scott, & Nelson, 2007). Methodologically, employing ERPs adds two distinctive advantages over merely relying on measures of speed and accuracy when measuring numerical magnitude comparison. Firstly, ERP measures provide a continuous picture of a cognitive process, making it possible to determine which stages of the process are being elicited and possibly modulated; and, secondly, ERPs provide an online measure of the processing of information, making it the “reaction time of the twenty-first century” (Luck, 2005). Another advantage of using ERPs is their ability to be employed across a wide age range, allowing for the developmental tracking of neural markers of numerical cognition. In fact, ERPs have been used to show that three-month-old infants can be sensitive to small changes in numerical ratios (Izard et al., 2008), and that seven-month-old infants show similar neural responses to adults in ratio dependency following Weber's Law (Hyde & Spelke, 2012).

Two main ERP components over posterior scalp sites are thought to reflect number processing: the N1, the first negative component peaking around 150 ms post-stimulus; and the P2p, the second posterior positivity peaking around 200 ms post-stimulus (Dehaene, 1996; Temple & Posner, 1998; Libertus, Woldorff, & Brannon, 2007; Hyde & Spelke, 2009; Rubinsten, Sury, Lavro, & Berger, 2013). The N1 component is associated with visuospatial attentional processing and reflecting the "notation effect"; i.e. perceiving the difference between verbal numerals –e.g., "six"– versus Arabic digits –e.g., "6". The N1 is followed by the P2p component in the encoding phase, and is currently considered to be a marker of approximate magnitude representation (Hyde & Spelke, 2012), since it reflects the distance effect modulated by close numbers to far numbers. Specifically, the amplitude of the P2p was greater for small distances than large distances (Dehaene, 1996; Temple & Posner, 1998; Pinel, Dehaene, Riviere, & LeBihan, 2001; Szucs & Csépe, 2004; Szűcs & Csépe, 2005; Libertus et al., 2007; Szũcs, Soltész, Jármi, & Csépe, 2007; Hyde & Spelke, 2009, 2012; Soltész et al., 2007; Soltesz et al., 2011).

Figure 7.1. shows the main ERP components modulated by numerical magnitude processing (A) over parietal areas (B).

A



B

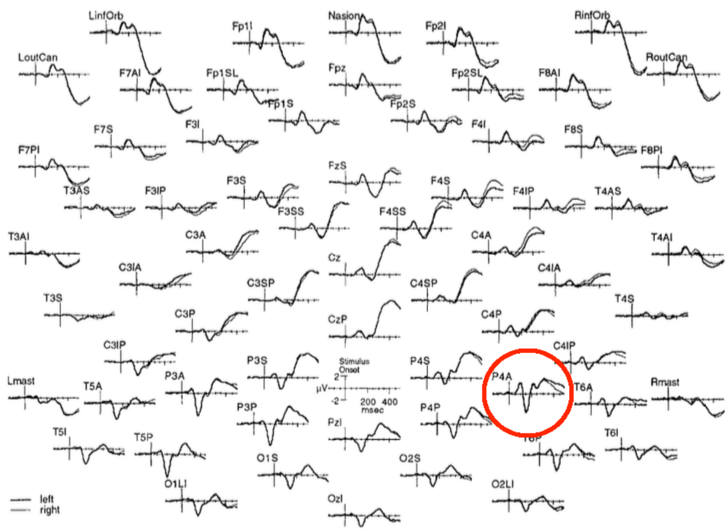


Figure 7.1: the main ERP components modulated by numerical magnitude processing (A) over parietal areas (B).

P2p. Left-hand side: the components involved in numerical magnitude processing, first identified by Dehaene (1996). Right-hand side: The P2p emerges over the parietal areas around 200 ms after stimulus presentation reflecting distance effect modulated by close numbers to far numbers.

Temple and Posner (1998) investigated differences between children and adults when performing non-symbolic and symbolic numerical magnitude comparisons, observing that children were more than three times slower than adults (a 480 ms response time in adults and a 1495 ms response time in children). Surprisingly, there was no such difference in the numerical distance effect, which was measured by ERPs. Both children and adults showed an ERP distance effect at around 200 ms after stimulus presentation. The ERPs, therefore, indicated that access to numerical representations *per se* appeared to be as fast in children as in adults (Soltész et al., 2007; Soltész & Szucs, 2009; Soltész et al., 2011). Temple and Posner suggested that the delayed magnitude judgements in children were therefore due to less well developed inhibitory skills.

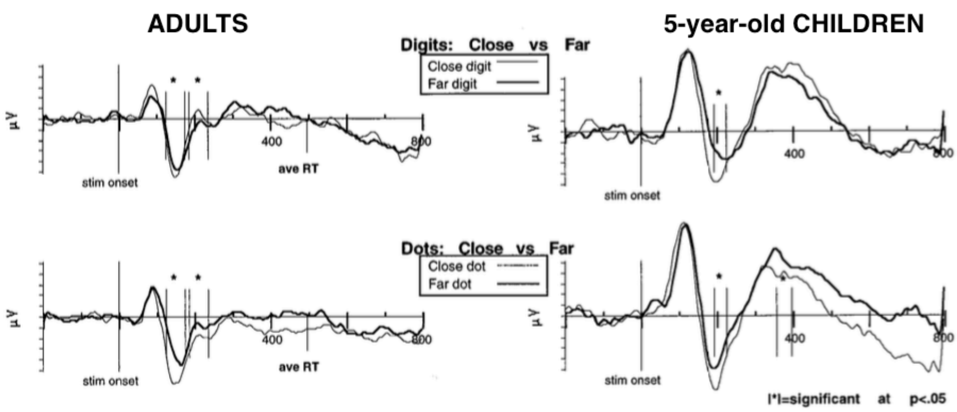


Figure 7.2: Numerical magnitude comparisons in adults and children.

Numerical magnitude comparisons elicited same posterior parietal ERP response in children and adults, despite significant differences observed in RTs (Temple and Posner, 1998).

Interestingly, only a handful of studies have used ERPs to explore the relationship between the processing of numerical magnitudes and mathematical learning disabilities (MLD). Soltész et al. (2007) evaluated adults and adolescents with, and without, developmental dyscalculia while performing symbolic comparison tasks. They found that the subjects with MLD showed no modulation of their brain activity in relation to changes in numerical distance, in contrast to those without MLD. Heine et al. (2013) evaluated eight-year-old children with, and without, MLD, and found that those with maths difficulties did not show a distance effect in a series of non-symbolic numerical comparison tasks. Gomez-Velazquez et al. (2015) investigated numerical magnitude comparison in a group of children with low mathematical achievement. Lower amplitudes in components modulated by number tasks were observed in the group of children with lower mathematical achievement, in agreement with previous studies. The results suggest a substantially diminished distance effect in children with mathematical learning disabilities, indicating less sensitivity to numerical changes.

We investigated the electrophysiological correlates of symbolic and non-symbolic numerical processing while comparing smaller versus larger distances in a group of school-aged children born prematurely compared to their term-born peers. We hypothesised that if VP children have imprecise numerical representations when performing non-symbolic numerical magnitude comparison, as has been suggested by previous studies (Hellgren et al., 2013; Libertus et al., 2017), they would be less sensitive to numerical changes, similar to children with MLD. If this was the case, smaller amplitudes would be expected in the P2p component, modulated by small distance in the non-symbolic numeric tasks. No differences between groups in the symbolic magnitude comparison would be expected, since previous behavioural studies have failed to demonstrate significant differences between term and preterm children (e.g. Guarini et al, 2014). Conversely, if VP children do not have impressive numerical representation, in line with previous findings (Guarini et al., 2014, Simms et al., 2015), no differences between groups would be observed either in the symbolic or non-symbolic tasks. This would reflect similar amplitudes in the P2p component.

7.3 Methods

7.3.1 Participants

The procedure adopted in this study was identical to the experimental tasks employed on Study 3 (for a detailed description of the methods, please see section 3.2.9.5 and 6.3.2.1). Here, we focused on the analysis of the electrophysiological data collected from the experimental tasks evaluating symbolic and non-symbolic magnitude comparison. Sixty-eight children participated in the study. Fourteen participants (eight preterm and six term) were not included in this study due to problems with equipment (n=4), not enough trials with good data recorded (n=5), and an additional five participants were excluded after visual inspection of the individual waveforms revealed a lack of the expected morphology. Thus, the final sample of this study comprised 54 out of the 68 children assessed. Thirty preterm children (10 VP, 9 males and 20 EP, 13 males) and 24 term-born children (19 males) joined the final sample of this study. Out of the total sample of the preterm group 69% were extremely preterm children. Demographics variables did not differ from Study 3.

7.3.2 Recording environment

The EEG was recorded in a quiet room at the Wolfson Assessment Centre. The participant sat behind an opaque divider to avoid distraction.

7.3.2.1 EEG recording

Recordings were obtained using NetStation software 5.3.0.1 (Electrical Geodesics Inc., OR) on a Mac OS 10.3.9 software. EEG was acquired at a digitising rate of 250 Hz with a bandwidth of 0.1-100 Hz using a 256-channel EEG Hydrocel Geodesic Sensor Net

(2008, EGI, Electrical Geodesics, Inc., Eugene, OR, USA). A vertex reference was used for recording (Cz). Display filters between 0.1 Hz and 100 Hz were used during the recording for better online visualisation of the recording.

The EEG was recorded using a GES 200 high-density, high impedance recording system with a NetAmps 200 amplifier and HydroCel Geodesic Sensor Nets with 256 channels (Electrical Geodesics Inc., OR). Suitable sizes were available for all participants. Saline solution was used as the conductivity medium for recording, in line with the technical manual of the recording system.

The amplifier was calibrated, and impedances were measured for each recording. Channels with an impedance higher than 50 kΩ were checked for good contact with the scalp and adjusted where necessary, following procedures described in the Geodesic Sensor Net Technical Manual (Electrical Geodesic Inc., OR). An electrooculogram (EOG) was also recorded through the EEG system using a pair of electrodes positioned above and below both eyes for the detection of eye-related artefacts in the EEG.

7.3.2.1 Data processing

The EEG recordings were exported to EEGLAB format for processing and analysis. EEG data were analysed using EEGLAB version 14.1.2b (Delorme & Makeig, 2004), a MATLAB toolbox (2017, The Mathworks, Inc., Massachusetts, USA), and the ERPLAB plug-in (Lopez-Calderon & Luck, 2014). Firstly, all EEG recordings were band-pass filtered (0.5-30 Hz). These filter settings are within the recommended range for EEG analyses with infants and children and have previously been employed in ERP studies investigating numerical cognition in children (Hoehl & Wahl, 2012).

A group of noisy channels[15] were permanently excluded from all of the recordings due to their position in the periphery of the head and neck that contribute to higher impedance, causing greater disturbance to the data. The removed electrodes marked in red are illustrated in Figure 3.4. The final montage had a total of 173 electrodes. From the remaining electrodes, channels with poor quality data, presenting flatlines, drifts or excessive noise, were removed by visual inspection. The signal was submitted to an Independent Component Analysis to identify and correct eye blink artefacts, followed by the previously removed bad channels' spherical spline interpolation. The signal was then re-referenced offline to the average of all electrodes. The data was segmented into epochs of 900 ms; 100 ms pre-stimulus onset and 800 ms post-stimulus onset. Data was visually inspected and those containing artefacts were rejected prior to averaging. After exclusion of all error trials, baseline corrected average ERPs were computed for each participant and stimulus category (baseline: 100-0 ms pre-stimulus). A participant was excluded from this study if less than twenty good trials per condition resulted from this process, in line with recommendations from Hoehl and Wahl (2012). Finally, the data were averaged across all trials per participant, and across all participants in order to obtain the grand-averaged waveforms, which were created for the two groups: term and preterm.

[15] EEG recording is highly susceptible to various forms and sources of noise, such as environmental (a large array of electronic equipment) or physiological noise (eyeball movement or muscular movement).

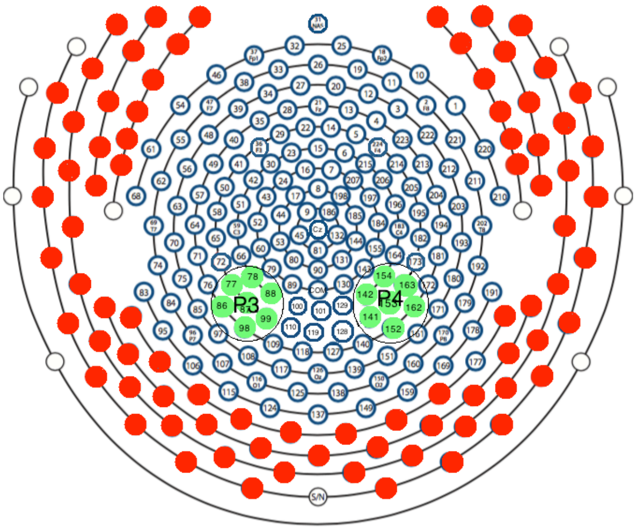


Figure 7.3: Schematic representation of the final montage of the Geodesic sensor net 256 channels.

The final montage included 173 electrodes (in blue). Electrodes coloured in red were permanently excluded due to artefacts. The channel cluster sites, represented in green, are located over the left (P3: 77, 78, 86, 87, 88, 98, 99) and right (P4:141, 142, 152, 153, 154, 162, 163) parietal areas. Channel selection was based on previous research using similar age ranges and paradigms (Soltesz et al., 2007; Heine et al., 2011; Gomez-Velazquez et al., 2015).

7.4 Data analysis

7.4.1 Behavioural analysis

A 2 x 2 x 2 mixed-design ANOVA was performed on proportion of correct responses (CR), with *Group* (preterm *vs* term children) as the between-subject factor, and *Task* (non-symbolic and symbolic) and *Distance* (small and large) as the within-subject factors. A second 2 x 2 x 2 mixed-design ANOVA with the same variables (*Group* (preterm *vs* term children), and *Task* (non-symbolic and symbolic) and *Distance* (small and large), both as within subject factors) was performed on reaction time (RT) for correct responses. The dependent variables were the proportion of correct responses

(CR) and reaction time (RT), respectively. CR was Arcsine transformed to meet assumptions of ANOVA.

7.4.2 EEG data analysis

Group-averaged ERP waveforms comprised two main ERP components:[16] a negative peaking component around 212 ms with a window between 200-300 ms (N1); and a positive peaking component around 336 ms with a window between 300-400 ms (P2p). The selection of the time windows for the different ERP components was data-driven, according to the grand averages and the main visually-detected changes. Regarding the choice for the electrodes, a group of electrodes located over the left (P3: 77, 78, 86, 87, 88, 98, 99) and right (P4:141, 142, 152, 153, 154, 162, 163) parietal areas were selected according to the EGI 256 channel geodesic map (see section 3.2.9.4). Channel selection was based on previous research using similar age ranges and paradigms (Dehaene, 1996; Temple and Postner, 1998; Soltesz et al., 2007, Hyde and Spelke, 2009; Heine et al., 2011; Gomez-Velazquez et al., 2015).

A visual inspection was performed for all data blind to group membership. A 2 x 2 x 2 x 2 mixed-design ANOVA was used to analyse the electrophysiological results, with *Group* (term and preterm) as the between-subject factor, and *Task* (non-symbolic and symbolic), *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors. Analyses were conducted with peak amplitudes and

[16] Unexpectedly, a noticeable delay was observed in the waveforms after stimulus onset across all tasks, conditions, sites and groups. Due to equipment failure, we were unable to run a timing test. Timing tests allow a comparison between the time when the stimulus presentation system states that the stimulus is presented, and the time that the AV device physically detects the presentation of the stimulus, as it would be presented to the subject. As such, the 100 ms delay between stimulus onset and the first component (P1) was interpreted as a delay in stimulus presentation. The delay was constant and therefore cannot account for differences between groups and conditions.

latencies as dependent variables across each component (N1, P2p), separately. Grenhouse-Geisser corrected results were used when the assumption of sphericity was violated.

7.5 Results

7.5.1 Behavioural data

Results regarding behavioural data were extensively discussed on Study 3 (see section 6.6). The results here replicated these previous findings, as follows below.

A 2 x 2 x 2 mixed-design ANOVA was performed to examine proportion of correct responses (CR) as dependent variable with *Group* (preterm *vs* term children) as the between-subject factor, and *Task* (non-symbolic and symbolic) and *Distance* (small and large) as the within-subject factors. CR was Arcsine transformed. After applying Bonferroni corrections for correct response (α = 0.006), a main effect was found for *Group* (F (1,52)=18.939, *p*=<0.001), with preterm children performing less accurately than term children (preterm: M_{CRs} = 87.1%, SE = 0.94, term: M_{CRs} = 93.2%, SE = 1.05). A main effect for *Task* (F (1,52)=88.784, *p*=<0.001) was observed, with higher accuracy for symbolic tasks (M_{CRs} = 93.8%, SE = 0.64) than non-symbolic tasks (M_{CRs} = 86.5%, SE = 0.96). A main effect for *Distance* (F (1,52)=297.842, *p*=<0.001) was also observed, with more accurate answers for large distance (M_{CRs} = 95.0%, SE = 0.60) than small distance (M_{CRs} = 85.3%, SE = 0.92). The only interaction observed was between *Group* and *Task* (F (1,52)=29.280, *p*=0.004), in which preterm children were less accurate in both the non-symbolic task (M_{CRs} = 81.6%, SE = 1.28) and symbolic task (M_{CRs} = 92.6%, SE = 0.85) than their term-born peers (non-symbolic: M_{CRs} = 91.5%, SE = 1.44; symbolic: 95.0%, SE = 0.95). To follow this up, we ran a post hoc paired t-test for the non-symbolic and symbolic task, for each group separately. This revealed a significant difference between tasks for both the term group (t(23)=2.710, *p*=0.012) and the preterm group (t(29)=11.717, *p*=<0.001), with a larger difference for the preterm group. This suggests that the interaction between *Group* and *Task* was driven by large differences in tasks among the preterm group.

Regarding RT, the ANOVA was the same 2 x 2 x 2 as above. It yielded a significant effect of *Task* (F (1,52)=35.413, *p*=<0.001), with slower answers elicited by the non-symbolic task (M_{RTs} = 1013 ms, SE = 24.9) in contrast to the symbolic task (M_{RTs} = 873 ms, SE = 28.20). A *Distance* effect was also observed, (F (1,52)=256.89, *p*=<0.001), with faster answers for large distance (M_{RTs} = 879 ms, SE = 22.80) than small distance (M_{RTs} = 1006 ms, SE = 25.58). The only interaction observed was between *Task* and *Distance* (F (1,52)=19.611, *p*=<0.001), in which small distance trials elicited slower answers (M_{RTs} = 1059 ms, SE = 25.47) in comparison to large distance (M_{RTs} = 967 ms, SE = 25.71) in the non-symbolic tasks, with mean difference of 92 seconds between the small and large distance (t(53)=-10.717, *p*=<0.001). Small distance trials also elicited slower answers (M_{RTs} = 955 ms, SE = 32.05) in comparison to large distance (M_{RTs} = 792 ms, SE = 25.04) in the symbolic tasks, with mean difference of 163 seconds between the small and large distance (t(53)=-12.316, *p*=<0.001). The results suggest that the bigger difference in the symbolic task explained the interaction. No other significant interactions were observed for reaction time. Given that we were mainly interested in group differences, we additionally carried out independent t-tests to explore differences between groups for the non-symbolic and symbolic tasks. The effect of differences in conditions (small and large distances) on accuracy (CRs) and RTs were explored. After Bonferroni corrections for multiple comparisons, significant differences between groups were observed for accuracy for non-symbolic representations, regardless of the condition. Significant differences between groups were observed in large distance (t(52)=1.231, p=0.001, d=1.0) and small distance (t(52)=2.125, p=<0.001, d=1.4). Longer RTs was also observed for the symbolic task, regardless of the condition in the preterm group. Significant differences were observed in large distance (t(52)=1.323, p=0.002, d=0.9) and small distance (t(52)=0.024, p=0.005, d=0.7).

Figure 7.4 illustrates the behavioural results regarding this numerical magnitude comparison. A summary of the results of the two ANOVAs on RTs and CRs is provided in table 7.1. Means for CRs (SD) and RTs (SD) according to group, task and distance are displayed in table 7.2.

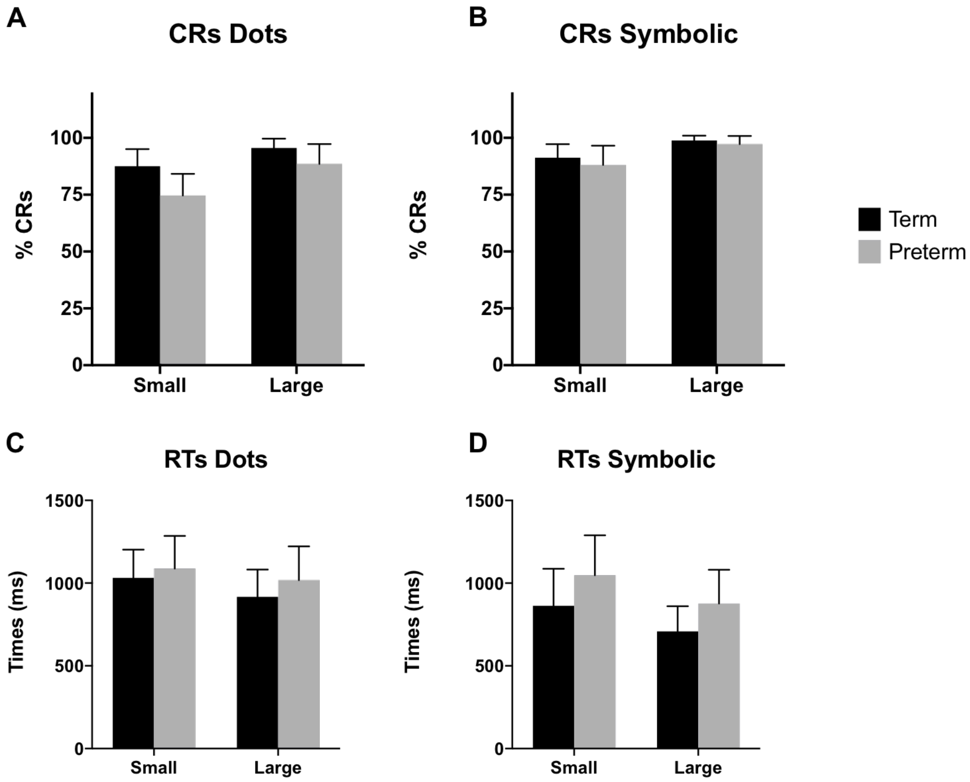


Figure 7.4: Behavioural results for numerical magnitude comparison

A) Term children were significantly more accurate than VP children in the non-symbolic task for both conditions. B) Preterm children were as accurate as their term counterparts in the symbolic task. C) No significant differences were found between groups in respect to response time for the symbolic task. D) Preterm children were slower than term children when performing the symbolic magnitude comparison task, but differences were not significant. CRs: correct responses; RTs: reaction times. Error bars represent standard deviation of the data.

Table 7.1: F, degrees of freedom, and values for both ANOVAs on CRs and RTs

	ANOVA CRs			ANOVA RTs		
	F	Df	*p*	F	Df	*p*
Group	18.939	1,52	<0.001[a,b,c]	7.075	1,52	0.010
Task	88.784	1,52	<0.001[a,b,c]	35.413	1,52	<0.001[a,c]
Distance	297.842	1,52	<0.001[a,b,c]	256.897	1,52	<0.001[a,b,c]
Group x Task	9.280	1,52	0.004[a,c]	4.326	1,52	0.042
Group x Distance	0.173	1,52	0.679	0.767	1,52	0.385
Task x Distance	1.184	1,52	0.282	19.611	1,52	<0.001[a]
Group x Task x Distance	1.226	1,52	0.273	3.708	1,52	0.060

CRs = correct responses, RTs = Reaction times. [a] Remains significant after applying Bonferroni correction (α = 0.007), [b] Remains significant after controlling for full composite IQ (ANCOVA).[c] Remains significant after controlling for SES (ANCOVA)

Table 7.2: Children's performance in the non-symbolic and symbolic tasks in respect to RTs and CRs

		Term			**Preterm**			**Difference between control and VP children**		
Task	Distance	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
Dots										
CRs (%)	Large	24	95.57	4.09	30	88.54	8.74	7.03 (3.14 to 10.92)	0.001[a]	1.0
RTs (ms)		24	917.22	164.61	30	1017.77	204.27	-100.55 (-203.73 to 2.63)	0.056	0.5
CRs (%)	Small	24	87.50	7.56	30	74.68	9.49	12.81 (8.15 to 17.47)	<0.001[a]	1.4
RTs (ms)		24	1031.72	170.32	30	1087.74	197.66	-56.01 (-158.27 to 46.23)	0.277	0.3
Symbolic										
CRs (%)	Large	24	98.82	2.12	30	97.20	3.60	1.59 (-0.075 to 3.26)	0.023	0.5
RTs (ms)		24	708.57	152.07	30	876.30	204.09	-167.72 (-268.24 to -67.20)	0.002[a]	0.9
CRs (%)	Small	24	91.25	5.95	30	88.07	8.44	3.18 (-0.75 to 7.12)	0.139	0.4
RTs (ms)		24	862.85	224.33	30	1047.25	241.49	-184.39 (-313.02 to -55.77)	0.005[a]	0.7

CRs= Correct responses; RTs= Reaction time; [a]Remains significant after applying Bonferroni correction (α = 0.006).

7.5.2 Electrophysiological data

7.5.2.1 N1

A 2 x 2 x 2 x 2 mixed-design ANOVA was used to analyse the amplitudes in the N1 component as the dependent variable, with *Group* (term and preterm) as the between-subject factor, and *Task* (non-symbolic and symbolic), *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors. The analysis based on the amplitudes of the earlier negative component (N1) revealed a main effect of *Task* ($F_{(1,52)}=44.167$, p=<0.001) and *Site* ($F_{(1,52)}=101.169$ $p=0.002$). Symbolic tasks elicited higher amplitudes (M_{amp}=-3.20 µV, SE = 0.35) than non-symbolic tasks (M_{amp}=-1.63 µV, SE = 0.26). Higher amplitudes were observed in the right hemisphere (M_{amp}=-3.00 µV, SE = 0.34) than in the left hemisphere (M_{amp}=-1.83 µV, SE = 0.33). An interaction was observed between *Task* and *Site* ($F_{(1,52)}=5.832$, $p=0.019$), in which the symbolic task elicited higher amplitudes in the right hemisphere (M_{amp} =-4.02 µV, SE = 0.43) than the left hemisphere (M_{amp}=-2.38 µV, SE = 0.40), in comparison to the non-symbolic task, which had higher amplitudes in the right hemisphere (M_{amp}=-1.98 µV, SE = 0.31) than the left hemisphere (M_{amp}=-1.28 µV, SE = 0.33). To follow this up, we ran a 2 x 2 ANOVA, with *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors for each task separately. This revealed a main effect of site only for the symbolic task ($F_{(1,53)}$ =12.366, $p=0.001$), in comparison to the non-symbolic tasks ($F_{(1,53)}$ = 3.096, $p=0.084$). Mean differences between amplitudes suggested that the interaction was driven by bigger differences between sites (right x left hemisphere) than by task (non-symbolic x symbolic).

A second interaction was observed between *Group, Task, Distance* and *Site* (F (1,52)=8.126, p=0.006). To follow this up, we ran a 2 x 2 x 2 mixed-design ANOVA with *Task* (non-symbolic and symbolic), *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors for each group separately. This revealed a 3 way interaction solely for the term group (F (1,23)=8.822, p=0.007), in comparison to the preterm group (F (1,29)=0.820, p=0.373). Such 3 way interaction was driven by a 2 way interaction between *Distance* and *Site* solely for the non-symbolic task (F (1,23)=8.241, p=0.009), in comparison to the symbolic task (F (1,29)=4.726, p=0.040). In turn, the 2 way interaction was explained by a difference between site solely for the small distance (t (23)=3.134, p=0.005), when compared to the large distance (t (23)=0.924, p=0.365). These results suggest that the 4 way interaction was driven by site differences between the conditions in the non-symbolic task for the term group.

A 2 x 2 x 2 x 2 mixed-design ANOVA was used to analyse the latencies in the N1 component as the dependent variable, with *Group* (term and preterm) as the between-subject factor, and *Task* (non-symbolic and symbolic), *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors. The analysis of latencies revealed that this component showed a main effect of *Task* (F (1,52) = 6.061, p=0.017), in which the latency period was slightly shorter for the non-symbolic task (non-symbolic task: M_{lat} = 223 ms, SE = 2.09; symbolic task: M_{lat} = 228 ms, SE = 1.62). No interactions were observed for latencies in this component. Table 7.3 summarises the results from the N1 component in respect to amplitudes and latencies.

Table 7.3: F, degrees of freedom, and values for both ANOVAs on amplitudes and latencies for the N1 component

	ANOVA on amplitudes			ANOVA on latencies		
	F	Df	*p*	F	Df	*p*
Group	0.291	1,51	0.592	0.191	1,51	0.663
Task	44.167	1,51	< 0.001[a]	6.061	1,51	0.017
Distance	0.0011	1,51	0.996	0.869	1,51	0.356
Site	101.169	1,51	0.002[a]	1.052	1,51	0.310
Task x Group	0.032	1,51	0.860	0.092	1,51	0.763
Group x Distance	0.018	1,51	0.893	0.639	1,51	0.428
Group x Site	1.495	1,51	0.227	0.400	1,51	0.530
Task x Distance	0.165	1,51	0.687	0.938	1,51	0.337
Task x Site	5.832	1,51	0.019	0.001	1,51	1.000
Distance x Site	1.266	1,51	0.266	0.373	1,51	0.544
Task x Distance x Site	2.734	1,51	0.102	2.709	1,51	0.106
Group x Task x Distance	3.909	1,51	0.053	0.095	1,51	0.759
Group x Task x Site	0.003	1,51	0.959	1.204	1,51	0.278
Group x Distance x Site	0.264	1,51	0.102	0.001	1,51	0.972
Group x Task x Distance x Site	8.126	1,51	0.006[b]	1.442	1,51	0.235

[a]Remains significant after applying Bonferroni correction ($\alpha = 0.003$).
[b]Remains significant after controlling for full composite IQ (ANCOVA).

7.5.3 P2p

A 2 x 2 x 2 x 2 mixed-design ANOVA was used to analyse the amplitudes in the P2p component as the dependent variable, with *Group* (term and preterm) as the between-subject factor, and *Task* (non-symbolic and symbolic), *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors. The amplitudes of the second positive component (P2p) revealed a main effect of *Task* (F (1,52)=28.765, p=<0.001), where higher amplitudes were observed in the non-symbolic task (M_{amp} = 5.50 µV, SE= 0.30) than in the symbolic task (M_{amp} = 3.97 µV, SE = 0.31). No other main effects or interactions were observed for amplitudes in this component.

A 2 x 2 x 2 x 2 mixed-design ANOVA was used to analyse the latencies in the P2p component as the dependent variable, with *Group* (term and preterm) as the between-subject factor, and *Task* (non-symbolic and symbolic), *Distance* (small and large) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors. The analysis of latencies revealed that this component showed a main effect of *Group* (F (1,52)=7.120, p=0.010), in which the latency period was shorter for the term group (M_{lat} = 325 ms, SE = 1.58) than the preterm group (M_{lat} = 330 ms, SE = 1.41). An interaction was observed between *Task*, *Site* and *Group* (F (1,52)=5.642, p=0.021). To explain such an interaction, we ran a 2 x 2 x 2 mixed-design ANOVA, with *Task* (non-symbolic and symbolic) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors, and *Group* (term and preterm) as the between-subject factor, for each distance separately. This revealed a main effect of *Group* (F (1,52)=4.988, p=0.030) as well as an interaction between *Task*, *Site* and *Group,* but only for the large distance (F (1,52)=4.119, p=0.048). No interaction was observed between *Task*, *Site* and *Group* for the small distance (F

(1,52)=2.257, *p*=0.139). We then ran 2 x 2 ANOVA with *Task* (non-symbolic and symbolic) and *Site* (P3, for left hemisphere, and P4 for right hemisphere) as the within-subject factors for each group separately, for the large distance. This revealed an interaction between *Task* and *Site* only for the preterm group (F (1,29)=5.244, *p*=0.029). The preterm group, when performing the non-symbolic task in the large distance condition, had longer latencies in the right hemisphere (M_{lat} = 333.105 ms, SE = 3.385) than the left hemisphere (M_{lat} = 329.250 ms, SE = 3.194; (F (1,29)=0.694, *p*=0.412), in comparison to the symbolic task, which had longer in the left hemisphere (M_{lat}=332.41 ms, SE = 2.437) than the right hemisphere (M_{lat}= 326.762 ms, SE = 3.206; (F (1,29)=2.217, *p*=0.147). No interaction was observed between *Task* and *Site* for the large distance in the term group (F (1,23)=0.712, *p*=0.408). No other interaction was observed in any other follow-up ANOVAs. This suggests that the interaction between *Task*, *Site* and *Group* was driven by large differences in latencies between hemispheres among the preterm group.

Table 7.4 summarises the results from the P2p component in respect to amplitudes and latencies. Table 7.5 shows descriptive results for latencies and amplitudes for each recording site, condition and group. Figure 7.5 shows grand-average waveforms from the term and preterm group in respect to the two comparisons tasks. Figure 7.6 shows bar charts of amplitude and latencies for each task according to group, component, distance and site. Topographic maps corresponding to the grand-averaged waveforms from the two groups for the two tasks are shown in figure 7.7.

Table 7.4: F, degrees of freedom, and values for both ANOVAs on amplitudes and latencies for the P2p component

	ANOVA on amplitudes			ANOVA on latencies		
	F	Df	*p*	F	Df	*p*
Group	1.210	1,51	0.276	7.120	1,51	0.010
Task	28.765	1,51	<0.001[a]	0.370	1,51	0.681
Distance	0.108	1,51	0.744	0.376	1,51	0.542
Site	0.026	1,51	0.873	0.538	1,51	0.467
Task x Group	0.638	1,51	0.428	0.171	1,51	0.681
Group x Distance	0.197	1,51	0.659	0.001	1,51	0.969
Group x Site	0.090	1,51	0.766	0.229	1,51	0.634
Task x Distance	0.030	1,51	0.862	0.111	1,51	0.926
Task x Site	1.118	1,51	0.295	0.022	1,51	0.882
Distance x Site	0.057	1,51	0.812	0.528	1,51	0.471
Task x Distance x Site	1.858	1,51	0.179	1.171	1,51	0.284
Group x Task x Distance	0.687	1,51	0.411	0.570	1,51	0.454
Group x Task x Site	0.024	1,51	0.878	5.642	1,51	0.021
Group x Distance x Site	0.977	1,51	0.327	0.241	1,51	0.626
Group x Task x Distance x Site	0.773	1,51	0.383	0.320	1,51	0.574

[a]Remains significant after applying Bonferroni correction ($\alpha = 0.003$).
No differences remained significant after controlling for full composite IQ (ANCOVA).

Table 7.5: Mean latencies (ms) and voltages (µV) in the N1 and P2p components separated by task, group, site and distance

			Non-symbolic							
			Small				Large			
			Latencies		Amplitudes		Latencies		Amplitudes	
Component	Group	Site	Mean	SD	Mean	SD	Mean	SD	Mean	SD
N1	Term	P3	220.71	17.91	-0.40	2.07	222.52	18.65	-1.46	2.09
		P4	221.42	20.87	-2.24	2.45	227.28	24.11	-1.90	1.95
	Preterm	P3	222.64	18.29	-1.78	2.83	224.13	23.48	-1.49	3.03
		P4	226.15	20.43	-1.98	2.59	226.93	22.89	-1.82	3.07
P2p	Term	P3	326.11	13.44	5.52	3.04	325.16	17.16	5.35	2.69
		P4	327.11	15.32	5.93	3.14	325.54	16.96	5.94	3.20
	Preterm	P3	332.26	14.52	5.04	3.60	333.10	18.54	5.42	3.22
		P4	329.29	16.73	5.56	3.46	329.25	17.49	5.19	3.32

			Symbolic							
			Small				Large			
			Latencies		Amplitudes		Latencies		Amplitudes	
Component	Group	Site	Mean	SD	Mean	SD	Mean	SD	Mean	SD
N1	Term	P3	222.52	20.26	-2.33	2.68	225.95	20.08	-1.62	2.29
		P4	232.71	21.40	-4.12	3.09	229.33	22.90	-4.02	2.94
	Preterm	P3	228.85	20.38	-2.55	3.15	229.21	20.64	-3.03	3.89
		P4	229.16	17.04	-3.93	3.72	227.12	17.65	-4.00	3.36
P2p	Term	P3	327.78	12.86	4.466	3.05	326.64	14.32	4.38	3.42
		P4	321.09	14.74	4.19	2.90	321.69	15.23	4.52	2.43
	Preterm	P3	332.22	18.56	3.95	2.60	326.76	17.55	3.44	3.32
		P4	331.23	16.30	3.47	3.14	332.41	13.34	3.36	3.76

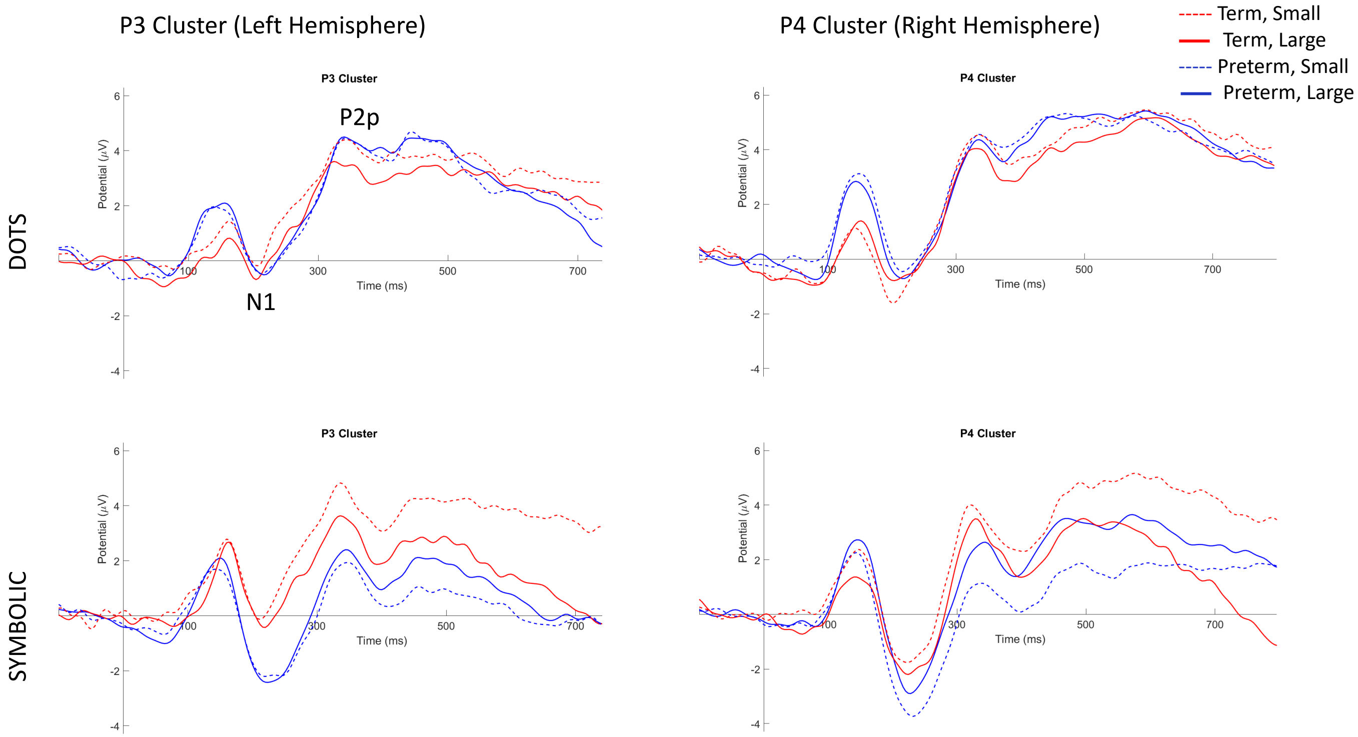


Figure 7.5: Grand-averaged waveforms obtained from the two comparisons tasks.
Blue lines preterm group and red lines represent the term group. Dotted lines represent small distance and continuous lines represent large distance. P3 cluster represents the cluster of channels from the left parietal areas and P4 cluster represents the cluster of channels from the right parietal areas. Voltage is plotted with negative going down.

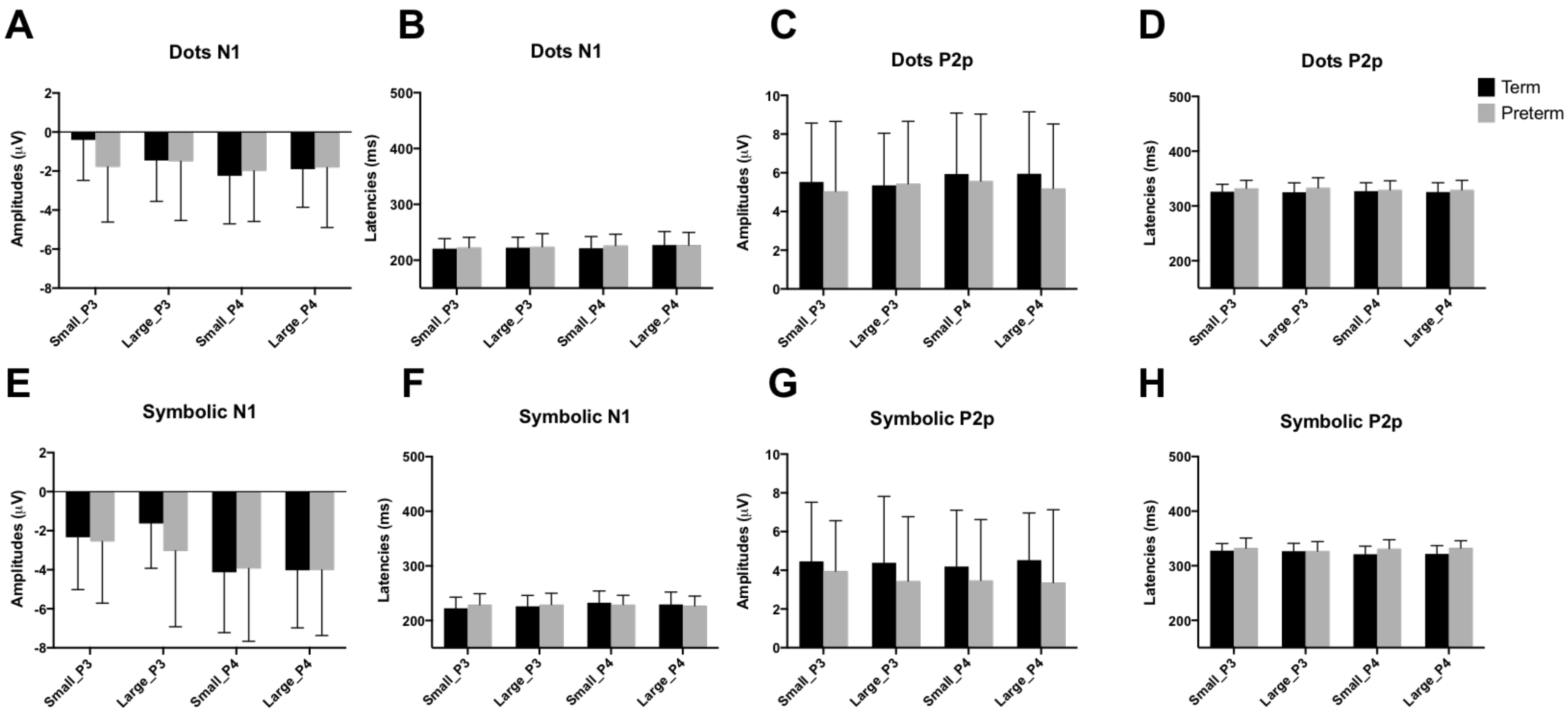


Figure 7.6: Bar charts of amplitudes and latencies for the two comparison tasks

Dots (top row) and symbolic (bottom row) according to group (term and preterm), component (N1, P2p), distance (small and large) and site (P3, for left hemisphere and P4, for right hemisphere). Marginal differences were observed between the term group and the preterm group in the N1 component for amplitudes in the non-symbolic task with a small distance in the left hemisphere (A). Significant differences were observed between groups in the P2p component for latencies in the symbolic task for both large and small distances in the right hemisphere (H). Error bars represent the standard deviation of the data.

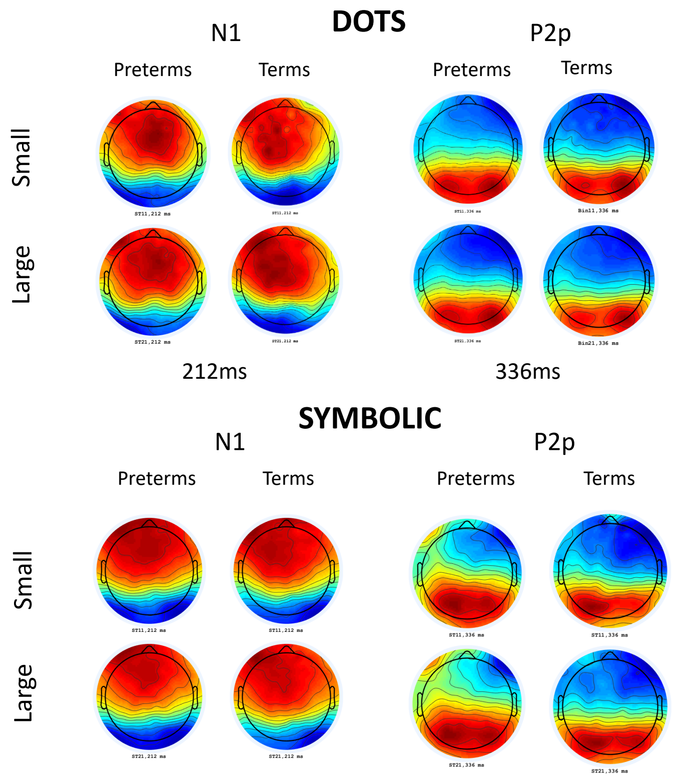


Figure 7.5: Topographic maps corresponding to the grand averaged waveforms from the different groups in the two comparison tasks.

These maps represent the topographic voltage while performing non-symbolic and symbolic magnitude comparison tasks from the different components (N1 at 212 ms and P2p at 336 ms).

7.6 Discussion

To date, this is the first study to investigate the neural correlates of numerical magnitude comparisons using ERPs in a group of VP children. Our goal was to investigate the electrophysiological characteristics of a group of children born prematurely, comparing their ability to determine symbolic and non-symbolic numerical magnitudes with that of their term-born peers. This approach allowed us to examine whether children born prematurely have deficiencies in non-symbolic and/or symbolic numerical magnitude representations. Two main parietal components were identified: the N1 at approximately 220 ms post stimulus presentation, and the P2p at approximately 300 ms post-stimulus presentation. The N1 is typically regarded as the brain's involuntary attention response and the P2p is considered to be a marker of the approximate magnitude determination (Hyde & Spelke, 2012). Our results showed that VP children recruit neuronal resources differently from their term-born counterparts when performing numerical magnitude comparisons. These differences, however, do not indicate that VP children have imprecise representation of symbolic or non-symbolic numerical magnitudes.

One important factor that might limit the scope of the present study is a noticeable delay (100 ms) in the waveforms after stimulus onset across all tasks, conditions, sites and groups. The delay between stimulus onset and the first component was interpreted as a delay in stimulus presentation. Usually, stimulus delay is measured by timing test allowing a comparison between the time when the stimulus presentation system states that the stimulus is presented, and the time that the AV device physically detects the presentation of the stimulus as it would appear to the subject. Due to an equipment fault, however, we were unable to support this hypothesis. As such, we interpreted the 100 ms delay as a delay in stimulus presentation and thus all components have a delay of 100 ms. For example, usually, the N1 component peaks at around 100 ms followed by the P2p component, peaking at around 200 ms. In our study, however, N1 peaked at 212 ms post stimulus presentation, and the P2p at approximately 336 ms post-stimulus presentation.

Another factor to weigh in the interpretation of our results is an unexpected finding regarding the P1 component. After visual inspection of the waveforms, the P1 component had higher amplitudes for the non-symbolic tasks in the preterm group. This component is typically interpreted as a marker of early, non-specific visual processing. Higher amplitudes in the P1 component for the non-symbolic task indicate that the preterm group found the visual stimuli to be more complex than the term group did, although those differences were not statistically significant. Differences were only evident in the non-symbolic task, suggesting that this type of stimulus (e.g., dots) is particularly challenging for children born very prematurely to process. Indeed, very preterm birth is highly linked with visual deficits and visual functions associated with the dorsal visual processing stream, such as global motion perception and visuomotor integration, and this may contribute to the difficulties in learning, particularly in reading and mathematics (Leung, Thompson, Black, Dai, & Alsweiler, 2018). Dot comparison tasks are the most commonly used task to measure the approximate number system in children (e.g., De Smedt et al., 2013), and it continues to be used with children born prematurely despite the fact that this population has high rates of visual difficulties (Hellgren et al., 2013; Guarini et al., 2014; Sims et al., 2015; Libertus et al., 2017). Previous studies have shown, however, that ANS can be assessed not just by visual tasks, such as a dot comparison but also by haptic stimuli[17] (e.g., Gallace, Tan,

[17] Stimuli based on the sense of touch.

& Spence, 2007). For instance, Gimbert et al. (2016) investigated the performance of school-aged children using both a visual and a haptic approximate number processing task. Their results suggest that the ANS can receive inputs from visual and haptic modalities, suggesting an amodal representation of approximate numerosities (Gimbert, Gentaz, Camos, & Mazens, 2016). In order to disentangle whether difficulties in the ANS in the preterm population are due to difficulties in visual processing, as suggested by our findings, future studies should employ other sensory modality tasks such as auditory and tactile tasks. This would make it possible to support whether poorer performance in non-symbolic tasks is associated with a specific modality.

Our analysis focused on two main components: N1 and P2p. The N1 component has been linked with attention and early perceptual processing of the stimuli (Griffin, Miniussi, & Nobre, 2002), but there are also reports that N1 might reflect perceptual discrimination between two different numerical categories; i.e., small and large digits (Xuan, Chen, He, & Zhang, 2009). In addition, when comparing small and large number processing in adults, N1 modulation has been interpreted as reflecting the distribution and maintenance of spatial attention (Hyde & Wood, 2011; Hyde & Spelke, 2012). Thus, our results seem to agree with the literature regarding the physiological role of N1. Regarding differences between groups, the early negative component (N1) showed higher amplitudes in the left hemisphere in the preterm group when performing the non-symbolic task with a small numerical distance, although those differences were only marginal. In general, non-symbolic task modulates activity moderately lateralised to the right hemisphere. Our results, therefore, might indicate that the VP group recruits more contralateral neural resources to successfully perform the non-symbolic task, particularly in more difficult trials, such as where there is a small numerical distance.

Since this component is intrinsically related to attentiveness, it is reasonable to interpret that the preterm group recruited more attentional resources to perform the non-symbolic tasks when there was small numerical distance as a result of a compensatory mechanism, i.e. higher amplitudes and bilateral recruitment.

Typically, the P2p component is a late posterior positivity originate from underlying sources in the inferior parietal regions reflecting domain-specific processing functions (Dehaene, 1996; Pinel et al., 2001), probably representing the recruitment of specific resources for the representation and manipulation of numerical quantities (Dehaene et al., 2003). Our findings revealed that amplitudes for this component did not differ significantly between groups. Previous studies have suggested that children with low mathematical attainment (Gomez-Velazquez et al., 2015) and young adults with mathematical learning disabilities (Soltesz et al., 2007) demonstrate lower amplitudes in this component when performing numerical magnitude, suggesting a failure to recruit adequate neural resources to perform number-related tasks. Our results, however, suggest similar amplitudes for the P2p component for both preterm and term groups. That said, significant differences were observed between groups in respect to latencies in the P2p component. The VP group had significantly longer latencies in the symbolic task in the right hemisphere. Commonly, shorter latencies are a marker of superior mental performance relative to longer latencies (Sur & Sinha, 2009). This might suggest that VP children encode information similarly to their term-born peers but have a delay in processing symbolic information. In addition, typically symbolic stimuli elicit left-lateralised activity (Cappelletti, Barth, Fregni, Spelke, & Pascual-Leone, 2007). The VP group, however, showed longer right-lateralised latencies. This supports the view that the VP group recruited wider neuronal areas to perform the symbolic tasks, implying, as previously discussed, a compensatory mechanism, i.e. bilateral recruitment. This is in line with neuroimaging studies showing wider areas being activated by number-related tasks. For example, Clark et al., (2017) revealed that adults born prematurely showed greater activations in fronto-parietal brain areas than their full-term peers when comparing non-symbolic magnitudes, but had similar performance in behavioural tasks.

Previous studies based on behavioural results have debated whether children born prematurely display deficits in non-symbolic numerical magnitudes due to difficulties in processing number-related information (Hellgren et al., 2013; Libertus et al., 2017) or whether the nature of their difficulties is associated with domain-general skills, such as processing speed (Guarini et al., 2014). In Study 3 (chapter 6), we extensively investigated different outcomes (accuracy, *w* and inverse efficient scores) to elucidate the underlying mechanisms of numerical representations in VP children. Children born prematurely significantly had lower scores compared to their term-born counterparts in all accuracy measures for the non-symbolic magnitude comparison task, regardless of the numerical distance, even after controlling for IQ and excluding participants with low IQ — implying that numerical representation is an area of difficulty in the very preterm group. Differences between groups were only evident in accuracy, but not in reaction time. Previous studies have suggested that VP children have poorer performance in non-symbolic tasks and that this was associated with difficulties in processing speed rather than numerical representations. For example, Guarini et al. (2014) revealed that extremely preterm children were as accurate as the term-born children, but were significantly slower, suggesting difficulties in processing speed. Our results suggest that VP children were as fast as their term-born peers in the non-symbolic task, but not as accurate. This might indicate that difficulties related to non-symbolic numerical representation could be explained by difficulties in inhibition rather than processing speed. The results from the ERPs seems to support this view. In the non-symbolic task, similar latencies were observed between groups, although the preterm group had significant poorer behavioural performance. This might suggest that the preterm group has difficulties in error monitoring, typically associated with inhibitory skills, at least in respect to the non-symbolic task. Conversely, our results indicate that later latencies were observed when the VP group was encoding symbolic information, but no significant differences between groups were observed in behavioural performance. The contrast between ERP and behavioural results in the symbolic tasks could potentially indicate that the VP group have better error monitoring in the symbolic task, implying later latencies reflecting more accurate answers.

One way of disentangling inhibitory skills and numerical processes is to employ the non-symbolic numerical Stroop paradigm. In this type of task, to ensure that participants solve dot comparison tasks based on the number of elements rather than visual characteristics, the task consists of both congruent and incongruent trials. In the congruent trials, visual cues such as the average dot size and convex hull of the array are positively correlated with numerosity, i.e. the array with more dots is made up of larger dots and covers a greater area. In incongruent trials, average dot size and the convex hull of the array are negatively correlated with numerosity, i.e. the array with fewer dots is made up of larger dots and covers a greater area. For a participant to respond accurately to an incongruent dot comparison task trial, they must inhibit the irrelevant and misleading visual information, such as dot size and convex hull, and respond solely based on numerosity estimations (Gilmore et al., 2013; Szucs et al., 2013; Clayton and Gilmore, 2015). Another paradigm that allows to disentangle inhibitory skills and numerical processes is the numerical Stroop paradigm (Szucs et al. 2007, Soltesz et al., 2011). In a numerical Stroop task, participants have to compare the physical magnitudes of the presented numbers (i.e., their font-size). Reaction times are slower and error rates higher whenever physical (i.e., font-size) and numerical magnitude information lead to opposing response biases (e.g., 5 3) (Klein et al., 2014). Studies investigating the non-symbolic numerical Stroop paradigm and the numerical Stroop task in the preterm population are sparse. Using fMRI, research has shown that lower gestational age was related to more activation in frontal cortex areas when performing the numerical Stroop paradigm, suggesting a compensatory neural recruitment (Klein et al., 2014). Compensatory neural recruitment reflects wider brain activations when similar behavioural responses are observed. Klein et al. (2018) demonstrated that while performing numerical Stroop tasks, VP children activated brain areas typically attributed to cognitive control. Future studies investigating the performance of the non-symbolic numerical Stroop tasks could elucidate whether difficulties in performing non-symbolic numerical magnitudes are associated with lower inhibitory skills in the preterm population.

One of the limitations of this study is that we only included correct trials in our analysis of the electrophysiological data. Including both correct and incorrect trials in this study would clarify whether the preterm population have atypical mechanisms in error monitoring typically associated with inhibitory skills. Error-related negativity (ERN or Ne) has been associated with the acts of monitoring and error, even when the participant is not explicitly aware of making the error (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001). It was unfeasible to investigate this component in our study, as it would require several incorrect trials to obtain a stable ERP waveform. This would involve including several trials in each task making the tasks intolerable to children. Another limitation of this study is the fact that our analysis was based on data obtained purely from parietal areas. It is true that previous studies investigating numerical magnitudes, including those investigating children with low mathematical achievement, have focused on parietal areas, since late posterior positivities are thought to originate from underlying sources in the inferior parietal regions (e.g., Gomez-Velazquez et al., 2015). However, an atypical fronto-parietal network thought to be responsible for number processing has been reported in children with mathematical learning disabilities in brain structural studies (Rykhlevskaia et al., 2009), functional magnetic resonance (Price et al., 2007; Kuican et al., 2011) and electrophysiological studies (Soltesz et al., 2007; Heine et al., 2013). In addition, recently, Gomez-Velazquez et al. (2017) demonstrated that different mathematical achievement levels are related to various degrees of frontoparietal connectivity. Thus, future studies would benefit from including the analysis of frontoparietal networks when investigating number processing in the preterm population.

Taken together, our data shows that the VP population has different electrophysiological patterns from those of their term-born peers when performing numerical magnitude comparisons. These differences, however, do not indicate that these children have imprecise numerical representations, similar to individuals with mathematical learning disabilities. Individuals with mathematical learning disabilities typically show lower amplitudes, in particular in the late positive component (P2p). Here, we observed, firstly: the preterm group appears to recruit more contralateral areas elicited by tasks. More areas are left lateralised in the non-symbolic task and right lateralised areas in the symbolic task. This suggest an increase of allocated neural resources, the use of less well lateralised regions and thus, the presence of compensatory mechanisms. Secondly: bigger amplitudes were observed in early components (P1 and N1), in particular to the non-symbolic task, implicating greater recruitment of neural resources prior to the encoding phase, related to visual processing and attentiveness, respectively. Thirdly: late components had longer latencies (P2p), particularly in the symbolic task, showing a rather slower cognitive process when encoding information related to tasks; this may be related to error monitoring. Taken together, our results indicate distinctive electrophysiological patterns of numerical magnitudes in a sample of children born prematurely. These differences, however, are not similar to activation patterns in those with imprecise numerical representation.

8 General Discussion

Mathematics skills have become increasingly important in modern jobs. Learning mathematics at an early age is fundamental to ensuring academic success in STEM disciplines (science, technology, engineering and mathematics) and for maximising future integration into professional life (Wang & Goldschmidt, 2003). Individuals with higher mathematical abilities are more likely to have higher incomes, better housing and better jobs in adulthood (Ritchie & Bates, 2013). Conversely, individuals at increased risk of having difficulties with maths are at risk of having low adult wealth. Preterm birth has been associated with decreased wealth in adulthood mediated by lower mathematics attainment in middle childhood (Basten et al., 2015). Research has been investigating contributing factors for difficulties in maths following preterm birth, and this is the focus of this thesis.

Research has suggested that the approximate number system (ANS), a cognitive system that supports the estimation of the magnitude of a group of elements without relying on language or symbols, plays a crucial role in the development of numerical abilities (Jordan, Glutting, & Ramineni, 2010). The precision of a child's ANS has been shown to predict subsequent mathematical achievement in school (e.g., Halberda, Mazzocco, & Feigenson, 2008). As a consequence, numerous studies have investigated the ANS as a potential measure to identify at-risk populations (Piazza et al., 2010).

Previous studies have shown that VP school-aged children had poorer performance compared to their term-born peers in tasks measuring the ANS (Helgren et al., 2013; Libertus et al., 2017). Yet, the exact nature of their difficulties in basic numerical representation remains and it is unclear whether their difficulties with numerical representation are directly associated with their poor performance in mathematical tasks. More work is needed to understand the nature of the difficulties in numerical skills following preterm birth and when those difficulties start to emerge. Numerical skills, however, are likely to emerge from a combination of multiple core abilities, both domain-specific and domain-general. While domain-general skills explain academic abilities in terms of cognitive factors, such as working memory, reasoning, processing speed in all domains of knowledge, domain-specific skills explain maths abilities by identifying numerical factors that underpin the development and execution of many but not all skills employed in mathematics, such as single digit processing, number system knowledge, number line estimation and numerical magnitude comparison.

This thesis contributes to a better understating of the cognitive precursors of formal mathematical skills in individuals born very preterm, exploring both domain-general and domain-specific skills in four different ways. Firstly, we investigated whether early difficulties in the ANS are already observed in the first year of postnatal life following preterm birth. This was achieved through cross-sectional studies assessing numerical sensitivity in infants aged six-months-old (Study 1, chapter 4) and twelve-months old (Study 2, chapter 5). Secondly, we explored the association of cognitive precursors of maths abilities in the first year of life following preterm birth. This was achieved through assessing numerical sensitivity and visual working memory in infants aged twelve-months-old (Study 2, chapter 5). Thirdly, we evaluated the cognitive mechanisms underpinning difficulties in mathematical abilities in VP school-aged children. This was achieved by assessing domain-general skills (intelligence and executive functions, including processing speed, working memory, attention, planning and inhibition), a set of domain-specific skills (symbolic and non-symbolic numerical magnitude comparison) and standardised measures of maths performance (Study 3, chapter 6). Finally, we determined the neural mechanisms underlying numerical magnitude comparison in VP children. This was achieved by employing ERPs to examine differences in the neural resources recruited by VP children when performing symbolic and non-symbolic numerical magnitude comparisons. The following sections will set out the main findings of the studies conducted within this thesis, examine how they contribute to the literature on the effects of prematurity on numerical cognition, and explain the limitations of this work and future directions in the field.

8.1 Number processing in VP infants

It has been proposed that human numerical development starts even before birth. This is supported by findings indicating that during the last trimester of pregnancy foetuses are already able to discriminate numerosity (Schleger et al., 2014). Moreover, de Hevia et al. (2014) found that, already during the first three days after birth, full-term-born infants are able to associate numerical and physical magnitude (de Hevia, Veggiotti, Streri, & Bonn, 2017). Models of neurodevelopment and brain plasticity propose that even subtle perturbations to brain development in early life, when neuronal connectivity is forming at a rapid rate, can lead to significantly altered developmental trajectories and emerging, specific deficits (Karmiloff-Smith, 1998). It is therefore unsurprising that infants who endure the stresses of preterm birth are at increased risk for a wide range of academic and cognitive difficulties, with very and extremely preterm (<32 weeks) infants being at higher risk. It seems reasonable to assume that premature birth might obstruct typical neural processes facilitating numerical development in the human brain and atypical trajectories of number processing would be expected. Studies investigating the development of numerical abilities in the preterm population are very limited. Thus, the first question I addressed in this thesis was: *Do infants born very prematurely show imprecise numerical representations?*

Employing a number familiarisation task, we explored VP infants' abilities to discriminate numerosities in infants aged six-months (Studies 1, chapter 3) and twelve-months (Study 2, chapter 5). This is the first study to investigate numerical sensitivity in VP infants. We hypothesised that if VP infants already display difficulties in the early stages of numerical development they would not be able to demonstrate numerical sensitivity. Our results showed that, to some extent, VP infants were sensitive to numerical differences and, and at least during the first year of postnatal life, VP infants show the ability to discriminate numerosities. Although our results generally did not replicate previous findings exploring numerical sensitivity in term-born infants (e.g., Xu & Spelke, 2000, Xu & Arriaga, 2007), it is plausible to conclude that VP infants did not display difficulties when discriminating numerosities. For example, in Study 1, VP infants looked significantly longer towards novel numerosities on the second test trial, indicating they were able to discriminate between two sets of elements with different quantities. In addition, VP infants displayed significantly longer looking time towards novel numerosities than term-born infants in Study 2. All things considered, our results indicate that preterm infants display numerical sensitivity within the first year of life.

The concept that infants display numerical abilities is based on empirical evidence from studies using paradigms assessing the ANS. Different modalities have been employed in these studies, such as visual paradigms using looking time measures, auditory paradigms testing sequences of tones, and cross-modal paradigms (e.g., auditory and visual, or visual and tactile) either concurrently or consecutively (Feigenson et al., 2004). A recent meta-analysis of studies with a range of modalities to the examination of numerical abilities in infants (Smyth & Ansari, 2020) revealed that infants can discriminate between both small and large unimodal and cross-modal numerosities. Studies investigating large numerosities are statistically underpowered, however. Thus, adequately powered replication studies are crucial to enable stronger inferences from the infant data to ground theories concerning the ontogenesis of numerical cognition. In line with this view, our results should be cautiously interpreted, given the small sample size and difficulties we faced replicating results from previous studies.

8.2 Number processing and domain-general skills in VP infants

Mathematical abilities emerge from a combination of multiple foundation abilities, both domain-specific and domain-general. Working memory and number sense have been pointed to as contributors to mathematical performance (e.g., Rotzer et al., 2009; Landerl, Bevan and Butterworth, 2004, respectively). A large body of research has revealed that visual working memory is a domain-general skill that plays a key role in the development of maths abilities (e.g., Friso-van den Bos et al., 2007; Rotzer et al., 2009). On the other hand, number sense, a domain-specific skill, has also been associated with the development of mathematical skills over time (Chen and Li, 2014, Fazio et al., 2014; Schneider et al., 2017). Recent research has investigated the associations between the different components of working memory and the performance in number sense in typically developing children. Among the components of working memory, the central executive, with its distinct functions, has been the most studied. The updating function is most commonly pointed to as a predictor of number sense (Kroesbergen et al., 2009; Lee et al., 2012), and a meta-analysis has also revealed an association between number sense and the update and inhibition functions (Friso-van den Bos et al., 2013). These two cognitive factors that play a crucial role in the development of mathematics skills have been unexplored in the VP population during infancy, however. Thus, my next question was: *Is numerical sensitivity in VP infants associated with their visual working memory capacities?*

Our results revealed that at twelve months old, the performance of VP infants in a numerical sensitivity task was not associated with their visual working memory abilities. Further, no significant differences between VP infants and their term-born peers were observed in the performance of visual working memory or numerical sensitivity. It might be plausible to assume that these two distinctive domains are dissociated at this stage of development, but that in later stages of development they become more integrated and play important roles in the development of mathematical abilities. In addition, their predictive value in tasks evaluating domain-specific and domain-general in mathematics performance is comprehensively influenced by the stage of development. For example, Gimbert et al. (2019) showed that numerical discrimination was a significant specific predictor of mathematics achievement only in five-year-olds and that working memory was a significant general predictor only in seven-year-olds, suggesting that working memory becomes a stronger predictor of mathematics achievement after entrance into formal schooling, whereas number sense acuity loses predictive power. Thus, future work should focus on the development of both domain-general and domain-specific abilities and how their interaction helps the development of numerical skills in the VP population.

8.3 Future studies on number processing in VP infants

The ability to discriminate numerosities increases during the first year of life and continues throughout childhood (Halberda &

Feigenson, 2008). Whereas new-borns require a 1:3 ratio to detect differences in numerosity (Izard et al., 2009), by six months of age babies can handle a 1:2 ratio, and by nine months a 2:3 ratio (Xu & Spelke, 2000; Lipton & Spelke, 2003; Libertus & Brannon, 2010). Here, we explored differences in numerosity in VP infants aged six and twelve months using the age equivalent ratio to detect numerical differences (1:2 ratio and 2:3 ratio, respectively). Other numerical abilities already observed during infancy and the early years of childhood, however, remain remarkably unexplored in the VP population. For example, by the age of eleven months, we know that infants are able to discriminate more and less, in an ability known as ordinality. Toddlers with Fragile X syndrome, a syndrome with high prevalence of numerical deficits, demonstrated a significant impairment in ordinality (Owen, Baumgartner and Rivera, 2013). Cardinality is an ability demonstrated by the age of three years, when toddlers are able to make one to one correspondence (Geary, 2000). Cardinality is another stage in the development of numerical abilities that has been remarkably unexplored in individuals born prematurely. Thus, other stages of the development of numerical abilities alongside domain-general skills need to be further explored in the VP population in order to identify the nature of their difficulties and when those difficulties start to emerge.

Although the ability to discriminate large numerosities might not be impaired in VP infants, as our results seem to indicate, the discrimination between numerosities is made in large and small numbers (Cordes & Brannon, 2008). Here, we just evaluated the infants' ability to discriminate large numerosities. The ability to discriminate small numerosities can be later related to the ability to enumerate small sets of up to three elements quickly in a feature called subitising. This area of research remains unexplored in the VP population and future studies should explore it.

Furthermore, future studies need to be longitudinal in nature in order to reveal unique predictors of mathematical difficulties of children born prematurely, as well as identifying when difficulties start to emerge. Since the trajectories of mathematical abilities are likely to depend on a combination of multiple foundations, both domain-specific and domain-general skills need to be systematically investigated in longitudinal studies. For example, Xenidou-Dervou et al. (2017) conducted a longitudinal study in typically developing children to assess their domain-general and mathematics-specific early cognitive abilities as well as their general mathematics achievement. It was found that a constellation of multiple cognitive abilities contributed to the children's mathematical performance. Latent growth modelling revealed that working memory abilities, IQ, counting skills, non-symbolic and symbolic approximate arithmetic and comparison skills explained individual differences in the children's mathematics achievement test. Unexpectedly, however, only one out of all the assessed cognitive abilities was a unique predictor of the children's individual growth rates in mathematics achievement: their performance in the symbolic approximate addition task. Xenidou-Dervou et al.'s (2017) results highlight the importance of identifying domain-general and mathematics-specific cognitive skills in children at risk of struggling with mathematics, and the benefit of employing longitudinal studies to identify the unique predictors that contribute to struggles with maths (Xenidou-Dervou, Molenaar, Ansari, van der Schoot, & van Lieshout, 2017).

8.4 Number processing in VP children

Of all academic subjects, children born prematurely are most likely to have difficulties with maths. While approximately 15% of term-born children have difficulties learning maths, a 2.3-fold increase is found in children born prematurely. Importantly, and in contrast to reading difficulties, maths difficulties are not explained by the very preterm children's lower IQ (Johnson et al., 2011). This suggests that there is something about maths that children born preterm find especially difficult. Thus, my following question was regarding the nature of maths difficulties in the preterm population: *Are difficulties in mathematical performance in VP children associated with domain-general or domain-specific skills?*

We explored group differences between VP and term groups in domain-general, domain-specific skills and mathematical outcomes. Our study replicated previous findings showing that VP children had lower scores compared to term-born children in several domains. Domain-general difficulties in VP children were evident in intelligence, visual and verbal working memory, processing speed, inhibition and planning. Domain-specific difficulties in VP children were evident in non-symbolic representations indexed both by accuracy, *w* and inverse efficiency score. Similar performance was observed between groups in respect to attention and all measures of accuracy for the symbolic comparison task. No significant differences were observed in measures for time of response for any task. As expected, VP children also had lower scores compared to their term-born peers in standardised measures for maths performance.

Although difficulties in numerical representation were evident in our group of VP children, they did not account for difficulties in maths performance. This was further explored by investigating whether difficulties with maths in VP children were driven by domain-general or domain-specific skills. Previous studies have been inconclusive about the nature of maths difficulties in the VP group. For example, Simms et al. (2015) revealed that difficulties in maths in a group of VP children were driven by deficits in domain-general EF. Conversely, Libertus et al. (2017) showed that the nature of problems with maths in children born prematurely were the result of imprecise numerical representations. Using different methods, our results are in line with the first of these contentions: that mathematics abilities in VP children are associated with domain-general skills (Simms et al., 2013b; Guarini et al., 2014; Simms et al., 2015; Tinelli et al., 2015). Our results indicate that domain-general skills including verbal and non-verbal intelligence, working memory, processing speed, attention, planning and inhibition are the best predictors of mathematical performance, beyond domain-specific skills. Simms et al. (2015) showed that domain-general skills contributed significantly to mathematical performance when assessing VP children, explaining 72% of the variance in mathematical performance. Our study replicates this, showing that domain-general skills explained 64.5% of the variance in mathematical performance. In fact, both studies suggest that domain-specific skills explain very little of the variance in mathematical performance in the VP group. For example, in our study, reaction times from symbolic and non-symbolic tasks were the main domain-specific outcomes that accounted for mathematical performance, with a decrease in reaction time explaining an increase in maths performance of approximately 11%. In contrast, the same variable only explained 6% of the performance on the term group. It is important to note, however, that only domain-general components were significant predictors of maths performance for both groups.

Another point we addressed in our study was regarding numerical representation in the VP population. We extensively explored measures of non-symbolic numerical magnitude comparisons in VP children (accuracy, w and IES). This revealed that VP children were as fast as their term-born peers, but not as accurate.

Differences between groups were particularly notable in the non-symbolic task, but no significant differences were found in the symbolic task. This might suggest that, in our study, difficulties related to numerical representation could be explained by difficulties in inhibition, implying faster answers but not as accurate as the term group. This suggests that inhibition, a domain-general skill, plays a particularly important role in non-symbolic comparison performance in the VP population. In line with previous results (Simms et al., 2015; Guarini et al., 2014), we suggested that VP children do not have imprecise numerical representations, but rather difficulties in domain-general skills that affect their performance in tasks assessing numerical representations.

The outcomes from our study have direct implications for educational settings. The ultimate aim of this study was to provide evidence regarding the underlying mechanisms of mathematical difficulties in children born prematurely and what are the better strategies to provide interventions. Based on the current results and previous literature, interventions targeting general cognitive problems, rather than numerical representations should be considered (Simms et al., 2015). The type of interventions that would most benefit VP children are discussed further in the next section.

8.5 Future studies on number processing in VP children

8.5.1 Improving mathematical performance in VP children

It remains unclear when VP children would benefit most from interventions taking place. Developmental models of neurodevelopmental disorders and brain plasticity would indicate that the earlier a targeted intervention is commenced, the greater the potential for altering atypical trajectories (Karmiloff-Smith, 1998). It is argued that if we can detect early difficulties before the full range of difficulties associated with mathematical difficulties has emerged, it may be possible to develop interventions that limit the impact of early difficulties on mathematical attainment. Future studies investigating the impact of prematurity in numerical cognition would therefore benefit from investigating when difficulties start to emerge and different types of interventions in this population struggling with maths.

Interventions based on domain-general skills have been proposed for individuals with mathematical difficulties. For example, interventions focused on training EFs may be a useful method of improving mathematical achievement. Thorell et al. (2009) showed that preschool children who received five weeks of working memory training improved inhibitory control in the trained abilities. No differences were observed in other inhibitory control tasks, however, nor to other executive tasks including, e.g., working memory or problem-solving tasks (Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009). As such, studies training merely on generic EF tasks have been criticised as improving performance on tasks that are rarely transferred to non-trained tasks (Spierer, Chavan, & Manuel, 2013). EF training based on tasks unrelated to real context are unlikely to be effective (Bryck & Fisher, 2012; Moreau & Conway, 2014; Jaroslawska, Gathercole, Allen, & Holmes, 2016). It has been proposed that this lack of transfer from EF training may reflect the domain-specific ways in which information is processed. Training domain-general skills (such as working memory) may not have as much impact on the control of knowledge as training these EF skills within a target domain (such as mathematics). Thus, interventions based on domain-general skills should consider embedding general domains training within the learning domain. A recent study demonstrated transferable skills after training in domain-general skills (inhibition) embedded within the learning domain, providing evidence that domain-specific inhibitory control intervention contributes to benefits that are transferable to academic achievement, including mathematical performance (Wilkinson et al., 2019). VP children with difficulties in mathematics could benefit from similar interventions. Future studies should explore whether this type of intervention would help VP children struggling with maths.

Another type of intervention that VP children could benefit from is home numeracy. Home numeracy has been defined as parent-child interactions that include experiences with numerical content in daily-life settings and it is supposed to have a positive impact on calculation or mathematical ability in general (Yıldız, Sasanguie, De Smedt, & Reynvoet, 2018). Previous studies of children in the general population have shown that parent-child interaction and the home numeracy environment both affect children's maths skills (Zippert & Rittle-Johnson, 2018). This area has not been explored in relation to preterm birth and represents a possible route to intervention.

8.5.2 Educators' knowledge of the impact of preterm birth on mathematical achievement

In order to identify VP children at risk of presenting with mathematical difficulties, it is imperative that educators understand the impact of preterm birth on academic achievement. A study based in the UK revealed that teachers and educational psychologists have poor knowledge of the behavioural and academic challenges experienced by children born preterm (Johnson, Gilmore, Gallimore, Jaekel, & Wolke, 2015). Surprisingly, when identifying potential domains where VP present with difficulties, maths was the area most under-recognised by teaching staff and educational psychologists. In addition, education professionals feel unprepared to support VP children in school. Similar results were also found in a recent Canadian study (Church, Cavanagh, Lee, & Shah, 2019). Their results showed that educators are unprepared to address the academic challenges for the preterm child, and training is needed, suggesting that parents and providers need to be prepared to advocate. Since teachers have primary responsibility for providing long-term support for children born preterm, this is of significant public health and educational concern. In order to improve teachers' knowledge of preterm birth, Johnson et al. (2019) developed an interactive e-learning resource designed to improve education professionals' knowledge of long term outcomes following preterm birth, and strategies that can be used to support children's learning (www.pretermbirth.info). Teachers' confidence in supporting children born preterm was also significantly improved after using the resource. It is important to note, however, that many parents are reluctant to inform teachers that their child had a very preterm delivery, because this may single out their child for different treatment (Marlow & Johnson, 2007). Taken together, future work should focus on helping parents, teachers and educational psychologists to understand the impact of prematurity in academic life and translate findings to educational settings.

8.6 Neural correlates of number processing in VP children

Findings from behavioural studies were inconclusive demonstrating that VP children have deficits in numerical representation. A few studies have advocated that VP children have imprecise numerical representation (e.g., Hellgren et al., 2013; Libertus et al., 2019); but other studies have shown that difficulties in numerical representation are due to deficits in domain-general skills such as processing speed (e.g., Guarini et al., 2014) or that VP children do not have imprecise numerical representations (Simms et al., 2015). Thus, the nature of numerical representations in VP children remains unclear. A technique that elucidates the underlying neural mechanisms of numerical representations in VP children is ERPs. Thus, my final question was: *Do VP children have atypical neural correlates of numerical representation (symbolic and non-symbolic)?*

Our results showed that VP children displayed similar neural signatures when encoding numerical information either related to symbolic or non-symbolic magnitude representation, except from delays processing information related to symbolic measures. Sensory and attention resources allocated prior to the encoding phase were significantly different between the term and the preterm groups. VP children showed wider activations in right inferior parietal regions involved in numerical magnitude comparisons. Wider activations, supported by topographic maps, denote that the VP group recruit more neural resources, thus indicating a less localised area, and potentially, a more immature system.

Our findings converge with previous results from fMRI studies suggesting a re-organisation of the functional neural systems that support magnitude processing in individuals born prematurely. For example, wider frontal activation has been associated with lower gestational ages when VP children perform symbolic magnitude comparisons (Klein et al. 2014; Klein et al. 2018). Another study has revealed that adults born prematurely show greater activations in fronto-parietal brain areas than their full-term peers when comparing non-symbolic magnitudes, despite similar behavioural performance (Clark et al., 2017). Based on previous studies and the results presented within this work, the current assumption is that individuals born prematurely present a developmental delay in the specialisation of neural regions recruited related number processing. According to our results, it is plausible to infer that VP children had wider activations, indicating more controlled and effortful, and less automatic (parietal activation), processing of numerical information. This is in line with the fronto-parietal activation shift (Ansari, 2008), whereby children recruit parietal regions, in particular the IPS, to a lesser extent, and frontal regions to a greater extent compared to adults (Ansari et al., 2005; Ansari & Dhital, 2006; Cantlon et al., 2006; Kucian et al., 2008; Holloway & Ansari, 2010). This conclusion is only speculative, however, since we only investigated school-aged children in our experiment, and not number processing in other VP groups (i.e. pre-schoolers or adults). An alternative interpretation is that the allocation of wider neural recourses can be an indication of a compensatory effect, whereby high-functioning individuals born preterm draw on alternative neural regions to maintain a level of behavioural performance equivalent to their full-term peers, similar to what previous fMRI studies reported (e.g., Clark et al., 2017). These interpretations are not mutually exclusive and further studies could support these hypotheses regarding brain mechanisms of numerical processing in VP individuals by assessing different time points.

8.7 Future directions of research into neural correlates of number processing in VP children

The literature surrounding the neural mechanisms of number processing in individuals born prematurely is very limited. Typically, studies are underpowered due to the difficulties in recruiting this population using neuroimaging techniques. In addition, differences in the methodologies employed make comparisons between studies difficult. Thus, the field faces great challenges to increase our current understanding of neural correlates of number processing in VP children.

Cognitive development is thought to depend on the refinement and specialisation of functional circuits over time (Johnson, 2011). Yet little is known about how this process unfolds over the course of childhood and what goes awry in the VP brain. Future studies should focus on growth trajectories of functional brain networks, focusing on the maturation of parietal circuits associated with number processing. For example, studies with a typically developing population suggest an interactive specialisation in which brain regions start with a broad range of functionality and gradually become more specialised for specific functions. Studies employing fMRI have revealed that the involvement of the IPS in numerical representation can be found in children as young as four years old (Cantlon et al., 2006). The recruitment of the IPS during numerical tasks increases during the course of development for symbolic (Ansari et al., 2005) and non-symbolic (Ansari & Dhital, 2006) numerical magnitudes. When comparing adults and six- to seven-year-old children in symbolic and non-symbolic numerical comparison, adults showed greater activity in the superior parietal lobe relative to children for both symbolic and non-symbolic numerical processing (Cantlon et al., 2008). In contrast, children showed greater activity in fronto-parietal networks when number magnitude processing (Kaufmann et al., 2009). Future studies employing similar approaches would allow a better insight into the longitudinal maturation of functional parietal circuits linked with number processing in the VP brain. To date, only one study has investigated volumetric differences in grey matter in the IPS, an area associated with number-related information in VP population. The results showed that adolescents born prematurely have significantly less grey matter in the IPS (Isaacs et al., 2001). It remains unclear, however, whether differences between term and preterm individuals are already observed in early stages of development.

Another technique that would greatly benefit the understanding of neural correlates of number processing in VP children is diffusion tensor imaging (DTI). This technique employs specific MRI sequences, as well as software that generates images from the resulting data that uses the diffusion of water molecules to generate contrast in MR images. As a result, DTI has been used extensively to map white matter tractography in the brain. Studies with typically developing children have revealed connectivity between the parietal and prefrontal cortex decreasing over time. In contrast, connectivity within posterior brain regions, including intra-hemispheric and inter-hemispheric parietal connectivity, as well as parietal connectivity with ventral temporal occipital cortex regions increased over time (Battista et al., 2018). A few studies have suggested that connectivity of the fronto-parietal network is altered in dyscalculics. Two functional connectivity studies have shown hyperconnectivity between the intraparietal sulcus and frontal regions in children with dyscalculia compared to typically developing children, whereas Kucian et al. (2014) showed reduced structural connectivity in the superior longitudinal fasciculus. Given the fact that individuals born prematurely show systematically aberrant structural connectivity, future studies should investigate the structural connectivity networks associated with number processing in VP individuals.

8.8 Neuroconstructivism

Within the neuroconstructivist approach, the basis of cognitive development can be characterised by mutually induced changes between the neural and cognitive levels (Westermann et al., 2007). In other words, the neurodevelopment of infants born prematurely is constrained by underlying brain structures which are, in turn, affected by experience-dependent processes. Thus, the preterm population is not characterised by an initial delay that recovers during development, but by atypical developmental trajectories. The neuropsychological profiles of very preterm infants show a great heterogeneity, depending on neonatal immaturity, medical complications, environmental, relational and social factors. In particular, in the first weeks of life, the sensory development and behaviour of the preterm infant are negatively affected by neonatal characteristics and morbidities, the stressful environment of the neonatal intensive care unit, and social factors which may influence later neurodevelopment leading to complications such as motor delays, global cognitive impairment, visual perception problems, executive functioning deficits, and learning difficulties in school, especially mathematical learning (Sansavini et al., 2011)

According to the neuroconstructivist framework, assessing atypical development in terms of the cascading developmental effects of small perturbations early in the developmental trajectory should result in a better understanding of the impact of prematurity in the numerical cognition. This implies tracing back to infancy the origins of number deficit. These might not be in the number domain directly, but could, for example, be a deficit in processing information. A slower processing speed would affect other domains but to a lesser degree, meaning that these other domains could look normal in subsequent development but may camouflage subtle deficits (Farran & Karmiloff-Smith, 2012). Neuroconstructivism has been employed to understand atypical trajectories of different conditions and numerous genetic disorders (e.g., Williams syndrome) and, to a lesser extent, the impact of prematurity in certain domains. For example, Vandormael et al. (2019) investigated language development in a preterm population (Vandormael et al., 2019). Language difficulties of very preterm children are often associated with early perceptual, cognitive, communicative and motor problems that may have cascading effects on later more complex abilities (Sanvani, Guarani and Caselli, 2011). An important role of speed processing and executive functions is associated with atypical trajectories in language and literacy skills in the preterm population (Guarini, Zuccarini and Sansavini, 2019). In fact, the biggest contribution of this work was to trace back the atypical trajectories of numerical development in the preterm population, an area markedly unexplored.

In summary, our results showed that, during infancy, VP babies are able to discriminate numerosities, but that they take longer time to explore new numerical stimuli. This might reflect a slower processing speed or even a deficit in the visual system in scanning arrays of objects. Later in development, school-aged children who were born prematurely demonstrated difficulties in processing numerical information. We found significant differences between terms and VP children when processing non-symbolic numerical magnitudes. This might be associated with deficits in sensory (more demanding visual stimuli) and attention resources (deficits in inhibitory skills and processing speed). We speculated that faster, but less accurate responses were associated with difficulties in inhibitory skills. In contrast, when processing symbolic information (a visually less demanding stimulus), neural resources were allocated more slowly, but similar behavioural results were observed. Our ultimate goal was to understand whether domain-general or domain-specific difficulties accounted for difficulties in maths. Processing number-related information accounted very little for the performance in mathematics, with different patterns emerging from VP and term children.

Taken together, our results corroborate the neuroconstructivist approach, whereby the maths difficulties of very preterm children are potentially associated with early perceptual and cognitive problems that may have cascading effects on later more complex abilities related to maths. These difficulties may become more apparent after school entrance, at a time when more complex tasks require the allocation of more domain-general executive functions to succeed.

8.9 Strengths of the Study

Our current understanding of numerical abilities in VP individuals has been gradually increasing in recent years. Yet, several aspects remain unclear. There were numerous characteristics of this study that were novel in helping to elucidate the underlying mechanisms of numerical cognition in the VP population. One of the key strengths of the present study was its attempt to identify early difficulties in number processing after preterm birth (Study 1 and 2). To our knowledge, this was the first study to measure the numerical abilities of a high-risk preterm (i.e. VP/EP) infant sample. In addition, in Study 2 we examined the early foundations of numerical abilities exploring domain-specific and domain-general skills during infancy in VP infants. This was the first time a study of this type has been carried out. Furthermore, despite the evidence showing that VP children have deficits in numerical representations, to the best our knowledge, our study was the first to employ ERPs to explore symbolic and non-symbolic numerical comparisons in VP children (Study 4). Neuroimaging techniques allow us to provide additional evidence about where VP recruits neural recourses differently from their term-born peers. Finally, given the fact that we currently face a replication crisis in different subjects of science (Munafò et al., 2017), including in psychology, it could be argued that our results from Study 3, replicating previous results, is a strength of this work.

8.10 Limitations of the study

The current investigation had several limitations that restrict some of the conclusions that can be drawn from the results. Some of these have already been stated in respect to the individual studies. Next, I discuss limitations associated with participant and cohort characteristics as well some methodological considerations.

8.10.1 Participants

The representativeness of our current sample demands cautious interpretation of findings. For instance, VP children were mostly recruited through UCLH. UCLH has close links to UCL, with therefore elevated base rates of parents with higher education degrees and high socio-economic status. For this reason, our sample is also likely to be biased towards higher functioning children with less academic and/or behavioural difficulties. Thus, the sample obtained at this site cannot be regarded as fully representative of a cross-section of the UK population. For instance, the term group had an inflated cognitive performance when compared to the standard population mean as illustrated by an average 111 points on the cognitive scale of the Bayley-III and an average performance of 114 points in an IQ test, when the standard population mean is estimated to be 100 points. Likewise, the performance of the preterm group was also inflated given the results of previous reports. For instance, VP infants had an average of 102 points in the Bayley-III performing at a similar level to the standard population mean. VP children also had similar performance in IQ, with an average performance of 98 points Together, the results indicate inflated cognitive performance in the term and the VP group, both in the infant and the school-aged cohort. Despite this, we observed an average absolute difference of 12.5 points between groups both in developmental measures (Bayley-III) and intelligence measures (WISC-IV), which is in line with results from previous studies. Thus, despite the elevated cognitive performance observed in our sample, the differences between groups remained similar to what previous studies reported and subsequently interpretations should not be ignored.

One notable limitation of the study was the small sample size comprising the infant studies. In line with previous studies, we anticipated that part of our infant data would be unusable due to behaviour such as fussiness. Studies investigating preterm infants' looking time estimated attrition rates of approximately 10% (Rose et al., 2002). Higher attrition rates are observed in infant studies investigating numerical cognition, with one third of data being unusable (e.g., Hyde & Spelke, 2011). Here, our attrition rates were higher than anticipated with approximately 50% of data being dismissed. It is reasonable to speculate that infants born prematurely would display more inattentive behaviour (Brogan et al., 2014), which could reflect in higher rates of missing data. Thus, those VP infants who were unable to perform the task could also be the ones most at risk of showing difficulties in numerical skills later in life. Alternatively, we also speculated that having additional tasks in our experiments increased our attrition rates, with infants being less attentive in some tasks.

8.10.2 Methodological considerations

8.10.2.1 Infant studies

One important caveat in our studies investigating numerical sensitivity in VP infants is the design of tasks. Stimuli designed for tasks tapping into the ANS necessarily are controlled for perceptible variables that otherwise could be confounders. Infants, however, typically do not engage with this type of stimulus (e.g., array of dots), and more often than desired sessions needed to be resumed after infants showed fussiness and we were unable to collect data. More ecological friendly stimuli could prevent loss of data and provide a more enjoyable experience for the infants joining similar studies. This remains a challenge in the field of numerical cognition, however, since the need to control for visual confounders is inevitable.

8.10.2.2 Controversies surrounding the ANS

A major goal of this work was to investigate, in the VP population, the primitive and rudimentary cognitive system that allow humans and other species to discriminate quantities without relying on language, the ANS. In Study 1 (chapter 4) and 2 (chapter 5) we assessed the ANS in VP infants using looking time measures testing numerical sensitivity. In studies 3 (chapter 6) and 4 (chapter 7) we assessed the ANS in VP children employing non-symbolic numerical magnitudes, using behavioural or electrophysiological measures, respectively. Much recent research has focused on understanding the ANS, given that it has been claimed that it is a cognitive system underlying human mathematical competence. Recently, however, issues have been raised regarding the assessment of the ANS, specifically: issues in controlling continuous properties of the ANS; and a lack of evidence for a causal association between the ANS and maths performance.

The most common method of indexing ANS acuity is to use a non-symbolic dot comparison task. Currently, however, there is no standard protocol for creating the dot array stimuli and it is unclear whether tasks that control for different visual cues, such as cumulative surface area and convex hull size, measure the same cognitive constructs. Clayton, Gilmore and Inglis (2019) have recently investigated how the accuracy of non-symbolic magnitude judgements is influenced by visual controls. Their results showed no significant correlation between participants' accuracy scores in trials created with the protocols tested (the Panamath program and Gebuis and Reynvoet's script), suggesting that tasks employing these protocols may measure different cognitive constructs. When generating our stimuli and designing our experiments to assess ANS, we cautiously controlled for visual properties using the script created by Piazza et al. (2010). This method is very widely employed in the literature investigating numerical cognition but, given the considerations discussed here, our results should be interpreted cautiously in line with this recent concern about assessing the ANS.

Another issue related in respect to the ANS is the lack of evidence for a causal association between the ANS and maths performance. In Study 3, one of our goals was to investigate associations between the ANS, measured by non-symbolic magnitude comparison, and maths performance. Our hypothesis was driven by previous studies indicating a causal association between these two measures. A recent *p-curve* was carried out investigating the evidence for the causal links between the ANS and maths performance (Inglis, Batchelor, Gilmore, & Watson, 2017). Their findings indicated that the published literature did not contain evidence of a causal link between performance in ANS tasks and standardised mathematics tests. The authors argued that the lack of evidence for such a link is potentially due to substantially underpowered studies. Addressing this issue, a recent large cross-sectional study (≃1200 children) has been undertaken (Caviola, Colling, Mammarella, & Szűcs, 2020). This, however, also found a lack of evidence for an association between non-symbolic magnitude comparison measures and mathematics achievement. Measures of symbolic number comparison accuracy and spatial working memory were, however, specifically associated with mathematical performance (Caviola et al., 2020). Given the sample size in our studies, our results should also be interpreted cautiously, although, generally our results seem to agree with the literature in pointing to a lack of causal association between ANS and maths performance, either in the preterm group and the term group.

8.10.2.3 Domain-general, domain-specific skills and mathematical outcomes

In chapter 2, I discussed the role of several domain-general skills in the performance of maths skills, such as working memory. For example, visuospatial sketchpad capacity is the best predictor for nonverbal arithmetic tasks in pre-schoolers (Levine, Jordan and Huttenlocher, 1992; McKenzie, Bull and Gray, 2003; Rasmussen and Bisanz, 2005; Simmons, Chris and Horne, 2008). Older children, however, increasingly rely on the phonological loop, a working memory component found to be the best predictor of performance in verbal mathematics problems (Rasmussen and Bisanz, 2005). Given the specific contributions of the three core components of working memory (central executive, phonological loop and visuospatial sketchpad) to the development of mathematical skills, future studies should investigate them during the developmental of mathematical abilities in VP children. Visuospatial skills are important domain-general skills to the development of math skills (Gillian et al., 2017) and are poorly investigated in VP children (but see Simms et al., 2015) and not explored in our study. In relation to domain-specific skills, we only explored numerical magnitude comparison, but other measures might be important in the development of numerical skills and are largely unexplored in the VP population. Finally, different components of maths scores were investigated (numerical operations and mathematical reasoning from WIAT-II), but we generally reported findings from combined scores, rather than disentangling potential differences in performance in these distinctive subcomponents of maths performance.

8.11 Conclusion

The present study investigated the impact of prematurity in numerical cognition in infants and children. The results of the study show that at early stages of development, there was no evidence that VP infants had difficulties discriminating numerosities. In childhood, when children's difficulties with formal mathematical abilities start to emerge, difficulties with numerical representation are already noticeable, especially with non-symbolic representations. These difficulties, however, explained very little of the variance in mathematical performance. Domain-general abilities were the better predictors for mathematical performance, beyond either symbolic or non-symbolic numerical representations. In addition, VP children recruited neural resources similar to their term-born peers when encoding numerical information. Sensory and attentional resources, however, were allocated distinctively from their term-born peers.

Together, our results indicate that difficulties in maths are potentially associated with early perceptual and cognitive problems that may have cascading effects on later more complex abilities related to mathematical performance. Mathematical difficulties are likely to become noticeable only after school entrance, when more complex tasks require the allocation of more domain-general executive functions to succeed. Parental advocacy is needed to engage teachers in the potential problems found among VP children, in order that they may be recognised as different from those of the general population so more appropriate strategies may be taken to support maths education, an important skill for adult life.

Appendix

List of Contents

Appendix 3.1. Ethics approval REC and R&D (Infant Studies)

University College London Hospitals NHS
NHS Foundation Trust

Joint Research Office

Office Location:

FINAL R&D APPROVAL

04/11/2011

Professor Neil Marlow

Dear Professor Neil Marlow,

Project ID: 10/0312 (Please quote in all correspondence)
REC Ref: 10/H0720/80
UKCRN ID: 57812
Title: The UCH Preterm Development Project: growing up after very preterm birth.

Thank you for registering the above study with the UCL/UCLH/RF Joint Research Office (UCLH Site). I am pleased to inform you that your study now has local R&D approval to proceed and recruit participants at University College London Hospitals NHS Foundation Trust.

Please note that all documents received have been reviewed and this approval is granted on the basis of the key documents provided which are ethically approved by the Research Ethics Committee:

Document	Date
REC approval and approved documents	15/02/2011

As Chief/Principal Investigator you are required to ensure that your study is conducted in accordance with the Department of Health's Research Governance Framework for Health and Social Care (2nd edition 2005) and that all members of the research team are aware of their responsibilities under the Framework.

This R&D approval is conditional upon you complying with all requirements of the Research Ethics Committee notice of favourable opinion and other any relevant regulatory bodies.

Please find attached the conditions of the R&D approval and a reminder of your responsibilities as a researcher and ensure that both yourself and the research team are familiar with and understand the roles and responsibilities both as a team and individually.

UCL HOSPITALS

UCL Hospitals is an NHS Foundation Trust comprising: The Eastman Dental Hospital, The Heart Hospital, Hospital for Tropical Diseases, National Hospital for Neurology and Neurosurgery, The Royal London Hospital for Integrated Medicine and University College Hospital (incorporating the former Middlesex and Elizabeth Garrett Anderson Hospitals).

National Research Ethics Service
North West London REC 2

15 February 2011

Professor Neil Marlow

Study Title:	**The UCH Preterm Development Project: growing up after very preterm birth**
REC reference number:	**10/H0720/80**

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

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

Confirmation of ethical opinion

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

Ethical review of research sites

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

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

Conditions of the favourable opinion

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

Appendix 3.2. Parent's information sheet (Infant Studies)

UCL

Institute for Women's Health

University College London Hospitals NHS Foundation Trust

PDP
UCH Preterm Development Project

THE UCH PRETERM DEVELOPMENT PROJECT

Testing infant's attention over the first year

Parent information leaflet

Please read this leaflet carefully – it explains the research study we are currently doing and asks for your help. If you require any further information, please contact Lara Platten-Brown or Charlotte Sanderson.

Taking part is purely voluntary and whether or not you agree to join the study will not change the care you get in any way.

Sparks

Version 2.0 9 October 2011

UCL

The UCH Preterm Development Project

We are sorry to approach you at such a difficult time following the birth of your premature baby. Here at UCH for many years now we have had a very active research programme to help make our care the best it can be. In particular we have been studying the development of babies who are born very early. We routinely follow up all babies from intensive care for at least 2 years after going home to ensure that their development is progressing nicely or we arrange help if there are problems. This is part of our normal service.

In this project we are studying how the development of your baby after birth is related to how your baby grows up after s/he has gone home.

We would be very grateful for your help with this research study.

To help us work out what things affect your baby as he or she grows up, we need to better understand how the brain grows and develops after premature birth, in terms of physical growth, how it is organised and how your baby remembers things and interacts with you as s/he grows up in the first year.

Why this is important –

We know that babies who are born before 32 weeks of gestation (more than 8 weeks early) can develop perfectly normally, but for a significant proportion development is slow and babies may need extra help to develop as well as possible. What we want to know is which aspects of their early development are important as we can then target our treatments better.

We have carried out a lot of studies on older children, which show us how babies develop later in childhood and why some babies develop learning problems. We are working hard to try to develop programs to improve these outcomes. However, these programmes may be much more effective if we can identify the earliest signs of problems. Then we can start working on improving development from when the baby is in the neonatal unit or soon after – to give all premature babies the best start.

What does it mean for you and your baby?

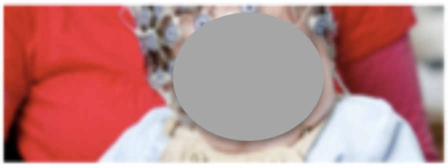
This part of the study starts after you take your baby home. We have developed some measures of memory and of interaction with adults, which we think underlie the problems that many premature children have when they are older. We very much want to find out whether short-term memory and the speed with which babies interpret information are different from those of full term babies.

Your care after discharge is provided locally and also by our follow up service led by Dr Angela Huertas, who you may have met in the neonatal unit. In addition to the routine checks on health and development we would like to carry out some tests of memory, looking and to observe how your baby interacts with adults and computer pictures.

The tests usually continue only as long as your baby continues to show interest in looking. We would also make a video recording of your baby to record where she or he is looking during the recording. Once the study analysis is completed these images are destroyed.

Memory and interpreting information:

These tests are done simply by sitting your baby in front of a television screen and observing how long s/he looks at objects shown there. This is highly automated and runs on a computer programme. During some of the tests we attach small discs on your baby's scalp through which we can record an EEG (or brain wave) trace while the test is running, much like we are with the baby in the picture on the right. Much of this



process is automated and the whole episode takes about 30-40 minute extra during your visit. The EEG sensors are applied using a soft non-irritant jelly, which is easily cleaned off after the test is complete. We will do this on two occasions over the first year using slightly more complex tests when your baby is older, together with some games to work out how well she or he solves simple problems.

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

When your baby is around 12 months of age, we would like to observe how your baby interacts with you and with other adults. We do this in three ways.

First, one of our research team will have a short play session with the child. During this session the researcher will try to prompt your baby to make eye contact and track moving toys. The games are designed to allow observation of your baby's natural behaviour and responses in a social situation. After this, we will ask you to play with your child for a few minutes, in any way you normally would. This will show us how your baby's behaviour changes when interacting with someone they know well compared to less familiar, adults. These observations will give us a better understanding of your baby's general social and communication development.

How does your baby respond to eye contact and faces?

Babies can detect eye contact from a very young age, and this ability is thought to be an important building block for their development and learning. Difficulties detecting eye contact may underlie some later difficulty in social skills.

Typically, babies are attracted to adults' faces if they are looking directly at them and show special brain responses to eye contact. We want to know if babies born preterm are just as sensitive to eye contact.

To do this, we will simply show your baby pictures of faces that we have made on a computer and observe where your baby looks. An example of the pictures we use is shown on the right. While this is happening we would also like to make recordings of your baby's brain waves when looking at the faces and other objects in the same way as we have done over the previous tests.

How does your baby switch their attention?

Finally, we are interested in how flexibly your baby switches his/her attention between objects and events that s/he finds interesting, because the way in which babies visually explore their environment is very important for learning and social development.

To measure this, your baby sits in front of a television screen that will show cartoon pictures. Other moving pictures will appear on the screen, either to the left or to the right of the cartoon. We will then measure how long it takes the child to move their attention between the animations. This will tell us how flexibly babies can switch their attention.

After taking part, we will then discuss later follow up with you, as we would like to discuss later assessments. We will ask your permission to do this at that time.

All of this information is used to help us understand preterm development, and to use when we talk to you and other parents in the future.

What do we need to do now?

After you have spoken to one of the research team and read this leaflet, please read and complete the consent form. The person you have spoken to will also need to sign the consent form to confirm they have spoken to you about the study - so please don't sign the form until you are happy you know what the study entails. We will then speak to you about scheduling an appointment at the EGA Wing (University College Hospital).

Some Frequently Asked Questions – FAQ:

Are there disadvantages or risks to taking part in the study?
There are no risks associated with EEG. The EEG tasks will only continue as long as your baby remains interested in the task.

Are there any benefits to taking part?
This study will not be of any direct benefit to your baby, although we will use the information from the tests to tell you how we think your baby is progressing. If it seems to us that your baby needs extra support we will discuss with you how we provide it.

How can you find out more about the study and about very preterm babies?
At any time you can ask to speak to one of the team working on the study and explain or discuss the study further. More very helpful information about premature babies is available from support organisations such as BLISS, whose address is at the end of this leaflet.

Where will the study take place?
The study will take place in the Elizabeth Garrett Anderson (EGA) Wing of UCLH. You will be reimbursed for travel expenses to and from the hospital.

Will my taking part in this study be kept confidential?
Yes. All information that is collected about you and your baby during the research will be kept strictly confidential. Neither you nor your baby can be identified through the study results.

What will happen to the results of the research?
The results of this research will be published on the website, will be described to other doctors at scientific meetings and will be published in medical (peer-reviewed) journals. No individual children or families will be identifiable in any of this material.

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

Who can give me further information?

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

Who is organising and funding the research?
The Study Director is Professor Neil Marlow.
Also involved in the project are Miss. Charlotte Sanderson (UCL) and Lara Platten-Brown (UCL).
This study is funded by a national children's medical research charity, **SPARKS**.
In addition, the Wellcome Trust supports the research into EEG analysis. Some of the researchers are partly funded by the UCLH/UCL Clinical Biomedical Research Centre.

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

Appendix 3.3. Written consent (Infant Studies)

University College London Hospitals NHS
NHS Foundation Trust

PDP
UCH Preterm Development Project

Project ID: 10/0312
REC Ref: 10/H0720/80
UKCRN ID: 57812

Patient Identification Number for this trial: __________________

Name of Researcher: Professor Neil Marlow

FORM FOR PARENTAL CONSENT

Please initial each box

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

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

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

Version 4.0 March 2015

UCL HOSPITALS

UCL Hospitals is an NHS Foundation Trust comprising: The Eastman Dental Hospital, The Heart Hospital, Hospital for Tropical Diseases, National Hospital for Neurology and Neurosurgery, The Royal London Homoeopathic Hospital and University College Hospital (incorporating the former Middlesex and Elizabeth Garrett Anderson Hospitals).

Appendix 3.4. Parent's questionnaire (Infant Studies)

PDP
UCH Preterm Development Project

Study Number : ______________
Infant ID: ______________
Infant Gender: ______________
Date completed: ______________

Brain development after very preterm birth

Questionnaire for parents -
About your family

Now that you are going home, we will need to keep in touch with you for appointments. Please could you complete this form so we can get in touch if we need change appointments?

It would also be really helpful if you could also give us some details about your family at home to use in our analysis.

We realise these are personal data.

1. **All the information will be treated in the strictest confidence and will not be seen by anyone outside study.**
2. **The family information will be coded and will be used anonymously in all our analysis**
3. **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 Centre: 020 7679 6060

Thank you very much for your help

Office use only:

Entry 1

Entry 2

Please continue on next page

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): ______________________

Please continue on next page

Section A: Your address and contact details

Please tell us your Home Address	
Postcode	
Your telephone number at home	
Your mobile or cellphone 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	

Please continue on next page

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 cellphone number	
Their email address	

Please continue on next page

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

9 **What is your ethnicity? (Please tick as appropriate).**

	You		Your partner	
White				
English/Welsh/Scottish/Northern Irish/British	☐	1	☐	1
Irish	☐	2	☐	2
Gypsy or Irish Traveller	☐	3	☐	3
Any other white background	☐	4	☐	4
Mixed/Multiple Ethnic Groups				
White and Black Caribbean	☐	5	☐	5
White and Black African	☐	6	☐	6
White and Asian	☐	7	☐	7
Any Other Mixed/Multiple Ethnic Background	☐	8	☐	8
Asian/Asian British				
Indian	☐	9	☐	9
Pakistani	☐	10	☐	10
Bangladeshi	☐	11	☐	11
Chinese	☐	12	☐	12
Any Other Asian Background	☐	13	☐	13
Black/African/Caribbean/Black British				
African	☐	14	☐	14
Caribbean	☐	15	☐	15
Any Other Black/African/Caribbean Background	☐	16	☐	16
Other Ethnic Group				
Arab	☐	17	☐	17
Any Other Ethnic Background	☐	18	☐	18

Please continue on next page

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

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.

Please continue on next page

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

Please continue on next page

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 3.5. Invitation to VP families to join children's study

uclh

Subject: Study invitation

Date

Dear Parents of

As you will be aware UCLH neonatal service has been at the forefront of research over the past 30 years. Much of the research we are currently involved with is about how children grow up after being born prematurely. We have a series of studies running at the moment and would like to invite you and xxxx to join us in a study of the effect of premature birth on mathematical skills. We are interested in this area as it seems to be the area in which ex-premature children struggle most at school and also it seems that they have a different set of problems with maths compared to those born at full term.

I am enclosing with this letter an information sheet and consent form for this contact, which we are running with our colleagues in the Institute of Child Health. Because xxx was born very prematurely, we would like to invite you to take part. We would be grateful if you would return the enclosed form in the freepost envelope provided even if you decide not to take part. If we have not heard from you within a month we will write again.

I am sorry it means a day away from school or holidays but, if you want to take part, we will try to be very flexible about the timings of this; the research fellow, Merari Ferreira, will explain that when she rings you.

I do hope that the leaflet and information therein is self explanatory but please feel free to contact me or Merari if you have any questions.

Yours sincerely,

Neil Marlow
Consultant Neonatologist
Professor of Neonatal Medicine

Appendix 3.6. Consent to contact VP families to join study (Children's Studies)

CONSENT FOR CONTACT

☐ I have read the study invitation and AGREE to the research team to contact myself to arrange a visit to the university with my child.

☐ I have read the study invitation and DON'T AGREE to the research team to contact myself to arrange a visit to the university with my child.

Please provide the following details and return this form if you want us to contact you to arrange the study assessment with your child. We have provided an envelope with a freepost.

My child's name: ………………………………………………………………………

My child's date of birth: …………………………………………………………………

Parent or Guardian name: …………………………………………………………………

Relationship to child: ……………………………………………………………………

My home telephone number: ………………………………………………………………

My mobile number: ………………………………………………………………………

My email: ……………………………………………………………………………

Thank you for your time

Appendix 3.7. Parent's information sheet (Children's Studies)

Neil Marlow DM FMedSci
Professor of Neonatal Medicine

PARENT INFORMATION SHEET

PRETERM'S NUMBER PROCESSING RESEARCH

Researchers: Merari Ferreira, Dr Michelle de Haan and Professor Neil Marlow

We are contacting you to invite your child to take part in the Preterm's Number Processing Research study. This letter tells you about the study and what will happen if you give permission for your child to take part in it. Before you decide to give your permission, it is important for you to understand why the study is being done and what it will involve. We are aware that your child was not born preterm but we would be grateful if you could read this letter and decide if you can help us.

What is the Preterm's Number Processing Research?
Each year, around 10,000 babies are born very premature (before 32 weeks of gestation, or more than 8 weeks early) in England. Being born very premature can have long lasting effects on a child's development, particularly with regard to their learning and their achievement in school. One area of schooling that many premature children find especially difficult is mathematics. To help us find new ways to help provide appropriate support for premature children in this area we need to understand the kinds of difficulties they may have in more detail. This is what this research aims to do.

We are interested in finding out more about the specific kinds of difficulties premature children have with maths and identifying the underlying causes of these difficulties. From earlier studies, we suspect that these difficulties might be related to the child's memory, attention and the speed with which they process new information, but we are as yet unsure. From these results we can develop ways in which teachers can help them improve their achievement in school.

We would like to invite you and your child to contribute to the evaluation of this project. We are looking for children between 8 and 10 years of age who were not born prematurely (i.e. born at full **term)** to act as a comparison group. We would be grateful if you would consider taking part in this study. More details are overleaf.

INFOPARENTS_CONTROL_VERSION1_20.01.15

Why has my child been invited to take part in the study?
We are inviting primary school children who were born very premature (before 32 weeks of gestation) to take part in the study.

Alongside these children we would also like children who were not born premature to take part. This is so that we can find out if the skills that are related to maths are the same in premature and non-premature children.

Do I have to take part?
It is up to you and your child to decide if you want to take part. If you and your child choose to take part we will ask you to sign a consent form to say you and your child have agreed to take part. You can withdraw at any stage without giving a reason.

What will happen if my child takes part in the study?
If you decide to join the study, you would need to bring your child to a 1-day assessment meeting at our lab locate in the Wolfson Centre, part of the Institute of Child Health. The study will be conducted in a relaxed and friendly atmosphere. During the day your child will be asked to take part in number of different activities that are a mixture of tests to assess his/her maths skills, memory, attention. All the tests are like games and quick puzzles and are similar to the kids of activities your child usually does at school. We have designed a computer based test, where we show your child different pictures made up of dots and numbers. During this we also would record your child's brain waves using an electroencephalograph or EEG. The EEG involves your child wearing a stretchy cap with built in small sensors called electrodes. The cap is quick and easy to fit, and is comfortable to wear; these electrodes measure the brain activity on the scalp surface.

We estimate that the whole assessment meeting will last about 4 hours, including breaks. We will to some activities in the morning and some in the afternoon. If your child becomes tired s/he will be given breaks whenever necessary. You do not need to be present at the assessment, but you are welcome to come and watch if you would like to. Participants will receive £10 voucher and travel expenses.

Will I be given the results of my child's tests?
Yes, we will send you a letter detailing your child's test results around 2-3 weeks after our visit.

Will I need to do anything?
In addition, we will ask you to complete a questionnaire to tell us about how your child is getting on at school, such as which are areas of strength and where there are more difficulties, and short questionnaires about behaviour and attention. We will also ask you some questions about your home and family that are relevant to your child's education.

InfoParents_ControL_version1_20.01.15

What will you do with all the information?

All of the information we collect about you and your child will be treated in the strictest confidence. It will be seen only by our study team and we keep all personal information separate from the research results so that you or your child will not be able to be identified. When the study is finished, we will destroy the original paperwork and any videos we have made. When we publish the results we will not mention children by name and no child will be able to be identified from these reports.

Is there any benefit to taking part in this study?

There is no direct benefit to you or your child for taking part in this study. However, we hope that the results of the study will help us identify ways of supporting premature children in school in the future. You will receive a detailed assessment of his/her maths skills and other abilities related to learning. You may find this information useful to give to your child's teacher. We are happy to explain any questions you have about the results at any time.

Has anyone approved this study?

This study has been reviewed by experts in the field and has been approved by the NHS Research Ethics Committee (REC 15/LO/1687).

What do I need to do now?

Please take some time to decide whether you would like your child to take part in the study. Participation in the study is entirely voluntary. When you have decided, please complete the attached form and return it in the freepost envelope provided to tell us whether or not you would like your child to take part.

Useful Contacts

Chief Investigator:
Professor Neil Marlow

PhD researcher:
Merari Ferreira

This project is funded by:

InfoParents_ControL_version1_20.01.15

Appendix 3.8. Ethics approval REC and R&D (Children's Studies)

UCL
UCL INSTITUTE OF CHILD HEALTH

Great Ormond Street Hospital for Children NHS
NHS Foundation Trust

Joint Research and Development Office
Division of Research and Innovation

18/02/2016

Dear Dr Michelle de Haan,

Project Title	Neural Correlates of Number Processing in infants and children born preterm
R&D Number	14NP02
Protocol version	1
Protocol date	16 September 2015
Funder	Science without Frontiers
Sponsor	University College London (UCL)

This project has been granted Management Approval by the Joint Research & Development Office.

Approval Conditions:

- An ICH Risk Assessment must be submitted to Lesley Alterman, ; for approval within 30 days of the date of this letter. Please contact Lesley for advice.
- NHS Permissions must be in place at UCLH site(s) before research activity can commence.
- Permissions must be sought from Head Teachers or Principal (as appropriate)at Participating Schools
- You must submit an annual report which will be sent to you by the Joint R&D Office when it is due.
- The PI must inform the Joint R&D Office of any changes to the start and end dates of the project, or if there are any changes to the protocol or personnel. At the end of the study the PI will be sent a final report form to complete and return to the Joint R&D Office.

Please be aware that although you have been granted R&D approval you will not be authorised to spend against your award unless there is a signed contract with the research funder / lead site.

Please contact the Joint R&D Office if you require any further guidance or information on any matter mentioned above. We wish you every success in your research.

Yours sincerely,

Manju Agarwal
Research Management and Governance Officer
Joint Research and Development Office

The child first and always

Page 1 of 1 ICH (Non-NHS) R&D Approval v1.2 14NP02

NHS
Health Research Authority
Research Ethics Service
London - Hampstead Research Ethics Committee

20 October 2015

Professor Neil Marlow

Dear Professor Marlow

Study title:	**Neural Correlates of Number Processing in children born preterm**
REC reference:	**15/LO/1687**
Protocol number:	**15/0411**
IRAS project ID:	**167502**

The Research Ethics Committee reviewed the above application at the meeting held on 14 October 2015.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Manager Rachel Eltringham, Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below. .

Conditions of the favourable opinion

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

Appendix 3.9. Written consent (Children's Studies)

UCL

University College London Hospitals NHS
NHS Foundation Trust

PRETERM'S NUMBER PROCESSING RESEARCH

Child name: <<child name>>
DOB: << DOB>>
Study number: <<studynumber>>

CONSENT FORM

Please initial boxes

• I have read the study information leaflet (v1, dated 06.05.2015) and I have had the opportunity to ask questions. I understand why the study is being carried out and what is involved.	☐
• I agree to my child taking part in the study.	☐
• I agree to the research team contacting my child's school to arrange to carry out an assessment with my son/daughter's classmate **(just for parents of children born preterm)**.	☐
• I agree to my child's teacher being approached in confidence for information about my child's school progress.	☐
• I understand that I will be invited to meet the researcher and that after the assessment I will receive a confidential report of the assessment results.	☐
• I understand that no information will be disclosed to anyone outside the study without my permission.	☐
• I agree to part of the assessment being video recorded when required for the purposes of scoring and checking the assessment process.	☐
• I give permission for the study team to access my child's medical notes for information that is routinely collected about my child's care on the neonatal intensive care unit.	☐
• I understand that my child's participation is voluntary and that I am free to withdraw him/her at any time, without giving reason.	☐
• I agree to the study team contacting me in the future to invite my child to take part in a follow-up study.	☐

OR

• I do not agree to my child taking part in the study.	☐

Parent/Guardian name . . . .

Parent/Guardian Signature . . . .

Date . . . . Relationship to child . . . .

Consent parents_v1_06.06.15

Appendix 3.10. Parent's questionnaire (Children's Studies)

PRETERM'S NUMBER PROCESSING RESEARCH

Participant ID: ________________
Participant Gender: ________________
Date completed: ________________
Age of Participant: ________________

Questionnaire for parents - About your family

Thank you ever so much for coming to see us in our lab. Please could you complete this form and bring it along to your visit with us in lab.

We realise these are personal data.

1. **All the information will be treated in the strictest confidence and will not be seen by anyone outside the study.**
2. **The family information will be coded and will be used anonymously in all our analysis**
3. **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: 02079052652

Thank you very much for your help

Please continue on next page

To begin, please can you provide us with the following information:

i) Name of participant 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): ______________________

Please continue on next page

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	

Please continue on next page

Section D: 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 E: 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.

Please continue on next page

2. Has your child ever been assessed by Educational Psychologist? If yes, please specify why he/she has been referred and main outcomes

____

____

____

____

3. Does your child have learning difficulties (e.g. dyslexia, dyscalculia, concerns raised at parents' evening)? If yes, please specify.

____

____

4. Does your child receive special help at school (if yes, please specify how many hours)?

____

____

Please continue on next page

____

____

2.4 Mood or sleep disorders (e.g. depression, excessive sleeping, prone to isolation or withdrawal, persistent eating problems)? If yes, please specify

____

____

2.5 Other chronic medical conditions (e.g. Major: cystic fibrosis, poorly controlled asthma, diabetes, eczema)? If yes, please specify

____

____

2.6 Any sensory deficit (e.g. vision or hearing loss)? If yes, please specify

____

____

3. Did your child ever receive or is currently receiving (if yes, please specify why and when):

   3.1 Speech and language therapy? ____

   3.2 Physiotherapy?____

   3.3 Occupational therapy? ____

   3.4 Psychological or psychiatrist help? ____________________________________

Section H: About your child's education

1. What current year is your child? ________________________

Please continue on next page

11 **Did your child experience any infections or seizures?**

12 **Did your have retinopathy of prematurity or any vision problems? If so, did your child receive any treatment for vision correction?**

Section G: About your child's health

1. How would you describe the general health of your child?

____

____

2. Has your child ever had:

   2.1 Any major injuries or accidents? If yes, please specify

   ____

   ____

   2.2 Neurological condition (e.g. epilepsy, cerebral palsy, stroke)? If yes, please specify

   ____

   ____

   2.3 Developmental disorders (e.g. autism, serious learning and developmental problems ADD/ADHD, Asperger syndrome, dyspraxia)? If yes, please specify

Please continue on next page

Part 1: Questions about your child'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 child's birthweight?** ☐

7 **Did your child need to be admitted to the neonatal unit for any reason?**

Yes ☐ 1
No ☐ 2

If yes, please tell us why

____

8 **If yes, how much time did they spend in the Neonatal Intensive Care Unit? (In days/weeks/months)**

____

9 **Did your child experience any bleeding (Intraventricular hemorrhage -IVH) in the brain?**

____

10 **Did your child need any operation or intervention when he/she was a baby? If so could you explain us why.**

Please continue on next page

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 F: About your child

Please continue on next page

Section C: 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)

5 **Which language is the most predominately spoken in the home?**

Please give us a percentage of the time spoken in the language (rough estimate)

Please continue on next page

Yes	☐	1
No	☐	2

7 **What is your current age?** ☐ Years

8 **What is your partner's current age?** ☐ Years

9 **What is your ethnicity? (Please tick as appropriate).**

	You		Your partner	
White				
English/Welsh/Scottish/Northern Irish/British	☐	1	☐	1
Irish	☐	2	☐	2
Gypsy or Irish Traveller	☐	3	☐	3
Any other white background	☐	4	☐	4
Mixed/Multiple Ethnic Groups				
White and Black Caribbean	☐	5	☐	5
White and Black African	☐	6	☐	6
White and Asian	☐	7	☐	7
Any Other Mixed/Multiple Ethnic Background	☐	8	☐	8
Asian/Asian British				
Indian	☐	9	☐	9
Pakistani	☐	10	☐	10
Bangladeshi	☐	11	☐	11
Chinese	☐	12	☐	12
Any Other Asian Background	☐	13	☐	13
Black/African/Caribbean/Black British				
African	☐	14	☐	14
Caribbean	☐	15	☐	15
Any Other Black/African/Caribbean Background	☐	16	☐	16
Other Ethnic Group				
Arab	☐	17	☐	17
Any Other Ethnic Background	☐	18	☐	18

Please continue on next page

Section B: 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

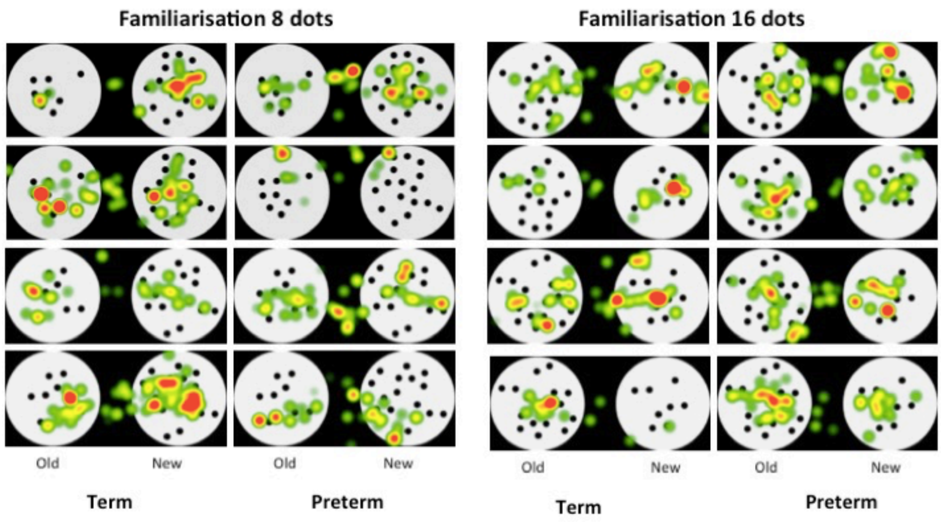
5 **Are you:**

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?** ___

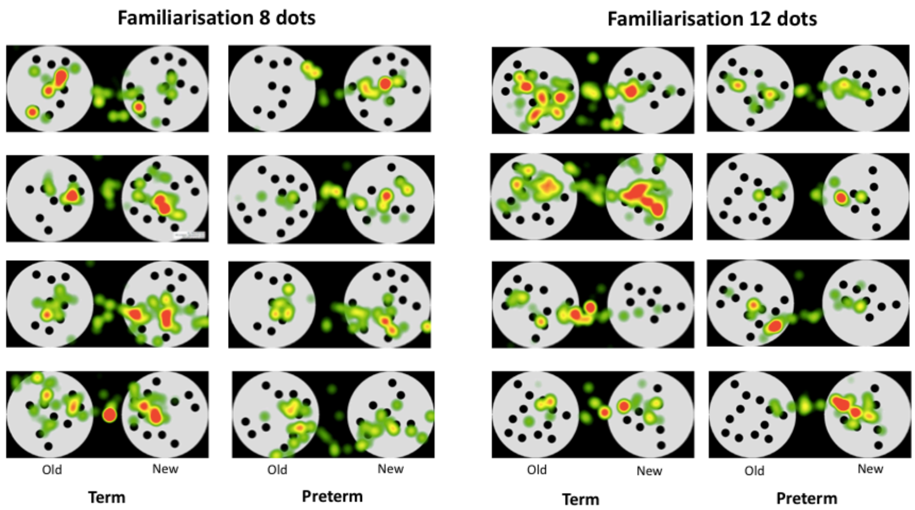
Please continue on next page

Appendix 4.1. Heat maps from eye-tracking data (Study 1)



Eye-tracking heat maps of test trials (chapter 4). Test trials on left-hand side represent looking time from infants who were familiarised with eight dots. Test trials on right-hand side represent fixation looking time from infants who were familiarised with sixteen dots. Test trials were paired between term and preterm groups. The order of the test trials represented here does not represent the real order in the task. Half of the test trials were rotated for illustration purposes.

Appendix 5.1. Heat maps from eye-tracking data (Study 2)



Eye-tracking heat maps of test trials (chapter 5). Test trials on the left-hand side represent looking time from infants who were familiarised with eight dots. Test trials on the right-hand side represent fixation looking time from infants who were familiarised with twelve dots. Test trials were paired between term and preterm groups. The order of test trials represented here does not represent the real order in the task. Half of test trials were rotated for illustration purposes.

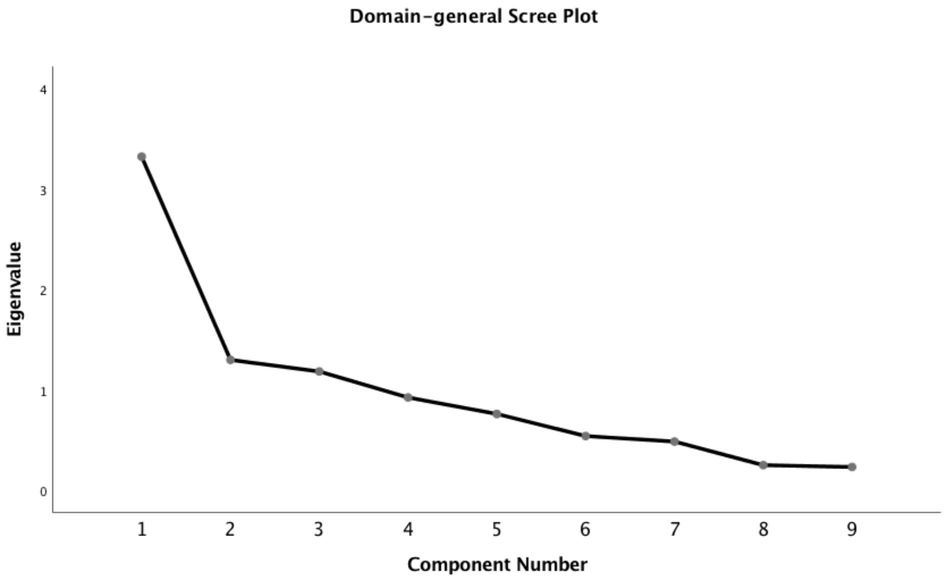
Appendix 6.1. Missing data

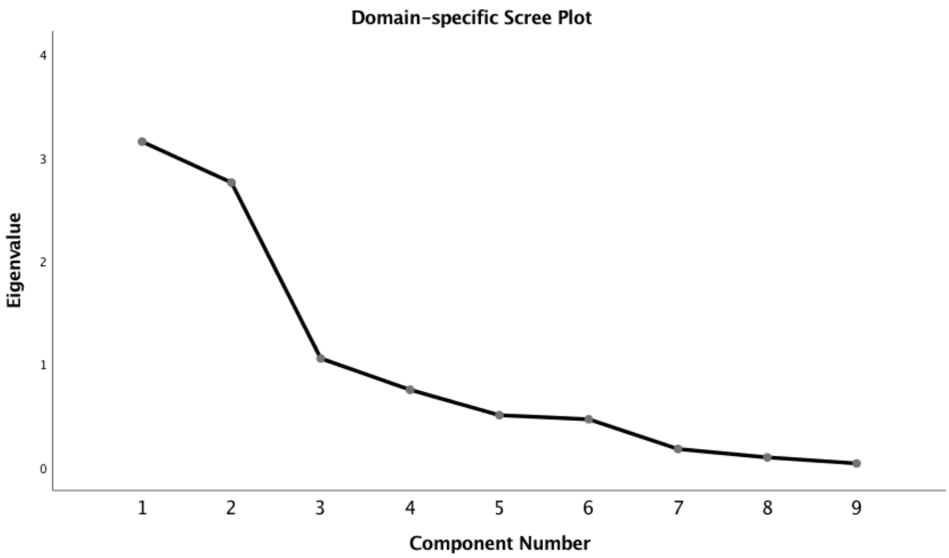
All participants completed the core subtests of the WISC-IV, except for one participant in the preterm group who was schooled in a different language than English and unable to complete the subtests comprising the Verbal Comprehension Index. The participant was not excluded from the sample as schools in the UK teaching in a different language other than English are still obligated to teach a minimum number of hours in English, hence the participant had a good understand of English to performance other tasks, but was unfamiliar with the verbal content comprising the Verbal Comprehension Index. A considerable number of participants in the preterm group (n=13) were not able to performance the Creature Counting from the Tea-Ch. Likewise, a great amount of data (n=20) from both groups were missing from the AWMA due to either equipment fault or license expiry. Due to time constraints not all children completed Single Spelling and Reading from the WIAT-II. Due to equipment fault, one dataset missing in the preterm group from the dot magnitude comparison task was missing. Table below illustrates number of missing datasets for each assessment reported as percentage according to group.

Missing datasets for each assessment.

Measure	**% FT children**	**% VP children**
WISC-IV	0	0
Tea-Ch Sky Search	0	0
Tea-Ch Score!	0	0
Tea-Ch Creature Counting	5.0%	43.0%
Tea-Ch Sky Search DT	0	6.0%
D-KEFS The Colour-Word Interference	2.6%	6.0%
D-KEFS Tower Test	0	0
AWMA	36.0%	20.0%
WIAT-II Maths	0	0
WIAT-II Reading	52.0%	16.0%
WIAT-II Spelling	55.0%	23.0%
Non-symbolic Magnitude Comparison	2.6%	0
Symbolic Magnitude Comparison	2.6%	3.3%

Appendix 6.2. Scree Plots from PCAs





Appendix 6.3. Preterm group performance on domain-general abilities and educational attainment compared to the standard population mean

	Preterm			Difference between VP children and standard population mean			
Test	N	Mean	SD	Mean difference (95% CI)	T	Df	*P*
WISC-IV							
VCI	37	104.46	15.93	4.45 (-0.85 to 9.77)	1.703	36	0.097
PRI	38	99.00	13.58	-1 (-5.47 to 3.47)	-0.454	37	0.653
WMI	38	94.85	13.20	-5.18 (-9.52 to -.084)	-2.42	37	0.021
PSI	38	93.79	13.70	-6.21(-10.77 to -1.65)	-2.75	37	0.009
Full Scale	37	98.86	15.97	-1.13 (-6.46 to 4.19)	-0.432	36	0.668
Tea-Ch							
Sky Search	38	7.89	3.20	-2.105 (-3.16 to -1.05)	-4.052	37	<0.001*
Score!	38	8.79	3.61	-1.211 (-2.40 to -0.02)	-.2.067	37	0.046
Creature Counting	25	9.12	2.63	-0.880 (-1.97 to 0.21)	-1.670	24	0.108
Sky Search DT	36	6.06	3.77	-3.944 (-5.22 to -2.67)	-6.262	35	<0.001*
D-KEFS							
CWIT	35	8.29	3.15	-1.714 (-2.80 to -0.63)	-3.211	34	0.003
Tower Test	38	8.71	2.16	-1.289 (-2.00 to -.58)	-3.667	37	0.001*
AWMA							
Verbal WM	32	47.31	30.62	-2.68 (-13.73 to 8.35)	-0.496	31	0.623
Visual WM	32	45.83	32.06	-4.17 (-15.74 to 7.39)	-0.737	31	0.467
WIAT-II							
Mathematics	38	99.24	19.37	-0.763 (-7.13 to 5.61)	-0.243	37	0.810
Numerical Operations	38	103.76	18.06	3.76 (-2.18 to 9.70)	1.284	37	0.207
Mathematical Reasoning	38	95.21	17.88	-4.78 (-10.67 to 1.09)	-1.651	37	0.107
Word Reading	32	103.82	14.00	3.81 (-1.15 to 8.79)	1.566	32	0.127
Spelling	31	101.19	14.06	1.19 (-3.97 to 6.35)	0.472	30	0.640

*Remains significant after applying Bonferroni correction (α = 0.002).

Appendix 6.4. Term group performance on domain-general abilities and educational attainment compared to the standard population mean

		Term		Difference between term children and standard population mean			
Test	N	Mean	SD	Mean difference (95% CI)	T	Df	*P*
WISC-IV							
Verbal Comprehension	30	115.77	10.19	15.767 (11.96 to 19.57)	8.472	29	<0.001*
Perceptual Reasoning	30	111.47	15.40	11.467 (5.72to 17.22)	4.0778	29	<0.001*
Working Memory	30	104.93	11.24	4.933 (0.73 to 9.13)	2.403	29	0.023
Processing Speed	30	106.40	17.89	6.400 (-0.28 to 13.08)	1.959	29	0.060
Full Scale	30	114.80	12.41	14.80 (10.17 to 19.43)	6.532	29	<0.001*
Tea-Ch							
Sky Search	30	9.60	3.081	-0.4 (-1.55 to 0.75)	-0.711	29	0.483
Score!	30	10.17	3.312	0.167 (-10.7 to 1.40)	0.276	29	0.785
Creature Counting	28	10.36	3.188	0.357 (-0.88 to 1.59)	0.593	27	0.558
Sky Search DT	30	7.47	3.340	-2.533 (-3.78 to -1.29)	-4.155	29	<0.001*
D-KEFS							
Colour-Word Interference Test	29	11.34	2.224	1.345 (0.50 to 2.19)	3.256	28	0.003
Tower Test	30	11.10	2.426	1.100 (0.19 to 2.01)	2.483	29	0.019
AWMA							
Verbal WM	16	64.31	29.30	14.313 (-1.30 to 29.93)	1.953	15	0.070
Visual WM	16	78.44	21.24	28.438 (17.12 to 39.76)	5.354	15	<0.001*
WIAT-II							
Maths	30	124.03	17.46	24.03 (17.51 o 30.55)	7.53	29	<0.001*
Numerical Operations	30	124.70	15.68	24.70 (18.84 to 30.56)	8.62	29	<0.001*
Mathematical Reasoning	30	116.03	15.21	16.03 (10.39 to 21.68)	5.80	29	<0.001*
Word Reading	10	116.60	5.77	16.60 (12.47 to 20.73)	9.08	9	<0.001*
Spelling	9	117.22	11.37	17.22 (8.48 to 25.97)	4.54	8	0.002

*Remains significant after applying Bonferroni correction (α = 0.002).

Appendix 6.5. Children's performance on domain-general and educational attainment excluding participants with IQ <85

	Term			Preterm			Difference between control and VP children		
Test	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
WISC-IV									
VCI	30	115.77	10.19	29	109.93	12.64	5.836 (-0.143 to 11.81)	0.056	0.5
PRI	30	111.47	15.40	29	104.86	7.16	6.60 (0.306 to 12.90)	0.040	0.5
WMI	30	104.93	11.24	29	99.45	9.04	5.48 (0.153 to 10.81)	0.044	0.5
PSI	30	106.40	17.89	29	97.66	11.93	8.74 (0.788 to 16.70)	0.032	0.5
Full Scale	30	114.80	12.41	29	105.86	8.93	8.93 (3.28 to 14.59)	0.002	0.8
Tea-Ch									
Sky Search	30	9.6	3.08	29	8.31	2.88	1.29 (-0.266 to 2.84)	0.102	0.4
Score!	30	10.17	3.31	29	9.52	3.52	0.649 (-1.13 to 2.43)	0.468	0.1
Creature Counting	28	10.36	3.18	22	9.23	2.79	1.13 (-0.601 to 2.86)	0.196	0.3
Sky Search DT	30	7.47	3.34	27	6.63	3.79	0.837 (-1.05 to 2.73)	0.380	0.2
D-KEFS									
CWIT	29	11.34	2.22	28	8.50	3.37	2.84 (1.33 to 4.35)	<0.001*	0.9
Tower Test	30	11.10	2.42	29	8.76	2.08	2.34 (1.16 to 3.52)	<0.001*	1.1
AWMA									
Verbal WM	16	64.31	29.30	23	58.17	27.49	6.13 (-12.48 to 24.76)	0.508	0.2
Visual WM	16	78.44	21.24	23	53.37	30.92	25.07 (6.99 to 43.15)	0.008	0.9
WIAT-II									
Mathematics	30	124.03	17.46	29	106.69	13.65	17.34 (9.15 to 25.53)	<0.001*	1.1
Numerical Operations	30	124.70	15.68	29	109.48	14.94	15.21 (7.22 to 23.20)	<0.001*	0.9
Mathematical Reasoning	30	116.03	15.21	29	102.66	11.51	13.37 (6.35 to 20.40)	<0.001*	0.9
Word Reading	10	116.60	5.77	25	107.84	10.22	8.76 (1.71 to 15.80)	0.016	1.0
Spelling	9	117.22	11.37	25	103.36	13.40	13.86 (3.62 to 24.10)	0.010	1.1

*Remains significant after applying Bonferroni correction (α = 0.002).

Appendix 6.6. Children's performance on domain-specific tasks on CRs, RTs, w and IES excluding participants with IQ <85

		Term			**Preterm**			**Difference between control and VP children**		
Task	Distance	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
Dots										
CRs (%)	Large	30	95.34	3.93	29	90.30	7.47	5.04 (1.94 to 8.14)	0.006*	0.7
RTs (ms)		30	936.42	160.03	29	1000.66	199.56	-64.24 (-158.38 to 29.90)	0.177	0.3
CRs (%)	Small	30	86.30	7.97	29	76.76	7.96	9.54 (5.38 to 13.69)	<0.001*	1.2
RTs (ms)		30	1067.83	184.26	29	1064.47	183.11	3.36 (-92.43 to 99.15)	0.944	0.01
w		30	0.05	0.03	29	0.12	0.10	-0.07 (-0.11 to -0.03)	0.001*	0.9
IES	Large	30	9.82	1.65	29	11.14	2.33	-1.32 (-2.37 to -0.26)	0.015	0.6
	Small	30	12.52	2.80	29	14.00	2.89	-1.48 (-2.9 to 0.003)	0.05	0.5
Symbolic										
CRs (%)	Large	29	98.80	2.06	29	97.97	3.48	0.82 (-0.68 to 2.33)	0.222	0.3
RTs (ms)		29	743.81	176.03	29	847.13	179.93	-103.31 (-196.95 to -9.68)	0.031	0.5
CRs (%)	Small	29	91.45	5.77	29	89.86	7.15	1.58 (-1.83 to 5.006)	0.322	0.2
RTs (ms)		29	906.81	240.98	29	1034.45	231.19	-127.64 (-251.86 to -3.41)	0.044	0.5
IES	Large	29	7.53	1.84	29	8.63	1.78	-1.09 (-2.04 to -0.13)	0.026	0.6
	Small	29	9.92	2.67	29	11.52	2.57	-1.59 (-2.97 to -0.217)	0.024	0.6

CRs= Correct responses; RTs= Reaction time; *w* = Weber fraction; IES = Inverse Efficacy Score. *Remains significant after applying Bonferroni correction (α = 0.003).

Appendix 6.7. Children's performance (*term vs VP)* on domain-specific tasks on RTs, CRs, *w* and IES

		Term			**VP**			**Difference between control and VP children**		
Task	Distance	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
Dots										
CRs (%)	Large	30	95.3	3.9	11	90.34	4.77	5.00 (2.03 to 7.97)	0.010[a]	1.1
RTs (ms)		30	936.42	160.03	11	1005.99	118.83	-69.56 (-176.90 to 37.77)	0.198	0.4
CRs (%)	Small	30	86.3	7.97	11	80.43	8.07	5.85 (0.16 to 11.57)	0.040[a]	0.7
RTs (ms)		30	1067.83	184.26	11	1087.09	113.86	-19.25 (-139.77 to 101.25)	0.748	0.1
w		30	0.05	0.03	11	0.09	0.05	-0.40 (-0.07 to -0.10)	0.010[a]	0.9
IES	Large	30	9.82	1.65	11	11.17	1.53	-1.34 (-2.50 to -0.19)	0.024[a]	0.8
	Small	30	12.52	2.80	11	13.63	2.01	-1.11 (-2.98 to 0.75)	0.236	0.4
Symbolic										
CRs (%)	Large	29	98.8	2.06	11	98.36	2.09	0.43 (-1.00 to 1.86)	0.389	0.2
RTs (ms)		29	743.81	176.03	11	909.15	166.15	-165.34 (-285.66 to -45.02)	0.008[a]	0.9
CRs (%)	Small	29	91.45	5.77	11	91.50	5.40	-0.052 (--3.99 to 3.88)	0.832	0.01
RTs (ms)		29	906.81	240.98	11	1047.50	163.81	-140.69 (-294.78 to 12.40)	0.072	0.6
IES	Large	29	7.53	1.84	11	9.27	1.87	-1.75 (-3.02 to -0.44)	0.010[a]	0.9
	Small	29	9.92	2.67	11	11.54	2.30	-1.61 (-3.40 to 0.16)	0.075[a]	0.6

CRs= Correct responses; RTs= Reaction time; *w* = Weber fraction; IES = Inverse Efficacy Score.
No variable remained significant after applying Bonferroni correction (α = 0.003).
Given the small sample for the VP group, we carried out Mann-Whitney. [a]Significant differences found between groups, but did not remained significant after Bonferroni corrections (α = 0.002).

Appendix 6.8. Children's performance (*term vs EP)* on domain-specific tasks on RTs, CRs, *w* and IES

		Term			EP			Difference between control and VP children		
Task	Distance	N	Mean	SD	N	Mean	SD	Mean difference (95% CI)	*p*	Effect size (Cohen's d)
Dots										
CRs (%)	Large	30	95.3	3.9	26	87.74	9.93	7.60 (3.65 to 11.55)	0.001*	1.0
RTs (ms)		30	936.42	160.03	26	1011.72	244.30	-75.30 (-184.59 to 33.98)	0.173	0.3
CRs (%)	Small	30	86.3	7.97	26	72.54	9.52	13.57 (9.06 to 18.44)	<0.001*	1.5
RTs (ms)		30	1067.83	184.26	26	1064.81	229.20	3.01 (-107.80 to 113.84)	0.957	0.01
w		30	0.05	0.03	26	0.22	0.27	-0.166 (-0.27 to -0.63)	0.002*	0.8
IES	Large	30	9.82	1.65	26	11.64	2.94	-1.82 (-3.08 to -0.56)	0.005	0.7
	Small	30	12.52	2.80	26	14.86	3.51	-2.34 (-4.03 to -0.64)	0.008	0.7
Symbolic										
CRs (%)	Large	29	98.8	2.06	25	96.43	4.81	2.36 (0.392 to 4.33)	0.011	0.6
RTs (ms)		29	743.81	176.03	25	868.77	237.18	-124.95 (-238.05 to -11.85)	0.031	0.5
CRs (%)	Small	29	91.45	5.77	25	86.72	8.86	4.72 (0.69 to 8.76)	0.027	0.6
RTs (ms)		29	906.81	240.98	25	1063.54	302.14	-156.73 (-305.11 to -8.36)	0.039	0.5
IES	Large	29	7.53	1.84	25	9.00	2.45	-1.46 (-2.64 to -0.28)	0.016	0.6
	Small	29	9.92	2.67	25	12.29	3.50	-2.36 (-4.05 to -0.67)	0.007	0.7

CRs= Correct responses; RTs= Reaction time; *w* = Weber fraction; IES = Inverse Efficacy Score.
*Remains significant after applying Bonferroni correction (α = 0.002).

Appendix 6.9. Mathematical performance: actual scores versus predicted scores

We explored differences between the actual scores obtained for mathematics from the WIAT-II and predicted scores from the WISC-IV. Paired t-test between actual scores from the WIAT-II on Numerical Operations, Mathematical Reasoning and Mathematics and predicted scores from the WISC-IV on the same measures were carried out for each group separately. Predicted scores from the WISC-IV were significantly different from the actual scores from the WIAT-II for the term group, showing that the term group over performed the mentioned measures according to their expected IQ. In contrast, the preterm group performed according to what was predicted from the scores obtained from the WISC-IV. The table below illustrates the predicted scores from the WISC-IV for the WIAT-II and the actual scores obtained from the WIAT-II.

Predicted scores from the WISC-IV for the WIAT-II and the actual scores from the WIAT- II

	Predicted scores WISC-IV			**Actual Scores WIAT-II**			
	N	Mean	SD	Mean	SD	Mean difference (95% CI)	*p*
Term							
Mathematics	30	110.63	9.02	124.03	17.84	-15.46 (-20.12 to -10.81)	<0.001*
Numerical Operations	30	109.23	7.84	124.70	15.68	-13.40 (-18.20 to -8.59)	<0.001*
Mathematical Reasoning	30	110.57	8.92	116.03	15.21	-5.46 (-9.45 to -1.48)	<0.001*
Preterm							
Mathematics	37	99.19	12.56	99.54	19.55	-4.51 (-9.13 to 0.110)	0.055
Numerical Operations	37	99.24	11.06	103.76	18.31	-0.351 (-4.58 to 3.88)	0.867
Mathematical Reasoning	37	99.19	12.45	95.73	17.83	3.45 (-0.051 to 6.970)	0.053

*Remains significant after applying Bonferroni correction (α = 0.01).

References

Aarnoudse-Moens, Cornelieke S H, Smidts, D. P., Oosterlaan, J., Duivenvoorden, H. J., & Weisglas-Kuperus, N. (2009). Executive function in very preterm children at early school age. *Journal of Abnormal Child Psychology*, *37*(7), 981–993. https://doi.org/10.1007/s10802-009-9327-z

Aarnoudse-Moens, Cornelieke Sandrine Hanan, Weisglas-Kuperus, N., Duivenvoorden, H. J., van Goudoever, J. B., & Oosterlaan, J. (2013). Executive Function and IQ Predict Mathematical and Attention Problems in Very Preterm Children. *PLoS ONE*, *8*(2), 1–7. https://doi.org/10.1371/journal.pone.0055994

Aarnoudse-Moens, Cornelieke Sandrine Hanan, Weisglas-Kuperus, N., Van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics*, *124*(2), 717–728. https://doi.org/10.1542/peds.2008-2816

Adrian, J., Haist, F., & Akshoomoff, N. (2019). *Mathematics Skills and Executive Functions Following Preterm Birth: A Longitudinal Study of 5-to 7-Year Old Children*. 10.31234/osf.io/j42at

Agrillo, C., & Bisazza, A. (2017). Understanding the origin of number sense: a review of fish studies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *373* (1740), 20160511. https://doi.org/10.1098/rstb.2016.0511

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2009). Use of number by fish. *PloS One*, *4*(3), e4786–e4786. https://doi.org/10.1371/journal.pone.0004786

Ahlheim, C., & Love, B. C. (2018). Estimating the functional dimensionality of neural representations. *NeuroImage*, *179*, 51-62. https://doi.org/10.1016/j.neuroimage.2018.06.015

Akoglu H. (2018). User's guide to correlation coefficients. *Turkish* *journal of emergency medicine*, *18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

Akshoomoff, N., Joseph, R. M., Taylor, H. G., Allred, E. N., Heeren, T., O'Shea, T. M., & Kuban, K. C. K. (2017). Academic Achievement Deficits and Their Neuropsychological Correlates in Children Born Extremely Preterm. *Journal of Developmental and Behavioral Pediatrics*, *38*(8), 627–637. https://doi.org/10.1097/DBP.0000000000000479

Allotey, J., Zamora, J., Cheong-See, F., Kalidindi, M., Arroyo-Manzano, D., Asztalos, E., … Thangaratinam, S. (2018). Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64 061 children. *An International Journal of Obstetrics and Gynaecology*, *125*(1), 16–25. https://doi.org/10.1111/1471-0528.14832

Alloway, T. (2007). Automated Working Memory Assessment (AWMA) London. *UK: Pearson Assessment*.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Anderson, P., Doyle, L. W., Callanan, C., Carse, E., Casalaz, D., Charlton, M. P., … Yu, V. (2003). Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. *JAMA: Journal of the American Medical Association*.289(24),23264-72.
https://doi.org/10.1001/jama.289.24.3264

Anderson, P. J. (2008). Towards a developmental model of executive function. In *Executive functions and the frontal lobes: A lifespan perspective.* (pp. 3–21). Philadelphia, PA, US: Taylor & Francis.

Anderson, P. J. (2014). Neuropsychological outcomes of children born very preterm. *Seminars in Fetal and Neonatal Medicine*, *19*(2), 90–96. https://doi.org/10.1016/j.siny.2013.11.012

Anderson, P. J., & Burnett, A. (2017). Assessing developmental delay in early childhood - concerns with the Bayley-III scales. *The Clinical Neuropsychologist*, *31*(2), 371–381. https://doi.org/10.1080/13854046.2016.1216518

Anderson, P. J., Cheong, J. L. Y., & Thompson, D. K. (2015). The predictive validity of neonatal MRI for neurodevelopmental outcome in very preterm children. *Seminars in Perinatology*, *39*(2), 147–158. https://doi.org/10.1053/j.semperi.2015.01.008

Anderson, P. J., De Luca, C. R., Hutchinson, E., Spencer-Smith, M. M., Roberts, G., & Doyle, L. W. (2011). Attention problems in a representative sample of extremely preterm/extremely low birth weight children. *Developmental Neuropsychology*, *36*(1), 57–73.https://doi.org/10.1080/87565641.2011.540538

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews. Neuroscience*, *9*(4), 278–291. https://doi.org/10.1038/nrn2334

Ansari, D., & Dhital, B. (2006). Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: an event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *18*(11),1820–1828.
https://doi.org/10.1162/jocn.2006.18.11.1820

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *Neuroreport*, *16*(16), 1769–1773. https://doi.org/10.1097/01.wnr.0000183905.23396.f1

Ansari, D., & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: a neuroconstructivist perspective. *Trends in Cognitive Sciences*, *6*(12), 511–516. https://doi.org/10.1016/s1364-6613(02)02040-5

Arsalidou, M., & Taylor, M. J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*,

54(3),2382–2393.
https://doi.org/10.1016/j.neuroimage.2010.10.009

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Association, A. P., & others. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5).* American Psychiatric Publishing.

Aylward, G. P. (2005). Neurodevelopmental outcomes of infants born prematurely. *Journal of Developmental and Behavioral Pediatrics :JDBP*,*26*(6),427–440.
https://doi.org/10.1097/00004703-200512000-00008

Aylward, G. P. (2020). Is It Correct to Correct for Prematurity? Theoretic Analysis of the Bayley-4 Normative Data. *Journal of Developmental and Behavioral Pediatrics*, *41*(2), 128–133. https://doi.org/10.1097/DBP.0000000000000739

Back, S. A., & Miller, S. P. (2014). Brain injury in premature neonates: A primary cerebral dysmaturation disorder? *Annals of Neurology*,*75*(4), 469–486. https://doi.org/10.1002/ana.24132

Baddeley. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/s1364-6613(00)01538-2

Baddeley, A. D. (1993). Verbal and visual subsystems of working memory. *Current Biology : CB*, *3*(8), 563–565. https://doi.org/10.1016/0960-9822(93)90059-w

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, Alan. (1986). Working memory. *Working Memory.* New York, NY, US: Clarendon Press/Oxford University Press.

Baddeley, Alan. (2003). Working memory: looking back and looking forward. *Nature Reviews. Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Balakrishnan, U., Amboiram, P., Ninan, B., Chandrasekharan, A., Rangaswamy, R., & Subramanian, L. (2018). MRI at term equivalent age for predicting long-term neurodevelopmental outcome in preterm infants–a cohort study. *Journal of Maternal-Fetal and Neonatal Medicine*, *7058*. https://doi.org/10.1080/14767058.2018.1532498

Bartelet, D., Ansari, D., Vaessen, A., & Blomert, L. (2014). Cognitive subtypes of mathematics learning difficulties in primary education. *Research in Developmental Disabilities*, *35*(3), 657–670. https://doi.org/10.1016/j.ridd.2013.12.010

Bartelet, D., Vaessen, A., Blomert, L., & Ansari, D. (2014). What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology*, *117*(1), 12–28. https://doi.org/10.1016/j.jecp.2013.08.010

Basten, M., Jaekel, J., Johnson, S., Gilmore, C., & Wolke, D. (2015). Preterm Birth and Adult Wealth: Mathematics Skills Count. *Psychological Science*, *26*(10), 1608–1619. https://doi.org/10.1177/0956797615596230

Battista, C., Evans, T. M., Ngoon, T. J., Chen, T., Chen, L., Kochalka, J., & Menon, V. (2018). Mechanisms of interactive specialization and emergence of functional brain circuits supporting cognitive development in children. *NPJ Science of Learning*, *3*, 1. https://doi.org/10.1038/s41539-017-0017-2

Bayless, S., & Stevenson, J. (2007). Executive functions in school-age children born very prematurely. *Early Human Development*,*83*(4),247–254. https://doi.org/10.1016/j.earlhumdev.2006.05.021

Bayley, N. (2006). *Bayley scales of infant and toddler development*. PsychCorp, Pearson.

Bellman, M., Byrne, O., & Sege, R. (2013). Developmental assessment of children. *BMJ*, *345*(7891), 1–9. https://doi.org/10.1136/bmj.e8687

Beres, A. M. (2017). Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain

Potentials (ERPs) in Language Research. *Applied Psychophysiology Biofeedback*, *42*(4), 247–255. https://doi.org/10.1007/s10484-017-9371-3

Berg, D. H. (2008). Working memory and arithmetic calculation in children: the contributory roles of processing speed, short-term memory, and reading. *Journal of Experimental Child Psychology*, *99*(4), 288–308. https://doi.org/10.1016/j.jecp.2007.12.002

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA*, *288*(6), 728–737. https://doi.org/10.1001/jama.288.6.728

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A. B., … Lawn, J. (2013). Born Too Soon: The global epidemiology of 15 million preterm births. *Reproductive Health*, *10*(SUPPL. 1), 1–14. https://doi.org/10.1186/1742-4755-10-S1-S2

Bobadilla-Suarez, Sebastian. Webers fraction. Retrieved from: https://github.com/bobaseb/webers_fraction, last accessed on 21st May 2019).

Bogale, B., Kamata, N., Mioko, K., & Sugita, S. (2011). Quantity discrimination in jungle crows, Corvus macrorhynchos. *Animal Behaviour*, *82*, 635–641. https://doi.org/10.1016/j.anbehav.2011.05.025

Böhm, B., Smedler, A.-C., & Forssberg, H. (2004). Impulse control, working memory and other executive functions in preterm children when starting school. *Acta Paediatrica, 93*, 1363–1371. https://doi.org/10.1080/08035250410021379

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: evidence from the preschool years. *Journal of Experimental Child Psychology*, *114*(3), 375–388. https://doi.org/10.1016/j.jecp.2012.09.015

Botting, N., Powls, A., Cooke, R. W., & Marlow, N. (1998). Cognitive and educational outcome of very-low-birthweight children in early adolescence. *Developmental Medicine and Child Neurology*, *40*(10), 652–660. https://doi.org/10.1111/j.1469-8749.1998.tb12324.x

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240. https://doi.org/10.1016/S0010-0277(02)00005-7

British Association of Perinatal Medicine. (October 2019). Retrieved from www.bapm.org/resources/80-perinatal-management-of-extreme-preterm-birth-before-27-weeks-of-gestation-2019

Brogan, E., Cragg, L., Gilmore, C., Marlow, N., Simms, V., & Johnson, S. (2014). Inattention in very preterm children: Implications for screening and detection. *Archives of Disease in Childhood*, *99*(9), 834–839. https://doi.org/10.1136/archdischild-2013-305532

Brósch-Fohraheim, N., Fuiko, R., Marschik, P. B., Resch, B., & Liu, J. (2019). The influence of preterm birth on expressive vocabulary at the age of 36 to 41 months. *Medicine*, *98*(6). https://doi.org/10.1097/MD.0000000000014404

Brouwer, M. J., Kersbergen, K. J., Van Kooij, B. J. M., Benders, M. J. N. L., Van Haastert, I. C., Koopman-Esseboom, C., … Groenendaal, F. (2017). Preterm brain injury on term-equivalent age MRI in relation to perinatal factors and neurodevelopmental outcome at two years. *PLoS ONE*, *12*(5), 1–13. https://doi.org/10.1371/journal.pone.0177128

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, *51*(1), 5–13. https://doi.org/10.5334/pb-51-1-5

Bryck, R. L., & Fisher, P. A. (2012). Training the brain: practical applications of neural plasticity from the intersection of cognitive neuroscience, developmental psychology, and prevention science. *The American Psychologist*, *67*(2), 87–100. https://doi.org/10.1037/a0024657

Brydges, C. R., Landes, J. K., Reid, C. L., Campbell, C., French, N., & Anderson, M. (2018). Cognitive outcomes in children and adolescents born very preterm: a meta-analysis. *Developmental Medicine and Child Neurology*, *60*(5), 452–468. https://doi.org/10.1111/dmcn.13685

Bugden, S., & Ansari, D. (2011). Individual differences in children's mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition*, *118*(1), 32–44. https://doi.org/10.1016/j.cognition.2010.09.005

Bugden, S., & Ansari, D. (2016). Probing the nature of deficits in the "Approximate Number System" in children with persistent Developmental Dyscalculia. *Developmental Science*, *19*(5), 817–833. https://doi.org/10.1111/desc.12324

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*(3), 273–293. https://doi.org/10.1207/S15326942DN1903_3

Bull, R., Espy, K. A., & 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), 205–228. https://doi.org/10.1080/87565640801982312

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*. *65*(1):1-24.

https://doi.org/10.1006/jecp.1996.2358

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology*. *15* (3), 421-442. https://doi.org/10.1080/87565649909540759

Bull, R., & Lee, K. (2014). Executive Functioning and Mathematics Achievement. *Child Development Perspectives*, *8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Burnett, A. C., Cheong, J. L. Y., & Doyle, L. W. (2018). Biological and Social Influences on the Neurodevelopmental Outcomes of Preterm Infants. *Clinics in Perinatology*, *45*(3), 485–500. https://doi.org/10.1016/j.clp.2018.05.005

Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.

Butterworth, B.. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *46*(1), 3–18. https://doi.org/10.1111/j.1469-7610.2004.00374.x

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences*, *14*(12), 534–541. https://doi.org/10.1016/j.tics.2010.09.007

Butterworth, B, Reeve, R., Reynolds, F., & Lloyd, D. (2008). Numerical thought with and without words: Evidence from indigenous Australian children. *Proceedings of the National Academy of Sciences*, *105*(35), 13179–13184. https://doi.org/10.1073/pnas.0806045105

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nature Reviews. Neuroscience*, *9*(8), 613–625. https://doi.org/10.1038/nrn2459

Cangöz, B., Olkun, S., Altun, A., & Salman, F. (2018). Comparing two cut-off based criteria while investigating the discriminatory characteristics of a tablet-based dyscalculia screening battery for 5–9 age group. *Noropsikiyatri Arsivi*, *55*(2), 177–182. https://doi.org/10.5152/npa.2017.19255

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2006.01719.x

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology*, *4*(5), 844–854. https://doi.org/10.1371/journal.pbio.0040125

Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009). The neural development of an abstract concept of number. *Journal of Cognitive Neuroscience*, *21*(11), 2217–2229. https://doi.org/10.1162/jocn.2008.21159

Cappelletti, M., Barth, H., Fregni, F., Spelke, E. S., & Pascual-Leone, A. (2007). rTMS over the intraparietal sulcus disrupts numerosity processing. *Experimental Brain Research*, *179*(4), 631–642. https://doi.org/10.1007/s00221-006-0820-0

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*. *33* (3), 386-404.
https://doi.org/10.1016/0022-0965(82)90054-6

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PloS One*, *7*(4), 24–25. https://doi.org/10.1371/journal.pone.0033832

Caviola, S., Colling, L. J., Mammarella, I. C., & Szűcs, D. (2020). Predictors of mathematics in primary school: magnitude comparison, verbal and spatial working memory measures. *Developmental Science*, (2), e12957. https://doi.org/10.1111/desc.12957

Chawanpaiboon, S., Vogel, J. P., Moller, A. B., Lumbiganon, P., Petzold, M., Hogan, D., … Gülmezoglu, A. M. (2019). Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *The Lancet Global Health*, *7*(1), e37–e46. https://doi.org/10.1016/S2214-109X(18)30451-0

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, *148*, 163–172. https://doi.org/10.1016/j.actpsy.2014.01.016

Cho, S., Ryali, S., Geary, D. C., & Menon, V. (2011). How does a child solve 7 + 8? Decoding brain activity patterns associated with counting and retrieval strategies. *Developmental Science*, *14*(5), 989–1001. https://doi.org/10.1111/j.1467-7687.2011.01055.x

Choo, Y. Y., Agarwal, P., How, C. H., & Yeleswarapu, S. P. (2019). Developmental delay: Identification and management at primary care level. *Singapore Medical Journal*, *60*(3), 119–123. https://doi.org/10.11622/smedj.2019025

Church, P. T., Cavanagh, A., Lee, S. K., & Shah, V. (2019). Academic challenges for the preterm infant: Parent and educators' perspectives. *Early Human Development*, *128*, 1–5. https://doi.org/10.1016/j.earlhumdev.2018.09.016

Cipolotti, L., & van Harskamp, N. (2001). *Disturbances of number processing and calculation.* In R. S. Berndt (Ed.), *Handbook of neuropsychology. Handbook of neuropsychology: Language and aphasia* (p. 305–331). Elsevier Science Publishers B.V.

Clark, C. A. C., Liu, Y., Wright, N. L. A., Bedrick, A., & Edgin, J. O. (2017). Functional neural bases of numerosity judgments in healthy adults born preterm. *Brain and Cognition*, *118*, 90–99. https://doi.org/10.1016/j.bandc.2017.07.011

Clark, C. A. C., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology*, *46*(5), 1176–1191. https://doi.org/10.1037/a0019672

Clayton, S., Inglis, M., & Gilmore, C. (2019). Developmental differences in approaches to nonsymbolic comparison tasks. *Quarterly Journal of Experimental Psychology*, *72*(3), 436–445. https://doi.org/10.1177/1747021818755296

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. *Neuron*, *53*(2), 307–314. https://doi.org/10.1016/j.neuron.2006.12.025

Colmar, S., & Double, K. (2017). Working Memory Interventions with Children: Classrooms or Computers? *Journal of Psychologists and Counsellors in Schools*, *27*(2), 264–277. https://doi.org/10.1017/jgc.2017.11

Cordes, S., & Brannon, E. M. (2009). The relative salience of discrete and continuous quantity in young infants. *Developmental Science*, *12*(3), 453–463. https://doi.org/10.1111/j.1467-7687.2008.00781.x

Costa, H. M., Nicholson, B., Donlan, C., & Van Herwegen, J. (2018). Low performance on mathematical tasks in preschoolers: the importance of domain-general and domain-specific abilities. *Journal of Intellectual Disability Research*, *62*(4), 292–302. https://doi.org/10.1111/jir.12465

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*. https://doi.org/10.1016/j.tine.2013.12.001

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, *162*, 12–26. https://doi.org/10.1016/j.cognition.2017.01.014

Csibra G, Hernik M, Mascaro O, Tatone D, Lengyel M. Statistical treatment of looking-time data (2016). *Dev Psychol.* 52(4):521-36. doi: 10.1037/dev0000083.

Culham, J. C., & Valyear, K. F. (2006). Human parietal cortex in action. *Current Opinion in Neurobiology*, *16*(2), 205–212. https://doi.org/10.1016/j.conb.2006.03.005

de Hevia, M. D., Veggiotti, L., Streri, A., & Bonn, C. (2017). At Birth, Humans Associate "Few" with Left and "Many" with Right. *Current Biology*, *27 (24), 3879-3884e2.* https://doi.org/10.1016/j.cub.2017.11.024

De Rose, P., Albamonte, E., Lagana, V., Sivo, S., Pisoni, S., Gallini, F., … Ricci, D. (2013). Perceptual-motor abilities in pre-school preterm children. *Early Human Development*, *89*(10), 809–814. https://doi.org/10.1016/j.earlhumdev.2013.07.001

De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology*, *108*(2), 278–292. https://doi.org/10.1016/j.jecp.2010.09.003

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103*, 186–201. https://doi.org/10.1016/j.jecp.2009.01.004

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/https://doi.org/10.1016/j.tine.2013.06.001

Dehaene, S. (1996). The organization of brain activations in number comparison: event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience*, *8*(1), 47–68. https://doi.org/10.1162/jocn.1996.8.1.47

Dehaene, S, & Changeux, J. P. (1993). Development of elementary numerical abilities: a neuronal model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407. https://doi.org/10.1162/jocn.1993.5.4.390

Dehaene, S, & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*, 83–120.

Dehaene, S, Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science (New York, N.Y.)*, *284*(5416), 970–974. https://doi.org/10.1126/science.284.5416.970

Dehaene, S. (1997). *The Number Sence*. New York, NY: Oxford University Press.

Dehaene, S. (2001). Précis of The Number Sense. *Mind & Language*, *16*, 16–36. https://doi.org/10.1111/1468-0017.00154

Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics*. New York, NY: Oxford University Press.

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology*, *14*(2), 218–224. https://doi.org/10.1016/j.conb.2004.03.008

Dehaene, S.., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3), 487–506. https://doi.org/10.1080/02643290244000239

Wilson, A. J., & Dehaene, S. (2007). *Number sense and developmental dyscalculia.* In D. Coch, G. Dawson, & K. W Fischer (Eds.), *Human behavior, learning, and the developing brain: Atypical development* (p. 212–238). The Guilford Press.

Delazer, M, Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic--an fMRI study. *Brain Research. Cognitive Brain Research*, *18*(1), 76–88. https://doi.org/10.1016/j.cogbrainres.2003.09.005

Delazer, Margarete, Domahs, F., Lochy, A., Bartha, L., Brenneis, C., & Trieb, T. (2004). The acquisition of arithmetic knowledge

- an FMRI study. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *40*(1), 166–167. https://doi.org/10.1016/s0010-9452(08)70936-5

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). Delis-Kaplan Executive Function System.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

DePaolis, R. A., Keren-Portnoy, T., & Vihman, M. (2016). Making Sense of Infant Familiarity and Novelty Responses to Words at Lexical Onset. *Frontiers in Psychology*, *7*, 715. https://doi.org/10.3389/fpsyg.2016.00715

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *The British Journal of Educational Psychology*, *82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

Diamond, A. (1985). Development of the Ability to Use Recall to Guide Action, as Indicated by Infants' Performance on AB̄. *Child Development*, *56*(4), 868–883. https://doi.org/10.2307/1130099

Diamond, A. (2001). Looking closely at infants' performance and experimental procedures in the A-not-B task. *Behavioral and Brain Sciences*, *24*(1), 38–41. https://doi.org/DOI: 10.1017/S0140525X01253916

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dietrich, J. F., Huber, S., & Nuerk, H. C. (2015). Methodological aspects to be considered when measuring the approximate number system (ANS) - a research review. *Frontiers in Psychology*, *6*, 1–14. https://doi.org/10.3389/fpsyg.2015.00295

Ding, S., Lemyre, B., Daboval, T., Barrowman, N., & Moore, G. P. (2019). A meta-analysis of neurodevelopmental outcomes at 4–10 years in children born at 22–25 weeks gestation. *Acta Paediatrica, International Journal of Paediatrics*, *108*(7), 1237–1244. https://doi.org/10.1111/apa.14693

Donlan, C. (2007). *Mathematical development in children with specific language impairments*. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (p. 151–172). Paul H. Brookes Publishing Co.

Dowker, A. (2004). What Works for Children with Mathematical Difficulties? Sudbury: Department of Educational Skills.

Downes, M., Kelly, D., Day, K., Marlow, N., & de Haan, M. (2018). Visual attention control differences in 12-month-old preterm infants. *Infant Behavior and Development*, 50, 180–188. https://doi.org/10.1016/j.infbeh.2018.01.002

Duncan, G. J., Magnuson, K., Kalil, A., & Ziol-Guest, K. (2012). The importance of early childhood poverty. *Social Indicators Research*. 108, 87-98. https://doi.org/10.1007/s11205-011-9867-9

Edwards, L. A., Wagner, J. B., Simon, C. E., & Hyde, D. C. (2016). Functional brain organization for number processing in pre-verbal infants. *Developmental Science*, *19*(5), 757–769. https://doi.org/10.1111/desc.12333

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465–486. https://doi.org/10.1207/s15326942dn2601_6

Fanari, R., Meloni, C., & Massidda, D. (2019). Visual and Spatial Working Memory Abilities Predict Early Math Skills: A

Longitudinal Study. *Frontiers in Psychology*. 10, 2460. https://www.frontiersin.org/article/10.3389/fpsyg.2019.02460

Fantz, R. L. (1964). Visual experience in infants: Decreased attention familiar patterns relative to novel ones. *Science. 146* (3644): 668-70. https://doi.org/10.1126/science.146.3644.668

Farran, E. K., & Karmiloff-Smith, A. (2012). *Neurodevelopmental disorders across the lifespan: A neuroconstructivist approach*. Oxford University Press.

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*(1), 53–72. https://doi.org/10.1016/j.jecp.2014.01.013

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, *110*(41), 16616 – 16621. https://doi.org/10.1073/pnas.1315235110

Feigenson, L., Dehaene, S., & Spelke, E. (2004a). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Feigenson, L., Dehaene, S., & Spelke, E. (2004b). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Development Perspectives*, *7*(2), 74–79. https://doi.org/10.1111/cdep.12019

Fias, W., Menon, V., & Szucs, D. (2013). Multiple components of developmental dyscalculia. *Trends in Neuroscience and Education*, *2*(2), 43–47. https://doi.org/10.1016/j.tine.2013.06.006

Foster-Cohen, S. H., Friesen, M. D., Champion, P. R., & Woodward, L. J. (2010). High prevalence/low severity language delay in preschool children born very preterm. *Journal of Developmental and Behavioral Pediatrics, 31*(8), 658–667. https://doi.org/10.1097/DBP.0b013e3181e5ab7e

Frank, Michael. (2020). *'ManyBabies' Theory paper*. Retrieved from https://manybabies.github.io/projects/

Frey, H. A., & Klebanoff, M. A. (2016). The epidemiology, etiology, and costs of preterm birth. *Seminars in Fetal and Neonatal Medicine*, *21*(2), 68–73. https://doi.org/10.1016/j.siny.2015.12.011

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/https://doi.org/10.1016/j.edurev.2013.05.003

Friso-Van Den Bos, I., Van Der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Fuchs, L. S., Compton, D. L., Fuchs, D., Powell, S. R., Schumacher, R. F., Hamlett, C. L., … Vukovic, R. K. (2012). Contributions of domain-general cognitive resources and different forms of arithmetic development to pre-algebraic knowledge. *Developmental Psychology*, *48*(5), 1315–1326. https://doi.org/10.1037/a0027475

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29

Fuhs, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: contributions of inhibitory control. *Developmental Science*, *16*(1), 136–148. https://doi.org/10.1111/desc.12013

Gale, C., Statnikov, Y., Jawad, S., Uthaya, S. N., Modi, N., Modi, N., … Cowan, F. (2018). Neonatal brain injuries in England: Population-based incidence derived from routinely recorded clinical data held in the National Neonatal Research Database. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *103*(4), F301–F306. https://doi.org/10.1136/archdischild-2017-313707

Gallace, A., Tan, H. Z., & Spence, C. (2007). Multisensory numerosity judgments for visual and tactile stimuli. *Perception & Psychophysics*, *69*(4), 487–501. https://doi.org/10.3758/BF03193906

Gallistel, C. R. (1990). Representations in animal cognition: an introduction. *Cognition*, *37*(1–2), 1–22. https://doi.org/10.1016/0010-0277(90)90016-d

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74. https://doi.org/10.1016/0010-0277(92) 90050-R

Garfield, C. F., Karbownik, K., Murthy, K., Falciglia, G., Guryan, J., Figlio, D. N., & Roth, J. (2017). Educational Performance of Children Born Prematurely. *JAMA Pediatrics*, *171*(8), 764–770. https://doi.org/10.1001/jamapediatrics.2017.1020

Geary, D. C. (2000). From infancy to adulthood: The development of numerical abilities. *European Child and Adolescent Psychiatry*, *9* (2), 11–17. https://doi.org/10.1007/s007870070004

Geary, D C, Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: a longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology*, *77*(3), 236–263. https://doi.org/10.1006/jecp.2000.2561

Geary, D C, & Moore, A. M. (2016). Cognitive and brain systems underlying early mathematical development. *Progress in Brain Research*, *227*, 75–103. https://doi.org/10.1016/bs.pbr.2016.03.008

Geary, D. C. (2004). Mathematics and Learning Disabilities. *Journal of Learning Disabilities*, *37*(1), 4–15. https://doi.org/10.1177/00222194040370010201

Geary, D. C. (2007). An evolutionary perspective on learning disability in mathematics. *Developmental Neuropsychology*, *32*(1), 471–519. https://doi.org/10.1080/87565640701360924

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Developmental Psychology*, *47*(6), 1539–1552. https://doi.org/10.1037/a0025510

Geary, D. C. (2015). The Classification and Cognitive Characteristics of Mathematical Disabilities in Children. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition* (pp. 1–25). Oxford Handbooks Online. https://doi.org/10.1093/oxfordhb/9780199642342.013.017

Geary, D. C, Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology*, *27*(5), 787–797. https://doi.org/10.1037/0012-1649.27.5.787

Geary, D. C, & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. *Handbook of Mathematical Cognition*. New York, NY, US: Psychology Press.

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003

Gebuis, T., Herfs, I. K., Kenemans, J. L., De Haan, E. H. F., & Van

Der Smagt, M. J. (2009). The development of automated access to symbolic and non-symbolic number knowledge in children: An ERP study. *European Journal of Neuroscience*, *30*(10), 1999–2008. https://doi.org/10.1111/j.1460-9568.2009.06994.x

Geldof, C., Wassenaer-Leemhuis, A., de Kieviet, J., Kok, J., & Oosterlaan, J. (2011). Visual perception and visual-motor integration in very preterm and/or very low birth weight children: A meta-analysis. *Research in Developmental Disabilities*, *33*, 726–736. https://doi.org/10.1016/j.ridd.2011.08.025

Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number.* Cambridge, MA: Harvard University Press.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*(4), 293–304. https://doi.org/10.1177/00222194050380040301

Gilligan, K., Flouri, E., & Farran, E. (2017). The contribution of spatial ability to mathematics achievement in middle childhood. *Journal of Experimental Child Psychology*, *163*, 107–125. https://doi.org/10.1016/j.jecp.2017.04.016

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., … Inglis, M. (2013). Individual Differences in Inhibitory Control, Not Non-Verbal Number Acuity, Correlate with Mathematics Achievement. *PLoS ONE*, *8*(6), 1–9. https://doi.org/10.1371/journal.pone.0067374

Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *Quarterly Journal of Experimental Psychology*, *64*(11), 2099–2109. https://doi.org/10.1080/17470218.2011.574710

Gilmore, C., Clayton, S., Cragg, L., McKeaveney, C., Simms, V., & Johnson, S. (2018). Understanding arithmetic concepts: The role of domain-specific and domain-general skills. *PLoS ONE*, *13*(9), 1–20. https://doi.org/10.1371/journal.pone.0201724

Gilmore, C., & Cragg, L. (2018). The Role of Executive Function Skills in the Development of Children's Mathematical Competencies. In A. Henik & W. Fias (Eds.), *Heterogeneity of Function in Numerical Cognition* (263–286). Academic Press. https://doi.org/10.1016/B978-0-12-811529-9.00014-5

Gilmore, C., Göbel, S. M., & Inglis, M. (2018). *An Introduction to Mathematical Cognition*. Oxon: Routledge.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*(3), 394–406. https://doi.org/10.1016/j.cognition.2010.02.002

Gilmore, R., & Johnson, M. (1995). Working memory in infancy: six-month-olds' performance on two versions of the oculomotor delayed response task. *Journal of Experimental Child Psychology*, *59*, 397–418.

Gimbert, F., Camos, V., Gentaz, E., & Mazens, K. (2019). What predicts mathematics achievement? Developmental change in 5- and 7-year-old children. *Journal of Experimental Child Psychology*, *178*, 104–120. https://doi.org/10.1016/j.jecp.2018.09.013

Gimbert, F., Gentaz, E., Camos, V., & Mazens, K. (2016). Children's Approximate Number System in Haptic Modality. *Perception*, *45*(1–2), 44–55. https://doi.org/10.1177/0301006615614448

Glass, H. C., Costarino, A. T., Stayer, S. A., Brett, C. M., Cladis, F., & Davis, P. J. (2015). *Outcomes for extremely premature infants*. *Anesthesia and Analgesia*, *120*(6), 1337-1351. https://doi.org/10.1213/ANE.0000000000000705

Gobel, S. M., & Rushworth, M. F. S. (2004). Cognitive neuroscience: acting on numbers. *Current Biology*, *14*(13), 517-519. https://doi.org/10.1016/j.cub.2004.06.042

Gómez-Velázquez, F. R., Berumen, G., & González-Garrido, A. A. (2015). Comparisons of numerical magnitudes in children with different levels of mathematical achievement. An ERP study.

Brain Research, *1627*, 189–200. https://doi.org/10.1016/j.brainres.2015.09.009

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*(5695), 496–499. https://doi.org/10.1126/science.1094492

Gredebäck, G., Johnson, S., & Hofsten, C. (2010). Eye Tracking in Infancy Research. *Developmental Neuropsychology*, *35*, 1–19. https://doi.org/10.1080/87565640903325758

Griffin, I. C., Miniussi, C., & Nobre, A. C. (2002). Multiple mechanisms of selective attention: differential modulation of stimulus processing by attention to space or time. *Neuropsychologia*, *40*(13), 2325–2340. https://doi.org/10.1016/s0028-3932(02)00087-8

Gruber, O., Indefrey, P., Steinmetz, H., & Kleinschmidt, A. (2001). Dissociating neural correlates of cognitive components in mental calculation. *Cerebral Cortex*, *11*(4), 350–359. https://doi.org/10.1093/cercor/11.4.350

Grunau, R. E., Whitfield, M. F., & Davis, C. (2002). Pattern of learning disabilities in children with extremely low birth weight and broadly average intelligence. *Archives of Pediatrics & Adolescent Medicine*, *156*(6), 615–620. https://doi.org/10.1001/archpedi.156.6.615

Grunewaldt, K., Løhaugen, G., Austeng, D., Brubakk, A.-M., & Skranes, J. (2013). Working Memory Training Improves Cognitive Function in VLBW Preschoolers. *Pediatrics*, *131*, e747-54. https://doi.org/10.1542/peds.2012-1965

Guarini, A., Sansavini, A., Fabbri, M., Alessandroni, R., Faldella, G., & Karmiloff-Smith, A. (2014). Basic numerical processes in very preterm children: A critical transition from preschool to school age. *Early Human Development*, *90*(3), 103–111. https://doi.org/10.1016/j.earlhumdev.2013.11.003

Hack, M. (2009). Adult outcomes of preterm children. *Journal of Developmental and Behavioral Pediatrics*, *30*(5), 460–470.

https://doi.org/10.1097/DBP.0b013e3181ba0fba

Hack, M., Flannery, D. J., Schluchter, M., Cartar, L., Borawski, E., & Klein, N. (2002). Outcomes in young adulthood for very-low-birth-weight infants. *The New England Journal of Medicine*, *346*(3), 149–157. https://doi.org/10.1056/NEJMoa010856

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive S ciences*, *13*(2), 65–73. https://doi.org/10.1016/j.tics.2008.11.003

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116 – 11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Harel-Gadassi, A., Friedlander, E., Yaari, M., Bar-Oz, B., Eventov-Friedman, S., Mankuta, D., & Yirmiya, N. (2018). Developmental assessment of preterm infants: Chronological or corrected age? *Research in Developmental Disabilities*, *80*, 35–43. https://doi.org/10.1016/j.ridd.2018.06.002

Heine, A., Tamm, S., Wißmann, J., & Jacobs, A. M. (2011). Electrophysiological correlates of non-symbolic numerical magnitude processing in children: Joining the dots. *Neuropsychologia*, *49*(12), 3238–3246. https://doi.org/10.1016/j.neuropsychologia.2011.07.028

Heine, A., Wißmann, J., Tamm, S., De Smedt, B., Schneider, M., Stern, E., … Jacobs, A. M. (2013). An electrophysiological investigation of non-symbolic magnitude processing: Numerical distance effects in children with and without mathematical learning disabilities. *Cortex*, *49*(8), 2162–2177. https://doi.org/10.1016/j.cortex.2012.11.009

Hellgren, K., Halberda, J., Forsman, L., Ådén, U., & Libertus, M. (2013). Compromised approximate number system acuity in extremely preterm school-aged children. *Developmental Medicine and Child Neurology*, *55*(12), 1109–1114. https://doi.org/10.1111/dmcn.12206

Henik, A., Rubinsten, O., & Ashkenazi, S. (2015). Developmental dyscalculia as a heterogeneous disability. *The Oxford Handbook of Numerical Cognition.* New York, NY, US: Oxford University Press.

Hessels, R. S., & Hooge, I. T. C. (2019). Eye tracking in developmental cognitive neuroscience – The good, the bad and the ugly. *Developmental Cognitive Neuroscience*, *40*, 100710. https://doi.org/https://doi.org/10.1016/j.dcn.2019.100710

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology*, *37*(3), 187–209. https://doi.org/10.1080/87565641.2011.627958

Holloway, I. D., & Ansari, D. (2010). Developmental specialization in the right intraparietal sulcus for the abstract representation of numerical magnitude. *Journal of Cognitive Neuroscience*, *22*(11), 2627–2637. https://doi.org/10.1162/jocn.2009.21399

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*(3), 339–366. https://doi.org/10.1080/01443410500341056

Hornman, J., Kerstjens, J. M., de Winter, A. F., Bos, A. F., & Reijneveld, S. A. (2013). Validity and internal consistency of the Ages and Stages Questionnaire 60-month version and the effect of three scoring methods. *Early Human Development*, *89*(12), 1011–1015. https://doi.org/10.1016/j.earlhumdev.2013.08.016

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy*

Research, 5, 69–95.

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology, 19*(3), 338–352. https://doi.org/10.1037/0012-1649.19.3.338

Hutchinson, E. A., De Luca, C. R., Doyle, L. W., Roberts, G., & Anderson, P. J. (2013). School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics*, *131*(4). https://doi.org/10.1542/peds.2012-2311

Hyde, D. C., Boas, D. A., Blair, C., & Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. *NeuroImage*, *53*(2), 647–652. https://doi.org/10.1016/j.neuroimage.2010.06.030

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition*, *131*(1), 92–107. https://doi.org/10.1016/j.cognition.2013.12.007

Hyde, D. C., & Spelke, E. S. (2009). All Numbers Are Not Equal: An Electrophysiological Investigation of Small and Large Number Representations. *Journal of Cognition and Development Cognition*, *21*(6), 1039–1053. https://doi.org/10.1162/jocn.2009.21090

Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: Evidence for two core systems underlying numerical cognition. *Developmental Science*, *14*(2), 360–371. https://doi.org/10.1111/j.1467-7687.2010.00987.x

Hyde, D. C., & Spelke, E. S. (2012). Spatiotemporal dynamics of processing nonsymbolic number: An event-related potential source localization study. *Human Brain Mapping*, *33*(9), 2189–2203. https://doi.org/10.1002/hbm.21352

Hyde, D. C., & Wood, J. N. (2011). Spatial attention determines the nature of nonverbal number representation. *Journal of*

Cognitive Neuroscience, *23*(9), 2336–2351. https://doi.org/10.1162/jocn.2010.21581

Imbo, I., & LeFevre, J.-A. (2010). The role of phonological and visual working memory in complex arithmetic for Chinese- and Canadian-educated adults. *Memory & Cognition*, *38*(2), 176–185. https://doi.org/10.3758/MC.38.2.176

Inder, T. E., Wells, S. J., Mogridge, N. B., Spencer, C., & Volpe, J. J. (2003). Defining the nature of the cerebral abnormalities in the premature infant: A qualitative magnetic resonance imaging study. *Journal of Pediatrics*, *143*(2), 171–179. https://doi.org/10.1067/S0022-3476(03)00357-3

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin and Review*, *18*(6), 1222–1229. https://doi.org/10.3758/s13423-011-0154-1

Inglis, M., Batchelor, S., Gilmore, C., & Watson, D. G. (2017). Is the ANS linked to mathematics performance? *Behavioral and Brain Sciences*, *14*(e174). https://doi.org/10.1017/S0140525X16002120

Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*(1), 147–155. https://doi.org/10.1016/j.actpsy.2013.11.009

Isaacs, E. B., Edmonds, C. J., Lucas, A., & Gadian, D. G. (2001). Calculation difficulties in children of very low birthweight: A neural correlate. *Brain*, *124*(9), 1701–1707. https://doi.org/10.1093/brain/124.9.1701

Iuculano T. Neurocognitive accounts of developmental dyscalculia and its remediation. *Prog Brain Res.* 2016;227:305-33. https://doi: 10.1016/bs.pbr.2016.04.024.

Iuculano, T., & Kadosh, R. C. (2014). Preliminary evidence for performance enhancement following parietal lobe stimulation in Developmental Dyscalculia. *Frontiers in Human*

Neuroscience, *8*(1), 1–10. https://doi.org/10.3389/fnhum.2014.00038

Iuculano, T., Rosenberg-Lee, M., Richardson, J., Tenison, C., Fuchs, L., Supekar, K., & Menon, V. (2015). Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities. *Nature Communications*, *6*(1), 8453. https://doi.org/10.1038/ncomms9453

Iuculano, T., Tang, J., Hall, C. W. B., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science*, *11*(5), 669–680. https://doi.org/10.1111/j.1467-7687.2008.00716.x

Izard, V., Dehaene-Lambertz, G., & Dehaene, S. (2008). Distinct cerebral pathways for object identity and number in human infants. *PLoS Biology*, *6*(2), 0275–0285. https://doi.org/10.1371/journal.pbio.0060011

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Jaekel, J., Baumann, N., & Wolke, D. (2013). Effects of gestational age at birth on cognitive performance: a function of cognitive workload demands. *PloS One*, *8*(5), e65219. https://doi.org/10.1371/journal.pone.0065219

Jaekel, J., & Wolke, D. (2014). Preterm birth and dyscalculia. *Journal of Pediatrics*, *164*(6), 1327–1332. https://doi.org/10.1016/j.jpeds.2014.01.069

Jaroslawska, A. J., Gathercole, S. E., Allen, R. J., & Holmes, J. (2016). Following instructions from working memory: Why does action at encoding and recall help? *Memory & Cognition*, *44*(8), 1183–1191. https://doi.org/10.3758/s13421-016-0636-5

Johansson, S., & Cnattigius, S. (2010). Epidemiology of preterm birth. *Neurodevelopmental Outcomes of Preterm Birth: From*

Childhood to Adult Life, 1–16. https://doi.org/10.1017/CBO9780511712166.002

Johnson, M H. (2001). Functional brain development in humans. *Nature Reviews. Neuroscience*, *2*(7), 475–483. https://doi.org/10.1038/35081509

Johnson, Mark H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21. https://doi.org/10.1016/j.dcn.2010.07.003

Johnson, S., Bamber, D., Bountziouka, V., Clayton, S., Cragg, L., Gilmore, C., … Wharrad, H. J. (2019). Improving developmental and educational support for children born preterm: Evaluation of an e-learning resource for education professionals. *BMJ*, *9*(6), 1–8. https://doi.org/10.1136/bmjopen-2019-029720

Johnson, S., Fawke, J., Hennessy, E., Rowell, V., Thomas, S., Wolke, D., & Marlow, N. (2009). Neurodevelopmental disability through 11 years of age in children born before 26 weeks of gestation. *Pediatrics*, *124*(2). https://doi.org/10.1542/peds.2008-3743

Johnson, S., Gilmore, C., Gallimore, I., Jaekel, J., & Wolke, D. (2015). The long-term consequences of preterm birth: What do teachers know? *Developmental Medicine and Child Neurology*, *57*(6), 571–577. https://doi.org/10.1111/dmcn.12683

Johnson, S., Hennessy, E., Smith, R., Trikic, R., Wolke, D., & 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*, *94*(4). https://doi.org/10.1136/adc.2008.152793

Johnson, S., & Marlow, N. (2011). Preterm birth and childhood psychiatric disorders. *Pediatric Research*, *69* (5), 22–28. https://doi.org/10.1203/PDR.0b013e318212faa0

Johnson, S., & Marlow, N. (2014). Growing up after extremely preterm birth: Lifespan mental health outcomes. *Seminars in Fetal and Neonatal Medicine*, *19*(2), 97–104. https://doi.org/10.1016/j.siny.2013.11.004

Johnson, S., & Marlow, N. (2017). Early and long-term outcome of infants born extremely preterm. *Archives of Disease in Childhood*, *102*(1), 97–102. https://doi.org/10.1136/archdischild-2015-309581

Johnson, S., Moore, T., & Marlow, N. (2014). Using the Bayley-III to assess neurodevelopmental delay: Which cut-off should be used? *Pediatric Research*, *75*(5), 670–674. https://doi.org/10.1038/pr.2014.10

Johnson, S., Strauss, V., Gilmore, C., Jaekel, J., Marlow, N., & Wolke, D. (2016). Learning disabilities among extremely preterm children without neurosensory impairment: Comorbidity, neuropsychological profiles and scholastic outcomes. *Early Human Development*, *103*, 69–75. https://doi.org/10.1016/j.earlhumdev.2016.07.009

Johnson, S., Wolke, D., Hennessy, E., & Marlow, N. (2011). Educational outcomes in extremely preterm children: Neuropsychological correlates and predictors of attainment. *Developmental Neuropsychology*, *36*(1), 74–95. https://doi.org/10.1080/87565641.2011.540541

Johnson, S., Wolke, D., & Marlow, N. (2008). Developmental assessment of preterm infants at 2 years: validity of parent reports. *Developmental Medicine and Child Neurology*, *50*(1), 58–62. https://doi.org/10.1111/j.1469-8749.2007.02010.x

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The Importance of Number Sense to Mathematics Achievement in First and Third Grades. *Learning and Individual Differences*, *20*(2), 82–88. https://doi.org/10.1016/j.lindif.2009.07.004

Joseph, R. M., O'Shea, T. M., Allred, E. N., Heeren, T., Hirtz, D., Jara, H., … Kuban, K. C. K. (2016). Neurocognitive and academic outcomes at age 10 years of extremely preterm newborns. *Pediatrics*, *137*(4). https://doi.org/10.1542/peds.2015-4343

Julie, N., Alain, C., & Jacqueline, L. (2013). Impact of Language Abilities on Exact and Approximate Number Skills Development: Evidence From Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *56*(3), 956–970. https://doi.org/10.1044/1092-4388(2012/10-0229)

Kadosh, R. C., & Dowker, A. (2015). *The Oxford handbook of numerical cognition*. Oxford: Oxford Library of Psychology.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398. https://doi.org/10.1016/s1364-6613(98)01230-3

Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-Analyses of Developmental fMRI Studies Investigating Typical and Atypical Trajectories of Number Processing and Calculation. *Developmental Neuropsychology*, *36*, 763–787. https://doi.org/10.1080/87565641.2010.549884

Kavšek, M., & Bornstein, M. H. (2010). Visual habituation and dishabituation in preterm infants: A review and meta-analysis. *Research in Developmental Disabilities*, *31*(5), 951–975. https://doi.org/10.1016/j.ridd.2010.04.016

Keller, L., & Libertus, M. (2015). Inhibitory control may not explain the link between approximation and math abilities in kindergarteners from middle class families. *Frontiers in Psychology*, *6*, 1–11. https://doi.org/10.3389/fpsyg.2015.00685

Kelly, M. M. (2016). Educational Implications of Preterm Birth: A National Sample of 8- to 11-Year-Old Children Born Prematurely and Their Full-Term Peers. *Journal of Pediatric Health Care*, *30*(5), 464–470. https://doi.org/10.1016/j.pedhc.2015.11.001

Kerr-Wilson, C. O., MacKay, D. F., Smith, G. C. S., & Pell, J. P. (2012). Meta-analysis of the association between preterm delivery and intelligence. *Journal of Public Health (United Kingdom)*, *34*(2), 209–216. https://doi.org/10.1093/pubmed/fdr024

Kiechl-Kohlendorfer, U., Ralser, E., Pupp Peglow, U., Pehboeck-Walser, N., & Fussenegger, B. (2013). Early risk predictors for impaired numerical skills in 5-year-old children born before 32weeks of gestation. *Acta Paediatrica, International Journal of Paediatrics*, *102*(1), 66–71. https://doi.org/10.1111/apa.12036

Kilbride, H. W., Aylward, G. P., & Carter, B. (2018). What Are We Measuring as Outcome? Looking Beyond Neurodevelopmental Impairment. *Clinics in Perinatology*, *45*(3), 467–484. https://doi.org/10.1016/j.clp.2018.05.008

Klein, E., Moeller, K., Huber, S., Willmes, K., Kiechl-Kohlendorfer, U., & Kaufmann, L. (2018). Gestational age modulates neural correlates of intentional, but not automatic number magnitude processing in children born preterm. *International Journal of Developmental Neuroscience*, *65*, 38–44. https://doi.org/10.1016/j.ijdevneu.2017.10.004

Klein, E., Moeller, K., Kiechl-Kohlendorfer, U., Kremser, C., Starke, M., Kadosh, R. C., … Kaufmann, L. (2014). Processing of intentional and automatic number magnitudes in children born prematurely: Evidence from fMRI. *Developmental Neuropsychology*, *39*(5), 342–364. https://doi.org/10.1080/87565641.2014.939179

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*(6), 781–791. https://doi.org/10.1076/jcen.24.6.781.8395

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction*, *25*, 95–103. https://doi.org/https://doi.org/10.1016/j.learninstruc.2012.12.001

Kovachy, V. N., Adams, J. N., Tamaresis, J. S., & Feldman, H. M. (2015). Reading abilities in school-aged preterm children: A review and meta-analysis. *Developmental Medicine and Child Neurology*, *57*(5), 410–419. https://doi.org/10.1111/dmcn.12652

Kovack-Lesh, K., Horst, J., & Oakes, L. (2008). The Cat is out of the Bag: The Joint Influence of Previous Experience and Looking Behavior on Infant Categorization. *Infancy*, *13*, 285–307. https://doi.org/10.1080/15250000802189428

Kovas, Y., Giampietro, V., Viding, E., Ng, V., Brammer, M., Barker, G. J., … Plomin, R. (2009). Brain correlates of non-symbolic numerosity estimation in low and high mathematical ability children. *PLoS ONE*, *4*(2). https://doi.org/10.1371/journal.pone.0004587

Kroesbergen, E., Van der Ven, S., Kolkman, M., Luit, J. E. H., & Leseman, P. P. M. (2009). Executive functions and the development of (preparatory) Math skills. *Pedagogische Studien*, *86*, 334–349.

Kucian, K., & Kaufmann, L. (2009). A developmental model of number representation. *Behavioral and Brain Sciences*, *32*(3–4), 340–341. https://doi.org/DOI: 10.1017/S0140525X09990069

Kucian, K., Loenneker, T., Dietrich, T., Dosch, M., Martin, E., & von Aster, M. (2006). Impaired neural networks for approximate calculation in dyscalculic children: a functional MRI study. *Behavioral and Brain Functions : BBF*, *2*, 31. https://doi.org/10.1186/1744-9081-2-31

Kucian, K., Loenneker, T., Martin, E., & von Aster, M. (2011). Non-symbolic numerical distance effect in children with and without developmental dyscalculia: a parametric fMRI study. *Developmental Neuropsychology*, *36*(6), 741–762. https://doi.org/10.1080/87565641.2010.549867

Kucian, K., & von Aster, M. (2015). Developmental dyscalculia. *European Journal of Pediatrics*, *174*(1), 1–13. https://doi.org/10.1007/s00431-014-2455-7

Kyser, K. L., Morriss, F. H. J., Bell, E. F., Klein, J. M., & Dagle, J. M. (2012). Improving survival of extremely preterm infants born between 22 and 25 weeks of gestation. *Obstetrics and Gynecology*, *119*(4), 795–800. https://doi.org/10.1097/AOG.0b013e31824b1a03

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: a cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677–692. https://doi.org/10.1016/j.jecp.2010.11.001

Landerl, K., & Kölle, C. (2009). Typical and atypical development of basic numerical skills in elementary school. *Journal of Experimental Child Psychology*, *103*(4), 546–565. https://doi.org/10.1016/j.jecp.2008.12.006

Larroque, B., Ancel, P. Y., Marchand-Martin, L., Cambonie, G., Fresson, J., Pierrat, V., ... the Epipage Study group. (2011). Special care and school difficulties in 8-year-old very preterm children: The Epipage Cohort study. *PLoS ONE*, *6*(7). https://doi.org/10.1371/journal.pone.0021361

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., &

Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: executive functioning, patterns, numeracy, and arithmetic. *The British Journal of Educational Psychology*, *82*(Pt 1), 82–99. https://doi.org/10.1111/j.2044-8279.2010.02016.x

LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767. https://doi.org/10.1111/j.1467-8624.2010.01508.x

Leibovich, T., & Ansari, D. (2016). The Symbol-Grounding Problem in Numerical Cognition: A Review of Theory, Evidence, and Outstanding Questions. *Canadian Journal of Experimental Psychology*, *70*(1), 12–23. https://doi.org/10.1037/cep0000070

Leung, M. P., Thompson, B., Black, J., Dai, S., & Alsweiler, J. M. (2018). The effects of preterm birth on visual development. *Clinical & Experimental Optometry*, *101*(1), 4–12. https://doi.org/10.1111/cxo.12578

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*, *13*(6), 900–906. https://doi.org/10.1111/j.1467-7687.2009.00948.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*(6), 1292–1300. https://doi.org/10.1111/j.1467-7687.2011.01080.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences*, *25*(410), 126–133. https://doi.org/10.1016/j.lindif.2013.02.001

Libertus, M. E., Forsman, L., Adén, U., & Hellgren, K. (2017). Deficits in approximate number system acuity and mathematical abilities in 6.5-year-old children born extremely preterm. *Frontiers in Psychology*, *8*(JUL), 1–12. https://doi.org/10.3389/fpsyg.2017.01175

Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and Brain Functions : BBF*, *3*, 1. https://doi.org/10.1186/1744-9081-3-1

Linsell, L., Johnson, S., Wolke, D., O'Reilly, H., Morris, J. K., Kurinczuk, J. J., & Marlow, N. (2018). Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: A prospective, population-based cohort study. *Archives of Disease in Childhood*, *103*(4), 363–370. https://doi.org/10.1136/archdischild-2017-313414

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396–401. https://doi.org/10.1111/1467-9280.01453

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, *5*(3), 271–290. https://doi.org/10.1207/s15327078in0503_2

Litt, J. S., Gerry Taylor, H., Margevicius, S., Schluchter, M., Andreias, L., & Hack, M. (2012). Academic achievement of adolescents born with extremely low birth weight. *Acta Paediatrica (Oslo, Norway : 1992)*, *101*(12), 1240–1245. https://doi.org/10.1111/j.1651-2227.2012.02790.x

Litt, J., Taylor, H. G., Klein, N., & Hack, M. (2005). Learning disabilities in children with very low birthweight: Prevalence, neuropsychological correlates, and educational interventions. *Journal of Learning Disabilities*, *38*(2), 130–141. https://doi.org/10.1177/00222194050380020301

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, *22*(4), 395–410. https://doi.org/10.3758/BF03200866

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences*, *109*(46), 18737 LP – 18742. https://doi.org/10.1073/pnas.1207212109

Lowe, J., MacLean, P. C., Shaffer, M. L., & Watterberg, K. (2009). Early Working Memory in Children Born With Extremely Low Birth Weight: Assessed by Object Permanence. *Journal of Child Neurology*, *24*(4), 410–415. https://doi.org/10.1177/0883073808324533

Luck, S. (2005). *An Introduction to The Event-Related Potential Technique. Cambridge.*

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*(2), 256–261. https://doi.org/10.1016/j.cognition.2011.07.009

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, *17*(5), 714–726. https://doi.org/10.1111/desc.12152

Mangham, L. J., Petrou, S., Doyle, L. W., Draper, E. S., & Marlow, N. (2009). The cost of preterm birth throughout childhood in England and Wales. *Pediatrics*, *123*(2). https://doi.org/10.1542/peds.2008-1827

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., & Robertson, I. H. (2001). The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *The Journal of Child Psychology and Psychiatry* *and Allied Disciplines*, *42*(8), 1065–1081.

Marlow, N. (2015). Is survival and neurodevelopmental impairment at 2 years of age the gold standard outcome for neonatal studies? *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *100*(1), F82–F84. https://doi.org/10.1136/archdischild-2014-306191

Marlow, N., Hennessy, E. M., Bracewell, M. A., & Wolke, D. (2007). Motor and executive function at 6 years of age after extremely preterm birth. *Pediatrics*, *120*(4), 793–804. https://doi.org/10.1542/peds.2007-0440

Marlow, N., & Johnson, S. (2007). What the teacher needs to know. *Archives of Disease in Childhood*, *92*(11), 945. https://doi.org/10.1136/adc.2007.127969

Mathewson, K. J., Chow, C. H. T., Dobson, K. G., Pope, E. I., Schmidt, L. A., & Van Lieshout, R. J. (2017). Mental health of extremely low birth weight survivors: A systematic review and meta-analysis. *Psychological Bulletin*, *143*(4), 347–383. https://doi.org/10.1037/bul0000091

Mathiasen, R., Hansen, B. M., Nybo Andersen, A. M. N., Forman, J. L., & Greisen, G. (2010). Gestational age and basic school achievements: A national follow-up study in Denmark. *Pediatrics*, *126*(6). https://doi.org/10.1542/peds.2009-0829

Mazzocco, M. M M, Feigenson, L., & Halberda, J. (2011a). Impaired Acuity of the Approximate Number System Underlies Mathematical Learning Disability (Dyscalculia). *Child Development*, *82*(4), 1224–1237. https://doi.org/10.1111/j.1467-8624.2011.01608.x

Mazzocco, M. M M, Feigenson, L., & Halberda, J. (2011b). Preschoolers' Precision of the Approximate Number System Predicts Later School Mathematics Performance. *PLOS ONE*, *6*(9), e23749. Retrieved from https://doi.org/10.1371/journal.pone.0023749

Mazzocco, M. M M, & Myers, G. F. (2003). Complexities in

Identifying and Defining Mathematics Learning Disability in the Primary School-Age Years. *Annals of Dyslexia*, *53*(1), 218–253. https://doi.org/10.1007/s11881-003-0011-7

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, Panthera leo.. *Animal Behaviour*. Netherlands: Elsevier Science. https://doi.org/10.1006/anbe.1994.1052

McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children's arithmetic performance. *Educational and Child Psychology*, *20*, 93–108.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, *74*(3), 240–260. https://doi.org/10.1006/jecp.1999.2516

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology. Animal Behavior Processes*, *9*(3), 320–334.

Menon, V. (2016). Working memory in children's math learning and its disruption in dyscalculia. *Current Opinion in Behavioral Sciences*, *10*, 125–132. https://doi.org/10.1016/j.cobeha.2016.05.014

Merkley, R., Thompson, J., & Scerif, G. (2016). Of huge mice and tiny elephants: Exploring the relationship between inhibitory processes and preschool math skills. *Frontiers in Psychology*, *6*(JAN), 1–14. https://doi.org/10.3389/fpsyg.2015.01903

Michel, C. M., & Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *NeuroImage*, *61*(2), 371–385. https://doi.org/10.1016/j.neuroimage.2011.12.039

Miller, S. P., Ferriero, D. M., Leonard, C., Piecuch, R., Glidden, D. V., Partridge, J. C., … Barkovich, A. J. (2005). Early brain injury in premature newborns detected with magnetic resonance imaging is associated with adverse early neurodevelopmental outcome. *Journal of Pediatrics*, *147*(5), 609–616. https://doi.org/10.1016/j.jpeds.2005.06.033

Mithyantha, R., Kneen, R., McCann, E., & Gladstone, M. (2017). Current evidence-based recommendations on investigating children with global developmental delay. *Archives of Disease in Childhood*, *102*(11), 1071–1076. https://doi.org/10.1136/archdischild-2016-311271

Moeller, K., Willmes, K., & Klein, E. (2015). A review on functional and structural brain connectivity in numerical cognition. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2015.00227

Moll, K., Gobel, S. M., & Snowling, M. J. (2015). Basic number processing in children with specific learning disorders: Comorbidity of reading and mathematics disorders. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *21*(3), 399–417. https://doi.org/10.1080/09297049.2014.899570

Monette, S., Bigras, M., & Guay, M.-C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, *109*(2), 158–173. https://doi.org/10.1016/j.jecp.2011.01.008

Moreau, D., & Conway, A. R. A. (2014). The case for an ecological approach to cognitive training. *Trends in Cognitive Sciences*, *18*(7), 334–336. https://doi.org/10.1016/j.tics.2014.03.009

Moster, D., Lie, R. T., & Markestad, T. (2008). Long-term medical and social consequences of preterm birth. *New England Journal of Medicine*, *359*(3), 262–273. https://doi.org/10.1056/NEJMoa0706475

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Development of executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology*, *34*(4), 393–421. https://doi.org/10.1080/87565640902964524

Mulder, H., Pitchford, N. J., & Marlow, N. (2010). Processing speed and working memory underlie academic attainment in very preterm children. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *95*(4), 267–272. https://doi.org/10.1136/adc.2009.167965

Mulder, H., Pitchford, N. J., & Marlow, N. (2011). Processing speed mediates executive function difficulties in very preterm children in middle childhood. *Journal of the International Neuropsychological Society*, *17*(3), 445–454. https://doi.org/10.1017/S1355617711000373

Müller, J. B., Hanf, M., Flamant, C., Olivier, M., Rouger, V., Gascoin, G., … Nusinovici, S. (2019). Relative contributions of prenatal complications, perinatal characteristics, neonatal morbidities and socio-economic conditions of preterm infants on the occurrence of developmental disorders up to 7 years of age. *International Journal of Epidemiology*, *48*(1), 71–82. https://doi.org/10.1093/ije/dyy240

Mullin, A. P., Gokhale, A., Moreno-De-Luca, A., Sanyal, S., Waddington, J. L., & Faundez, V. (2013). Neurodevelopmental disorders: Mechanisms and boundary definitions from genomes, interactomes and proteomes. *Translational Psychiatry*, *3*(12), e329-6. https://doi.org/10.1038/tp.2013.108

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Mürner-Lavanchy, I., Ritter, B. C., Spencer-Smith, M. M., Perrig, W. J., Schroth, G., Steinlin, M., & Everts, R. (2014). Visuospatial working memory in very preterm and term born children - Impact of age and performance. *Developmental*

Cognitive Neuroscience, *9*, 106–116. https://doi.org/10.1016/j.dcn.2014.02.004

Mussolin, C., Mejias, S., & Noël, M. P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition*, *115*(1), 10–25. https://doi.org/10.1016/j.cognition.2009.10.006

Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2011). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education*, *1*, 21–31. https://doi.org/10.1016/j.tine.2012.09.003

Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education*, *1*(1), 21–31. https://doi.org/https://doi.org/10.1016/j.tine.2012.09.003

Mutaf Yıldız, B., Sasanguie, D., De Smedt, B., & Reynvoet, B. (2018). Frequency of Home Numeracy Activities Is Differentially Related to Basic Number Processing and Calculation Skills in Kindergartners. *Frontiers in Psychology*, *9*, 340. https://doi.org/10.3389/fpsyg.2018.00340

Nadeau, L., Boivin, M., Tessier, R., Lefebvre, F., & Robaey, P. (2001). Mediators of behavioral problems in 7-year-old children born after 24 to 28 weeks of gestation. *Journal of Developmental and Behavioral Pediatrics : JDBP*, *22*(1), 1–10. https://doi.org/10.1097/00004703-200102000-00001

National Institute for Health Care Excellence. (August 2017). *Developmental follow-up of children and young people born preterm.* Retrieved from www.nice.org.uk/guidance/ng72

Navarro, J. I., Aguilar, M., Concepcion, A., Gonzalo, R., Esperanza, M., & Menacho, I. (2011). Inhibitory Processes, Working Memory, Phonological Awareness, Naming Speed, and Early Arithmetic Achievement. *The Spanish Journal of Psychology*, *14*(2), 7–32. https://doi.org/10.5209/rev_SJOP.2011.v14.n2.6 Copyright

Nguyen The Tich, S., Anderson, P. J., Hunt, R. W., Lee, K. J., Doyle, L. W., & Inder, T. E. (2011). Neurodevelopmental and perinatal correlates of simple brain metrics in very preterm infants. *Archives of Pediatrics and Adolescent Medicine*, *165*(3), 216–222. https://doi.org/10.1001/archpediatrics.2011.9

Nieder, A. (2016). The neuronal code for number. *Nature Reviews. Neuroscience*, *17*(6), 366–382. https://doi.org/10.1038/nrn.2016.40

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*. *297*(5587), 1708–1711. https://doi.org/10.1126/science.1072493

Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(19), 7457–7462. https://doi.org/10.1073/pnas.0402239101

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*(5), 752–760. https://doi.org/10.1111/1469-8986.3850752

Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A Two-Minute Paper-and-Pencil Test of Symbolic and Nonsymbolic Numerical Magnitude Processing Explains Variability in Primary School Children's Arithmetic Competence. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0067918

Nys, J., Ventura, P., Fernandes, T., Querido, L., Leybaert, J., & Content, A. (2013). Does math education modify the approximate number system? A comparison of schooled and unschooled adults. *Trends in Neuroscience and Education*, *2*(1), 13–22. https://doi.org/https://doi.org/10.1016/j.tine.2013.01.001

O'Reilly, H., Johnson, S., Ni, Y., Wolke, D., & Marlow, N. (2020). Neuropsychological outcomes at 19 years of age following extremely preterm birth. *Pediatrics*, *145*(2). https://doi.org/10.1542/peds.2019-2087

O'Shea, T. M., Downey, L. C., & Kuban, K. K. C. (2013). Extreme prematurity and attention deficit: epidemiology and prevention. *Frontiers in Human Neuroscience*, *7*, 578. https://doi.org/10.3389/fnhum.2013.00578

Oakes, L. M. (2010). Using Habituation of Looking Time to Assess Mental Processes in Infancy. *Journal of Cognition and Development Cogn*, *11*(3), 29255–2682. https://doi.org/doi:10.1080/15248371003699977

Odd, D., Evans, D., & Emond, A. (2013). Preterm birth, age at school entry and educational performance. *PloS One*, *8*(10), e76615–e76615. https://doi.org/10.1371/journal.pone.0076615

Orrantia, J., Romualdo, S., Sánchez, R., Matilla, L., Munez, D., & Verschaffel, L. (2018). Numerical magnitude processing and mathematics achievement. *Revista de Educación (Madrid).*

Oudgenoeg-Paz, O., Mulder, H., Jongmans, M. J., van der Ham, I. J. M., & Van der Stigchel, S. (2017). The link between motor and cognitive development in children born preterm and/or with low birth weight: A review of current evidence. *Neuroscience and Biobehavioral Reviews*, *80*(January), 382–393. https://doi.org/10.1016/j.neubiorev.2017.06.009

Owen, E. R., Baumgartner, H. A., & Rivera, S. M. (2013). Using infrared eye-tracking to explore ordinal numerical processing in toddlers with Fragile X Syndrome. *Journal of Neurodevelopmental Disorders*, *5*(1), 1–10. https://doi.org/10.1186/1866-1955-5-1

Panamath. *What is a Weber Fraction?* (October 2011) Retrieved from http://panamath.org/wiki/index.php?title=What_is_a_Weber_Fraction%3F.

Park, J., Bermudez, V., Roberts, R. C., & Brannon, E. M. (2016). Non-symbolic approximate arithmetic training improves math performance in preschoolers. *Journal of Experimental Child Psychology*, *152*, 278–293. https://doi.org/10.1016/j.jecp.2016.07.011

Park, J., & Brannon, E. M. (2013). Training the Approximate Number System Improves Math Proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944.

Park, J., & Starns, J. J. (2015). The Approximate Number System Acuity Redefined: A Diffusion Model Approach. *Frontiers in Psychology*, *6*, 1955. https://doi.org/10.3389/fpsyg.2015.01955

Pascal, A., Govaert, P., Oostra, A., Naulaers, G., Ortibus, E., & Van den Broeck, C. (2018). Neurodevelopmental outcome in very preterm and very-low-birthweight infants born over the past decade: a meta-analytic review. *Developmental Medicine and Child Neurology*, *60*(4), 342–355. https://doi.org/10.1111/dmcn.13675

Passolunghi, M Chiara, & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology*, *88*(4), 348–367. https://doi.org/10.1016/j.jecp.2004.04.002

Passolunghi, M. Chiara, & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology*, *82*(1), 42–63. https://doi.org/10.1111/j.2044-8279.2011.02039.x

Peng, P., & Fuchs, D. (2016). A Meta-Analysis of Working Memory Deficits in Children With Learning Difficulties: Is There a Difference Between Verbal Domain and Numerical Domain? *Journal of Learning Disabilities*, *49*(1), 3–20.

https://doi.org/10.1177/0022219414521667

Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical substrates of arabic number processing, numerical comparison, and simple addition: a PET study. *Journal of Cognitive Neuroscience*, *12*(3), 461–479. https://doi.org/10.1162/089892900562273

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., … Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293–305. https://doi.org/10.1016/j.neuron.2006.11.022

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pinel, P, Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage*, *14*(5), 1013–1026. https://doi.org/10.1006/nimg.2001.0913

Pinel, Philippe, Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983–993. https://doi.org/10.1016/s0896-6273(04)00107-2

Pisch, M. (2015). *A longitudinal study of infant sleep and its effects* *on cognitive development*. (Doctoral dissertation). Retrieved from: https://pdfs.semanticscholar.org/f29f/b6e78b9be446eebd03bcab43fc2e196c57c9.pdf

Poulsen, G., Wolke, D., Kurinczuk, J. J., Boyle, E. M., Field, D., Alfirevic, Z., & Quigley, M. A. (2013). Gestational age and cognitive ability in early childhood: A population-based cohort study. *Paediatric and Perinatal Epidemiology*, *27*(4), 371–379. https://doi.org/10.1111/ppe.12058

Price, G. R., Holloway, I., Räsänen, P., Vesterinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, *17*(24), 1729–1732. https://doi.org/10.1016/j.cub.2007.10.013

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*(1), 50–57. https://doi.org/10.1016/j.actpsy.2012.02.008

Pritchard, V. E., Clark, C. A. C., Liberty, K., Champion, P. R., Wilson, K., & Woodward, L. J. (2009). Early school-based learning difficulties in children born very preterm. *Early Human Development*, *85*(4), 215–224. https://doi.org/10.1016/j.earlhumdev.2008.10.004

Quinn, J. A., Munoz, F. M., Gonik, B., Frau, L., Cutland, C., Mallett-Moore, T., … Buttery, J. (2016). Preterm birth: Case definition & guidelines for data collection, analysis, and presentation of immunisation safety data. *Vaccine*, *34*(49), 6047–6056. https://doi.org/10.1016/j.vaccine.2016.03.045

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalils, O. (2002). Representation of the gender of human faces by infants: a preference for female. *Perception*, *31*(9), 1109–1121. https://doi.org/10.1068/p3331

Raghubar, K., Barnes, M., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Reynolds, G. D. (2015). Infant visual attention and object recognition. *Behavioural Brain Research*, *285*, 34–43. https://doi.org/10.1016/j.bbr.2015.01.015

Reynvoet, B., & Sasanguie, D. (2016). The symbol grounding problem revisited: A thorough evaluation of the ans mapping account and the proposal of an alternative account based on symbol-symbol associations. *Frontiers in Psychology*, *7*(OCT), 1–11. https://doi.org/10.3389/fpsyg.2016.01581

Richmond, J. L., Zhao, J. L., & Burns, M. A. (2015). What goes where? Eye tracking reveals spatial relational memory during infancy. *Journal of Experimental Child Psychology*, *130*, 79–91. https://doi.org/10.1016/j.jecp.2014.09.013

Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: evidence from eye tracking. *Developmental Science*, *12*(4), 549–556. https://doi.org/10.1111/j.1467-7687.2009.00795.x

DeBoer, T., Scott, L. S., & Nelson, C. A. (2007). *Methods for acquiring and analyzing infant event-related potentials*. In M. de Haan (Ed.), *Studies in developmental psychology. Infant EEG and event-related potentials* (p. 5–37). Psychology Press.

Ritchie, S. J., & Bates, T. C. (2013). Enduring Links From Childhood Mathematics and Reading Achievement to Adult Socioeconomic Status. *Psychological Science*, *24*(7), 1301–1308. https://doi.org/10.1177/0956797612466268

Ritter, B. C., Nelle, M., Perrig, W., Steinlin, M., & Everts, R. (2013). Executive functions of children born very preterm - deficit or delay? *European Journal of Pediatrics*, *172*(4), 473–483. https://doi.org/10.1007/s00431-012-1906-2

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex,* *15*(11), 1779–1790. https://doi.org/10.1093/cercor/bhi055

Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, 1(4), 491-507. https://doi.org/10.1207/S15327078IN0104_9

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2002). Processing speed in the 1st year of life: a longitudinal study of preterm and full-term infants. *Developmental Psychology*, *38*(6), 895–902. https://doi.org/10.1037/0012-1649.38.6.895

Rose, S. A., Feldman, J. F., & 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*, *14*(5), 1161–1175. https://doi.org/10.1111/j.1467-7687.2011.01068.x

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2008). A Cognitive Cascade in Infancy: Pathways from Prematurity to Later Mental Development. *Intelligence*, *36*(4), 367–378. https://doi.org/10.1016/j.intell.2007.07.003

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology,* *18*(5), 704–713. https://doi.org/10.1037/0012-1649.18.5.704

Rosenberg-Lee, M., Ashkenazi, S., Chen, T., Young, C. B., Geary, D. C., & Menon, V. (2015). Brain hyper-connectivity and operation-specific deficits during arithmetic problem solving in children with developmental dyscalculia. *Developmental Science*, *18*(3), 351–372. https://doi.org/10.1111/desc.12216

Ross, G., Tesman, J., Auld, P. A., & Nass, R. (1992). Effects of subependymal and mild intraventricular lesions on visual attention and memory in premature infants. *Developmental Psychology, 28*(6), 1067–1074. https://doi.org/10.1037/0012-1649.28.6.1067

Rotzer, S., Kucian, K., Martin, E., von Aster, M., Klaver, P., & Loenneker, T. (2008). Optimized voxel-based morphometry in children with developmental dyscalculia. *NeuroImage*, *39*(1), 417–422. https://doi.org/10.1016/j.neuroimage.2007.08.045

Rotzer, S., Loenneker, T., Kucian, K., Martin, E., Klaver, P., & von Aster, M. (2009). Dysfunctional neural network of spatial working memory contributes to developmental dyscalculia. *Neuropsychologia*, *47*(13), 2859–2865. https://doi.org/10.1016/j.neuropsychologia.2009.06.009

Rourke, B. P., & Conway, J. A. (1997). Disabilities of arithmetic and mathematical reasoning: perspectives from neurology and neuropsychology. *Journal of Learning Disabilities*, *30*(1), 34–46. https://doi.org/10.1177/002221949703000103

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, *102*(3), 361–395. https://doi.org/10.1016/j.cognition.2006.01.005

Rubinsten, O., Sury, D., Lavro, D., & Berger, A. (2013). Processing ordinality and quantity: ERP evidence of separate mechanisms. *Brain and Cognition*, *82*, 201–212. https://doi.org/10.1016/j.bandc.2013.04.008

Rykhlevskaia, E., Uddin, L. Q., Kondos, L., & Menon, V. (2009). Neuroanatomical correlates of developmental dyscalculia: combined evidence from morphometry and tractography.

Frontiers in Human Neuroscience, *3*, 51. https://doi.org/10.3389/neuro.09.051.2009

Saigal, S, Hoult, L. A., Streiner, D. L., Stoskopf, B. L., & Rosenbaum, P. L. (2000). School difficulties at adolescence in a regional cohort of children who were extremely low birth weight. *Pediatrics*, *105*(2), 325–331. https://doi.org/10.1542/peds.105.2.325

Saigal, S., Den Ouden, L., Wolke, D., Hoult, L., Paneth, N., Streiner, D. L., … Pinto-Martin, J. (2003). School-age outcomes in children who were extremely low birth weight from four international population-based cohorts. *Pediatrics*, *112*(4), 943–950. https://doi.org/10.1542/peds.112.4.943

Sánchez-Pérez, N., Castillo, A., López-López, J. A., Pina, V., Puga, J. L., Campoy, G., … Fuentes, L. J. (2018). Computer-based training in math and working memory improves cognitive skills and academic achievement in primary school children: Behavioral results. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02327

Sansavini, A, Savini, S., Guarini, A., Broccoli, S., Alessandroni, R., & Faldella, G. (2011). The effect of gestational age on developmental outcomes: a longitudinal study in the first 2 years of life. *Child: Care, Health and Development*, *37*(1), 26–36. https://doi.org/10.1111/j.1365-2214.2010.01143.x

Sansavini, A., Guarini, A., & Caselli, M. C. (2011). Preterm birth: Neuropsychological profiles and atypical developmental pathways. *Developmental Disabilities Research Reviews*, *17*(2), 102–113. https://doi.org/10.1002/ddrr.1105

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *The British Journal of Developmental Psychology*, *30*(Pt 2), 344–357. https://doi.org/10.1111/j.2044-835X.2011.02048.x

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014).

The approximate number system is not predictive for symbolic number processing in kindergarteners. *Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, *114*(3), 418–431. https://doi.org/10.1016/j.jecp.2012.10.012

Sasanguie, D., Van den Bussche, E., & Reynvoet, B. (2012). Predictors for Mathematics Achievement? Evidence From a Longitudinal Study. *Mind, Brain, and Education*, *6*(3), 119–128. https://doi.org/10.1111/j.1751-228X.2012.01147.x

Schleepen, T. M. J., Van Mier, H. I., & De Smedt, B. (2016). The Contribution of Numerical Magnitude Comparison and Phonological Processing to Individual Differences in Fourth Graders' Multiplication Fact Ability. *PloS One*, *11*(6), e0158335–e0158335. https://doi.org/10.1371/journal.pone.0158335

Schleger, F., Landerl, K., Muenssinger, J., Draganova, R., Reinl, M., Kiefer-Schmidt, I., … Preissl, H. (2014). Magnetoencephalographic Signatures of Numerosity Discrimination in Fetuses and Neonates. *Developmental Neuropsychology*, *39*(4), 316–329. https://doi.org/10.1080/87565641.2014.914212

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12372

Schonhaut, L., Armijo, I., Schönstedt, M., Alvarez, J., & Cordero, M. (2013). Validity of the ages and stages questionnaires in term and preterm infants. *Pediatrics*, *131*(5). https://doi.org/10.1542/peds.2012-3313

Schonhaut, L., Pérez, M., Armijo, I., & Maturana, A. (2020). Comparison between Ages & Stages Questionnaire and Bayley Scales, to predict cognitive delay in school age. *Early Human Development*, *141*(November 2019), 104933. https://doi.org/10.1016/j.earlhumdev.2019.104933

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development*. United Kingdom: Blackwell Publishing. https://doi.org/10.2307/1128664

Serenius, F., Ewald, U., Farooqi, A., Fellman, V., Hafström, M., Hellgren, K., … Källén, K. (2016). Neurodevelopmental outcomes among extremely preterm infants 6.5 years after active perinatal care in Sweden. *JAMA Pediatrics*, *170*(10), 954–963. https://doi.org/10.1001/jamapediatrics.2016.1210

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences, 44*(3), 535–551. https://doi.org/10.1016/j.paid.2007.09.015

Shum, D., Neulinger, K., O'Callaghan, M., & Mohay, H. (2008). Attentional problems in children born very preterm or with extremely low birth weight at 7-9 years. *Archives of Clinical Neuropsychology*, *23*(1), 103–112. https://doi.org/10.1016/j.acn.2007.08.006

Shuman, M., & Kanwisher, N. (2004). Numerical Magnitude in the Human Parietal Lobe: Tests of Representational Generality and Domain Specificity. *Neuron*, *44*(3), 557–569. https://doi.org/https://doi.org/10.1016/j.neuron.2004.10.008

Siegler, R., & Ramani, G. (2009). Playing Linear Number Board

Games-But Not Circular Ones-Improves Low-Income Preschoolers' Numerical Understanding. *Journal of Educational Psychology*, *101*, 545–560. https://doi.org/10.1037/a0014239

Simmons, F., Soto-Calvo, E., Adams, A.M., Francis, H., Patel H. & Giofre, D. (2020). *Understanding the influence of cognition and the home learning environment on early number skills*. Retrieved from: https://www.ljmu.ac.uk/~/media/files/ljmu/microsites/lens/lens-final-report-for-publication.pdf?la=en

Simms, V., Cragg, L., Gilmore, C., Marlow, N., & Johnson, S. (2013a). Mathematics difficulties in children born very preterm: Current research and future directions. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *98*(5), 457–463. https://doi.org/10.1136/archdischild-2013-303777

Simms, V., Gilmore, C., Cragg, L., Clayton, S., Marlow, N., & Johnson, S. (2015). Nature and origins of mathematics difficulties in very preterm children: A different etiology than developmental dyscalculia. *Pediatric Research*, *77*(2), 389–395. https://doi.org/10.1038/pr.2014.184

Simms, V., Gilmore, C., Cragg, L., Marlow, N., Wolke, D., & Johnson, S. (2013b). Mathematics difficulties in extremely preterm children: Evidence of a specific deficit in basic mathematics processing. *Pediatric Research*, *73*(2), 236–244. https://doi.org/10.1038/pr.2012.157

Smyser, C. D., Kidokoro, H., & Inder, T. E. (2012). Magnetic resonance imaging of the brain at term equivalent age in extremely premature neonates: To scan or not to scan? *Journal of Paediatrics and Child Health*, *48*(9), 794–800. https://doi.org/10.1111/j.1440-1754.2012.02535.x

Smyth, R. E., & Ansari, D. (2020). Do infants have a sense of numerosity? A p-curve analysis of infant numerosity discrimination studies. *Developmental Science*, *23*(2), 1–11.

https://doi.org/10.1111/desc.12897

Soltesz, F., Goswami, U., White, S., & Szucs, D. (2011). Executive function effects and numerical development in children : behavioural and ERP evidence from a numerical Stroop paradigm. *Learning and Individual Differences*, *21*(6), 662–671. https://doi.org/10.1016/j.lindif.2010.10.004

Soltész, F., Szucs, D., Dékány, J., Márkus, A., & Csépe, V. (2007). A combined event-related potential and neuropsychological investigation of developmental dyscalculia. *Neuroscience Letters*, *417*(2), 181–186. https://doi.org/10.1016/j.neulet.2007.02.067

Spierer, A., Royzman, Z., & Kuint, J. (2004). Visual acuity in premature infants. *Ophthalmologica*, *218*(6), 397–401. https://doi.org/10.1159/000080943

Spierer, L., Chavan, C. F., & Manuel, A. L. (2013). Training-induced behavioral and brain plasticity in inhibitory control. *Frontiers in Human Neuroscience*, *7*(JUL), 1–9. https://doi.org/10.3389/fnhum.2013.00427

Squires, J., Bricker, D. D., & Twombly, E. (2009). *Ages & stages questionnaires*. Paul H. Brookes Baltimore, MD, USA:

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology (2006)*, *59*(4), 745–759. https://doi.org/10.1080/17470210500162854

Starke, M., Kiechl-Kohlendorfer, U., Kucian, K., Pupp Peglow, U., Kremser, C., Schocke, M., & Kaufmann, L. (2013). Brain structure, number magnitude processing, and math proficiency in 6- to 7-year-old children born prematurely: A voxel-based morphometry study. *NeuroReport*, *24*(8), 419–424. https://doi.org/10.1097/WNR.0b013e32836140ed

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood.

Proceedings of the National Academy of Sciences of the United States of America, *110*(45), 18116–18120. https://doi.org/10.1073/pnas.1302751110

Sun, J., & Buys, N. (2011). Prefrontal lobe functioning and its relationship to working memory in preterm infants. International Journal of Child and Adolescent Health, *4*(1), 13–17.

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70–73. https://doi.org/10.4103/0972-6748.57865

Swanson, H., & Beebe-Frankenberger, M. (2004). The Relationship Between Working Memory and Mathematical Problem Solving in Children at Risk and Not at Risk for Serious Math Difficulties. *Journal of Educational Psychology*, *96*, 471–491. https://doi.org/10.1037/0022-0663.96.3.471

Swanson, H. L. (2006). Cognitive processes that underlie mathematical precociousness in young children. *Journal of Experimental Child Psychology*, *93*(3), 239–264. https://doi.org/10.1016/j.jecp.2005.09.006

Szkudlarek, E., & Brannon, E. M. (2017). Does the approximate number system serve as a foundation for symbolic mathematics? *Language Learning and Development*, *13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Szűcs, D. (2016). Subtypes and comorbidity in mathematical learning disabilities: Multidimensional study of verbal and visual memory processes is key to understanding. In Marinella Cappelletti, Wim Fias (Eds) *The Mathematical Brain Across the Lifespan.* (Vol. 227, pp. 277–304). https://doi.org/10.1016/bs.pbr.2016.04.027

Szũcs, D., Soltész, F., Jármi, É., & Csépe, V. (2007). The speed of magnitude processing and executive functions in controlled and automatic number comparison in children: an electro-encephalography study. *Behavioral and Brain Functions*, *3*(1),

23. https://doi.org/10.1186/1744-9081-3-23

Szucs, D., & Csépe, V. (2004). Access to numerical information is dependent on the modality of stimulus presentation in mental addition: A combined ERP and behavioral study. *Cognitive Brain Research*, *19*(1), 10–27. https://doi.org/10.1016/j.cogbrainres.2003.11.002

Szűcs, D., & Csépe, V. (2005). The effect of numerical distance and stimulus probability on ERP components elicited by numerical incongruencies in mental addition. *Brain Research. Cognitive Brain Research*, *22*, 289–300. https://doi.org/10.1016/j.cogbrainres.2004.04.010

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*(10), 2674–2688. https://doi.org/10.1016/j.cortex.2013.06.007

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2014). Cognitive components of a mathematical processing network in 9-year-old children. *Developmental Science*, *17*(4), 506–524. https://doi.org/10.1111/desc.12144

Szucs, D., & Goswami, U. (2013). Developmental dyscalculia: Fresh perspectives. *Trends in Neuroscience and Education*, *2*(2), 33–37. https://doi.org/10.1016/j.tine.2013.06.004

Szucs, D., & Soltesz, F. (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia*, *45*(14), 3190–3202. https://doi.org/10.1016/j.neuropsychologia.2007.06.013

Tatsuoka, C., McGowan, B., Yamada, T., Espy, K. A., Minich, N., & Taylor, H. G. (2016). Effects of extreme prematurity on numerical skills and executive function in kindergarten children: An application of partially ordered classification modeling. *Learning and Individual Differences*, *49*, 332–340. https://doi.org/10.1016/j.lindif.2016.05.002

Taylor, H. G., & Clark, C. A. C. (2016). Executive function in children born preterm: Risk factors and implications for outcome. *Seminars in Perinatology*, *40*(8), 520–529. https://doi.org/10.1053/j.semperi.2016.09.004

Taylor, H. G., Espy, K. A., & Anderson, P. J. (2009). Mathematics deficiencies in children with very low birth weight or very preterm birth. *Developmental Disabilities Research Reviews*, *15*(1), 52–59. https://doi.org/10.1002/ddrr.51

Temple, E, & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7836–7841. https://doi.org/10.1073/pnas.95.13.7836

Teplan, M. (2002). Fundamental of EEG Measurement. Measurement Science Review, 2(2).

Thevenot, C., Chazoule, G., Masson, S., Castel, C., & Fayol, M. (2016). Les compétences numériques chez de jeunes enfants prématurés. *L'Année Psychologique*, *116*(2), 227–248. https://doi.org/10.4074/S0003503316000361

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106–113. https://doi.org/10.1111/j.1467-7687.2008.00745.x

Tinelli, F., Anobile, G., Gori, M., Aagten-Murphy, D., Bartoli, M., Burr, D. C., … Concetta Morrone, M. (2015). Time, number and attention in very low birth weight children. *Neuropsychologia*, *73*, 60–69. https://doi.org/10.1016/j.neuropsychologia.2015.04.016

Toll, S. W. M., Van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, *44*(6), 521–532. https://doi.org/10.1177/0022219410387302

Twilhaar, E. S., De Kieviet, J. F., Aarnoudse-Moens, C. S. H., Van Elburg, R. M., & Oosterlaan, J. (2018). Academic performance of children born preterm: A meta-analysis and meta-regression. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *103*(4), 322–330. https://doi.org/10.1136/archdischild-2017-312916

Ullman, H., Spencer-Smith, M., Thompson, D. K., Doyle, L. W., Inder, T. E., Anderson, P. J., & Klingberg, T. (2015). Neonatal MRI is associated with future cognition and academic achievement in preterm children. *Brain*, *138*(11), 3251–3262. https://doi.org/10.1093/brain/awv244

Vallortigara, G., Regolin, L., Chiandetti, C., & Rugani, R. (2010). Rudiments of mind: Insights through the chick model on number and space cognition in animals. *Comparative Cognition & Behavior Reviews, 5,* 78–99. https://doi.org/10.3819/ccbr.2010.50004

van't Hooft, J., van der Lee, J. H., Opmeer, B. C., Aarnoudse-Moens, C. S. H., Leenders, A. G. E., Mol, B. W. J., & de Haan, T. R. (2015). Predicting developmental outcomes in premature infants by term equivalent MRI: Systematic review and meta-analysis. *Systematic Reviews*, *4*(1), 1–10. https://doi.org/10.1186/s13643-015-0058-7

Van De Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Verbal and visual-spatial working memory and mathematical ability in different domains throughout primary school. *Memory and Cognition*, *43*(3), 367–378. https://doi.org/10.3758/s13421-014-0480-4

Van De Weijer-Bergsma, E., Wijnroks, L., Boom, J., De Vries, L. S., Van Haastert, I. C., & Jongmans, M. J. (2010). Individual differences in developmental trajectories of A-not-B performance in infants born preterm. *Developmental Neuropsychology*, *35*(6), 605–621. https://doi.org/10.1080/87565641.2010.508545

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*(1), 100–119. https://doi.org/10.1111/j.2044-8279.2011.02035.x

Van Herwegen, J., Ansari, D., Xu, F., & Karmiloff-Smith, A. (2008). Small and large number processing in infants and toddlers with Williams syndrome. *Developmental Science*, *11*(5), 637–643. https://doi.org/10.1111/j.1467-7687.2008.00711.x

van Houdt, C. A., Oosterlaan, J., van Wassenaer-Leemhuis, A. G., van Kaam, A. H., & Aarnoudse-Moens, C. S. H. (2019). Executive function deficits in children born preterm or at low birthweight: a meta-analysis. *Developmental Medicine & Child Neurology*, *61*(9), 1015–1024. https://doi.org/10.1111/dmcn.14213

Vanbinst, K., Ansari, D., Ghesquiere, P., & De Smedt, B. (2016). Symbolic Numerical Magnitude Processing Is as Important to Arithmetic as Phonological Awareness Is to Reading. *PloS One*, *11*(3), e0151045. https://doi.org/10.1371/journal.pone.0151045

Vanbinst, K., Ceulemans, E., Ghesquiere, P., & De Smedt, B. (2015). Profiles of children's arithmetic fact development: a model-based clustering approach. *Journal of Experimental Child Psychology*, *133*, 29–46. https://doi.org/10.1016/j.jecp.2015.01.003

Vandormael, C., Schoenhals, L., Hüppi, P. S., Filippa, M., Borradori Tolsa, C., & Montirosso, R. (2019). Language in Preterm Born Children: Atypical Development and Effects of Early Interventions on Neuroplasticity. *Neural Plasticity*, *2019*. https://doi.org/10.1155/2019/6873270

VanMarle, K., & Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Developmental Science*, *5*, F41–F49.

Venkatraman, V., Ansari, D., & Chee, M. W. L. (2005). Neural correlates of symbolic and non-symbolic arithmetic. *Neuropsychologia*, *43*(5), 744–753.

https://doi.org/10.1016/j.neuropsychologia.2004.08.005

Vogel, J. P., Chawanpaiboon, S., Moller, A. B., Watananirun, K., Bonet, M., & Lumbiganon, P. (2018). The global epidemiology of preterm birth. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, *52*, 3–12. https://doi.org/10.1016/j.bpobgyn.2018.04.003

Volpe, J. J. (2009). Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. *The Lancet Neurology*, *8*(1), 110–124. https://doi.org/10.1016/S1474-4422(08)70294-1

Von Aster, M. (2000). Developmental cognitive neuropsychology of number processing and calculation: Varieties of developmental dyscalculia. *European Child and Adolescent Psychiatry*, *9*(SUPPL. 2). https://doi.org/10.1007/s007870070008

von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine and Child Neurology*, *49*(11), 868–873. https://doi.org/10.1111/j.1469-8749.2007.00868.x

Wade, Nicholas. (2010). Pioneers of Eye Movement Research. i-Perception. 1. 33-68. https://doi.org/10.1068/i0389.

Wang, J., & Goldschmidt, P. (2003). Importance of middle school mathematics on high school students' mathematics achievement. *Journal of Educational Research*, *97*(1), 3–17. https://doi.org/10.1080/00220670309596624

Wechsler, D. (2001). Individual Achievement Test-II (WIAT-II). *The Psychological Corporation*.

Wechsler D. (2005). *Wechsler Intelligence Scale for Children*, 4th Edition San Antonio, TX: PsychCorp.

Wei, W., Guo, L., Georgiou, G. K., Tavouktsoglou, A., & Deng, C. (2018). Different subcomponents of executive functioning predict different growth parameters in mathematics: Evidence from a 4-year longitudinal study with Chinese children. *Frontiers in Psychology*, *9*(JUN), 1–10.

https://doi.org/10.3389/fpsyg.2018.01037

Wei, W., Yuan, H., Chen, C., & Zhou, X. (2012). Cognitive correlates of performance in advanced mathematics. *British Journal of Educational Psychology*, *82*(1), 157–181. https://doi.org/10.1111/j.2044-8279.2011.02049.x

Westermann, G., Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., & Thomas, M. S. C. (2007). Neuroconstructivism. *Developmental Science*, *10*(1), 75–83. https://doi.org/10.1111/j.1467-7687.2007.00567.x

Whitfield, M. F., Grunau, R. V. E., & Holsti, L. (1997). Extremely premature (≤ 800 g) schoolchildren: Multiple areas of hidden disability. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *77*(2). https://doi.org/10.1136/fn.77.2.F85

Wilkinson, H. R., Smid, C., Morris, S., Farran, E. K., Dumontheil, I., Mayer, S., … Thomas, M. S. C. (2019). Domain-Specific Inhibitory Control Training to Improve Children's Learning of Counterintuitive Concepts in Mathematics and Science. *Journal of Cognitive Enhancement*. https://doi.org/10.1007/s41465-019-00161-4

Williams, A. N., & Essex, C. (2004, July). "Developmental delay or failure to arrive?". *Developmental Medicine and Child Neurology*. England. https://doi.org/10.1017/s0012162204240828

Wilson-Ching, M., Molloy, C. S., Anderson, V. A., Burnett, A., Roberts, G., Cheong, J. L. Y., … Anderson, P. J. (2013). Attention difficulties in a contemporary geographic cohort of adolescents born extremely preterm/extremely low birth weight. *Journal of the International Neuropsychological Society : JINS*, *19*(10), 1097–1108. https://doi.org/10.1017/S1355617713001057

Wolke, D., Johnson, S., & Mendonça, M. (2019). The Life Course Consequences of Very Preterm Birth. *Annual Review of Developmental Psychology*, *1*(1), 69–92.

https://doi.org/10.1146/annurev-devpsych-121318-084804

Wong, H. S., Santhakumaran, S., Cowan, F. M., & Modi, N. (2016). Developmental assessments in preterm children: A meta-analysis. *Pediatrics*, *138*(2). https://doi.org/10.1542/peds.2016-0251

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: numerical discrimination and its signature limits. *Developmental Science*, *8*(2), 173–181. https://doi.org/10.1111/j.1467-7687.2005.00404.x

Woodward, L., Edgin, J., Thompson, D., & Inder, T. (2005). Object working memory deficits by early brain injury and development in the preterm infant. *Brain : A Journal of Neurology*, *128*, 2578–2587. https://doi.org/10.1093/brain/awh618

Woodward, L. J., Anderson, P. J., Austin, N., Howard, K., & Inder, T. E. (2006). Neonatal MRI to predict neurodevelopmental outcome. *The New England Journal of Medicine*, *355*, 685–694. https://doi.org/10.1056/NEJMoa053792

Woodward, L. J., Moor, S., Hood, K. M., Champion, P. R., Foster-Cohen, S., Inder, T. E., & 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*, *94*(5), 339–344. https://doi.org/10.1136/adc.2008.146282

World Health Organisation. (19th February 2018). *Preterm Birth*. Retrieved from www.who.int/news-room/fact-sheets/detail/preterm-birth.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750. https://doi.org/10.1038/358749a0

Xenidou-Dervou, I., De Smedt, B., van der Schoot, M., & van Lieshout, E. C. D. M. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and*

Individual Differences, *28*, 119–129. https://doi.org/10.1016/j.lindif.2013.09.012

Xenidou-Dervou, I., Molenaar, D., Ansari, D., van der Schoot, M., & van Lieshout, E. C. D. M. (2017). Nonsymbolic and symbolic magnitude comparison skills as longitudinal predictors of mathematical achievement. *Learning and Instruction*, *50*, 1–13. https://doi.org/10.1016/j.learninstruc.2016.11.001

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology*, *25*(1), 103–108. https://doi.org/10.1348/026151005X90704

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), 1–11. https://doi.org/10.1016/S0010-0277(99)00066-9

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*(1), 88–101. https://doi.org/10.1111/j.1467-7687.2005.00395.x

Xuan, B., Chen, X.-C., He, S., & Zhang, D.-R. (2009). Numerical magnitude modulates temporal comparison: An ERP study. *Brain Research*, *1269*, 135–142. https://doi.org/10.1016/j.brainres.2009.03.016

Yaari, M., Mankuta, D., Harel-Gadassi, A., Friedlander, E., Bar-Oz, B., Eventov-Friedman, S., … Yirmiya, N. (2018). Early developmental trajectories of preterm infants. *Research in Developmental Disabilities*, *81*(November 2017), 12–23. https://doi.org/10.1016/j.ridd.2017.10.018

Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1–9. https://doi.org/10.1016/j.lindif.2012.10.004

Young, C. J., Levine, S. C., & Mix, K. S. (2018). The connection between spatial and mathematical ability across development. *Frontiers in Psychology*, *9*(JUN), 1–7.

https://doi.org/10.3389/fpsyg.2018.00755

Zhang, H., Chang, L., Chen, X., Ma, L., & Zhou, R. (2018). Working memory updating training improves mathematics performance in middle school students with learning difficulties. *Frontiers in Human Neuroscience*, *12*(April), 1–12. https://doi.org/10.3389/fnhum.2018.00154

Zippert, E., & Rittle-Johnson, B. (2018). The home math environment: More than numeracy. *Early Childhood Research Quarterly*. https://doi.org/10.1016/j.ecresq.2018.07.009