EXECUTIVE FUNCTIONING PROFILE OF A NATIONAL COHORT OF ADULTS BORN VERY LOW BIRTHWEIGHT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN CHILD AND FAMILY PSYCHOLOGY

MEG ALLIE JACKSON

UNIVERSITY OF CANTERBURY

2022

Acknowledgements

I wish to acknowledge the support and guidance of my supervisor, Prof. Lianne Woodward during the preparation of this thesis. Thank you for giving me the opportunity to work on this project, it has been a privilege to work with you and I appreciate the team environment you establish for your students. Your professional experience and advice have been invaluable, both challenging and motivating me to achieve the best I could. I want to also acknowledge Dr Samantha Lee for her contribution that I have greatly appreciated. Thank you for your guidance throughout my thesis and giving your time to assist me. I would like to thank Dr Megan Gath for her assistance with the statistical analyses, as well as Dr Alice Kim and Karelia Levin for their contribution and support in the beginning stages. I also want to acknowledge Prof. Brian Darlow and Prof. John Horwood for their roles in overseeing the New Zealand 1986 VLBW study.

I want to thank my fellow students and colleagues, Larissa, Reisha, Elizabeth, Melissa and Liz. Thank you for all your support and helpful advice along the way. A very big thank you goes to my family for always supporting my university endeavors and encouraging me to achieve my best. A special thank you to my close friends, especially Rameshi for all her positive words and advice, and to my partner Mike for his endless encouragement.

Finally, I would like to express my immense gratitude towards the participants in this study who made the completion of this thesis possible. Thank you for your willingness to share your lives and experiences with us. I can only hope that this study provides some insight into the challenges individuals born very low birthweight might experience, and how researching long-term outcomes are of utmost importance for these individuals, their families and society.

Abstract

Introduction. Increasing evidence shows that children born very preterm (<33 weeks; VPT) and/or very low birthweight (<1500g; VLBW) are at high risk for neurodevelopmental difficulties across a range of domains. One of the most common adverse outcomes in this population is cognitive or intellectual impairment however, due to the broad nature of this outcome, there has been increasing interest in identifying specific areas of cognitive impairment. This has led to research examining EF (executive function). To date studies have almost exclusively examined EF difficulties in children and, to a lesser extent, in adolescents born VPT/VLBW. Very few studies have examined whether earlier observed EF impairments in these individuals persist into adulthood. Accordingly, the aim of this thesis was to compare performance of adults born VLBW and full term (FT) on a series of tasks across three key domains of EF: inhibitory control, working memory and cognitive flexibility. Secondary aims were to examine the extent to which between-group differences in EF may be affected by processing speed and the extent to which decreasing birthweight impacts on EF performance.

Research Methods. All infants born VLBW in New Zealand during 1986 were enrolled in a study of acute retinopathy of prematurity *(n*=413) and followed prospectively from birth to age 26-30 years. At 26-30 years (*M*=28.5), 229 participants (77% of survivors) and 100 age-matched FT born adults (*M*=28.2) underwent a comprehensive neuropsychological assessment as part of a two-day evaluation. Participants completed five EF tasks. These included: inhibitory control (Hayling Sentence Completion Test), visuospatial working memory (computerized Sternberg-based task), auditory-verbal working memory (Elevator Counting subtests from the Test of Everyday Attention) and cognitive flexibility (Comprehensive Trail Making Test [CTMT] and the Brixton Spatial Anticipation Test). The Symbol Digit Modalities Test (SDMT) provided a measure of processing speed.

Results. Findings revealed a pervasive pattern of impairment across multiple measures of EF in adults born VLBW relative to FT controls. Specifically, adults born VLBW performed significantly poorer on measures of inhibitory control, visuospatial working memory, cognitive flexibility and on one of the auditory-verbal working memory tests (p<.001). Consistent with these findings, rates of EF impairment were also significantly higher (OR range 2.10 – 7.37) among VLBW adults relative to the control group. Adjusting for the effects of processing speed attenuated the between-group differences in EF performance; however, for most tasks differences remained statistically significant. A linear association was evident between birthweight and EF performance (p<.001), suggesting a higher risk of impairment on EF tasks with decreasing birthweight. A confirmatory factor analysis performed with the task measures supported a single factor model of EF.

Discussion. At age 28 years, adults born VLBW perform worse than control adults on EF tasks across various domains. The findings suggest that while EF difficulties can be explained in part by processing speed, in large the differences indicate a global pattern of EF impairment that places VLBW individuals at higher risk of difficulties for everyday functioning. These findings emphasize the importance of appropriate intervention strategies in the early childhood years to mitigate the potential impact of EF impairments on later life course opportunities in adulthood. With appropriate intervention and supports in adulthood, some EF skills could be malleable to change however, continuing to research intervention opportunities and EF outcomes in later adulthood, could help to ensure that a high quality of life follows for individuals born VLBW.

Table of Contents

List of Tables

List of Figures

Chapter 1

Overview of Preterm Birth

1.1 Defining Preterm Birth

Advances in modern medicine and neonatal care have greatly improved the survival of prematurely born infants (Cheong et al., 2020). As a consequence, greater numbers of infants are being born alive and surviving at increasingly younger gestational ages. Available estimates of the global incidence of preterm birth (<37 weeks gestation) were reported to be as high as 10.6% worldwide in 2014, equating to an estimated 14.84 million live births, compared to 9.6% in 2005 (Beck et al., 2010; Chawanpaiboon et al., 2019). However, despite these improvements in survival, rates of long-term neurodevelopmental challenges for these children remain high and have become an increasing focus of public health concern.

One issue with reviewing contemporary literature on the developmental outcomes of preterm birth is the varied range of criteria used to define this population. Across existing research, the extent of prematurity has been defined on the basis of gestational age (GA), birthweight or both. Early research conducted in the 1980's and 1990's typically defined prematurity on the basis of birthweight due to limitations in assessing GA. The normal weight range at birth is considered between 2500g – 4400g, with <2500g being used to identify infants born low birthweight (Ministry of Health, 2019). There has been a trend for studies to focus on individuals born very or extremely low birth weight. The criterion of <1500g has been used to identify infants born very low birthweight (VLBW; Pyhälä et al., 2011; Breeman et al., 2015; Darlow et al., 2020), and <1000g has been used to identify infants born extremely low birthweight (ELBW; Burnett et al., 2015; Linsell et al., 2018; O'Reilly et al., 2020). Although birthweight has previously been considered a reliable indicator of prematurity, the advent of ultrasonography in more recent times has helped to provide a more accurate measure of fetal maturity or length of gestation. This has resulted in preterm births being increasingly identified and defined on the basis of GA. With respect to GA, infants are considered preterm when born <37 weeks; with late preterm defined as between 34–37 weeks; moderate preterm 32–34 weeks; very preterm (VPT) 28–32 weeks and extremely preterm (EPT) <28 weeks gestation (Hee Chung et al., 2020). There has been extensive literature examining neurodevelopmental outcomes associated with VPT birth, particularly in children, and as GA and birthweight are generally highly correlated, the following literature review includes studies of infants born <1500g and/or <33 weeks GA.

1.2 Preterm Birth in New Zealand

Since accurate recording data became available in the 1980's, preterm birth rates in New Zealand have continued to rise. Rates of singleton infants born <37 weeks rose 37.2% over almost 20 years, from 4.3% in 1980 to 5.9% in 1999, with the highest rise occurring after 1990 (Craig et al., 2002). Considering more recent trends, infants born preterm and low birthweight have continued to increase between 2008 and 2017. In 2017, 7.5% (4,503) of infants were born <37 weeks, with 6.2% between 32–36 weeks and 1.3% <33 weeks (Ministry of Health, 2019). Infants born between 32–36 weeks accounted for 5.9–6.3% of all births annually between 2008-2017, with infants born <33 weeks accounting for 1.2-1.3%. Additionally, 6.1% (3,469) of infants were born low birthweight (<2500g), accounting for 5.9–6.1% of all infants born each year from 2008 to 2017 (Ministry of Health, 2019).

In terms of survival to hospital discharge, infants born <37 weeks and <2500g accounted for 68.9% of infant deaths in 2004 (New Zealand Health Information Service, 2007). Consistent with international trends, mortality rates have tended to decrease over time; with infants born <37 weeks and <2500g, accounting for 55.2% and 56.8% of deaths respectively in 2016 (Ministry of Health, 2016).

1.3 General Outcomes of Preterm Birth

Increasing survival and preterm birth rates over time have consequently resulted in a need for researchers to focus on the developmental outcomes of preterm infants. Considerable research has examined the developmental outcomes of children born VPT/VLBW, with findings consistently identifying VPT/VLBW birth as a risk factor for developmental impairments. These impairments span a range of domains and include increased rates of neurosensory impairment including cerebral palsy, profound hearing or visual impairment, and poorer functioning across cognitive, behavioural and educational outcomes when compared with full-term controls (Horwood et al., 1998; de Kieviet et al., 2009; Darlow et al., 2013; Aarnoudse-Moens 2009; Allotey et al., 2018).

One of the most common adverse outcomes associated with VPT/VLBW birth is cognitive or intellectual impairment (Pritchard et al., 2014; Wiebe & Karbach, 2018). Cognitive functioning and impairment are typically assessed using measures of general intelligence (IQ) which offer a global measure of cognitive functioning (Cheong et al., 2020). However, recent research has identified executive function (EF) as a specific area of cognitive functioning that is particularly impacted in VPT/VLBW children (Vollmer & Stålnacke, 2019). To date, studies have almost exclusively examined EF difficulties in children and, to a lesser extent, in adolescents. Very few studies have examined whether the EF impairments observed during childhood persist over time into adulthood. The following introductory chapters will review the available literature on EF difficulties in children, adolescents, and adults who were born VPT/VLBW, beginning with a description of EF.

Chapter 2

Executive Function

Executive functioning (EF) has significant implications across all aspects of everyday life. Executive functioning impacts on mental and physical health, educational outcomes, cognitive, social, and psychological development (Diamond, 2013). The term EF, or executive control, encompasses a collection of higher-order cognitive processes that allow for goal-directed actions (Anderson, 2002). There is general agreement in contemporary research that there are three core components of EF including inhibitory control, working memory and cognitive flexibility (Diamond, 2013; Wiebe & Karbach, 2018). These EF components have a protracted developmental course from infancy to early adulthood, over time this has been largely posited to the gradual development of the prefrontal cortex of the brain and the establishment of myelination of neural connections (Wiebe & Karbach, 2018). Over the last two decades, neuroimaging studies have highlighted the involvement of neural networks within the developing brain, to suggest there are certain brain regions implicated in EF activation such as the dorsal anterior cingulate cortex (dACC) and parietal cortices (Seeley et al., 2007; Niendam et al., 2012; McKenna et al., 2017).

In terms of conceptualizing EF, a universal theoretical model is currently lacking. Traditionally, EF has been conceptualized as a unitary construct, as the primary executive responsible for higher order skills (Shallice, 1988). However, other models have since conceptualized EF as a more integrative system of multiple inter-related and inter-dependent processes that function together (Anderson, 2002). While many different models have been proposed over time, the most prominent contemporary EF model is one put forward by Akira Miyake and his colleagues (Miyake et al., 2000; Miyake & Friedman., 2012; Miyake & Friedman., 2017). This model presents a useful framework for characterizing and understanding the mechanisms of EF.

2.1 Miyake and colleagues' Unity/Diversity Model of Executive Function

The *unity/diversity* model of EF proposed by Miyake & Friedman (2012) presents a conceptual framework of EF as a multifactorial construct. Their initial research contributing to this conceptualization focused on the assessment of individual differences in EFs, with the purpose of understanding the nature and organization of these differences. The research focused primarily on the most frequently studied EFs, in which Miyake et al. (2000) used the following terms: *inhibition, updating* (working memory) and *shifting* (cognitive flexibility). Miyake et al. (2000) examined how these three domains of EF related to one another and found that those individual differences in EFs emphasized both the unity and diversity of executive processes. The EF domains correlated with one another to represent a common underlying ability however, they were also unique to specific components of EF (Miyake & Friedman, 2012). The model is set up to represent 'common EF', what is common across all three EFs (*unity*) and 'task specific', what is unique to that component (*diversity*). The three EF domains, and how they operate within Miyake and colleagues' model, are described below.

2.1.1 Inhibition

Inhibition or inhibitory control is one's ability to deliberately inhibit or delay an action, response and/or emotions in support of goal-directed behaviour (Diamond, 2013). There are different kinds of inhibition, including inhibiting a response from interfering with an action that has not yet happened, or the ability to stop an action already in progress (Roberts & Pennington, 1996; Miyake et al. 2000). In the context of EF, inhibitory control is often assessed as the inhibition or suppression of one's own automatic or learned responses (Miyake et al., 2000). In relation to the unity/diversity model, Miyake et al. (2000) found there to be no task specific variance for inhibition after common EF was accounted for, which means that inhibition did not explain variance over and above the common factor. When considering how inhibition relates to the other EF domains, this finding suggested that inhibition may almost perfectly correlate with common EF and is a required component for completing EF tasks. Inhibition is the only domain within the model that does not have a unique task-specific component, with this finding replicated in studies of children and adolescents (Huizinga et al., 2006; Friedman et al., 2008; Friedman et al., 2011).

2.1.2 Working Memory

Miyake et al. (2000) use the term *updating* to refer to the updating and monitoring of working memory representations relevant to the task at hand. There are two types of working memory that are distinguished by content, including verbal and non-verbal (visual-spatial) working memory (Diamond, 2013). Although working memory has originally been conceptualized as a temporary storage system, it is now often measured by tasks that involve storage components alongside simultaneous information processing in performing complex tasks (Baddeley, 1992; Baddeley et al., 2019). Therefore, it further involves the active manipulation of information in working memory as Miyake et al. (2000) describes. Within the unity/diversity model, after common EF was accounted for, there was a unique variance left for working memory termed ‘updating-specific’ ability (Miyake et al, 2000). The model theorizes that this unique component of working memory could be related to the precision of the updating process or that controlled retrieval processes from long term memory, may contribute to this task-specific component (Miyake & Friedman, 2012; Friedman & Miyake, 2017).

2.1.3 Cognitive Flexibility

Set-shifting or cognitive flexibility is the ability to flexibly switch from one task to another, and easily shift perspective when solving a problem or working on multiple activities (Miyake & Friedman, 2012). This domain is also implicated in the division of attention and the ability to process different sources of information concurrently (Anderson, 2002). Like working memory, cognitive flexibility also produced a task-specific variance that is reflected in the current model. This domain produced a unique variance after common EF was accounted for, termed 'shifting-specific' ability. The model theorizes that individual differences in cognitive flexibility may be capturing the speed of goal replacement, this is because cognitive flexibility requires the involvement of rapidly switching between goals and the ability to flexibly transition between new task representations (Miyake & Friedman, 2012; Friedman & Miyake, 2017).

2.2 Justification of model

The *unity/diversity* model theorizes specific abilities each EF domain could be tapping into that represents common EF. Miyake & Friedman (2012) posit that a basic requirement of EF tasks across the three domains involves the ability to actively maintain task goals and goal related information which reflects the unity of the EF domains. Additionally, the model theorizes diversity or rather that there are unique task-specific components within the domains of working memory and cognitive flexibility. Miyake and colleagues' model was chosen as an appropriate contemporary framework for this thesis because it presents a useful framework for characterizing the nature of specific EFs among individuals born VPT/VLBW, while recognizing the potential for both commonalities across EF tasks and unique task-specific components.

2.3 Executive Function in Children & Adolescents Born Very Preterm/Very Low Birthweight

Research examining EF in children born VPT/VLBW is extensive due to deficits in cognitive impairment becoming increasingly identified among this population within education settings (Aarnoudse-Moens et al., 2013; Arpi et al., 2019). Several quantitative meta-analyses have summarized the research findings of studies examining EF outcomes of children and adolescents born VPT/VLBW (Bhutta et al., 2002; Aarnoudse-Moens et al., 2009; Allotey et al., 2018; Brydges et al., 2018; van Houdt et al., 2019). A recent meta-analysis conducted by van Houdt et al. (2019), examined EF deficits in childhood across the domains of working memory, cognitive flexibility and inhibition. Thirty-five studies were included in this analysis, resulting in a total sample of 3360 children born preterm or low birthweight and 2812 term-born controls (van Houdt et al., 2019). Study inclusion criteria consisted of a) participants born preterm (<37 weeks) and/or with low birthweight (<2500g), b) a same-age comparison group was also studied, c) the study sample was born in 1990 or later, and d) the age of assessment had to be at least 4 years old (van Houdt et al., 2019). The age range was 4–14 years for the working memory studies and 4–11 years for the inhibition studies (van Houdt et al., 2019).

Studies were also included if they reported on the administration of tasks specific to working memory, cognitive flexibility and inhibition. The specific measures used in the studies included in the meta-analysis were as follows: *Visual-spatial working memory* was assessed using the Cambridge Neuropsychological Test Automated Battery Spatial Working Memory Task; *Verbal working memory* was assessed with the Digit Span Task (DST) and Letter Number Sequencing Task. For *cognitive flexibility*, the Trail Making Test (TMT) was used in assessment as well as the Trails Preschool Revised, an adapted version used for younger children. For inhibition, *response inhibition* was assessed with the Go/No-Go Task and *inference control* was assessed using the Test of Everyday Attention Children Sky Search Task.

Results showed that preterm/low birthweight children performed more poorly, on average, than the full-term controls, with a standardized mean difference of 0.5 on working memory and cognitive flexibility, and 0.4 on inhibition tasks (van Houdt et al., 2019). Additionally, at older ages, difficulties in these areas did not improve even with advancements in medical care, suggesting that difficulties in executive functioning may remain stable across childhood and into adolescence (van Houdt et al., 2019).

Another meta-analysis conducted by Brydges et al. (2018), examined very preterm birth in childhood and adolescence and its association with outcomes in the following cognitive domains: intelligence, executive functioning and processing speed. Sixty studies were included in the analysis, consisting of 6163 children and adolescence born VPT and 5471 term-born controls (Brydges et al., 2018). Study inclusion criteria consisted of a) participants born VPT (<32 weeks), b) a same-age term-born control group, c) participants were aged between 4 to 17 years, and d) the studies used standardized measures to assess the EF domains of working memory, inhibition and cognitive flexibility.

The findings from this meta-analysis showed that, in comparison to term-born controls, children born VPT performed more poorly across all outcomes. Brydges et al. (2018) reported a standardized mean difference of 0.82 on intelligence tests, 0.51 on executive function measures and 0.49 SDs on measures of processing speed. Deficits in intelligence and executive functioning were not associated with age at assessment but were associated with gestational age and/or birthweight (Brydges et al., 2018). These findings suggest that cognitive impairments tend to increase in severity with decreasing gestational age at birth and lower birthweight. Given the persistence of problems, these findings also suggest that difficulties likely reflect impairment in executive functioning rather than a developmental delay (Brydges et al., 2018).

Collectively, these meta-analyses suggest neurocognitive deficits appear to remain stable across time in children and adolescents born VPT/VLBW. Although there is evidence of this stability over time into adolescence, research findings have largely focused on EF outcomes in childhood. EF outcomes, particularly in later adolescence from the age of 16, have been researched to a much lesser extent than childhood and early adolescence. It is therefore important to examine EF outcomes among VPT/VLBW cohorts at this important developmental stage, to determine if earlier difficulties continue to persist.

2.4 Executive Function in late adolescence

The development of EF skills in adolescence become increasingly important for everyday functioning. From the age of 15-16 years, optimal EF skills are required in a range of tasks commonly learned during this stage of increasing independence and autonomy. For example, visual-spatial working memory, cognitive flexibility and inhibitory control are all important cognitive processes involved in learning to drive. The use of these self-control processes become increasingly important as adolescents engage in more complex problem-solving and long-term planning as they approach adulthood (Diamond, 2013).

Studies of EF outcomes of VPT/VLBW individuals in late adolescence have reported findings consistent with those reported for VPT/VLBW children. Luu et al. (2011) examined the EF outcomes in VPT born adolescents at age 16 years on tasks of cognitive flexibility, inhibition, planning/organization, verbal fluency, and verbal and visuospatial working memory. VPT born adolescents displayed deficits across all measures and were found to perform on average 0.4 to 0.6 SD's lower than FT born adolescents across all tasks. This study also examined rates of impairment. Results showed that 6% to 18% of the VPT group exhibited significant impairment (< 2 SD below standard test mean) across all EF measures, compared to 1% to 3% of the term born controls (Luu et al., 2011). Similarly, Lundequist et al. (2015) examined the executive functioning of adolescents at 18 years of age who were born <1500g as part of the Stockholm Neonatal Project using a comprehensive neuropsychological test battery involving measures of cognitive flexibility, working memory and attention. They found significant group differences across all EF measures, with moderate to large effect sizes (.55–.97).

Two studies have examined the longitudinal development of EFs among adolescents born VPT or VLBW. Allin et al. (2008) examined the maturation of cognitive outcomes from age 15 years to 19 years in individuals born <33 weeks gestation. The study examined phonological and semantic verbal fluency, assessing the EF functions that enable response initiation, short term memory retrieval and mental flexibility (Allin et al., 2008). The results indicated that deficits remained for the VPT group over time from 15 to 19 years of age, as they demonstrated poorer performance than a control group at both assessment time points ($p<.05$). Additionally, semantic verbal fluency appeared to significantly improve over time for the control group ($p<.001$) but not for the VPT group ($p=.89$). These findings suggest that the control group developed higher semantic verbal fluency overtime as their ability to generate words based on a semantic cue (e.g colours) increased. In contrast, the semantic verbal fluency abilities of VPT adolescence remained relatively stable from 15 to 19 years.

Stålnacke et al. (2019) also examined EF development longitudinally in individuals born <31 weeks, at ages 5.5 and 18 years as part of the Stockholm Neonatal Project. The developmental stability of two core executive functions, working memory and cognitive flexibility, were investigated in a multiple mediator structural equation model in line with Miyake and colleagues' theoretical EF model. Findings indicated that working memory performance at age 5.5 years predicted working memory performance at 18 years, and this was the same for cognitive flexibility (Stålnacke et al., 2019).

Taken together, these two longitudinal studies of VPT survivors support the concept that executive deficits are relatively stable from early childhood into the late adolescent years, and significant catch ups to strengthen executive functioning in preterm children were not observed beyond preschool years (Stålnacke et al., 2019). Considering the VPT research in later adolescence and the stability of neurocognitive deficits over time, examining EF outcomes in VPT/VLBW populations beyond this age period is warranted to determine whether earlier observed EF impairments persist into adulthood.

Chapter 3

Executive Functioning Abilities of Adults Born Very Preterm/Very Low Birthweight

The literature examining EF outcomes in children and adolescents born VPT/VLBW suggest that EF impairments are stable over time. Therefore, adulthood is an important developmental period to examine to determine if these observed impairments continue to persist. Executive functioning skills are integral to maintaining a level of independent functioning in adulthood (Diamond, 2013). The roles and responsibilities of everyday life increase in adulthood across occupational, financial, educational, social, and family settings. As such, EF impairment across the domains of inhibition, working memory and cognitive flexibility can result in an impaired ability to exert self-control and increase problems with budgeting, meeting deadlines or goals, interacting with others, misplacing possessions, and reasoning and problem-solving (Diamond, 2013).

Studies to date have almost exclusively examined EF outcomes in children and to some extent in adolescents born VPT/VLBW. In contrast, very few studies exist that have examined EF in adults born VPT/VLBW. This chapter discusses the relevant studies examining the EF abilities of adults born VPT/VLBW. Due to the nature of EF research using multiple task measures, the studies' findings have been reported by each EF domain: inhibitory control, working memory and cognitive flexibility.

3.1 Literature Review Methods

First, a comprehensive database search was conducted to identify existing published studies concerned with EF outcomes of adults born VPT/VLBW. The databases PsychINFO and PubMed were searched for relevant articles. Study selection criteria included: 1) sample consisted of adults age 19+ years born VPT (<33 weeks) and/or VLBW (<1500g), 2) inclusion of a term born control group, 3) EF outcomes measured using standardised tools designed to assess the constructs of inhibitory control, visual/verbal working memory and cognitive flexibility.

Nine studies were identified using these criteria. The methodology and findings of these studies are summarized in Table 1, including details of the sampling frames, year of birth, measures, constructs assessed, results, strengths and limitations. Studies in adulthood are defined from the age of 19 years and above based on the consensus within the research considering this age as young adulthood. Studies where the mean age was 19 years have been included in the review. Each study also included a measure of global cognitive ability (IQ), however as EF outcomes were the focus of the current review, IQ results are only reported when relevant.

Table 1 *Summary of Studies of Executive Function in Adults Born Very Preterm and/or Very Low Birthweight*

Author	Sample	EF Measures/Other relevant measures	Results	Strengths, Limitations & Comments
Nosarti et al. (2007)	61 VPT (<33 weeks), 64 Control Age: 20.62-24.78 Year of Birth: 1979-1982 Hospital cohort: University College Hospital London, United Kingdom	HSCT: Response Initiation/Inhibition TMT: Cognitive Flexibility WASI: Global Cognitive Ability	Total score: MD= .07, $p<.05$ Trail A: MD= 7.8, $p<.01$, Trail B: MD= 6.5, $p=ns$ Trails A-B: MD=1.5, $p=ns$ Full Scale IQ: MD= 6.61, $p<.01$	Lower follow-up rate (64.4%) EF differences speculated to be underestimated as participants retained had higher IQ scores than non-participants Control group recruited by media advertisements, may not be representative of the general population
Allin et al. (2011)	80 VPT (<33 weeks), 41 Control Age: 17-22 Year of Birth: 1982-1984 Hospital cohort:	HSCT: Response Initiation/Inhibition CVLT: Verbal Learning, Memory WMS: Working Memory	Total score: MD=1.1, $p<.001$ MD=4.2, $p=.02$ WMS immediate: MD= 4.4, $p=.09$ WMS delayed: MD= 4.8. $p=.06$	Extended analysis to examine white matter microstructure in relation to EF Representativeness issues as FT group recruited by volunteer advertisements Systematic differences in age between VPT (*M*=19.1) and FT group (*M*=18.5) a potential limitation for comparison

Author	**Sample**	**EF Measures/Other relevant measures**	**Results**	**Strengths, Limitations & Comments**
	University College Hospital London, United Kingdom	WASI: Global Cognitive Ability	Full Scale IQ: MD=10.8, $p<.05$	
Pyhälä et al. (2011)	103 VLBW (<1500g), 105 Control Age: 21.4-29.7 Year of Birth: 1978-1985 Hospital cohort: Helsinki University Central Hospital, Finland	TMT: Cognitive Flexibility Stroop Test: Inhibition, Working Memory CPT II: Sustained Attention, Inhibition ROCF: Visual Working Memory WAIS-III: Global Cognitive Ability	Trail A: $p<.05$ Trail B: $p<.05$ Trails A-B: $p<.05$ MD = 0.3-0.4 SD units MD= 0.1-0.3 SD units, $p<.05$ MD= VLBW group 0.2 SD units above control group mean, $p=ns$ Copy: MD= 0.3 SD units, $p<.05$ Immediate recall: MD=0.4 SD units, $p<.05$ Delay: $p=ns$ Full Scale IQ: MD=0.57, $p<.001$	Matched control group at birth High retention rates for both groups, 71.1% for VLBW, 68.2% for FT Selection bias; healthier participants of cohort overrepresented (exclusion of all those with neurosensory impairments)

Author	Sample	EF Measures/Other relevant measures	Results	Strengths, Limitations & Comments
Sølsnes et al. (2014)	42 VLBW (<1500g), 63 Control Age: 19-20 Year of Birth: 1986-1988 Hospital cohort: Trondheim University Hospital, Norway	Colour-Word Test (Stroop): Response Initiation/Inhibition TMT: Cognitive Flexibility Tower Test: Problem Solving (planning/ organisation) Design and Verbal Fluency: Problem Solving (initiation/fluency) WAIS-III: Global Cognitive Ability	MD=1.22 – 1.78, *p*=.02-.04 MD=0.73 – 3.06, *p*<.001 MD =0.4 (total task average), *p*=.06-.73 Design: MD= 1.73 (total task average), *p*<.01 Verbal: MD=1.58 (total task average), *p*=.06-.11 Full Scale IQ: MD=0.1 SD units, *p*<.001	Participants studied at different time points, well-established sample Lower follow up and retention rate over time Neonatal care improvements; potential generalisability limitations of results to preterm populations born after 1990
Eryigit Madzwamuse et al. (2015)	217 VPT/VLBW (<32 weeks, <1500g), 197 Control Age: 26 Year of Birth: 1985-1986	Stroop Test: Selective Attention, Verbal Inhibition, Cognitive Flexibility WASI-III: Global Cognitive Ability	MD=.06-.08 SD units, *p*<.001 Full Scale IQ: MD=1.2 SD units, *p*<.001	Large sample size & matched control group from birth Inclusion of adults with severe cognitive impairment (proxy cases)

Author	Sample	EF Measures/Other relevant measures	Results	Strengths, Limitations & Comments
	Regional cohort: Bavaria, Germany			Standardized test scores to term controls to control for Flynn Effects
Østgård et al. (2016)	55 VLBW (<1500g), 81 Control Age: 19-20 Year of Birth: 1986-1988 Hospital cohort: Trondheim University Hospital, Norway	WMS-III: Working Memory Stroop Test: Inhibition TMT: Cognitive Flexibility WCST: Working Memory, Cognitive Flexibility CPT II: Selective Attention, Inhibition Tower Test: Problem Solving Design and Verbal Fluency: Problem Solving WAIS: Global Cognitive Ability	MD=4 - 7.5, *p*<.001 MD=15.8 – 17.3, *p*<.001 MD=8.6 – 32.3, *p*<.001 MD=4.6, *p*=.32, MD=0.6-21.4, *p*=0.3-.60 MD=1.1, *p*=.07 Design: MD=7.9, *p*<.001 Verbal: MD=0.6, *p*=.08 Full Scale IQ: MD=0.1 SD units, *p*<.001	Comprehensive neuropsychological test battery No differences found between groups after controlling for SES, gestational age, birthweight and maternal age at birth, reducing selection bias Tester blinded to participant group status, reducing results bias

Author	**Sample**	**EF Measures/Other relevant measures**	**Results**	**Strengths, Limitations & Comments**
Kroll et al. (2017)	122 VPT (<33 weeks), 79 Control Age: 28-34 Year of Birth: 1979-1984 Hospital cohort: University College Hospital London, United Kingdom	HSCT: Response Initiation/Inhibition CANTAB subtests SOC + IED: Spatial planning, Attention TMT: Cognitive Flexibility CCPT-EC: Attention, Response Inhibition WASI: Global Cognitive Ability	MD=.74, $p<.001$ SOC: MD=.49, p=*ns* IED: MD=7.15, $p<.001$ MD=17.87, $p<.01$ MD=.72, p=*ns* Full Scale IQ: MD=8.5, $p<.001$	Later age group (M=31) useful to examine VPT participants recruited from wealthier geographical area (mean IQ scores within average range) Examiners not blinded to participants' status, potential results bias Controls selected via community advertisements, potential representative issues
Suikkanen et al. (2021)	133 EPT (<34 weeks), 241 LPT (34-36 weeks), 348 Control Age: 23.3 Year of Birth: 1985-1989 Regional cohort:	Groton Maze Learning Test: Visuospatial Working Memory Continuous Paired Associate Learning Test: Visual Associative Memory One Back Test: Working Memory	Number of errors: MD= 2.2-4.5, p=.18 Speed moves/10s: MD = 0.6, p=<.001 MD: 3.2, p=.78-.79 MD: 1.6 -3.8, p=.05 - .15	Large participant sample from a wide geographical region Cogstate: computer based tests rather than standardised EF measures, potential reliability and validity issues

Author	Sample	EF Measures/Other relevant measures	Results	Strengths, Limitations & Comments
	Northern Finland			
Woodward et al. (2021)	229 VLBW (<1500g), 100 Control Age: 26-30 Year of Birth: 1986 National Cohort: New Zealand	Computerized Sternberg-based task: Visuospatial Working Memory WISC-IQ	p=.02 - .78, Accuracy MD = 0.1 – 3.9 Reaction Time MD = 81.7 – 104.6 Efficiency MD= 83.6 – 204.5 Full Scale IQ: MD=10, p<.001	Representative cohort both regionally and nationally High recruitment and retention rates at 28 years (71%)

Notes: MD= Mean Difference
CANTAB=Cambridge Neuropsychological Test Automated Battery, CPT II=Conners' Continuous Performance Test, CCPT-EC=Continuous Performance Test-Errors of Commission, CVLT=California Learning Verbal Test; IED=Intra-Extra Dimensional Set Shift, HSCT=Hayling Sentence Completion Test, SOC=Stockings of Cambridge, TMT=Trail Making Test, ROCF=Rey-Osterrieth Complex Figure, WASI=Wechsler Abbreviated Scale of Intelligence, WAIS=Wechsler Adult Intelligence Scale, WCST=Wisconsin Card Sorting Test, WISC-IQ=Wechsler Intelligence Scale for Children, WMS-III=Weschler Memory Scale Third Edition.

3.2 Inhibitory Control

One of the first studies to examine EF outcomes in VPT adults was conducted by Nosarti et al. (2007) who examined a cohort of 61 VPT adults at a mean age of 22 years (range 20.62–24.78) and compared outcomes with 64 comparison adults born to FT. The VPT adults were a hospital based cohort born at the University College Hospital London. The study used the Hayling Sentence Completion Test (HSCT) as the inhibitory control measure to assess response initiation and inhibition. In this test, participants are required to first complete sentences with words that makes sense and then complete sentences with unconnected words in the second part. The findings showed that VPT adults performed significantly more poorly on the HSCT compared to FT born adults, with a mean difference of 0.7 (p<.05) after adjustment for IQ, sex and age. This means that on average VPT adults took longer to complete the test and made more errors (Nosarti et al., 2007).

Two later studies who also used VPT samples from the University College Hospital London, found similar findings on the HSCT. Allin et al. (2011) examined inhibitory control among 80 VPT adults at a mean age of 19 years (range 17–22) in comparison to 41 term born adults. Similar to the Nosarti et al. (2007) study, VPT adults performed significantly more poorly on the HSCT than FT born adults with a mean difference of 1.1 (p<.001). This finding also aligns with Kroll et al. (2017) who used the HSCT as a measure of inhibitory control among 122 VPT adults at a mean age of 31 years (range 28–34) in comparison to 89 term born adults. There was a moderate effect size of d=.66 between the two groups, with the VPT adults performing significantly lower on the HSCT (p<.001). Collectively, the three studies show that adults born VPT had greater difficulty with response initiation and inhibition on the HSCT.

Kroll et al. (2017) also assessed outcomes on the Continuous Performance Test - Errors of Commission (CCPT-EC) that taps into response inhibition, however between groups differences did not reach statistical nor clinical significance with a small effect size (*d*=.18). This is consistent with the findings reported by Pyhälä et al. (2011) who used a similar measure of response inhibition called the Conner's Continuous Performance Test II (CPT-II). Pyhälä et al. (2011) compared outcomes among 103 adults born VLBW at a mean age of 25 (range 21.4–29.7) and 105 FT born adults. The VLBW adults were a hospital based cohort born at the Helsinki University Central Hospital in Finland as part of the Helsinki Study of VLBW adults. Like Kroll et al. (2017), Pyhälä et al. (2011) did not find statistically significant differences in response inhibition on the CPT-II, as the mean difference suggested that majority of the VLBW adults did not perform more poorly on this task compared to FT born adults (almost 0.2 SD units above the control group mean). As Kroll et al. (2017) suggested, these findings may be attributed to methodology, i.e. a function of the tasks used, or they may suggest that certain aspects of inhibitory control may be more affected in VPT individuals than others.

Other studies have reported significant differences in inhibition using variations of the Stroop Test that assesses the ability to inhibit verbal responses. Sølsnes et al. (2014) compared outcomes among 42 adults born VLBW and 63 term born adults at 19–20 years who were a hospital based cohort born at the Trondheim University Hospital in Norway. Sølsnes et al. (2014) found that VLBW adults performed significantly more poorly than FT adults on the Stroop Test with a moderate effect size of *d*=.60 (*p*=.002). Similarly, Eryigit Madzawamuse et al. (2015) examined verbal inhibition using the Stroop task among 217 VPT/VLBW adults and 197 term born controls at age 26 years from the Bavarian Longitudinal Study. They found that on average VPT/VLBW adults performed 0.6 SDs units below the control group mean on the Stroop task.

Taken together, the existing studies generally suggest that adults born VPT/VLBW experience moderate difficulties in the domain of inhibitory control in comparison to term born controls.

3.3 Working Memory

Five existing studies have examined EF outcomes among VPT/VLBW adults with measures tapping into aspects of working memory. Allin et al. (2011) used two measures including the California Verbal Learning Test (CVLT) and the Weschler Memory Scale (WMS). Adults born VPT performed significantly more poorly than FT born adults on the CVLT with a mean difference of 4.2 (p=.02), however no significant differences were obtained on the WMS for both immediate and delayed working memory recall (p=.06-.09). Although the results for the WMS are trending towards significance, the study did not report effect sizes therefore it is difficult to infer the magnitude of difference between the VPT and FT adults.

Østgård et al. (2016) also examined working memory outcomes using a later edition of the Weschler Memory Scale-III (WMS-III) among 55 adults born VLBW and 81 term born controls at age 19–20 years. This cohort is the same hospital based cohort of the Sølsnes et al. (2014) study recruited from Trondheim University Hospital. In contrast to Allin et al. (2011), Østgård et al. (2016) found that VLBW adults performed significantly more poorly on the WMS-III in comparison to FT born adults, with large effect sizes (p<.001, d=.85–.89). However, the specific subtests of the WMS-III assessed mental control and spatial span which examine aspects of verbal and visuospatial working memory.

The findings from Allin et al. (2011) and Østgård et al (2016) suggest that VLBW adults may have greater difficulties in areas pertaining to verbal and visuospatial working memory, in comparison to retrieval processes of working memory. Additionally, Pyhälä et al.

(2011) found a similar result when using the Rey Osterrieth Complex Figure Test (ROCF) measuring visual memory processes involved in copying, immediate and delayed recall. The findings showed that visual memory performance of VLBW adults deteriorated in copying and immediate recall on the ROCF, however their performance did not deteriorate after a delay which is in line with what Allin et al. (2011) found in relation to delayed working memory recall. Together the studies suggest, that VPT/VLBW adults may have more intact long term retrieval processes and have greater difficulties in areas of verbal and visuospatial working memory. Two recent studies have specifically examined visuospatial working memory performance among VPT/VLBW adults to further suggest this is a particular area of difficulty for this group.

Suikkanen et al. (2021) examined visuospatial working memory using the Groton Maze Learning Test (GML) among 133 adults born early preterm (EPT, <34 weeks) and 241 adults born late preterm (LPT, 34–36 weeks) at age 23 years, in comparison to 348 adults born to FT. The EPT/LPT adults are part of a regional cohort from Northern Finland, who were recruited as part of the ESTER Preterm Birth Study of Northern Finland. To maintain consistency with the current review, the findings are focused on the adults born <34 weeks and will be referred to as VPT rather than EPT. Suikkanen et al. (2021) found that VPT adults performed significantly slower on the GML than FT adults, with 0.6 fewer moves per 10 seconds ($p<.001$), indicating weaker spatial memory efficiency.

These findings align with the study conducted by Woodward et al. (2021), who assessed visuospatial working memory performance among 229 adults born VLBW at a mean age of 28.5 (range 26.4–30.6 years) and 100 FT born adults. This cohort is a national cohort recruited at birth as part of the New Zealand VLBW study. They were assessed using a four span/difficulty level computerized task adapted from the Sternberg working memory paradigm. Woodward et al. (2021) found that adults born VLBW generally performed more poorly compared to term born controls demonstrating a performance decrement across three measures assessing accuracy, reaction time and efficiency with small to moderate effects (p=.02–.78, d=.03–.40). Additionally, task performance across the four different set levels were examined, with span level 1 being the easiest (cognitive load is lowest) and level 4 the most cognitively challenging. Overall VLBW adults were less accurate, took longer to respond and were less efficient in their responses compared to their EF counterparts as the demands of the task increased in complexity, suggesting that VLBW individuals may have greater difficulty when cognitive load is higher. Collectively, these results suggest that VPT/VLBW adults appear to demonstrate persistent impairment in visuospatial working memory.

3.4 Cognitive Flexibility

Some existing studies have used the Trail Making Test (TMT) to examine outcomes in the domain of cognitive flexibility among VPT/VLBW adults. The TMT is made up of two parts. Part A consists of visual scanning tasks, where participants must draw a line, for example, to connect numbered circles in ascending order. Part B consists of alternate switching between different task sets where participants must connect circles labelled by numbers and letters in an alternating fashion. Trails A-B is calculated as the difference between in speed/trail completion time the two parts as the measure of cognitive flexibility.

Kroll et al. (2017) found that VPT adults performed significantly poorer on the TMT compared to full term controls, indicating they had greater difficulty flexibly switching and efficiently problem solving between different tasks (p<.01, d=.42). Similarly, Sølsnes et al. (2014) found a large effect size difference between VLBW adults and those born to FT ranging from d=.74–.88 across the TMT tasks (p<.001) and Pyhälä et al. (2011) reported VLBW adults performed slower with higher inaccuracy on this task with mean differences ranging from 0.3–0.5 SD units below the control group mean. In contrast to these findings however, Nosarti et al. (2007) found VPT adults were significantly slower than controls to complete Part A of the trails (p<.01), but there were no statistically significant group differences on Part B or the Trails A-B score (p>.05).

In another study, Eryigit Madzawamuse et al. (2015) used the Stroop Test consisting of three tasks that assessed various EF domains including cognitive flexibility. The findings showed that adults born VPT/VLBW scored significantly lower on this task, with scores ranging from 0.5–0.8 SD units below the control group mean. In summary, although there are some inconsistencies in the literature, the majority of the existing studies to date have found cognitive flexibility to be a domain of impairment for adults born VPT/VLBW.

3.5 Role of Processing Speed

Processing speed involves the integration of efficiency, speed and fluency which are necessary for an individual to process information and are particularly important requirements for success on timed tasks (Anderson, 2002; Pyhälä et al. 2011). Some EF tasks have a timed component, and it is important to consider the role of processing speed on EF task performance. Pyhälä et al. (2011) is one of the earlier studies to note that slower performance in the VLBW group on timed EF tasks, may be accounted for by slower processing speed. For example, the studies conducted by Østgård et al. (2016) and Sølsnes et al. (2014) found no significant differences between adults born VLBW and FT on tests with no time limits such as the Tower Test, a measure of planning and problem solving (p>.05). Effect sizes on the Tower Test (total correct and total time) for the study by Sølsnes et al. (2014) ranged from .07–.41, and for the study by Østgård et al. (2016) these ranged from .30–.32. In contrast, some of the previously mentioned EF tasks including the TMT, Stroop tests and computerized visuospatial working memory tasks, are all timed and generally

VPT/VLBW adults performed significantly more poorly across these tasks. The results, in conjunction with Østgård et al. (2016) and Sølsnes et al. (2014), may suggest that performance on EF tasks is impacted to some extent by processing speed ability of VPT/VLBW adults. Processing speed is therefore an important area to consider when examining EF performance among this population.

3.6 Executive functioning in everyday life

Three studies have reported on how EF difficulties could be implicated in the everyday functioning of adults born VPT/VLBW. In addition to examining performance across various EF tasks, Kroll et al. (2017) examined how performance on those tasks associated with real-life outcomes. They found that EF difficulties among the VPT born adults had a stronger positive association with poorer real-life outcomes compared to the FT group, including work status (employed vs. unemployed) and years in full-time education (Kroll et al., 2017). This finding indicates the importance of EF for other areas of functioning and how EF deficits may differentially impact VPT populations in terms of educational and socioeconomic success. In contrast, Sølsnes et al. (2014) used the Behaviour Rating Inventory of Executive Function for Adults (BRIEF-A), a self-report measure of EFs in everyday life. They found that VLBW adults at age 19 years do not report more EF problems in daily life compared to FT adults, despite lower performance found on tasks across the three EF domains. This finding aligns with a recent study by Kim et al. (in press) who also used the BRIEF-A measure and found an improvement in self-reported executive functioning among VLBW adults at age 28 years. Kim et al. analyzed the data from the New Zealand VLBW study (Woodward et al., 2021), but these results are not reported in Table 1 due to the preliminary findings. Collectively, the results of Sølsnes et al. (2014) and Kim et al. (in press) suggest that adults born VLBW tend to perceive themselves as functioning well in everyday contexts, despite potential EF difficulties they may have. More research is required to understand how VPT/VLBW adults may experience and perceive EF difficulties in their everyday lives.

3.7 Methodological limitations of the current literature

The studies reviewed in Table 1 provide useful findings for understanding EF outcomes among adults born VPT/VLBW, however they are not without limitations. Common limitations, as discussed below, include issues with generalizability, recruitment methods, sample representativeness and control groups.

3.7.1 Generalizability across preterm populations

Advances in neonatal care over time could mean that EF deficits present in adulthood, may not generalize across other VPT/VLBW populations. Most of the studies included participants born in the late 1980's although some included those born in the late 1970's and therefore their findings may not be representative of VPT/VLBW cohorts born in more recent years (Nosarti et al., 2007; Pyhälä et al., 2011; Kroll et al., 2017). Neonatal intensive care has improved substantially over time, therefore there needs to be some caution when generalizing results to VPT/VLBW populations born in later years.

Additionally, some studies excluded participants who had severe cognitive or neurosensory impairment due to inability to perform the neuropsychological tests (Pyhälä et al., 2011; Sølsnes et al., 2014; Suikkanen et al., 2021). For example, in the study conducted by Nosarti et al. (2007), individuals with cerebral palsy and hearing impairments completed part of the assessments but their results were excluded from the analyses. Whereas, Eryigit Madzwamuse et al. (2015) included adults with severe cognitive impairment as proxy cases, to give a better estimate of the burden of VPT/VLBW birth. Inclusion of all participants is important in understanding true population estimates of EF abilities, so as not to underestimate the level of impairment experienced by these individuals and the level of support that they may require. Due to the variation across the studies, there must be caution in generalizing the results across VPT/VLBW populations.

3.7.2 Recruitment methods

In the literature, there are various methods of recruitment that have been employed for examining VPT/VLBW populations. Cases include preterm infants that have been recruited from one specific hospital such as Helsinki University Central Hospital, Trondheim University Hospital and University College Hospital London, (Pyhälä et al., 2011; Østgård et al. 2016; Kroll et al., 2017). These establishments are major teaching hospitals encompassing geographical areas that represent a higher socioeconomic status (SES), therefore the participants recruited are representative of a particular region and may not necessarily represent the general VPT/VLBW population. Eryigit Madswamuse et al. (2015) and Suikkanen et al. (2021) recruited regional cohorts from the geographically defined areas of Southern Bavaria, Germany and Northern Finland respectively. Other regional longitudinal studies have examined EF outcomes in EPT (<28 weeks) populations, including the Arvo Ylppö Longitudinal Study of Uusimaa, Finland and the Victorian Infant Collaborative Study (VICS) of Victoria, Australia (Heinonen et al., 2018; Burnett et al., 2018). National studies include the New Zealand 1986 VLBW cohort study that recruited all infants born <1500g and the EPICure cohort study, consisting of ELBW infants born throughout the United Kingdom and Ireland in 1995 and 2006 (Woodward et al., 2021; O'Reilly et al., 2020). These various recruitment methods mean that some participant groups may be more representative of preterm populations than others.

3.7.3 Participant retention

Low participant retention rate is a common limitation among the existing longitudinal studies. Some studies reported reasonably high retention rates upwards of 60% for both the VPT/VLBW and FT groups (Pyhälä et al., 2011; Eryigit Madzwamuse et al., 2015; Woodward et al., 2021). Nosarti et al. (2007) reported a low follow-up rate attributed to sample loss and attrition; some individuals could not be traced, and a considerable part of the cohort did not want to participate. The original cohort study included *n*=224 infants born VPT, however at the follow-up assessment (mean age, 22.25), *n*=61 individuals participated (Nosarti et al., 2007). Sølsnes et al. (2014) had a similarly low follow-up rate, with *n*=42 individuals born VLBW participating at the 19-20 year follow up from an original cohort of *n*=121. Sample loss and attrition over time can contribute to sample selection bias. In these cases when not all individuals can be retained in the study, the analysis is conducted with the remaining participants who represent a subset of the data. The remaining participants therefore may not be representative of the sample intending to be analyzed and this can impact the generalizability of the study results. Additionally, lower retention rates can threaten the validity of the study by decreasing the sample size, thus reducing statistical power and the ability to detect significant associations.

3.7.4 Control group representativeness

Another limitation in the literature is that studies used various methods of recruiting a comparison group, with the representativeness of the general population therefore varying greatly. Some methods used to recruit controls were via media and community advertisements that volunteers responded to (Nosarti et al., 2007; Allin et al., 2011; Kroll et al., 2017). Other studies used a control group matched for sex, age and birth hospital (Pyhälä et al., 2011; Eryigit Madzwamuse et al., 2015). Comparison groups are important in a study to help control for threats to validity. Without an adequate control group, it can be difficult to determine if the outcomes are due to preterm birth exposure or if they are due to other risk factors. It can also be difficult to accurately determine the level of impairment in relation to the general or typically developing population, and therefore what intervention or support is likely to be required.

While it is important to acknowledge methodological limitations to help prevent these in future research, it appears that these limitations tend to not exaggerate the difficulties of adults born VLBW but rather underestimate them. The findings have suggested that adults born VPT/VLBW experience difficulties across a range of EF tasks, however limitations of existing research means that it is important to address the extent of EF difficulties among this population in future with methodologically rigorous studies.

3.8 Rationale for current study

It is thought that EF abilities can continue to develop into the third decade of life (Wiebe & Karbach, 2018). Therefore, there is an interest in determining whether the relative EF impairment observed in VPT/VLBW children and adolescents does persist into adulthood or might improve with time and age. Based on the findings of the current studies, although EF abilities may continue to develop into early adulthood, there is some suggestion that difficulties observed in childhood and adolescence may persist overtime for the VPT/VLBW population across the domains of inhibitory control, working memory and cognitive flexibility. Additionally, some of the findings indicated that processing speed may be implicated in EF task performance for VPT/VLBW individuals. However, the existing data are not hugely compelling for three main reasons, 1) very little EF research has been conducted among VPT/VLBW adults, 2) there are methodological limitations in the extant literature and 3) there are contradictory findings for some of the EF measures and domains.

As such, there are current gaps in the literature as to the extent of EF impairment among VPT/VLBW adults in the third decade of life.

The current study aims to investigate EF outcomes among a national cohort of adults born VLBW in New Zealand. The initial findings of EF outcomes among the New Zealand VLBW cohort have solely examined self-reported executive functioning and findings from a psychometric measure of visuospatial working memory. The preliminary findings by Kim et al. (in press) suggested there is an improvement in self-reported executive functioning among VLBW adults and Woodward et al. (2021) showed that visuospatial working memory appears to be an area of difficulty for VLBW adults. Currently, EF performance has not been examined across a range of measures among this cohort raising the question of whether VPT/VLBW adults display global EF difficulties across multiple domains of EF and to what extent are their impairments in these areas.

Chapter 4

Research Aims

EF outcomes have almost exclusively been examined in children and adolescents born VPT/VLBW and higher rates of EF impairment are consistently reported among this population compared to a FT comparison group. In contrast, existing literature reporting on EF outcomes in adulthood following VPT/VLBW birth is limited. Preliminary studies to date reporting on a narrow range of EF measures, suggest that earlier observed EF impairments may persist into adulthood. The current study is novel in assessing outcomes among a national cohort of adults born VLBW in New Zealand across the three key domains of EF, using a range of EF measures. Thus, the overall aim of the present study was to characterize an executive functioning profile of a national cohort of adults born VLBW, by examining how they perform across various EF measures administered as part of a comprehensive neuropsychological assessment.

The specific research aims were to:

1) Compare the EF abilities of a national cohort of adults born VLBW in New Zealand during 1986 and followed up at age 26–30 years with a control group of same-age individuals who were born FT. Executive functioning was assessed using a range of measures assessing three key constructs of EF. These included inhibitory control, working memory (visuospatial and auditory-verbal) and cognitive flexibility.
2) To examine the extent to which observed differences in EF task performance between VLBW and FT born adults might be explained, in part or in full, by the effects of processing speed.

3) To assess the effects of birthweight on EF task performance. Of interest was the extent to which risk of EF difficulties might increase with decreasing birthweight.

Chapter 5

Method

5.1 Overall Design of the National 1986 VLBW Study

This research draws on quantitative data from the VLBW study, a prospective, longitudinal cohort study of the health and developmental outcomes of a national cohort of adults who were born VLBW in 1986 and followed to age 26–30 years (Darlow et al., 2015). The participants in the VLBW sample included all live born infants with a birthweight of <1500g who were admitted to a neonatal intensive care unit in New Zealand during 1986 (*n*=413). All VLBW infants were enrolled in a prospective cohort study of acute retinopathy of prematurity (Darlow, 1988; Darlow et al., 1992). As shown in Figure 1, 338 survived to hospital discharge and have been followed from birth and were assessed in infancy, at 7–8 years, 22–23 years and 26–30 years. At 7–8 years, 298 participants were assessed (91.4% of survivors) and 230 were assessed at 22–23 years (71% of survivors). The focus of this thesis is on the executive functioning outcomes of adults born VLBW at the 26–30 year follow-up. The primary data for this thesis were collected from the cognitive and neuropsychological test battery that was administered as part of the comprehensive medical and psychological assessment of the New Zealand VLBW cohort.

The study sample and measures included in the present analysis are described in detail below. Further details of the VLBW study are available in prior publications (Darlow et al. 2015).

5.2 Participants

Adults born very low birthweight. By 26–30 years, 323 participants were alive and eligible for follow-up. Of these, 229 completed the full assessment, reflecting a sample retention rate of 71%. Reasons for sample loss at the 26–30 year follow up included participants who were unable to be traced (*n*=35), declined to participate (*n*=38) or completed only the questionnaire component of the full assessment (*n*=21). The mean age of the VLBW participants at assessment was 28.5 years (26.4–30.6 years). Further details of the VLBW group are presented in Figure 1 outlining the original recruitment procedure and sample retention rates from birth to adulthood.

Comparison adults. Alongside the VLBW cohort, a comparison group of 100 age-matched control participants born FT (>37 weeks GA) in 1986 and who were not admitted to a neonatal intensive care unit were assessed. These individuals were recruited as part of the 22–23 year follow-up of the main study cohort, supplemented by peer nominations at age 28. To ensure a balance across regional distribution, gender and ethnicity, the comparison group were recruited via random sampling from the electoral rolls (*n*=69) or through peer nomination by a VLBW cohort member. The mean age of the comparison group at follow-up was 28.2 years (26.6–30.1 years).

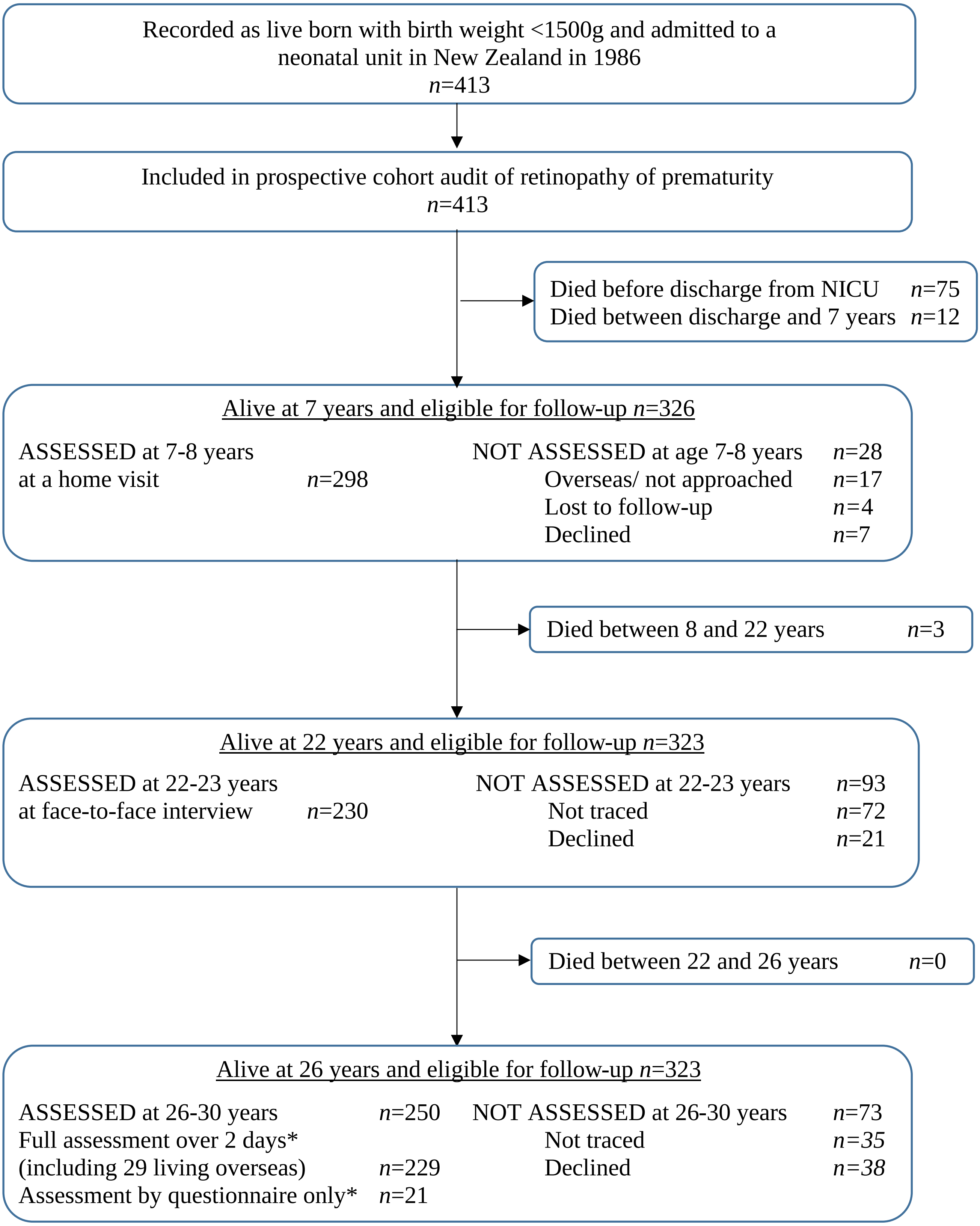

Figure 1. New Zealand 1986 VLBW Adult Follow-up Study: Cohort flow chart

* Between February 2013 and November 2016; NICU: Neonatal intensive care unit [Modified, with permission from Darlow BA, et al. Metabolic syndrome in very low birth weight young adults and controls: the New Zealand 1986 VLBW Study. *J Pediatric* 2019;206:128-33 (133.e1)]

5.3 Procedure

When the participants reached 26–27 years of age, they were contacted by telephone and invited to take part in a health and wellbeing multidisciplinary evaluation conducted over two days in Christchurch, New Zealand. The evaluation consisted of the following assessment components; cardioendocrine, respiratory, visual, MRI, psychological testing, and health and social functioning questionnaires. The focus of the present thesis was to characterize the executive function of the cohort using psychometric assessment measures that were selected to provide a comprehensive evaluation of a range of key executive abilities. Participants' accommodation, food and travel expenses were all covered by the study. Upon completion of the two-day assessment, participants were reimbursed $200 as compensation for their time.

The psychological assessment lasted approximately 2-3 hours, including breaks as needed. Assessments were administered by an experienced Clinical Psychologist, who was unaware of the participant's group status. All testing was conducted in a quiet room, with the participant sitting at either a desk or a computer screen. Informed, written consent was obtained from all participants. Ethical approval for the 26-30 year follow-up was obtained from the Upper South B Regional Ethics Committee (superseded by the Southern Health and Disability Committee) (Ref no. URB/12/05/015).

5.4 Measures

Five measures of executive function were administered. These measures were chosen to reflect the core executive domains identified in the Introduction. All measures were age appropriate/suitable for use within an adult population, had well-validated psychometric properties, and had been previously used to assess neuropsychological impairment in a diverse range of clinical and typically developing populations. For the current study, participant's performance on each executive function task were represented by raw, scaled, t-score or standard scores where appropriate. In addition, an individual's level of impairment on each test was calculated. To define impairment on each measure, the distributions of scores for the FT group were reviewed to determine appropriate cut-off points, in this case the lowest 10% of scores from the FT group. This was done to avoid the Flynn effect. The Flynn effect refers to the tendency for IQ scores to change overtime, and specifically the observed rise in IQ scores reflecting an apparent IQ increase among the general population (Flynn, 1987; Pietschnig & Voracek, 2015). This method was chosen as opposed to solely using test norms to avoid problems associated with the Flynn effect due to outdated test norms (Flynn, 1987). This is a well-established and accepted approach in the research literature (Trahan et al., 2014; Eryigit Madzwamuse et al., 2015; Woodward et al., 2021). In addition, this approach also helped to ensure cut-offs were consistent across measures. A description of each test, including its administration details, scoring and psychometric properties, is provided below for each EF domain.

5.4.1 Inhibitory Control

5.4.1.1 Hayling Sentence Completion Test

The Hayling Sentence Completion Test is a sentence completion task that measures inhibitory control. The test is entirely spoken, making it suitable for people with a wide range of impairments, including visual and motor function. The test consists of two sections, with each made up of a series of 15 sentences. In the first section (*Hayling 1*), the participant must complete the sentence the examiner reads aloud with a word that makes sense (Burgess & Shallice, 1997). In the second section (*Hayling 2*), the participant is asked to give a word that is unconnected to the sentence in any way, therefore inhibiting an expected response (Burgess & Shallice, 1997).

There are four scaled scores calculated for the Hayling test. Total time is calculated for both Part 1 (Initiation) and Part 2 (Inhibition), Category A (sentence is plausible) and Category B (somewhat connected to the meaning of the sentence) errors are combined to produce an error scaled score. An overall scaled score is calculated by summing the first three scaled scores. Hayling scaled scores range from 1 to 10, with 1= 'impaired', 6 = 'average' and 10 = 'very superior'. Any inhibitory control difficulty/impairment for this study was determined as a Hayling scaled score of 5, or a below average score, as noted in the test manual. This cut-point was further validated following an examination of the score distribution of the full term control group, for which the majority achieved a score of 6 or above.

The Hayling Test has good internal consistency, with split-half reliability coefficients of .93 for Hayling 1, .80 for Hayling 2 and .72 for Hayling Errors within an impaired control group with unilateral lesions (Burgess & Shallice, 1997). Test-retest reliability coefficients were .62 for Hayling 1 and .78 for Hayling 2 when retested on a group of 31 volunteers between two days and four weeks after first assessment (Burgess & Shallice, 1997). These findings demonstrate that participants with unilateral lesions will perform just as equally on both halves of the test, and that the Hayling measure has good test stability overtime. In terms of the Hayling Test's construct validity, moderate correlations were reported with other inhibitory control measures including the Six Elements Test in an adolescent population (0.40–0.65) and the Tower of London (.40–.64) (Clark et al., 2000; Andrés, & Van der Linden, 2000). Odhuba et al. (2005) considered the validity of the Hayling test in patients with brain-injury and reported strong associations with patient's disability ratings on the Dysexecutive Questionnaire (DEX Self). The questionnaire scores indicate executive functioning in daily life; the overall Hayling Score, Hayling 1 and Hayling 2 all demonstrated strong associations, supporting that the construct validity of the test is sensitive to difficulties experienced by patients with dysregulated executive functioning (Odhuba et al., 2005).

5.4.2 Working Memory

5.4.2.1 Sternberg Visuospatial Working Memory Task

This task is a novel computerised visuospatial working memory task adapted from the original Sternberg spatial working memory (SWM) paradigm (Sternberg, 1966; Woodward et al., 2021). The task was run through E-Prime (version 2.0). As depicted in Figure 3, during each SWM task trial a probe display is presented on a computer screen for 1000 ms. A visual array of objects to memorize is then presented for 3000 ms. Then, following a 5000 ms delay, participants are shown only one of the previously presented objects (the target stimulus). They must determine whether this target stimulus is in the same location as it was during its initial presentation.

Item stimuli are pictures of everyday objects. The number of stimuli in the visual array that participants attempted to memorise ranged from 1-4. Trials therefore varied in difficulty, depending on the number of objects the participant was required to memorise and recall the spatial location of (Woodward et al., 2021). There were 12 trials per set size, administered as four blocks and interspersed with rest breaks. In its development to be administered as part of the EF battery, participants received comprehensive training prior to testing, which covered all testing requirements. Feedback was provided until performance was 100% accurate. A further 12 computer-based practice trials were administered with feedback before commencing the official Sternberg task (Woodward et al., 2021).

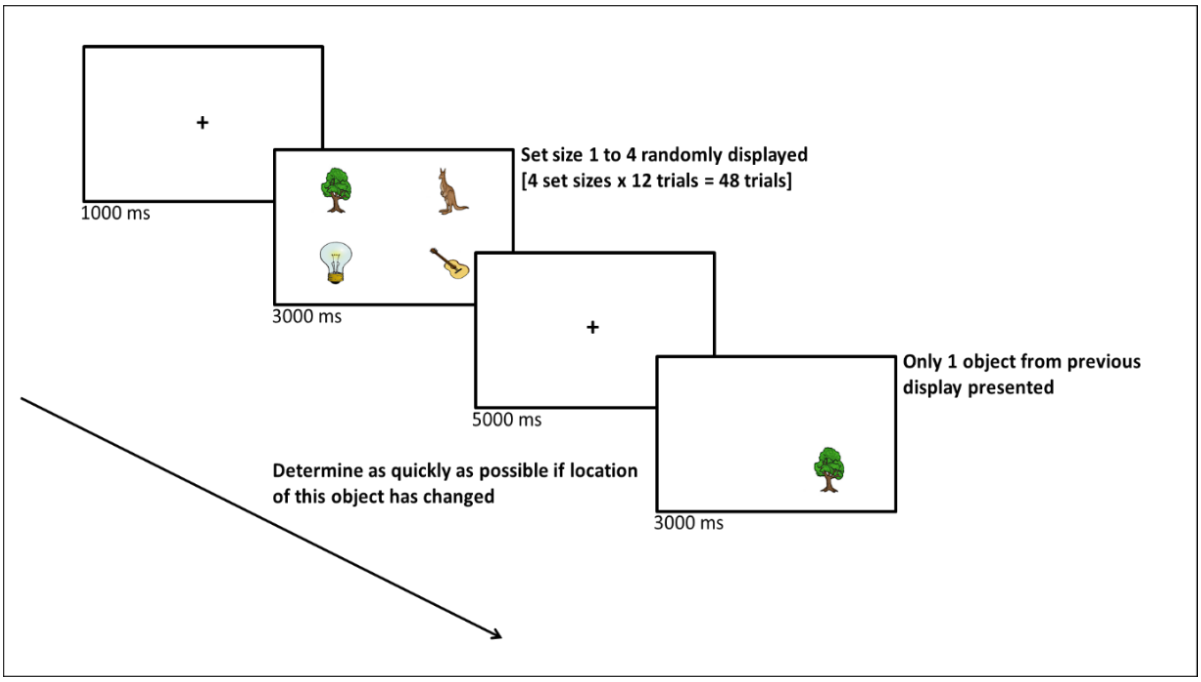

Figure 2. Example stimulus presentation trial in the Sternberg Visuospatial Working Memory Task (Woodward et al., 2021)

Task performance was assessed using the following variables: accuracy (% of correct trials), mean reaction time (on successful trials) and efficiency (ratio of reaction time to accuracy). These three variables were averaged across each trial set. Visuospatial working memory performance was assessed in the current analysis using "efficiency" as the primary outcome measure, given that it takes into consideration both the accuracy and mean reaction time of responses. Further, this Sternberg-based SWM task has previously been published including the analysis of results across all three variables for the same sample groups used in this thesis. For a full description refer to Woodward et al. (2021). Impairment on this task was defined by an efficiency score below the 10^{th} percentile of the FT comparison group on Set 4, with 2000ms used as the cut-off.

5.4.2.2 Auditory-Verbal Working Memory

Auditory-verbal working memory was assessed using two subtests from the Test of Everyday Attention (TEA); *Elevator counting with Distraction* and *Elevator Counting with Reversal*. Although tests of auditory selective attention, they also involve aspects of working memory. Specifically, they are sensitive to the components of manipulation and sequencing of auditory-verbal information that is relevant to working memory. Both subtests are presented to participants on an audio-tape. The scenario presented to participants is to pretend they are in an elevator with a non-functioning visual floor-indicator. Practice trials occur at the beginning to demonstrate the difference between high and low tones and allow the participant to practice counting.

Elevator counting with Distraction requires participants to count a series of low tones to establish which floor of a building they have arrived at, while excluding higher-pitched distractor tones (Robertson et al., 1994). *Elevator Counting with Reversal* involves the same scenario but with normal (medium pitched), low and high tones. Participants are required to count backwards upon hearing a low tone and upwards on hearing a high tone, while continuing to count in the same direction on hearing the normal tones. The nature of the task requires the participant to update and manipulate this auditory information in working memory.

For both subtests the overall score is the total number of correct responses out of 10. Following examination of the test manual and the FT group score distribution, a cut-point indicative of impairment was determined based on the lowest 10% of the FT group scores. A score of 7 or below was used to indicate impairment for the *Elevator Counting with Distraction* subtest and a score of 8 or below was used as the cut-off for *Elevator Counting with Reversal*.

The TEA Elevator counting tasks have good test-retest reliability; coefficients for one week between .66–.71 for healthy controls and 0.83 for stroke patients (Robertson et al., 1994). A principal components analysis was carried out to examine the construct validity of the TEA subtests. The two Elevator counting subtests loaded onto an Auditory-verbal working memory factor and correlated with the externally validated tasks; Backward Digit Span (.77) and the Paced Auditory Serial Addition Test (.58) (Robertson et al., 1994).

As these are subtests with an auditory component the validity has been tested to minimise the issue of unsuspected hearing deficits. A comparison of hearing impaired (M=5.0, SD=3.0) and control participants (M=4.7, SD=3.1) produced similar results on the Elevator Counting with Reversal task, to suggest the tests are sensitive to hearing impairment.

5.4.3 Cognitive Flexibility

5.4.3.1 Comprehensive Trail Making Test (CTMT)

The CTMT is based on the original Trail Making Test (TMT) that has been standardised for use with individuals ranging in age from 8 to 74 years. The test consists of five visual search and sequencing tasks. For each of the five tasks, participants must create a “trail” by connecting a series of stimuli in a specified order as quickly as possible (Reynolds, 2002). The five tasks increase in complexity through the addition of distractor stimuli and by requiring the participant to connect alternating numerical and lexical number stimuli targets (Moses, 2004). The CTMT takes around 5–10 minutes to administer. The CTMT is a widely used screening instrument in neuropsychological practice because it is highly sensitive to frontal lobe deficits including cognitive flexibility or set-shifting (Reynolds, 2002).

Each CTMT trail produces a raw score, reflecting the number of seconds required for the participant to complete the trail. Errors made on the trails are not converted to a scaled score because the examiner corrects the errors; these corrections are then added to the time needed to complete the trail. Raw scores are then converted to standardised T-scores, and a

Composite Index Score indicating performance across all five trails is obtained. Impairment on this test was defined as a score below the 10^{th} percentile of the comparison group.

The CTMT has demonstrated psychometric properties. Specifically, it has reasonable reliability with internal consistency coefficients for the five trail measures meeting or exceeding 0.70 (Moses, 2004). Test re-test reliability values for the five trails range from .70 to .78, and interrater reliability values ranging from .96 to .98 (Moses, 2004). In terms of test validity, Allen et al. (2009) examined the construct validity of the CTMT in children and adolescence with TBI. For convergent validity, the CTMT significantly correlated with tests of sustained attention, motor function and the Wechsler Index assessing processing speed and perceptual organisation (Allen et al., 2009). All are cognitive abilities required to complete the CTMT. In terms of discriminant validity, the CTMT had lower correlations with the verbal index from the Wechsler scales (Allen et al., 2009). Armstrong et al. (2008) has also examined the sensitivity of CTMT in adolescence with traumatic brain injury (TBI), confirming the findings of Allen et al. (2009). Additionally, the CTMT composite index scores yielded sensitivity and specificity estimates of .77 and .90 respectively, suggesting that the CTMT has the ability to differentiate between a clinical group of individuals with a TBI and a non-clinical group of healthy controls, with high degrees of sensitivity and specificity (Armstrong, 2008).

5.4.3.2 Brixton Spatial Anticipation Test

The Brixton Spatial Anticipation Test involves detecting rules in a sequence of stimuli. The test is presented to participants using a 56-page stimulus book. Each page presents ten circles, numbered 1 to 10, in two rows of five. One of the circles is filled in on each page and, as the participant is shown one page at a time, they are asked to indicate where the next filled circle will be (Burgess & Shallice, 1997). After a pattern is established, the rule changes and this continues throughout the stimulus book. The Brixton does not require a verbal response, making it suitable to administer in individuals who are non-verbal or have difficulties with speech production. The test is designed to assess executive functioning associated with frontal lobe lesions and is sensitive to the assessment of problems in dysexecutive patients, including rule detection (Burgess & Shallice, 1997).

The total number of errors made on the Brixton test is used to calculate a scaled score. The participant's response to the first item is always disregarded because it is a guess. The participant cannot know about the rule changes, therefore on these trials, the correct answer is where the blue circle would go had it not been for the rule change. The number of errors is added and converted to a scaled score between 1 – 10 (Burgess & Shallice, 1997). Impairment for this test was determined as a scaled score of 5, or a below average score, as noted in the test manual. This cut-point was further validated following an examination of the score distribution of the full-term control group, for which the majority scored an average of 6 or above.

Burgess and Shallice (1997) report adequate consistency reliability for the Brixton Spatial Anticipation test, with split-half reliability coefficients of .62 for a healthy control group. This was considered adequate given the nature of the task involving serial dependence between trials; meaning that what the participant has viewed previously will influence what they see and report next. Additionally, test re-test reliability was assessed on a group of 31 participants who were retested after initial testing at either two days, one week or four weeks, generating a value of .71 for all participants (Burgess & Shallice.,1997).

Vordenberg et al. (2014) administered the Brixton test to a group of 57 stroke patients with frontal lobe and subcortical lesions. Using the Functional Independence Measure (FIM), the Brixton test significantly correlated with the FIM cognitive subtotal and was unrelated to the FIM Motor subtotal, which coincides with the task not requiring motor ability (Vordenberg et al., 2014). In terms of convergent validity, the Brixton correlated with other tests measuring executive functioning including the Trail Making Test (TMT) measuring cognitive flexibility/set-shifting, and the Digit Span subtest measuring attention and working memory (Vordenberg et al., 2014).

5.4.4 Additional relevant measure: Processing speed

5.4.4.1 Symbol Digit Modalities Test (SDMT)

The SDMT is a commonly used measure of processing speed, originally developed to screen for cerebral dysfunction and has demonstrated sensitivity in identifying neurological impairment (Sheridan et al., 2006). Participants are provided with an A4 size response worksheet that presents rows of different symbols. There are nine different symbols in total, with each corresponding to the numbers 1–9, as presented in a reference key at the top of the worksheet. The participant has 90 seconds to manually pair the numbers with the symbols using this reference key (Sheridan et al., 2006). The overall score is the correct number of substitutions made in 90 seconds, with scores ranging between 0 and a maximum of 110.

Sheridan et al. (2006) supports the use of SDMT as a robust screening measure in a non-clinical community-based sample. They found that age, education, gender and income did not impact test performance on the SDMT, suggesting it is robust to the influence of social demographic factors (Sheridan et al., 2006). Based on the normative sample, the SDMT manual reports a test-retest reliability coefficient of .76 for the oral response version, and a correlation of .80 between the written and oral administrations in healthy samples (Sheridan et al., 2006).

The SDMT has been a common cognitive performance measure in multiple sclerosis patient populations, with test-retest reliability reported to be as high as .97 over 2 weeks and .74 over 2 years (Benedict, 2005; Drake et al., 2010). In terms of construct and predictive validity, the SDMT loaded onto factors of speed, executive function and memory (Benedict et al., 2017). This multimodal aspect of the SDMT accounts for high sensitivity to cognitive impairment in specific patient populations (Benedict et al., 2017).

5.5 Data Analysis

Data were analyzed using IBM SPSS Statistics 27. The analyses were conducted in five main steps. First, a data cleaning process was carried out which involved carefully reviewing the data against testing notes. This was done to check for missing/incomplete data, and reasons for test non-completion. Frequency distributions for each of the main outcome measures were examined to assess the extent of normality.

Secondly, between-group differences in EF task performance (VLBW vs. FT) were examined using independent samples t-tests for continuous variables. Mann-Whitney U *(U)* tests were also used where outcome variables were skewed, however the results did not change overall, and therefore results from the t-tests have been reported. Cohen's *d* was calculated as a measure of effect size, to indicate the magnitude of the between-group differences, using the following conventions suggested by Cohen (1988); small (*d*=0.2), medium (*d*=0.5) and large (*d*=0.8). To examine rates of impairment among VLBW and FT adults, cut-off points for impairment on each EF measure were calculated. As described in the measures section, impairment cut-points for each measure were determined based on the lowest 10% of scores from the FT group or test norms where appropriate. Odds ratios (OR) were then calculated to compare the risk of EF impairment of adults born VLBW compared to those adults born to FT.

Thirdly, the extent to which between-group differences could be explained by birthweight when controlling for processing speed was examined using a one-way analysis of covariance (ANCOVA). Partial eta squared (η_p^2) was calculated as a measure of effect size to determine the proportion of variance explained by birthweight on EF task performance, after accounting for the variance of processing speed. Lakens (2013) reported the following conventions for η_p^2: small (.01), medium (.06) and large (.14 and above).

The fourth step in the statistical analyses was to examine the effects of birthweight groupings on EF task performance using a linear one-way analysis of variance (ANOVA). For this analysis, study participants were further divided into three groups based on birthweight; FT (>1500g), VLBW (<1500g) and ELBW (<1000g). Of interest, was whether the effects of decreasing birthweight were associated with a decline in subsequent EF task performance in adulthood, since previous research undertaken with children and adolescents suggested that this may be the case. Eta squared (η^2) was calculated to measure the proportion of variance within an EF task that was accounted for by birthweight group.

The final step was to examine the relationships between the five EF measures. Bivariate correlations were assessed using Pearson's *r*. Following this, an exploratory Principal Components Analysis (PCA) was conducted to assess whether the measures loaded onto one or more EF factors. Based on those results, Confirmatory Factor Analysis (CFA) was performed using AMOS to create measurement models and compare the different model fits.

Chapter 6

Results

6.1 Characteristics of the sample

Table 2 presents the neonatal and social background characteristics of the two groups of participants born VLBW and FT. With respect to infant neonatal characteristics, as expected based on sample selection criteria, individuals in the VLBW group were born at a significantly lower mean birthweight than the FT group ($p<.001$). Both groups included similar proportions of male and female participants (*p*=).

Additional clinical data are reported in Table 2 for the VLBW group only, given the later recruitment of FT control participants and the fact that these conditions do not occur, or are extremely rare, in full term births. The mean gestational age for this group was 29 weeks (range 24-36 weeks). More than half of the group was exposed to antenatal steroids (56.3%), which are medications administered in pregnancy when expecting a preterm infant to accelerate the maturation of the fetuses' lungs (Jing et al., 2021). After birth, some individuals experienced medical complications. Specifically, 19.7% developed retinopathy of prematurity caused by abnormal development of the retinal blood vessels that can cause visual impairment (Tan et al., 2015), 22.3% had an infection in the bloodstream known as neonatal sepsis (Hornik et al., 2012), and a further 20.1% of infants were subject to respiratory problems associated with bronchopulmonary dysplasia requiring the need for supplemental oxygen support to at least 36 weeks GA (Gough et al., 2014).

Data describing the social background characteristics of the mothers of the VLBW and FT participants are also presented in Table 2. This is with the exception of missing data for three of the mothers of the VLBW group (n=226). Maternal age at birth was lower for mothers of the VLBW group compared to the FT group (p=.03). The socio-economic circumstances and educational attainment of the participants' families at birth were similar across the two groups (*ns*).

There were some between-group differences in terms of participant ethnicity (p=.03). The majority of participants across the two groups identified as NZ European (68.6% VLBW vs. 76.0% FT), however a greater proportion of VLBW participants identified as Māori (29.3% v. 16.0%). A small proportion of participants in each group identified as Pacific peoples, with slightly higher rates in the FT group (2.2% VLBW v. 8.0% FT).

Mothers of VLBW participants reported slightly higher rates of smoking during pregnancy than mothers of FT born participants, but this difference did not reach significance (42% v. 32%, p=.09). Significantly lower rates of breastfeeding were reported in the VLBW group, with 77% of mothers of VLBW adults reporting having breastfed their infant, in comparison to 88% in the FT group (p=.02).

Table 2 *Infant Neonatal and Social Background Characteristics of Participants Born Very Low Birthweight and Full Term*

Measure	VLBW (N=229)	FT (N=100)	x^2/t	*p*
Infant Clinical Characteristics				
M (SD) Birth weight, g	1132.9 (236.7)	3377.6 (584.3)	48.07	<.001
M (SD) Gestational age at birth, weeks	29.2 (2.5)			
% Male gender	44.5	37.0	1.62	.20
% Antenatal Steroids	56.3			
% Retinopathy of Prematurity	19.7			
% Neonatal Sepsis	22.3			
% Bronchopulmonary Dysplasia	20.1			
Social background				
M (SD) Maternal age at birth, years	25.8 (5.1)	27.2 (4.5)	2.15	.03
% Family of Professional/Managerial SES	33.7	34.0	12.95	.89
% Parent with Tertiary Education	53.1	63.0	15.15	.10
Ethnicity (self-report at 28 years)				
% Māori	29.3	16.0		
% Pacific Islander	2.2	8.0		
% NZ European/Other	68.6	76.0	11.35	.03
Maternal Smoking %	42.5	32.0	3.39	.09
Any breastfeeding %	77.2	88.0	5.09	.02

6.2 Neurodevelopment of the VLBW cohort

Table 3 describes the overall neurodevelopmental functioning of the VLBW and FT adults. This is reported because VPT/VLBW birth has been consistently identified as a risk factor for developmental impairments. Assessment measures included 1) their scores on the Wechsler Abbreviated Scale of Intelligence (WASI-II) administered at age 28 years, 2) rates of visual impairment assessed at 28 years and 3) a diagnosis of cerebral palsy (CP) by 7-8 years. Cerebral palsy was defined according to graded criteria used by Doyle (1995) as follows: i) severe: child is unlikely ever to walk and ii) moderate: child is able to walk but there is considerable limitation of movement (Darlow et al., 1997). Data for the WASI-II Full Scale IQ were missing for four VLBW participants.

As shown in Table 3, adults born VLBW had significantly lower full scale IQ scores than the FT group, indicating a lower average global cognitive ability (p<.001; 3 SDs below the control group mean). Reflecting their poorer overall IQ scores, VLBW participants had significantly higher rates of both mild (28.9% v. 14.0%) and severe (18.7% v. 0%) impairment on full scale IQ than FT participants. Additionally, a higher proportion of the VLBW group reported moderate to severe visual impairment in comparison to the FT group (22.7% v. 14.0%) and 2% of VLBW adults had cerebral palsy at ages 26–30 years.

Table 3 *Concurrent Neurodevelopmental Characteristics of Adults Born Very Low Birthweight and Full Term at 26-30 years*

Measure	VLBW (N=229)	FT (N=100)	*p*
M (SD) WASI-II Full Scale IQ Score	100.15 (14.39)	111.42 (11.31)	<.001
% Full Scale IQ Mild Impairment	28.9	14.0	.004
% Full Scale IQ Severe Impairment	18.7	0	-
% Moderate/Severe Visual Impairment	22.7	14.0	.11
% Cerebral Palsy	2	-	-

6.3 Executive Functioning Performance at age 28 years in Adults Born Very Low Birthweight and Full Term

As described in the Introduction and Method sections, three core executive functions identified based on previous research and theory were assessed. These included inhibitory control, working memory and cognitive flexibility. Inhibitory control was assessed using the widely used Hayling Sentence Completion Test. Working memory was assessed using the Sternberg Visuospatial Working Memory Task and the TEA Elevator Counting subtests. Finally, cognitive flexibility was assessed with the CTMT and Brixton Spatial Anticipation Test. The analysis process for EF task completion rates is explained first, followed by the results relevant to each EF domain. Within each domain, rates of task completion are reported first, followed by the examination of between-group differences in EF performance.

6.3.1 Rates of Task Completion and Analysis of Reasons for Non-completion

Before examining between group differences across the five tasks administered as part of the 28-year executive function battery, data were carefully checked and reviewed against testing notes. This was done to check for missing data or coding errors, as well as provide information on the proportion of individuals in each study group that were unable to complete each task, and most importantly to categorise reasons for task non-completion.

The first steps taken included reviewing the original testing notes regarding missing data for participants and cross-checking these with the data files, to confirm if the data was missing. Secondly, testing notes included the reasons for task non-completion, these were checked with notes included in the data files and collated together for each measure. Where uncertainties remained, these were discussed with the clinical psychologist who was responsible for administering and scoring all measures. Thirdly, all information was then collated to examine executive function task completion and the reasons for non-completion.

6.3.2 Inhibitory Control

6.3.2.1 Hayling Sentence Completion Test

The primary measure of Inhibitory Control was the Hayling Sentence Completion Test. Table 4 shows the completion rate and reasons for non-completion for adults born VLBW and FT on this measure of inhibitory control. This task had a non-completion rate of 3.5% (*n*=8) for the VLBW cohort and 0% for the FT group. The reasons for non-completion included participant time constraints (*n*=3), missing data (*n*=2), severe IQ impairment (*n*=1, score of <60 on the WASI-II) and unspecified (*n*=1). Additionally, a further participant had severe autism and was non-verbal, thus the Hayling test could not be administered due to the need for the participant to respond verbally to test prompts.

Table 4 *Task Completion Rate and Reasons for Non-Completion on the Hayling Sentence Completion Test*

Measure	VLBW (N=229)	FT (N=100)
% Completed task (*N*)	96.5 (221)	100.00 (100)
% Unable to complete task (*N*)	3.5 (8)	0
Reasons for non-completion (*N*)		
Missing data	2	
Time constraint	3	
Severe autism	1	
Severe IQ impairment	1	
Unspecified reasons	1	

Table 5 presents the inhibitory control scores of adults born VLBW and FT on the Hayling Sentence Completion Test, as well as the proportion of participants in each group whose performance met criteria for impairment. Performance on the task is represented by mean overall scaled scores assessing response initiation speed and response suppression. Results show there was a moderate between-group difference in scores, with the VLBW group performing more poorly than the FT group (p<.001, d=.56). That is, on average, adults in the VLBW group took longer to respond and made more errors compared to their FT counterparts. Further, VLBW participants as a group, were over seven times more likely than FT participants to meet criteria for impairment, with 18.6% compared to 3% scoring below the impairment cut-point (OR= 7.37, p<.001) on this inhibitory control measure.

Table 5 *Inhibitory Control Assessed using the Hayling Sentence Completion Test in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (N=221)	FT (N=100)	t/x^2	*p*	d/OR (95% CI)
M (SD) Hayling Overall Score	6.33 (1.64)	7.18 (1.20)	4.65	<.001	0.56
% Any Impairment	18.6	3.0	14.08	<.001	7.37 (2.22 – 24.40)

6.3.3 Working Memory

Two measures were completed to assess the visual and verbal working memory performance of adults born VLBW and FT. Visuospatial working memory was assessed using the Sternberg Visuospatial Working Memory task, with auditory-verbal working memory assessed with the TEA Elevator Counting subtests. Task completion rates and performance on each of these measures are reported below.

6.3.3.1 Sternberg Visuospatial Working Memory task

Table 6 presents the completion rate and reasons for non-completion for the Sternberg Visuospatial Working Memory task. This task had a non-completion rate of 7.4% (*n*=17) for the VLBW group and 0% for the FT group. As this task required participants to respond to visual stimuli, those who were blind (*n*=3) or had inadequate vision (*n*=1) were unable to do this task. Participants who failed the practice trials or did not complete the whole task were also excluded from the analysis (*n*=2). Other reported reasons for non-completion included fatigue and/or time constraints (*n*=10).

Table 6 *Task Completion Rate and Reasons for Non-Completion on the Sternberg Visuospatial Working Memory Task*

Measure	VLBW (N=229)	FT (N=100)
% Completed task (*N*)	92.6 (212)	100.00 (100)
% Unable to complete task (*N*)	7.4 (17)	0
Reasons for non-completion (*N*)		
Incomplete test	1	
Failed practice trial	1	
Inadequate vision	1	
Blind	3	
Severe autism	1	
Fatigue, time constraints or unspecified reasons	10	

Table 7 describes the performance of adults born VLBW and FT on the Sternberg based Visuospatial Working Memory task. As previously explained in the method section, for this analysis performance efficiency on this task is used as the primary outcome measure and results are reported for each of the four trial sets that varied in difficulty, with set 1 being the easiest and set 4 the most cognitively challenging.

As shown in Table 7, there were statistically significant between-group differences across the four levels, with small effect sizes (*ps*<.01, *ds* 0.31–0.37). Across all set levels, adults born VLBW responded less efficiently i.e they took longer to respond and made more incorrect responses than FT born adults. The greatest difference between the two groups was in the most difficult set, with the VLBW group performing 204.58 *ms* slower than the FT group. Given that set 4 has the highest working memory load, this finding suggests that as cognitive load increases, the performance decline of the VLBW group meant that they took longer and were more likely to make errors as they became more impaired or less efficient in their responses. VLBW participant's risk for having visuospatial working memory impairment, based on Set 4 performance, was twice that of FT participants, with 28.6% compared to 16% scoring above this cut-off (OR= 2.10, *p*=0.15).

Table 7 *Efficiency Assessed using the Sternberg Working Memory Task in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (*N*=212)	FT (*N*=100)	t/x^2	*p*	d/OR (95% CI)
M (SD) Efficiency (Reaction Time to Accuracy Trade- Off, ms)					
Set 1	1053.08 (259.08)	969.38 (248.08)	2.77	.006	0.32
Set 2	1445.01 (355.05)	1316.95 (342.54)	3.08	.002	0.36
Set 3	1553.06 (357.03)	1444.22 (336.41)	2.66	.008	0.31
Set 4	1870.68 (616.80)	1666.10 (507.46)	2.96	.002	0.37
% Any Impairment (Set 4)	28.6	16.0	5.86	0.15	2.10 (1.14 – 3.89)

6.3.3.2 Elevator Counting with Distraction

Table 8 shows the completion rate and reasons for non-completion for each group on the Elevator Counting with Distraction task, a measure of auditory-verbal working memory from the Test of Everyday Attention (TEA). In terms of completion rates 10.4% (*n*=23) of the VLBW group were unable to complete this task, compared to 2% (*n*=2) in the FT group. Reasons for non-completion in the VLBW group included participants reporting they could not manage the task or were unable to complete it (*n*=7). Deafness was another reason for non-completion (*n*=1), alongside severe autism (*n*=1) and participants being unable to distinguish the high tones used in this task (*n*=2). Additional reasons included factors such as time constraints (*n*=3) or feeling unwell (*n*=1). It is noted that several participants also had unspecified reasons for non-completion (*n*=8). Only 2 participants in the FT group did not complete the task, one due to being unable to distinguish the high tones (*n*=1) and another who failed to complete the test (*n*=1).

Table 8 *Task Completion Rate and Reasons for Non-Completion on the Elevator Counting with Distraction Task*

Measure	VLBW (N=229)	FT (N=100)
% Completed task (*N*)	89.6 (206)	98.0 (98)
% Unable to complete task (*N*)	10.4 (23)	2 (2)
Reasons for non-completion (*N*)		
Deaf	1	-
Unwell	1	-
Severe autism	1	-
Time constraints	3	-
Could not complete/manage	7	1
Could not distinguish high tones	2	1
Unspecified reasons	8	-

Table 9 describes the auditory-verbal working memory performance of adults born VLBW and FT on the Elevator Counting with Distraction task. Although the VLBW participants scored poorer than the FT group on this task, indicating that adults in the VLBW group made more incorrect responses compared to the FT group, this was a small magnitude of difference that did not quite reach statistical significance (p=.06, d=.34). Although examination of rates of auditory-verbal working memory impairment did suggest that VLBW participants were more than twice as likely as FT participants to experience difficulties in this domain, with 27.7% compared to 13.3% scoring below the impairment cut-point for this task (OR= 2.5, *p*=.005).

Table 9 *Auditory-Verbal Working Memory Assessed using the Elevator Counting with Distraction Task in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (N = 206)	FT (N = 98)	t/x^2	p	d/OR (95% CI)
M (SD) Elevator Counting with Distraction	9.77 (2.94)	10.72 (2.58)	2.74	.06	0.34
% Any Impairment	27.7	13.3	7.77	.005	2.5 (1.29 - 4.83)

6.3.3.3 Elevator Counting with Reversal

Table 10 presents the completion rates and reasons for non-completion on the Elevator Counting with Reversal subtest, an additional subtest of the TEA measuring auditory working memory. This task had the highest rate of non-completion across the EF measures used in the assessment, with 19.6% (*n*=45) of the VLBW group and 4% (*n*=4) in the FT group unable to complete it. It was reported in the testing notes that several participants in the VLBW group could not manage or complete this test as the below table shows (*n*=19). A wide array of reasons were given for non-completion including deafness (*n*=1), motor impairment (*n*=2), severe autism (*n*=1), fatigue (*n*=1), anxiety (*n*=1), illness (*n*=1), time constraints (*n*=5) or being unable to distinguish between the tones (*n*=3). For 11 participants this was not specified however the reasons may have been attributed to neurodevelopmental characteristics such as severe cognitive impairment or general reasons such as time constraints. Within the FT group, time constraints (*n*=2), being unable to distinguish the tones (*n*=1) and an incomplete test (*n*=1) were reasons for non-completion on this task.

Table 10 *Completion Rate and Reasons for Non-Completion on the Elevator Counting with Reversal Task*

Measure	VLBW (N=229)	FT (N=100)
% Completed task (*N*)	80.4 (184)	96.0 (96)
% Unable to complete task (*N*)	19.6 (45)	4 (4)
Reasons for non-completion (*N*)		
Motor impairment	2	-
Unwell	1	-
Fatigue	1	-
Anxiety	1	-
Severe autism	1	-
Deaf	1	-
Time constraints	5	2
Could not distinguish tones	3	1
Could not complete/manage	19	-
Incomplete	-	1
Unspecified reasons	11	-

Table 11 describes the performance of the VLBW and FT groups on the Elevator Counting with Reversal task. As shown, there was a moderate magnitude of effect between the two groups ($p<.001$, $d=.65$) with the VLBW group demonstrating greater difficulty in correctly identifying the rule for the tones and applying it while counting backwards and forwards, compared to the FT group. Additionally, VLBW participant's risk for having auditory-verbal working memory impairment, was over three times greater than FT participants, with 35.3% compared to 13.5% scoring below this cut-off (OR= 3.48, $p<.001$).

This task has the additional component of reverse counting compared to the Elevator Counting with Distraction test, suggesting that higher cognitive load leads to a more marked difference in performance between the VLBW and FT groups. Notably, given the significantly lower completion rate for this task (Table 10), it is possible that the results in Table 11 may, at least to some degree, underestimate the extent of impairment within this group.

Table 11 *Auditory Working Memory Assessed using the Elevator Counting with Reversal Task in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (N=184)	FT (N=96)	*t/x*^{2}	*p*	d/OR (95% CI)
M (SD) Elevator Counting with Reversal	9.17 (3.45)	11.21 (2.32)	5.21	<.001	0.65
% Any Impairment	35.3	13.5	14.89	<.001	3.48 (1.81 - 6.76)

6.3.4 Cognitive Flexibility

Two measures were completed to assess the cognitive flexibility performance of adults born VLBW and FT, including the Comprehensive Trail Making Test (CTMT) and the Brixton Spatial Anticipation Test. Task completion rates and performance on each of these measures are reported below.

6.3.4.1 Comprehensive Trail Making Test

Table 12 describes the completion rates and reasons for non-completion on the CTMT for the two study groups. This executive function task had the highest completion rate for the VLBW group, with only 7 participants (3.1%) being unable to complete the task. Since the CTMT requires participants to engage with visual stimuli and use fine motor skills (pencil to paper), blindness (*n*= 3) and motor impairment (*n*=2) were the main reasons for non-completion., in addition to severe autism (*n*=1) and severe IQ impairment (<60, *n*=2).

Table 12 *Completion Rate and Reasons for Non-Completion on the Comprehensive Trail Making Test*

Measure	VLBW (N=229)	FT (N=100)
% Completed task (*N*)	96.9 (222)	100.00 (100)
% Unable to complete task (*N*)	3.1 (7)	0
Reasons for non-completion (*N*)		
Severe autism	1	
Blind	3	
Severe IQ impairment	1	
Severe IQ & motor impairment	2	

Table 13 presents the cognitive flexibility scores of adults born VLBW and FT on the CTMT. This score assesses the time taken for the participant to complete the set of five trail tasks. There was a moderate between-group difference in CTMT scores, with VLBW participants taking significantly longer, on average, to correctly complete the trails than the FT group ($p<.001$, $d=.55$). Impairment on the CTMT was determined using a cut-off point of 189 or below, using a cut-off score equivalent to the lowest 10th percentile of the FT group score distribution. Based on this criterion, adults in the VLBW group were almost three times more likely than adults in the FT group to show impaired executive functioning on this cognitive flexibility task, with 38.3% compared to 18% scoring below this cut-off (OR=2.89, $p<.001$).

Table 13 *Cognitive Flexibility Assessed using the Comprehensive Trail Making Test in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (n = 222)	FT (n=100)	t/x^2	*p*	d/OR (95% CI)
M (SD) CTMT Score	201.43 (49.28)	226.76 (37.58)	4.57	<.001	0.55
% Any Impairment	38.3	18.0	13.04	<.001	2.89 (1.59 -5.04)

6.4.3.2 Brixton Spatial Anticipation Test

Table 14 presents the completion rate and reasons for non-completion on the Brixton Test in adults born VLBW and FT. In the VLBW group, 7.0% (*n*=16) were unable to complete the task compared to 0% within the FT group. The Brixton Spatial Anticipation test is based on identifying patterns using visual stimuli, therefore blindness was a reason for non-completion (*n*=3). Additionally, experiencing stress or anxiety (*n*=3) and feeling unwell (*n*=1) meant that the test was no longer administered to those participants. Other reasons for non-completion included severe autism (*n*=1), time constraints (*n*=2), and missing data (*n*=3) that could not be retrieved.

Table 14 *Task Completion Rate and Reasons for Non-Completion on the Brixton Spatial Anticipation Test*

Measure	VLBW (N=229)	FT (N=100)
% Completed task (*N*)	93.0 (213)	100.00 (100)
% Unable to complete task (*N*)	7.0 (16)	0
Reasons for non-completion (*N*)		
Severe autism	1	
Blind	3	
Missing data	3	
Time constraints	2	
Unwell	1	
Stress or anxiety	3	
Unspecified reasons	3	

Table 15 presents the results from the VLBW and FT born adults' performance on the Brixton Spatial Anticipation Test. The results demonstrate that the VLBW group performed significantly poorer than the FT group on this measure with a small to moderate effect size (p<.001, d=.41). Lower scores on the Brixton indicate higher impairment, therefore the results suggest that adults born VLBW had greater difficulty detecting rules and patterns in a sequence of stimuli, as they identified fewer patterns correctly compared to the FT group. Impairment on the Brixton was determined based on the lowest 10th percentile of the FT groups score distribution. Table 15 shows that 22.1% of VLBW adults compared to 8% of FT adults met criteria for impaired cognitive flexibility on the Brixton, with risk of impairment being over 3 times more likely for the VLBW group (OR= 3.26, p=.002).

Table 15 *Cognitive Flexibility Assessed using the Brixton Spatial Anticipation Test in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (N=213)	FT (N=100)	t/x^2	p	d/OR (95% CI)
M (SD) Brixton Overall Score	6.53 (2.05)	7.33 (1.65)	3.41	<.001	0.41
% Any Impairment	22.1	8.0	9.29	.002	3.26 (1.48-7.19)

6.4 Processing Speed at Age 28 years in Adults Born Very Low Birthweight and Full Term

Existing research suggested that observed differences in EF task performance between VLBW and FT born adults might be explained, in part or in full, by the effects of processing speed. This is because processing speed involves the integration of efficiency, speed and fluency which are particularly important for success on timed EF tasks. It was therefore important to examine the effects of processing speed on EF task performance to determine whether it is impacted to some extent by processing speed abilities of VPT/VLBW adults.

Table 16 describes the performance of adults born VLBW and FT on the Symbol Digit Modalities Test, a measure of processing speed or how quickly someone can complete a task. The findings below indicate that the VLBW group had significantly lower processing speed than the FT group (p<.001, d=.61). Adults born VLBW made fewer correct substitutions compared to the FT group, meaning that on average they took longer to respond and process information to complete the task. Reflecting this large between group difference, risk for processing speed impairment for adults born VLBW was six times greater than FT participants, with 40.1% compared to 10% scoring below this cut-off (OR=6.02, p<.001).

Considering the results of the previous EF tasks and those of processing speed performance, it is important to examine whether decline in EF performance persists after controlling for processing speed. This is to assess the extent to which processing speed difficulties may contribute to EF problems in VLBW survivors.

Table 16 *Processing Speed Assessed using the Symbol Digit Modalities Test in Adults Born Very Low Birthweight and Full Term at age 28 years*

Measure	VLBW (*N*=222)	FT (*N*=100)	t/x^2	*p*	d/OR (95% CI)
M (SD) SDMT Overall Score	50.52 (10.74)	56.71 (8.87)	5.04	<.001	0.61
% Any Impairment	40.1	10.0	29.32	<.001	6.02 (2.97 – 12.21)

6.4.1 Adjusting for the effects of processing speed on Executive Function performance

A one-way analysis of covariance (ANCOVA) was performed to assess the between-group differences in executive functioning performance for adults born VLBW and FT, following adjustment for processing speed (see Table 17). Following adjustment for processing speed, there was a statistically significant difference between VLBW and FT adults on Elevator Counting with Reversal $F(1, 275) = 12.41$, $p<.001$, $\eta_p^2 = .043$) and the Hayling test, $F(1, 314) = 9.52$, $p<.05$, $\eta_p^2 = .029$. The effect sizes suggest a small to moderate decline in performance continues to persist on these EF tasks for the VLBW group, after accounting for processing speed. After adjustment for processing speed, there were small between-group effects for the Sternberg task, $F(1, 306) = 3.79$, $p=.05$, $\eta_p^2 = .012$, CTMT, $F(1, 317) = 3.74$, $p=.05$, $\eta_p^2 = .012$, and the Brixton test, $F(1, 310) = 3.27$, $p=.07$, $\eta_p^2 = .010$ that did not quite reach statistical significance. As expected, there was no between-group differences on Elevator Counting with Distraction after adjusting for processing speed, $F(1, 296) = 1.33$, p =.25, η_p^2=.004.

Table 17 *Group means adjusted for Processing Speed on measures of Executive Function*

Measures	N[a]	VLBW	FT (*N*=100)	*F*	*p*	η_p^2
Inhibitory Control						
M (SE) Hayling	217	6.47 (.09)	7.01 (.15)	9.52	.002	.029
Working Memory						
M (SE) Sternberg *Set 4*	209	1858.84 (40.38)	1716.96 (59.16)	3.79	.05	.012
M (SE) Elevator Counting with Distraction	201	9.97 (.19)	10.37 (.28)[b]	1.33	.25	.004
M (SE) Elevator Counting with Reversal	182	9.42 (.21)	10.69 (.29)[c]	12.41	<.001	.043
Cognitive Flexibility						
M (SE) CTMT	220	206.58 (2.54)	215.58 (3.82)	3.74	.05	.012
M (SE) Brixton	213	6.65 (.12)	7.07 (.19)	3.27	.07	.010

Note. *SE*= Standard Error*;* η_p^2 = Partial Eta Squared
[a] = Number of VLBW participants who completed the SDMT and corresponding Executive Function measure; [b]= 98 FT participants; [c] = 96 FT participants.

6.5 Examining the effects of Birthweight on EF task performance

To examine the effects of birthweight on EF task performance, the VLBW group was divided into two groups: those born <1500g and those born <1000g to establish a subgroup of individuals born ELBW. A one-way analysis of variance (ANOVA) was performed to examine the effects of birthweight on each EF task. Table 18 presents the mean scores on each of these EF measure across the three groups: ELBW, VLBW and FT. A linear association was evident between birthweight and EF performance showing that adults born ELBW performed significantly lower across all EF tasks compared to those born VLBW and FT (p<.001), with medium to large effects. These findings indicate that the risk of pervasive impairment across domains of EF increases with decreasing birthweight. The largest difference between the three groups was on the Elevator Counting with Reversal task, $F(1, 258) = 37.39$, p<.001, η^2=.127. This task required multiple demands, suggesting that higher cognitive load leads to a more marked decline in performance for the ELBW group.

Table 18 *Executive functioning task performance of ELBW, VLBW and FT born adults at age 28 years*

Measure	ELBW[a]	VLBW[a]	FT (N=100)	*F*	*p*	η^2
Inhibitory Control						
M (SD) Hayling	6.02 (1.55)	6.47 (1.66)	7.04 (1.19)	16.34	<.001	.052
Working Memory						
M (SD) Sternberg *Set 4*	1962.95 (698.55)	1842.53 (587.12)	1634. 38 (371.69)	12.08	<.001	.041
M (SD) Elevator Counting with Distraction	8.97 (2.95)	10.10 (2.88)	10.72 (2.58)[b]	12.96	<.001	.046
M (SD) Elevator Counting with Reversal	7.83 (3.68)	9.68 (3.22)	11.20 (2.19)[c]	37.39	<.001	.127
Cognitive Flexibility						
M (SD) CTMT	188.48 (49.85)	206.71 (48.07)	224.28 (38.144)	21.41	<.001	.067
M (SD) Brixton	6.00 (1.69)	6.74 (2.14)	7.38 (1.64)	17.74	<.001	.058

Note. η^2 =Eta squared
[a] = The number of ELBW/VLBW participants varied across the EF measures, the average *n* across the tasks was 55 for the ELBW and 155 for the VLBW group; [b]= 98 FT participants; [c] = 96 FT participants.

6.6 Examining the relationships between measures of Executive Function

Results thus far indicate that adults born VLBW exhibit pervasive impairments across domains of EF spanning inhibitory control, visuospatial working memory, auditory-verbal working memory and cognitive flexibility. The largest difference in performance between adults born VLBW and FT, was on the Elevator Counting with Reversal subtest (*d*=.65). This was a difficult task that 19.6% of VLBW participants could not complete, suggesting that higher task demands lead to a more marked decline in performance for adults born VLBW. Results across the measures suggests that impairment is both persistent and pervasive across EF domains for VLBW adults, as opposed to difficulties being within one or two specific domains. This pattern of impairment across EF domains supports Miyake and colleagues' theoretical framework of EF as a construct that is both distinguishable within its respective domain, as well as a unified set of constructs with each EF operating in an integrative manner (Miyake & Friedman 2012). Therefore, as a final step in the analysis, we explored using factor analysis the extent to which these measures might load on a single common factor or might form a set of factors.

The key variables used in the analysis were as follows 1) Hayling Overall Scaled Score, 2) Set 4 Inverse Efficiency Score (Sternberg task), 3) Elevator Counting with Distraction Scaled Score, 4) Elevator Counting with Reversal Scaled Score, 5) CTMT T-Score and 6) Brixton Overall Scaled Score. Each of the variables, with the exception of the Sternberg, were standardised measures with the direction of scores in the same way, with a lower score indicating poorer performance. The scores for the Sternberg measure have been reverse scored for ease of interpretation. To examine the relationships between measures of EF using factor analysis, the following steps were taken: a) examining the preliminary correlations between measures, b) performing a principal components analysis of associations between tasks, and c) performing a confirmatory factor analysis to confirm the findings. These steps are described in detail below.

6.6.1 Correlations between measures of Executive Function

The first step involved computing Pearson correlation coefficients between the VLBW and FT groups performance on each of the EF measures. The correlations are shown for the total sample (*n*=329). As shown in Table 19, the EF measures were all significantly correlated (*ps*<.001), however correlations were weak to moderate ranging from *r*=.163–.541. The strongest correlation across the EF measures was observed as expected between the two TEA Elevator Counting tasks measuring auditory-verbal working memory (*r*=.54). Additionally, there was a tendency for the Sternberg visuospatial working memory task to be more weakly correlated with the other measures (*r*=.3 or below).

Table 19 *Correlations between measures of Executive Function*

	Inhibitory Control – Hayling Overall Scaled Score	Visuospatial WM - Efficiency score Set 4	Auditory-verbal WM - TEA Elevator Counting with Distraction Scaled Score	Auditory-verbal WM - TEA Elevator Counting with Reversal Scaled Score	Cognitive Flexibility – CTMT T-Score	Cognitive Flexibility – Brixton Overall Scaled Score
Inhibitory Control – Hayling Overall Scaled Score	-					
Visuospatial WM - Efficiency score Set 4	.196**	-				
Auditory-verbal WM - TEA Elevator Counting with Distraction Scaled Score	.226**	.163**	-			
Auditory-verbal WM - TEA Elevator Counting with Reversal Scaled Score	.258**	.258**	.541**	-		
Cognitive Flexibility – CTMT T-Score	.333**	.266**	.361**	.440**	-	
Cognitive Flexibility – Brixton Overall Scaled Score	.286**	.229**	.305**	.399**	.394**	-

** *p=<.001*

6.6.2 Principal Components Analysis of associations between tasks

A principal components factor analysis (PCA) was then used to assess the factor structure underlying the six observed measures of EF. However, prior to performing the PCA, examining the suitability of data for factor analysis was assessed. Examination of the previous correlation matrix revealed several coefficients of .3 and above. The Kaiser-Meyer-Olkin value was .79, exceeding the recommended value of .6 and Bartlett's Test of Sphericity reached statistical significance (p<.001), thus supporting factorability of the correlation matrix. The PCA was run for 1) the VLBW and FT groups combined, 2) the VLBW group only, and 3) the FT group only. The analysis was performed for each group to include all EF variables, for the Sternberg task it was necessary to select on key task measure, so the Set 4 inverse efficiency score was selected. The results explained below and presented in Table 20 are reflective of the whole sample. Results from the individual group analyses are not shown because the VLBW group showed largely the same results as the whole sample. While there were differences for the FT group, limitations within this cohort and low factor loadings for certain measures, indicated that a single factor structure was more viable.

The results of the PCA revealed the presence of one component with eigenvalues exceeding 1, explaining 43.2% of the variance. Inspection of the scree plot revealed a clear break after the first component therefore it was decided using Cattell's (1966) scree test, to retain one component for further investigation. The single-component solution explained a total of 43.2% of the variance, with all variables loading substantially on one factor including inhibitory control, visuospatial working memory, auditory/verbal working memory, and cognitive flexibility. Due to the single-factor nature, results for structure and pattern matrices could not be extracted. Collectively, results from both the bivariate correlational analyses and especially the PCA, indicated that tasks may be viewed as measuring a common latent EF construct.

6.6.3 Confirmatory Factor Analysis

To confirm the viability of a single factor model of executive function based on previous theoretical constructs, confirmatory factor analyses were performed across the whole sample. A path analysis diagram was created using the statistical software AMOS Graphics to examine the model fit of the EF measures. EF was used as the latent variable, and the EF tasks as the observed variables. The model shown in Figure 3 was created together with the standardised factor loadings. The regression weights show the correlations between the latent and observed variables ranging from .39 to .82. The remaining factor loadings reflect how much of the observed variable is explained by the latent variable ranging from .15 to .67. Errors of variance were needed for the observed variables because they explain the proportion of variance in each measurement that does not covary with the latent factor (Hu & Bentler, 1999).

The results for the analysis of this single factor model are shown in Table 20. Only a single factor model was achieved due to latent variables requiring the inclusion of ideally three measures to one latent variable for CFA. Therefore, it was not appropriate to create a latent variable for measures that had lower loadings, for example the Sternberg visuospatial working memory task. Similar to results from the PCA, the confirmatory analysis suggests this model of EF is acceptable. In Table 20 the chi-square is non-significant which implies a good model fit as it does not significantly differ from the 'true model' (Huu & Bentler, 1999). A good model fit is further demonstrated by the CMIN value being closer to 1, with CFI and RMSEA indices of >0.95 and <0.06 respectively. Additionally, the PCLOSE value is non-significant (0.58) and the TLI is >0.9 to further confirm that the single factor model of EF has the best goodness of fit.

These analyses provide support for the conceptualisation of these EF tasks as capturing inter-related domains of EF in adults born VLBW. Specifically, these tasks involve the maintenance of information in working memory, and the planning of responses based on visuospatial or auditory-verbal output as well as the ability to flexibly alter and inhibit responses according to visual or verbal stimuli.

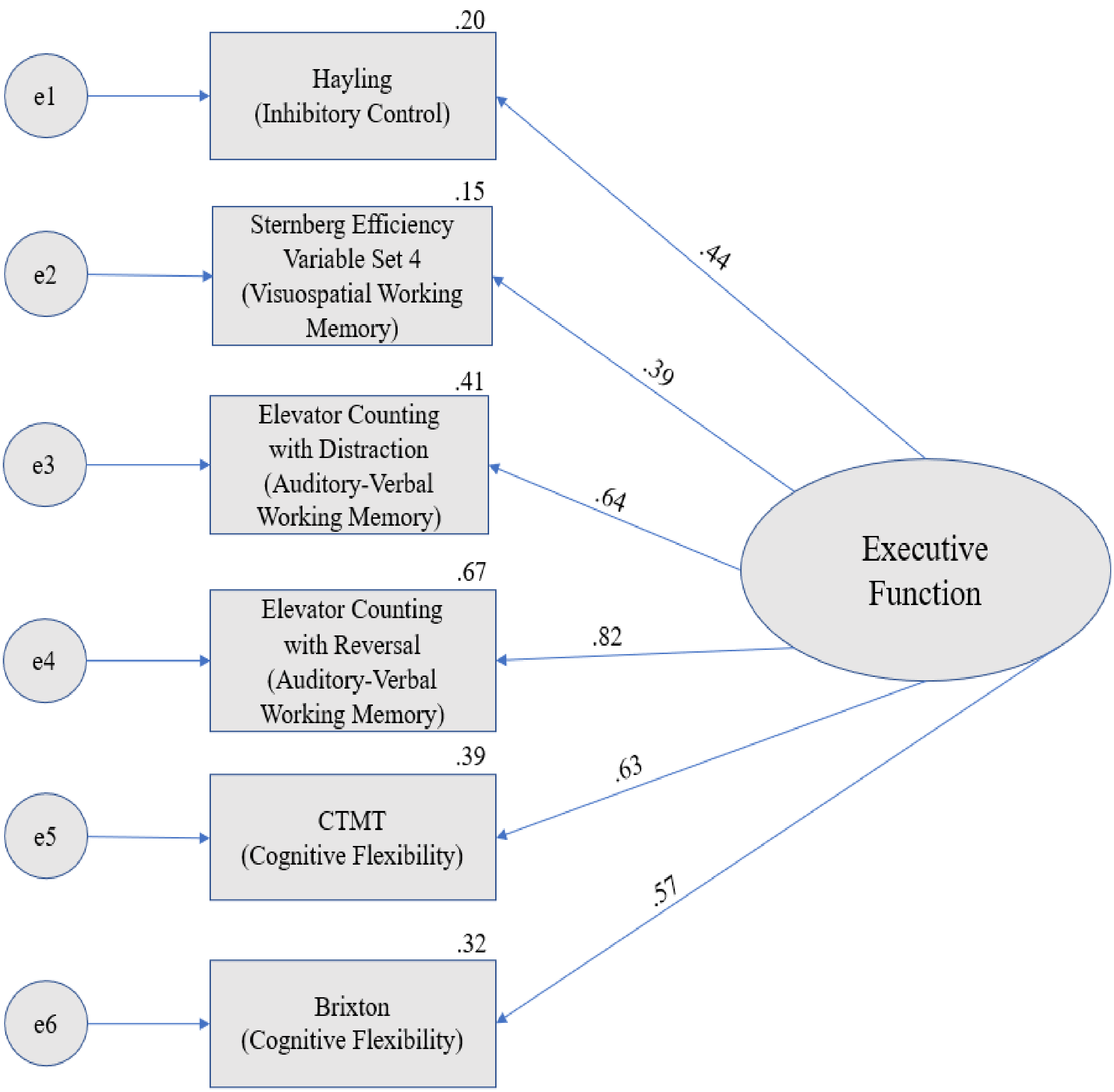

Figure 3. Single Model Factor of Executive Function Tested in Confirmatory Factor Analysis

Table 20 *Indexes Model of Fit for Confirmatory Factor Analysis of a Single Factor Executive Function Model for Adults Born Very Low Birthweight and Full Term at age 28 years*

	Total Sample (N=329)
Single Factor Executive Function Model	
Chi-square (df)	14.48 (9)
CMIN/DF	1.61
CFI	0.98
RMSEA	0.04
PCLOSE	0.58
TLI	0.96

Note. Chi Square (df)=x^2/degrees of freedom, non-significance indicates better fit; CMIN=Chi-square divided by degrees of freedom, closer to 1 indicates a better fit; CFI= Comparative Fit Index, value >0.95 indicates a satisfactory fit; RMSEA= Root mean square of approximation, value <0.06 indicates a satisfactory fit; PCLOSE = *p*-value of the null hypothesis that the estimate is below 0.05; TLI = Tucker Lewis Index, >0.9 indicates a satisfactory fit. (Hu & Bentler, 1999).

Chapter 7

Discussion

The overall goal of the present study was to assess executive functioning abilities of a national cohort of adults born VLBW across a range of important domains of EF. Research concerning individuals born VPT/VLBW have found this population to be at higher risk of EF impairment. Currently, research has almost exclusively studied EF outcomes in children and adolescence born VPT/VLBW, however relatively little is known about EF outcomes in adulthood. This study therefore addresses an important research gap in examining EF performance across a range of measures among a national cohort of adults born VLBW at 28 years. This was done by 1) comparing the performance of adults born VLBW and FT on a series of tasks across three key domains of EF, 2) examining if processing speed might account to some degree, for between-group differences in EF performance, 3) exploring the effects of birthweight to determine whether those with lower birthweight were at higher risk of EF impairment, 4) investigating the relationship between the five EF measures to determine the best model fit for this cohort.

There were several methodological strengths of the study. These include its prospective longitudinal design, high sample retention and the use of a comprehensive battery, including standardized measures and a novel computerized task purposefully designed to the well-established Sternberg paradigm. An added strength of the study included the size of the samples, with results obtained for 229 adults born VLBW and 100 comparison adults. Additionally, it is a national study, capturing a range of socioeconomic backgrounds and circumstances to be representative of the VLBW population in New Zealand. The key findings relating to each of the study aims are discussed below.

7.1 Executive Functioning Performance of VLBW adults at age 28 years

The first stage of the analysis examined the between-groups differences in EF performance of adults born VLBW and those born to FT. The study results are discussed below by EF domain.

Inhibitory Control. The Hayling Sentence Completion test was employed as a measure of inhibitory control, assessing response initiation speed and response suppression. The benefit of this test is that it is entirely verbal, making it suitable for people with impairments in visual and motor function, which are often associated with the neurodevelopment of VPT/VLBW groups. The high completion rate on this task (96.5%, *n*=221) meant that little data loss likely captured the level of impairment among VLBW adults, as individuals with visual and motor impairments were able to complete the task. Study results showed that adults born VLBW took longer to respond and made more errors compared to their FT counterparts. This indicates that on the Hayling test, adults born VLBW took more time generating words to complete sentences that either fit the context or were unconnected in any way. The higher number of errors meant that on average VLBW adults had greater difficulty inhibiting an expected response. In addition to these findings, VLBW adults in the present study were over 7 times more likely to meet criteria for any impairment on the Hayling test compared to FT adults (OR=7.37). The current findings are consistent with previous studies who also used the Hayling Test as a measure of inhibitory control (Nosarti et al., 2007; Allin et al., 2011; Kroll et al., 2017). The studies found VPT adults to perform significantly more poorly on the Hayling test, with mean differences ranging from .07–.74. The present study obtained a slightly higher mean difference of .85, which may be accounted for by the statistical power of the sample size being greater than the other studies.

The present findings also align with studies who used the Stroop Test to examine inhibitory control (Sølsnes et al., 2014; Eryigit Madzwamuse et al. 2015). Both studies found that adults born VPT/VLBW (range 19–29 years) performed significantly more poorly in inhibitory control compared to a FT comparison group. Sølsnes et al. (2014) reported a moderate effect size of d=.60, similar with the findings of the present study (d=.56). However, studies that used the Connors' Continuous Performance Test (CPT-II; Pyhälä et al., 2011) and the Continuous Performance Test -Errors of Commission (CCPT-EC; Kroll et al., 2017) found that between–group differences did not reach statistical nor clinical significance. This may be attributed to the function of those tasks, for example the task requirements are computerized and differ in comparison to the Hayling or Stroop tests that rely on more verbal responses. As the present findings are consistent with majority of the existing studies, it suggests that VPT/VLBW adults have greater difficulty on tasks where they must inhibit automatic verbal responses.

Working Memory. The computerized Sternberg-based task was used as a measure of visuospatial working memory performance that assessed performance efficiency. The task consisted of 4 sets of 10 trials varying in difficulty to reflect the number of objects the participant was required to memorise and recall the spatial location of. Results showed that adults born VLBW took longer to respond and made more incorrect responses as to the location of the target stimulus, than FT born adults. Notably, in a fixed time frame, the differences between the two groups widened with increasing task complexity. Specifically, between group differences were smaller on the first trial, with the VLBW adults being 83.7 *ms* slower on average than the FT group, whereas on the fourth trial, VLBW adults performed 204.58 *ms* slower. These findings indicate that as cognitive load increased, the VLBW group became on average increasingly less efficient in their responses. Notably, VPT/VLBW individuals experience greater attentional difficulties compared to individuals born to FT, suggesting that when EF demands are high, they may struggle to attend to these tasks or tire more easily (Woodward et al., 2021). It is possible that adults born VLBW were gradually becoming more inattentive throughout the trials and the less complex trials may have helped to attenuate the differences in performance. Nonetheless, VLBW adults were over twice as likely to have impairment on this visuospatial working memory task compared to FT adults (OR=2.10).

The present findings are consistent with the Suikkanen et al. (2021) study who employed the Groton Maze Learning Test as a measure of visuospatial working memory. This test involves navigating a hidden maze, requiring the participant to encode and manipulate spatial information as they are being timed (Pietrzak et al., 2008). The test uses efficiency index measures, that are sensitive to the participant's ability to store information in short term working memory and how effectively they use information about the maze to work through the following trials (Pietrzak et al., 2008). Suikkanen et al. (2021) found this particular task within their EF battery, to have the greatest difference between groups, in that VPT adults performed slower with on average 0.6 fewer moves per 10 seconds ($p<.001$). It relates to the current findings on the Sternberg task, in that they are both highly similar measures of visuospatial working memory and confirms that VPT/VLBW adults present with weaker spatial memory efficiency. Additionally, Østgård et al (2016) used the Weschler Memory Scale (WMS-III) and found differences within the mental control and spatial span subtests with large effect sizes of .85–.89. These are larger effect sizes compared to the current study (d=.31–.37) however this may be attributed to the methodology and function of the test used. Collectively, these results relate to the current study findings in suggesting that VPT/VLBW adults appear to demonstrate persistent impairment in visuospatial working memory. Both visual and verbal components of working memory were examined in the current study and the verbal tasks are examined below.

The TEA subtests, Elevator Counting with Distraction and Elevator Counting with Reversal, were used as measures of auditory-verbal working memory. These measures are sensitive to the assessment of selective attention and the manipulation of auditory-verbal information in working memory. Adults born VLBW did score lower overall than FT adults on the Elevator Counting with Distraction task. Although this difference did not reach statistical significance, VLBW adults were over twice as likely (OR=2.5) to meet criteria for impairment on this task. In contrast, the Elevator Counting with Reversal task was particularly difficult for the VLBW group which is confirmed by the significantly lower completion rate on this task (80.4%) compared to the other EF tasks. Several participants at the time stated that they 'could not manage' the task and it is likely that the extent of impairment within the VLBW group is underestimated in the results. Adults born VLBW obtained significantly lower scores on the Elevator Counting with Reversal task than FT adults, thus suggesting they had greater difficulty in correctly identifying the rule for the tones and applying it while counting backwards and forwards, compared to the FT group. This task demanded higher cognitive load because it involved the additional component of reverse counting which may account for its greater sensitivity to working memory difficulties. Additionally, adults born VLBW were over three times more likely to have impairment on this task (OR=3.48), indicating that VLBW adults are at higher risk of impairment in auditory-verbal working memory.

The present findings are consistent with the study conducted by Allin et al. (2011) who used the California Verbal Learning Test (CVLT) which measures components of verbal working memory and found that adults born VPT performed significantly more poorly than FT born adults by a mean difference of 4.2 (p<.05). This means that VPT adults had greater difficulty listening to a series of words and recalling the terms/categories that they belong too. The CVLT and Elevator Counting tasks combine the components of selective attention and working memory and therefore both involve high surveillance and high cognitive load. This may further indicate why the Elevator Counting with Reversal task was particularly difficult for some adults in the VLBW group to manage. The results from the present study and Allin et al. (2011) suggest that when tasks require a high level of surveillance, attentional difficulties among VPT/VLBW adults may contribute to poor performance on working memory tasks as demands (cognitive load) increase. There are very limited studies that have specifically examined auditory/verbal working memory in VPT/VLBW adults, as more have focused on the visuospatial components of working memory.

Collectively, it is clear that adults born VLBW show greater levels of impairment in visuospatial and auditory-verbal working memory. This reinforces present concerns that working memory is an on-going area of difficulty for VLBW adults, particularly when cognitive load and attentional demands are high.

Cognitive Flexibility. The CTMT was used as a measure of cognitive flexibility. It assesses the extent to which the participant is able to flexibly switch and adapt between strategies when problem solving. The present findings demonstrated that adults born VLBW took significantly longer to correctly complete the trail tasks in comparison to FT born adults. Additionally, VLBW adults were almost 3 times more likely to show impaired performance on this task (OR=2.89). This is consistent with previous studies who used the original version of the CTMT test, known as the Trail Making Test (TMT). This test is made up of two parts in which Part A consists of visual scanning tasks and Part B involving the alternate switching between different task sets of numbers and letters. Kroll et al. (2017) and Sølsnes et al. (2014) found that VPT adults performed significantly more poorly on the TMT compared to FT born adults, reporting moderate to large effect sizes of .42 and .74–.88 respectively. The latter reported by Sølsnes et al. (2014) is somewhat larger than Kroll et al. (2017) and the current findings (*d*=.55). Although, Kroll et al. (2017) measured only Part B of the TMT and there are methodological differences between the TMT and CTMT, which may explain the variation in the magnitude of effect sizes. In contrast, Nosarti et al. (2007) found that while VPT adults were significantly slower than controls to complete Part A of the trails ($p<.01$), there were no statistically significant group differences on Part B or the Trails A-B score. For the latter this means that no significant differences were found between speed/trail completion time across Part A and Part B. The present findings in relation to the existing studies, generally indicate that the ability to flexibly switch and problem solve between different tasks, is an area of difficulty for VPT/VLBW adults.

The Brixton Spatial Anticipation Test was also used as a measure of cognitive flexibility, assessing the ability to flexibly detect patterns in a series of different sequences. The findings showed that adults born VLBW had greater difficulty in detecting rules and patterns in a sequence of stimuli, as they identified fewer patterns correctly compared to the FT group (*d*=.41). Additionally, VLBW adults were over three times more likely to meet criteria for impairment on the Brixton test (OR=3.26). Kroll et al. (2017) employed two measures that tap into similar abilities including rule acquisition and spatial planning: the Intra-Extra Dimensional Set Shift (IED) and the Stockings of Cambridge (SOC). The IED measures mental shifting and flexibility of attention, in which participants must categorise stimuli and respond to changes while maintaining and shifting attention when required. The SOC is a task that assesses rule acquisition and spatial planning, in which participants are required to shift coloured circles between locations while planning to do it in as few moves as possible. On the IED measure, they found that VPT born adults performed significantly more poorly than FT adults, with a moderate effect size similar to the present finding (*d*=.52).

However, Kroll et al. (2017) found no statistically significant group differences on the SOC measure, with a smaller effect size of *d*=.29. To an extent this contradicts the current finding on the Brixton measure, as both tests involve rule acquisition and spatial planning, however they are fundamentally different measures, suggesting this result may be attributed to methodological differences (i.e the function of the task chosen). Collectively, although the use of different measures needs to be considered when interpreting the findings, results from the CTMT and Brixton tests suggest that VLBW adults are at greater risk of impairment in cognitive flexibility.

7.2 Executive Functioning Performance following covariate adjustment for Processing Speed

The present findings demonstrated that adults born VLBW had significantly poorer processing speed performance in comparison to the FT group (*d*=.61). This raises the possibility that some EF difficulties may potentially reflect processing speed abilities as earlier findings had suggested (Anderson, 2002; Pyhälä et al., 2011; Østgård et al., 2016). To examine this issue, between-group differences in EF performance were examined following covariate adjustment for the measure of processing speed. Significant differences remained on the Hayling ($p<.05$) and Elevator Counting with Reversal ($p<.001$) tests measuring inhibitory control and auditory-verbal working memory. These findings suggest that decline in performance continues to persist on these EF tasks for the VLBW group, after accounting for processing speed. Results almost reached statistical significance across the Sternberg ($p=.05$), CTMT ($p=.05$) and Brixton measures ($p=.07$). These findings suggest that between-group differences in EF task performance are attenuated when processing speed is taken into account statistically, but nonetheless remain significant or close to significance. Therefore, although processing speed may partly explain between-group differences, EF difficulties appear to continue to persist even after adjusting for effects of processing speed. Previous studies found no significant differences between adults born VLBW and FT on tasks that had no time limits including the Tower Tests measuring planning and problem solving, and the Wisconsin Card Sorting Test (WCST), purported to measure a range of EF abilities including cognitive flexibility, attention and working memory (Sølsnes et al., 2014; Østgård et al., 2016). The studies suggested that because processing speed abilities are required for success on timed tasks, it may influence EF task performance particularly among VLBW adults with poorer processing speed abilities. Although, the significance levels ranged from p=.06–.71 on the Tower Tests and p=.06–.96 on the WCST, indicating that some results were marginal and cannot completely be explained in full by effects of processing speed. In conjunction with the current study findings, although processing speed may in part explain between-group differences, EF difficulties appear to persist across EF domains.

7.3 Effects of Birthweight on Executive Functioning Performance

The present findings showed that adults born ELBW performed significantly more poorly than adults born VLBW and FT across all EF measures (p<.001). A linear association was therefore evident between birthweight and EF task performance, to indicate that decreasing birthweight increased the risk of impairment on EF tasks. The results suggest that there is pervasive impairment across the domains of inhibitory control, working memory and cognitive flexibility for individuals born ELBW as they demonstrated a more marked decline in performance compared to adults born VLBW or FT born adults. These findings are consistent with previous studies in childhood and adolescence who suggested that individuals born ELBW are most at risk of executive difficulties (Anderson & Doyle, 2004; Burnett et al., 2015). Overall, the findings indicate that risk of pervasive impairment across domains of EF increases with decreasing birthweight.

7.4 Relationships between measures of Executive Function

Although not a specific study aim, given the generally pervasive pattern of impaired performance across EF tasks, the relationship between EF measures were examined. Initial analyses indicated correlations between tasks and further PCA results tended to suggest these measures might load on a single latent construct. While this latent construct did not explain the majority of variance in task performance (43.2%), CFA was performed to confirm the viability of a single factor model. This analysis tended to suggest that the best model fit involved just one component including the Hayling test, Sternberg visuospatial working memory task, Elevator Counting subtests, CTMT and the Brixton test. This is in line with the PCA conducted by Kroll et al. (2017), who also found their EF measures of inhibitory control, spatial planning, cognitive flexibility, verbal fluency and attention shared considerable variance (43%) on a single component. These analyses generally support the conceptualisation of these tasks as a set of inter-related EF skills in line with the unity aspect of Miyake and colleagues' theoretical framework. It theorizes that the three EF domains, inhibitory control, working memory and cognitive flexibility are unified in the way that they correlate with one another to represent a common underlying ability. Effective EF task completion requires components from each of the three domains, thus utilizing multiple executive skills.

7.5 Theoretical Implications

The findings from the present study are in line with developmental theory in relation to how executive processes emerge and develop across the life span. Executive functioning has a protracted developmental course from infancy to early adulthood. Over time the gradual development of the prefrontal cortex of the brain and the establishment of myelination of neural connections results in the enhanced integration of cognitive processes that improves information processing and executive control (Anderson, 2002; Wiebe & Karbach, 2018).

This means that as children and adolescents emerge into adulthood, they are increasingly able to coordinate several pieces of information at a more complex level, as they are able to direct certain EF processes to how they need to be utilized (Wiebe & Karbach, 2018). For example, this could be freeing up working memory capacity to devote to attentional processes that are required to perform the task. If there has not been this rapid development of prefrontal connections, then differences in neural maturation may be likely to continue to constrain EF resources for adults born VLBW, resulting in less efficient EF task performance. This is particularly relevant to the current findings which suggest that EF difficulties associated with preterm birth persisted across EF domains and across development into adulthood.

7.6 Applied Implications

The present findings emphasize that there are persistent and pervasive problems in executive functioning for adults born VLBW and this identifies a number of clinical and research implications. Firstly, in terms of clinical implications, the findings highlight the importance of monitoring and supporting the development of EF at younger ages, particularly in childhood. EF skills in childhood are potentially more malleable and responsive to intervention strategies, therefore intervening earlier may help in mitigating the long-term effects of EF impairment in adulthood. Stålnacke et al. (2019) highlighted the importance of identification of executive deficits before school entry as EF deficits are likely to remain stable over time. While early monitoring is important, some results in the present study were fairly modest in terms of clinical significance. For example, effect sizes on the Sternberg task were on the smaller scale (d=.31–.37), suggesting that even in adulthood EF skills could be malleable to improvement with the right intervention and supports in place.

In terms of research implications, there is the need to continue to investigate the longitudinal development of EF within this population. Research beyond early adulthood is needed to better understand how EF continues to develop and change among VPT/VLBW individuals over the life span. While previous findings have shown that EF performance in this population is relatively stable from childhood into early adulthood, research would benefit from exploring EF outcomes in middle and later adulthood, particularly as the effects of ageing become apparent. Additionally, there is a need for research to investigate how to implement strategies and interventions to support EF skills in VLBW adults. EF skills are essential in everyday life and in adulthood these skills allow individuals to maintain a level of independent functioning and have been associated with improved self-reported quality of life in older age (Diamond, 2013). The practical implications of EF outcomes need to be considered in all areas of functioning including occupational, financial, educational and social relationships. Awareness around these real-life implications will be important in developing new or adapted strategies that reflect the roles and responsibilities of adulthood, as well as considering the needs and circumstances of VLBW individuals and the specific support they may require.

The findings from this study have the potential to contribute to research in informing theoretical and intervention models. However, there are limitations of the present study that should be taken into consideration in regards to future research and the interpretation of the study findings.

7.7 Limitations of the current study

Although efforts were made to remedy the methodological difficulties characteristic of longitudinal EF studies, the current study is not without its own limitations. First, a methodological issue of the current study is that the FT control group were recruited at a later stage and there are potential issues with sample representation. The control group was recruited at the 23–24 year follow-up based on electoral roles, supplemented by peer nomination by the VLBW group at 26 years (Darlow et al., 2015). As the group were recruited at a later stage of the study, and through various methods, there is the possibility the control group is not necessarily representative of the general typically developing population. Secondly, cut-offs for impairment were, with only one exception, determined based on both normative test data and distributions of the FT groups average scores. As impairment rates for the Sternberg task were based solely on the FT average score distribution, it creates the assumption the FT group is representative of a typically developing population and may potentially underestimate impairment in the VLBW group. While efforts were made to maintain consistency across rates of impairment, there still must be caution when interpreting the findings because of the potential limitations surrounding the control group.

Thirdly, the measures used in the assessment presented limitations for some participants with motor and neurosensory impairments which attributed to missing data. While the measures used are all standardized and ecologically validated in assessing EF performance, some tests were unable to be administered to participants. For example, individuals with visual and significant fine motor impairments were unable to complete the CTMT as it requires those functions to perform the task. Often in EF assessments, participants with severe neurosensory impairments are unable to perform the tests and are therefore excluded from the results (Pyhälä et al., 2011; Sølsnes et al., 2014; Suikkanen et al. 2021). While they are sometimes able to complete other measures within the EF battery, it does present a limitation in terms of understanding the true population estimates of EF abilities.

7.8 Directions for future research

Despite the limitations noted above, this study does advance understanding as being one of the few prospective longitudinal studies to follow individuals born VLBW from birth to adulthood. The national recruitment of this study has provided a valuable insight into long term outcomes for the VLBW population in New Zealand and is adding considerably to what is currently known about the neurocognitive development of individuals born VLBW. However, concerns and questions remain for VLBW populations that future research should address.

First, continuing to track the long-term development of individuals born VLBW will be important. Existing studies, including the present one, have examined EF outcomes up to 30 years of age, researching these outcomes into the third decade of life and beyond will be important in identifying whether neurocognitive deficits continue to stabilize overtime as VLBW individuals age. Secondly, the effects of perinatal and social predictors on EF outcomes in adulthood are yet to be fully explored in research. A couple of existing studies have looked at predictors of general cognitive ability (IQ). Eryigit Madzwamuse et al. (2015) found family socioeconomic status (SES; parents' education and occupation) to have a strong impact on IQ for both VLBW and FT adults, and Darlow et al. (2020) found the strongest predictors of IQ to be maternal education and birthweight. Predictors were not examined in the current study as it was beyond the scope of this thesis, therefore it would be beneficial for future research to examine predictors of EF performance in adults born VPT/VLBW.

Thirdly, the clinical significance of EF outcomes on VPT/VLBW birth would be important to examine in future research. Kroll et al. (2017) found there to be a strong positive association between EF and life achievement measures in the VLBW group compared to FT adults. Specifically, VLBW adults had less education and poorer social and occupational functioning. EF skills are crucial to everyday functioning and are likely to have a direct consequence on adulthood life course outcomes and achievements as initial findings by Kroll et al. (2017) have indicated. Impairments in EF that are found in statistical analyses could have a very significant impact on the everyday real-world functioning for some individuals born VLBW, therefore the clinical significance of those statistical findings needs to be further explored in adulthood.

Fourthly, the exploration of an EF profile for VPT/VLBW adults requires further replication. Issues attributed to this is that EF tasks are rarely pure measures of the construct they pertain to measure (Miyake & Friedman, 2012), and studies often use varying EF measures which can add to the difficulty of constructing an accurate EF profile for VPT/VLBW populations. For example, planning is a component of EF that has been less researched and preliminary findings among VLBW adults have reported no significant between groups differences on planning tasks (Sølsnes et al. 2014; Kroll et al., 2017). There has also been some evidence of contradictory findings across existing studies reporting non-significant results on EF tasks (Nosarti et al., 2007; Allin et al., 2011; Pyhälä et al. 2011). Further replication is needed to understand whether these findings may be attributed to variance in methodology or if they are suggestive of a VPT/VLBW profile in which certain aspects of EF are more affected than others.

7.9 Concluding remarks

Researching long term EF outcomes is important for understanding the nature of challenges that adults born VLBW may face and the extent of EF impairment overtime. The overall aim of the study was to characterise an executive functioning profile of a national cohort of adults born VLBW at age 26-30 years. Specific aims of this study were to examine the EF performance of a national cohort of adults born VLBW and adults born to FT across a range of measures, as well as examining the effects of processing speed and birthweight on EF performance. Findings of this study indicated that adults born VLBW perform worse on EF tasks and are at higher risk of impairment across a range of EF domains including inhibitory control, working memory (visuospatial and auditory-verbal) and cognitive flexibility, compared to FT born adults. Between-group differences in EF task performance were attenuated to some extent after controlling for processing however differences remained significant or close to significance across almost all EF tasks. Findings also suggested that risk for EF impairment increases with decreasing birthweight, as individuals born <1000g had the poorest EF performance across the tasks. Given the generally pervasive pattern of impaired performance across EF tasks, a factor analysis was conducted in which a single factor component was found to be most viable. This was line with Miyake and colleagues' framework theorizing that EF processes are unified in how they represent common underlying abilities.

To conclude, the findings of this study have important research and clinical implications. It is hoped that longitudinal research will continue to be undertaken with VPT/VLBW populations into later adulthood to greater understand long-term EF outcomes. The study emphasizes the importance of intervention and monitoring in early childhood before EF impairments potentially impact later life course opportunities in adulthood. However, positive findings suggest that with appropriate intervention and supports, EF skills could be malleable to changes in adulthood. Although adults born VLBW are at higher risk of EF impairment, continuing to research long-term outcomes and opportunities for interventions, could help to ensure that a high quality of life follows for individuals born VLBW.

References

Aarnoudse-Moens, C. S. H., 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 (Evanston), 124*(2), 717-728. https://doi.org/10.1542/peds.2008-2816

Aarnoudse-Moens, C. S. H., 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), e55994-e55994. https://doi.org/10.1371/journal.pone.0055994

Allen, D. N., Haderlie, M., Kazakov, D., & Mayfield, J. (2009). Construct and Criterion Validity of the Comprehensive Trail Making Test in Children and Adolescents with Traumatic Brain Injury. *Child neuropsychology, 15*(6), 543-553. https://doi.org/10.1080/09297040902748234

Allin, M., Walshe, M., Fern, A., Nosarti, C., Cuddy, M., Rifkin, L., Murray, R., Rushe, T., & Wyatt, J. (2008). Cognitive maturation in preterm and term born adolescents. *Journal of Neurology, Neurosurgery and Psychiatry, 79*(4), 381-386. https://doi.org/10.1136/jnnp.2006.110858

Allin, M. P., Kontis, D., Walshe, M., Wyatt, J., Barker, G. J., Kanaan, R. A., McGuire, P., Rifkin, L., Murray, R. M., & Nosarti, C. (2011). White matter and cognition in adults who were born preterm. *PloS one*, *6*(10), e24525. https://doi.org/10.1371/journal.pone.0024525

Allotey, J., Zamora, J., Cheong-See, F., Kalidindi, M., Arroyo-Manzano, D., Asztalos, E., van der Post, J. A. M, Mol, B. W., Moore, D., Birtles, D., Khan, K. S., & Thangaratinam, S. (2018). Cognitive, motor, behavioural and academic performances of children born preterm: A meta-analysis and systematic review involving 64 061 children. *BJOG: An International Journal of Obstetrics and Gynaecology, 125*(1), 16-25. https://doi.org/10.1111/1471-0528.14832

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71-82. https://doi.org/10.1076/chin.8.2.71.8724

Anderson, P. J., Doyle, L. W., Victorian Infant Collaborative Study Group, & Victorian Infant Collaborative Study Group. (2004). Executive functioning in school-aged children who were born very preterm or with extremely low birth weight in the 1990s. *Pediatrics (Evanston), 114*(1), 50-57. https://doi.org/10.1542/peds.114.1.50

Andrés, P., & Van der Linden, M. (2000). Age-related differences in supervisory attentional system functions. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 55*(6), P373-P380.

Armstrong, C. M., Allen, D. N., Donohue, B., & Mayfield, J. (2008). Sensitivity of the comprehensive trail making test to traumatic brain injury in adolescents. *Archives of clinical neuropsychology, 23*(3), 351-358. https://doi.org/10.1016/j.acn.2007.11.004

Arpi, E., D'Amico, R., Lucaccioni, L., Bedetti, L., Berardi, A., & Ferrari, F. (2019). Worse global intellectual and worse neuropsychological functioning in preterm-born children at preschool age: A meta-analysis. *Acta Paediatrica, 108*(9), 1567-1579. https://doi.org/10.1111/apa.14836

Baddeley, A. (1992). Working memory. *Science (American Association for the Advancement of Science), 255*(5044), 556-559. https://doi.org/10.1126/science.1736359

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018;2019;). From short-term store to multicomponent working memory: The role of the modal model. *Memory & Cognition, 47*(4), 575-588. https://doi.org/10.3758/s13421-018-0878-5

Beck, S., Wojdyla, D., Say, L., Betran, A. P., Merialdi, M., Requejo, J. H., Rubens, C., Menon, R., & Van Look, P. F. A. (2010). The worldwide incidence of preterm birth: A systematic review of maternal mortality and morbidity. *Bulletin of the World Health Organization, 88*(1), 31-38. https://doi.org/10.2471/BLT.08.062554

Benedict, R. H. (2005). Effects of using same- versus alternate-form memory tests during short-interval repeated assessments in multiple sclerosis. *Journal of the International Neuropsychological Society, 11*(6), 727-736. https://doi.org/10.1017/S1355617705050782

Benedict, R. H., DeLuca, J., Phillips, G., LaRocca, N., Hudson, L. D., Rudick, R., & Multiple Sclerosis Outcome Assessments Consortium. (2017). Validity of the symbol digit modalities test as a cognition performance outcome measure for multiple sclerosis. *SAGE Publications*. https://doi.org/10.1177/1352458517690821

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *JAMA: The Journal of the American Medical Association, 288*(6), 728-737. https://doi.org/10.1001/jama.288.6.728

Breeman, L. D., Jaekel, J., Baumann, N., Bartmann, P., & Wolke, D. (2015). Preterm Cognitive Function Into Adulthood. *Pediatrics (Evanston), 136*(3), 415-423. https://doi.org/10.1542/peds.2015-0608

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

Burnett, A. C., Scratch, S. E., Lee, K. J., Cheong, J., Searle, K., Hutchinson, E., De Luca, C., Davey, M., Roberts, G., Doyle, L. W., Anderson, P. J., & Victorian Infant Collaborative Study Group. (2015). Executive function in adolescents born <1000 g or <28 weeks: A prospective cohort study. *Pediatrics (Evanston), 135*(4), e826-e834. https://doi.org/10.1542/peds.2014-3188

Burnett, A. C., Anderson, P. J., Lee, K. J., Roberts, G., Doyle, L. W., Cheong, J. L. Y., Callanan, C., Carse, E., Charlton, M. P., Davis, N., Duff, J., Hutchinson, E., Hayes, M., Kelly, E., McDonald, M., Opie, G., Watkins, A., Williamson, A., Woods, H., . . . for the Victorian Infant Collaborative Study Group. (2018). Trends in executive functioning in extremely preterm children across 3 birth eras. *Pediatrics (Evanston), 141*(1), e20171958. https://doi.org/10.1542/peds.2017-1958

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests.* Bury St Edmunds: Thames Valley Test Company.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Chawanpaiboon, S., Vogel, J. P., Moller, A., Lumbiganon, P., Petzold, M., Hogan, D., Landoulsi, S., Jampathong, N., Kongwattanakul, K., Laopaiboon, M., Lewis, C., Rattanakanokchai, S., Teng, D. N., Thinkhamrop, J., Watananirun, K., Zhang, J., Zhou, W., & 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

Cheong, J. L., Spittle, A. J., Burnett, A. C., Anderson, P. J., & Doyle, L. W. (2020). Have outcomes following extremely preterm birth improved over time? *Seminars in Fetal & Neonatal Medicine*, *25*(3),101114-101114. https://doi.org/10.1016/j.siny.2020.101114

Clark, C., Prior, M., & Kinsella, G. J. (2000). Do executive function deficits differentiate between adolescents with ADHD and oppositional defiant/conduct disorder? A neuropsychological study using the Six Elements Test and Hayling Sentence Completion Test. *Journal of abnormal child psychology*, *28*(5), 403-414.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Craig, E. D., Thompson, J. M. D., & Mitchell, E. A. (2002). Socioeconomic status and preterm birth: New Zealand trends, 1980 to 1999. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 86*(3), F142-F146. https://doi.org/10.1136/fn.86.3.F142

Darlow, B. A. (1988). Incidence of retinopathy of prematurity in new zealand. *Archives of Disease in Childhood, 63*(9), 1083-1086. https://doi.org/10.1136/adc.63.9.1083

Darlow, B. A., Horwood, L. J., & Clemett, R. S. (1992). Retinopathy of prematurity: Risk factors in a prospective population-based study. *Paediatric and Perinatal Epidemiology, 6*(1), 62.

Darlow, B. A., Horwood, L. J., Pere-Bracken, H. M., & Woodward, L. J. (2013). Psychosocial Outcomes of Young Adults Born Very Low Birth Weight. *Pediatrics (Evanston), 132*(6), e1521-e1528. https://doi.org/10.1542/peds.2013-2024

Darlow, B. A., Horwood, L. J., Woodward, L. J., Elliott, J. M., Troughton, R. W., Elder, M. J., Epton, M. J., Stanton, J. D., Swanney, M. P., Keenan, R., Melzer, T. R., McKelvey, V. A., Levin, K., Meeks, M. G., Espiner, E. A., Cameron, V. A., & Martin, J. (2015). The new zealand 1986 very low birth weight cohort as young adults: Mapping the road ahead. *BMC Pediatrics, 15*(1), 90-90. https://doi.org/10.1186/s12887-015-0413-9

Darlow, B. A., Woodward, L. J., Levin, K. J., Melzer, T., & Horwood, L. J. (2020). Perinatal and childhood predictors of general cognitive outcome at 28 years in a very-low-birthweight national cohort. *Developmental Medicine & Child Neurology, 62*(12), 1423-1428. https://doi.org/10.1111/dmcn.14649

de Kieviet, J. F., Piek, J. P., Aarnoudse-Moens, C. S., & Oosterlaan, J. (2009). Motor development in very preterm and very low-birth-weight children from birth to adolescence: a meta-analysis. *JAMA*, *302*(20), 2235–2242. https://doi.org/10.1001/jama.2009.1708

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Doyle, L. W. (1995). Outcome to five years of age of children born at 24-26 weeks' gestational age in victoria. the victorian infant collaborative study group. *Medical Journal of Australia, 163*(1), 11.

Drake, A., Weinstock-Guttman, B., Morrow, S., Hojnacki, D., Munschauer, F., & Benedict, R. (2010). Psychometrics and normative data for the multiple sclerosis functional composite: Replacing the PASAT with the symbol digit modalities test. *Multiple Sclerosis, 16*(2), 228-237. https://doi.org/10.1177/1352458509354552

Eryigit Madzwamuse, S., Baumann, N., Jaekel, J., Bartmann, P., & Wolke, D. (2015). Neuro-cognitive performance of very preterm or very low birth weight adults at 26 years. *Journal of Child Psychology and Psychiatry, 56*(8), 857-864. https://doi.org/10.1111/jcpp.12358

Flynn, J. R. (1987). "massive IQ gains in 14 nations: What IQ tests really measure": Correction to flynn. *Psychological Bulletin*, *101*(3), 427-427. https://doi.org/10.1037/h0090408

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology. General, 137*(2), 201-225. https://doi.org/10.1037/0096-3445.137.2.201

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology, 47*(5), 1410-1430. https://doi.org/10.1037/a0023750

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86,* 186-204. https://doi.org/10.1016/j.cortex.2016.04.023

Gough, A., Linden, M., Spence, D., Patterson, C. C., Halliday, H. L., & McGarvey, L. P. A. (2014). Impaired lung function and health status in adult survivors of bronchopulmonary dysplasia. *The European Respiratory Journal, 43*(3), 808.

Hee Chung, E., Chou, J., & Brown, K. A. (2020). Neurodevelopmental outcomes of preterm infants: A recent literature review. *Translational Pediatrics, 9*(Suppl 1), S3-S8. https://doi.org/10.21037/tp.2019.09.10

Heinonen, K., Lahti, J., Sammallahti, S., Wolke, D., Lano, A., Andersson, S., Pesonen, A. K., Eriksson, J. G., Kajantie, E., & Raikkonen, K. (2018). Neurocognitive outcome in young adults born late-preterm. *Developmental Medicine & Child Neurology, 60*(3), 267-274. https://doi.org/10.1111/dmcn.13616

Hornik, C. P., Fort, P., Clark, R. H., Watt, K., Benjamin, D. K., Smith, P. B., Manzoni, P., Jacqz-Aigrain, E., Kaguelidou, F., & Cohen-Wolkowiez, M. (2012). Early and late onset sepsis in very-low-birth-weight infants from a large group of neonatal intensive care units. *Early Human Development*, *88*(Suppl 2), S69-S74. https://doi.org/10.1016/S0378-3782(12)70019-1

Horwood, L. J., Mogridge, N., & Darlow, B. A. (1998). Cognitive, educational, and behavioural outcomes at 7 to 8 years in a national very low birthweight cohort. *Archives of disease in childhood. Fetal and neonatal edition, 79*(1), F12-F20. https://doi.org/10.1136/fn.79.1.F12

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huizinga, M., Dolan, C. V., & van der Molen, Maurits W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017-2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Jing, J., Dai, Y., Li, Y., Zhou, P., Li, X., Mei, J., Zhang, C., Sangild, P. T., Tang, Z., Xu, S., Su, Y., He, X., & Zhu, Y. (2021). Single-course antenatal corticosteroids is related to faster growth in very-low-birth-weight infant. *BMC Pregnancy and Childbirth, 21*(1), 50-50. https://doi.org/10.1186/s12884-020-03510-w

Kim, H. M., Horwood, J., Harris, S., Bora, S., Darlow, B., & Woodward, L. (in press). Self-reported executive function abilities in adults born at very low birthweight. *Paediatric and Perinatal Epidemiology*.

Kroll, J., Karolis, V., Brittain, P. J., Tseng, C.-E. J., Froudist-Walsh, S., Murray, R. M., & Nosarti, C. (2017). Real-life impact of executive function impairments in adults who were born very preterm. *Journal of the International Neuropsychological Society, 23*(5), 381-389. https://doi.org/10.1017/S1355617717000169

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4,* 1-12. https://doi.org/10.3389/fpsyg.2013.00863

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

Lundequist, A., Böhm, B., Lagercrantz, H., Forssberg, H., & Smedler, A. (2015). Cognitive outcome varies in adolescents born preterm, depending on gestational age, intrauterine growth and neonatal complications. *Acta Paediatrica*, *104*(3), 292-299. https://doi.org/10.1111/apa.12864

McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: A meta-analysis of functional neuroimaging data. *Frontiers in Human Neuroscience, 11*, 154-154. https://doi.org/10.3389/fnhum.2017.00154

Ministry of Health. (2016). *Fetal and Infant Deaths 2016*. Wellington: Ministry of Health. https://www.health.govt.nz/publication/fetal-and-infant-deaths-2016

Ministry of Health. (2019). *Report on Maternity 2017*. Wellington: Ministry of Health. https://www.health.govt.nz/publication/report-maternity-2017

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Moses, J. J. A. (2004). Test review-Comprehensive Trail Making Test (CTMT). *Archives of clinical neuropsychology, 19*(5), 703-708.

New Zealand Health Information Service. (2007). *Fetal and Infant Deaths 2003 & 2004*. Wellington: Ministry of Health. https://www.health.govt.nz/publication/fetal-and-infant-deaths-2003-2004

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 241-268. https://doi.org/10.3758/s13415-011-0083-5

Nosarti, C., Giouroukou, E., Micali, N., Rifkin, L., Morris, R. G., & Murray, R. M. (2007). Impaired executive functioning in young adults born very preterm. *Journal of the International Neuropsychological Society, 13*(4), 571-581. https://doi.org/10.1017/S1355617707070725

Odhuba, R. A., van den Broek, M. D, & Johns, L. C. (2005). Ecological validity of measures of executive functioning. *British Journal of Clinical Psychology, 44*(2), 269-278. https://doi.org/10.1348/014466505X29431

O'Reilly, H., Johnson, S., Ni, Y., Wolke, D., & Marlow, N. (2020). Neuropsychological outcomes at 19 years of age following extremely preterm birth. *Pediatrics (Evanston), 145*(2), e20192087. https://doi.org/10.1542/peds.2019-2087

Østgård, H. F., Sølsnes, A. E., Bjuland, K. J., Rimol, L. M., Martinussen, M., Brubakk, A., Håberg, A. K., Skranes, J., & Løhaugen, G. C. C. (2016). Executive function relates to surface area of frontal and temporal cortex in very-low-birth-weight late teenagers. *Early Human Development, 95*, 47-53. https://doi.org/10.1016/j.earlhumdev.2016.01.023

Pietrzak, R. H., Maruff, P., Mayes, L. C., Roman, S. A., Sosa, J. A., & Snyder, P. J. (2008). An examination of the construct validity and factor structure of the groton maze learning test, a new measure of spatial working memory, learning efficiency, and error monitoring. *Archives of Clinical Neuropsychology, 23*(4), 433-445. https://doi.org/10.1016/j.acn.2008.03.002

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the flynn effect (1909-2013). *Perspectives on Psychological Science, 10*(3), 282-306. https://doi.org/10.1177/1745691615577701

Pritchard, V. E., Bora, S., Austin, N. C., Levin, K. J., & Woodward, L. J. (2014). Identifying very preterm children at educational risk using a school readiness framework. *Pediatrics (Evanston), 134*(3), e825-e832. https://doi.org/10.1542/peds.2013-3865

Pyhälä, R., Lahti, J., Heinonen, K., Pesonen, A. K., Strang-Karlsson, S., Hovi, P., Järvenpää, A.L., Eriksson, J. G., Andersson, S., Kajantie, E., & Räikkönen, K. (2011). Neurocognitive abilities in young adults with very low birth weight. *Neurology*, *77*(23), 2052–2060. https://doi.org/10.1212/WNL.0b013e31823b473e

Reynolds, C. R. (2002). *Comprehensive trail making test (CTMT).* Austin, TX: Pro-Ed, 408.

Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology, 12*(1), 105-126. https://doi.org/10.1080/87565649609540642

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. J. S. A. (1994). *The test of everyday attention (TEA).* Bury St. Edmunds, UK: Thames Valley Test Company, 197-221.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge University Press. https://doi.org/10.1017/CBO9780511526817

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of Neuroscience, 27*(9), 2349-2356. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Sheridan, L. K., Fitzgerald, H. E., Adams, K. M., Nigg, J. T., Martel, M. M., Puttler, L. I., ...

& Zucker, R. A. (2006). Normative Symbol Digit Modalities Test performance in a community-based sample. *Archives of Clinical Neuropsychology, 21*(1), 23-28.

Sølsnes, A. E., Skranes, J., Brubakk, A., & Løhaugen, G. C. C. (2014). Executive functions in very-low-birth-weight young adults: A comparison between self-report and neuropsychological test results. *Journal of the International Neuropsychological Society, 20*(5), 506-515. https://doi.org/10.1017/S1355617714000332

Stålnacke, J., Lundequist, A., Böhm, B., Forssberg, H., & Smedler, A. C. (2019). A longitudinal model of executive function development from birth through adolescence in children born very or extremely preterm. *Child neuropsychology: a journal on normal and abnormal development in childhood and adolescence*, *25*(3), 318–335. https://doi.org/10.1080/09297049.2018.1477928

Sternberg, S. (1966). High-speed scanning in human memory. *Science (American Association for the Advancement of Science), 153*(3736), 652-654. https://doi.org/10.1126/science.153.3736.652

Suikkanen, J., Miettola, S., Heinonen, K., Vääräsmäki, M., Tikanmäki, M., Sipola, M., Matinolli, H., Järvelin, M., Räikkönen, K., Hovi, P., & Kajantie, E. (2021;2020;). Reaction times, learning, and executive functioning in adults born preterm. *Pediatric Research, 89*(1), 198-204. https://doi.org/10.1038/s41390-020-0851-4

Tan, Z., Chong, C., Darlow, B., & Dai, S. (2015). Visual impairment due to retinopathy of prematurity (ROP) in new zealand: A 22-year review. *British Journal of Ophthalmology, 99*(6), 801-806. https://doi.org/10.1136/bjophthalmol-2014-305913

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The flynn effect: A meta-analysis. *Psychological Bulletin, 140*(5), 1332-1360. https://doi.org/10.1037/a0037173

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 and Child Neurology, 61*(9), 1015-1024. https://doi.org/10.1111/dmcn.14213

Vollmer, B., & Stålnacke, J. (2019). Young adult motor, sensory, and cognitive outcomes and longitudinal development after very and extremely preterm birth. *Neuropediatrics, 50*(4), 219.

Vordenberg, J. A., Barrett, J. J., Doninger, N. A., Contardo, C. P., & Ozoude, K. A. (2014). Application of the brixton spatial anticipation test in stroke: Ecological validity and performance characteristics. *Clinical Neuropsychologist, 28*(2), 300-316. https://doi.org/10.1080/13854046.2014.881555

Wiebe, S. A., & Karbach, J. (2018). *Executive function: Development across the life span.* Routledge.

Woodward, L. J., Horwood, J. L., Darlow, B. A., & Bora, S. (2021). Visuospatial working memory of children and adults born very preterm and/or very low birth weight. *Pediatric Research.*

APPENDIX A:
ETHICAL APPROVAL FOR THE NEW ZEALAND 1986 VLBW STUDY:
26-30 YEAR FOLLOW-UP

Health
and
Disability
Ethics
Committees

Upper South B Regional Ethics Committee
c/- Ministry of Health
6 Hazeldean Road, Level 1 Montgomery Watson Building
Addington, Christchurch
PO Box 3877
Phone: (03) 974 2305
Email: uppersouthb_ethicscommittee@moh.govt.nz

14 May 2012

Professor Brian Darlow
Department of Paediatrics
Christchurch School of Medicine
University of Otago
PO Box 4345
Christchurch

Dear Professor Darlow

Re: Ethics ref: **URB/12/05/015** (please quote in all correspondence)
Study title: The New Zealand 1986 VLBW cohort as young adults: mapping the road ahead
Investigators: Professor Brian Darlow, Associate Professor John Horwood, Professor Lianne Woodward, Associate Professor John Elliot, Associate Professor Richard Troughton

This study was given ethical approval by the Upper South B Regional Ethics Committee on 7 May 2012. A list of members of the Committee is attached.

Approved Documents

- Part 4
- Form A
- Part 5
- Part 6
- Consent form dated 10 April 2012
- Information sheet and consent form for genetic profiles of heart disease, dated 10 April 2012
- Letter of support from Neonatal Trust

This approval is valid until 31 December 2015, provided that Annual Progress Reports are submitted (see below).

Access to ACC

For the purposes of section 32 of the Accident Compensation Act 2001, the Committee is satisfied that this study is not being conducted principally for the benefit of the manufacturer or distributor of the medicine or item in respect of which the trial is being

Upper South B Regional Ethics Committee
c/- Ministry of Health
6 Hazeldean Road, Level 1 Montgomery Watson Building
Addington, Christchurch
PO Box 3877
Phone: (03) 974 2305
Email: uppersouthb_ethicscommittee@moh.govt.nz

carried out. Participants injured as a result of treatment received in this trial will therefore be eligible to be considered for compensation in respect of those injuries under the ACC scheme.

Amendments and Protocol Deviations

All significant amendments to this proposal must receive prior approval from the Committee. Significant amendments include (but are not limited to) changes to:

- the researcher responsible for the conduct of the study at a study site
- the addition of an extra study site
- the design or duration of the study
- the method of recruitment
- information sheets and informed consent procedures.

Significant deviations from the approved protocol must be reported to the Committee as soon as possible.

Annual Progress Reports and Final Reports

The first Annual Progress Report for this study is due to the Committee by 7 May 2013. The Annual Report Form that should be used is available at www.ethicscommittees.health.govt.nz. Please note that if you do not provide a progress report by this date, ethical approval may be withdrawn.

A Final Report is also required at the conclusion of the study. The Final Report Form is also available at www.ethicscommittees.health.govt.nz.

Requirements for the Reporting of Serious Adverse Events (SAEs)

SAEs occurring in this study must be individually reported to the Committee within 7-15 days only where they:

- are *unexpected* because they are not outlined in the investigator's brochure, and
- are not defined study end-points (e.g. death or hospitalisation), and
- occur in patients located in New Zealand, and
- if the study involves blinding, result in a decision to break the study code.

There is no requirement for the individual reporting to ethics committees of SAEs that do not meet all of these criteria. However, if your study is overseen by a data monitoring committee, copies of its letters of recommendation to the Principal Investigator should be forwarded to the Committee as soon as possible.

Health and Disability Ethics Committees

Upper South B Regional Ethics Committee
c/- Ministry of Health
6 Hazeldean Road, Level 1 Montgomery Watson Building
Addington, Christchurch
PO Box 3877
Phone: (03) 974 2305
Email: uppersouthb_ethicscommittee@moh.govt.nz

Please see www.ethicscommittees.health.govt.nz for more information on the reporting of SAEs, and to download the SAE Report Form.

Statement of compliance
The committee is constituted in accordance with its Terms of Reference. It complies with the *Operational Standard for Ethics Committees* and the principles of international good clinical practice.

The committee is approved by the Health Research Council's Ethics Committee for the purposes of section 25(1)(c) of the Health Research Council Act 1990.

We wish you all the best with your study.

Yours sincerely

Diana J. Whipp

Mrs Diana Whipp
Administrator Upper South B Regional Ethics Committee
Email: uppersouthb_ethicscommittee@moh.govt.nz

APPENDIX B:

ONGOING APPROVAL FOR THE NEW ZEALAND 1986 VLBW STUDY

Health and Disability Ethics Committees
Ministry of Health133
Molesworth Street
PO Box 5013
Wellington0800 4
ETHICS
hdecs@health.govt.nz

06 March 2020

Professor Brian Darlow Department of
Paediatrics University of Otago
ChristchurchPO Box 4345
Christchurch 8140

Dear Professor Darlow,

Re:	**Ethics ref:**	**URB/12/05/015/AM08**
	Study title:	The New Zealand 1986 VLBW cohort as young adults: mapping the road ahead

I am pleased to advise that this annual progress report has been *approved*, following review by the Chairperson of the Southern Health and Disability Ethics Committee on 20January 2020. Existing approval remains valid.

Your next progress report is due by 12 February 2021.

Please don't hesitate to contact the HDEC secretariat for further information. We wishyou all the best for your study.

Yours sincerely,

Dr Sarah Gunningham
Chairperson
Southern Health and Disability Ethics Committee

Encl: appendix A: documents submitted

Appendix A
Documents submitted and approved

Document	Version	Date
Post Approval Form	AM08	16 January 2020