

CHILDREN BORN PREMATURELY: COGNITIVE OUTCOMES AND
PRELIMINARY FINDINGS FOR SUBSEQUENT INTERVENTION

Helen Hoi Lam Ko
PostGradDip in Psychology
School of Psychology
Murdoch University

This thesis is presented for the degree of Doctor of Psychology (Clinical) of
Murdoch University, 2015

I declare that this thesis is my own account of my research and contains as its main content, work which has not previously been submitted to a degree at any tertiary education institution.

Helen Hoi Lam Ko

ABSTRACT

Well-established evidence shows that children born preterm/low birth weight (LBW) are at increased risk of academic difficulties (Lee, Yeatman, Luna, & Feldman, 2011; Pritchard et al., 2009) and, despite global IQ scores within the normal range, nonetheless display lower academic performance than their same age peers (Bhutta, Cleves, Casey, Craddock, & Anand, 2002; Kerr-Wilson, Mackay, Smith, & Pell, 2011). This is not fully understood and previous attempts to improve these circumstances through means of cognitive intervention have met with little success. Therefore, the current thesis investigates possible underlying mechanisms of this intellectual disparity and tests the effectiveness of one potential intervention. In doing so, the studies presented focus specifically on fluid intelligence (Taub, 2002). The investigation through fluid intelligence is relatively novel in the current literature and therefore worthy of further exploration. Normal individual differences in fluid intelligence have been explained with reference to information processing parameters. Previous studies have shown that children born preterm/LBW have impairments in basic processes identified with executive function (Aarnoudse-Moens, Smidts, Oosterlaan, Duivenvoorden, & Weisglas-Kuperus, 2009; Mulder, Pitchford, Hagger, & Marlow, 2009). However, the current study is the first to test whether differences in fluid intelligence, as measured by the Cattell Culture Fair Tests, between preterm ($n = 217$) and typically developing children ($n = 145$) could be accounted for by differences in working memory and cognitive flexibility, as measured by the digit span tasks and the Wisconsin Card Sorting Test respectively. Results indicate that the seven to nine years old preterm cohort performed less well on measures of fluid intelligence than their peers across all age groups and their differences were partially mediated by both working memory and cognitive flexibility in a multiple mediation analysis. It also identified at

least one year of developmental delay in fluid intelligence between the clinical group and their peers.

Provided with evidence from Study 1 and parallel research suggesting that computerized working memory training may enhance working memory and fluid intelligence in non-clinical groups (Jaeggi, Buschkuhl, Jonides, & Perrig, 2008; Klingberg, Forssberg, & Westerberg, 2002; Studer et al., 2009), the second goal of this thesis was to conduct a preliminary study to investigate the feasibility of cognitive training for children born preterm/LBW. Therefore, in the second study, the utility of a brief adaptive working memory span training program (Buschkuhl, Jaeggi, Kobel, & Perrig, 2008) was tested in typically developing children. Sixty-three children, aged seven to nine years, were randomly assigned to one of three groups: Intervention, active control and passive control. The intervention group was trained in the adaptive version of the working memory span task and the active control group was trained in the non-adaptive version. Both groups trained for 15 minutes each day for a duration of 20 days. Participants in the passive control group participated only in pre and post assessments. All participants were assessed using the digit span and spatial span tasks for measuring working memory, the Stroop task for measuring executive control, a reaction time task for measuring processing speed and the Raven's Standard Progressive Matrices for measuring fluid intelligence. Results indicate that children in the intervention group improved on their trained task and demonstrated significant far transfer effects on the assessment of fluid intelligence compared to both control groups. However, no near transfer to other measures was found. The reason behind the occurrence of far transfer effect without evidence of near transfer effects was unclear. However, given that the adaptive complex working memory training task was not in any way similar to the fluid intelligence measure, significant differences in fluid intelligence gains were unlikely to

have been a consequence of practice or general familiarity effects but, rather, a consequence of the training.

Although Study 1 identified that working memory and cognitive flexibility partially mediate birth status-related differences in G_f , the impact of these variables on academic performance in children born preterm is still unknown. Nonetheless, current evidence of far transfer to fluid intelligence after adaptive complex working memory span training provides support for the utility of WM training and modifiability in G_f . This in turn provides a preliminary evidence-base approach for psychologists to work toward providing neuro-remediation treatment options to targeted clinical groups, such as those born preterm with fluid intelligence deficiencies. In combination, the outcomes of these two studies provide both a theoretical contribution to our understanding of the deficits observed in children born preterm and an applied contribution to beginning the process of developing appropriate intervention programmes suitable for this clinical group in the future, with hopeful prospects for improving cognitive outcomes.

TABLE OF CONTENTS

DECLARATION	2
ABSTRACT	3
TABLE OF CONTENTS	6
LIST OF TABLES AND FIGURES	11
ACKNOWLEDGEMENTS	13
CHAPTER 1 – General Overview	14
Aims of the Current Studies	19
Overview of Chapters	20
CHAPTER 2 - Literature Review: Children Born Preterm/ Low Birth Weight	22
Introduction	22
Defining Prematurity	22
Incidence of Preterm Birth	24
Commonly Associated Neurological Disorders in Preterm Birth/LBW	25
Prematurity and Academic Performance	27
Outcomes of Standardized Tests and Teacher Ratings	27
Special Educational Needs and Academic Attainment	30
Summary	31
CHAPTER 3 – Literature Review: Cognitive Outcomes of Preterm Birth	33
Introduction	33
Intelligence Quotient (IQ) and Education Attainment	35
Prematurity and IQ	36
Do Cognitive Differences Measured by IQ Tests Persist into Adulthood?	39
Outcomes in Domain-Specific Abilities	40
Impact of Confounding Factors	45

Prematurity, Cognitive Abilities & Intervention

Fluid Intelligence (G_f) and Education Attainment	47
Prematurity and Fluid Intelligence	49
Explaining Individual Differences in Fluid Intelligence through Information Processing Parameters	51
Executive Function (EF)	54
Inhibitory Control	55
Cognitive Flexibility	55
Working Memory	56
Normal Developmental Trajectories of EF and Methodological Concerns	60
Prematurity and Information Processing Parameters of EF	62
Inhibitory control	63
Cognitive Flexibility	67
Working Memory	69
Confounding Variables	74
Relationships between Working Memory and Fluid Intelligence	75
Summary	80
CHAPTER 4 - Study 1: Effects of preterm birth on fluid intelligence: Investigating working memory and cognitive inflexibility as potential mediators	82
Introduction	82
The Current Study	85
Hypotheses	85
Method	86
Participants	86
Measures	88
Procedure	90

Results	91
Data Preparation	92
Birth Group Differences	93
Mediation Analysis	99
Discussion	104
Birth Group Differences	105
Mediation Analysis	108
Limitations	111
Clinical Implications and Future Directions	113
Summary	115
CHAPTER 5 – Literature Review: Computerized Working Memory Training	117
Introduction	117
Types of Computerized Working Memory Training	118
Variations in Training Regime	123
Near Transfer and Far Transfer Effects	125
Near Transfer to Working Memory	132
Far Transfer to Other Cognitive Tasks	138
Far Transfer to Fluid Intelligence	141
Does Training Gain Persist?	148
The Importance of the <i>Adaptive</i> Nature of Training	150
Other Factors Affecting Learning and Possible Transfer of Learning	152
Summary of the Various Training Regimes	153
Cogmed Training	153
Dual <i>n</i> -back Training	155
Complex Working Memory Span Training	156

Ideal Characteristics of Working Memory Training Studies	156
Summary	158
CHAPTER 6 - Study 2: Adaptive working memory span task training to increase fluid intelligence in typically developing children	160
Introduction	160
The Current Study	163
Hypotheses	164
Method	164
Participants	164
Measures	168
Procedure	175
Results	178
Data Preparation	178
Specific Training Effects: Intervention Group	180
Individual Differences and Training Performance	184
Between Group Comparisons at Pre-test	185
Transfer Effects: Between Group Comparisons at Post-test	186
Discussion	191
Implications of Training Performance for the Intervention Group	191
Transfer to Working Memory	192
Transfer to Other Cognitive Tests	193
Transfer to G_f	195
Parallel Versions of Assessments	198
Limitations	199
Clinical Implications	203

Future Directions	204
Summary	206
CHAPTER 7 – Summary and General Discussion	208
Summary of Current Research	208
Summary of Findings	209
Summary of Study 1	209
Similar Developmental Trajectories in G_f	210
Different Developmental Trajectories in WM	210
Different Developmental Trajectories in Cognitive Flexibility	212
Summary of Study 2	213
Results with Higher Validity than Existing Studies	214
Theoretical Contributions	214
Integrated Practical Implications for Children Born Preterm/LBW	215
Future Directions	216
Conclusions	217
REFERENCES	219
APPENDICES	258
Appendix A: Human Research Ethics Approval Letter for Study 1	258
Appendix B: Participating Schools: Letter of Approval	259
Appendix C: Participant Invitation Letter	260
Appendix D: Parent Consent Form	262
Appendix E: School-aged Participants' Consent Form	263
Appendix F: Human Research Ethics Approval Letter for Study 2	264
Appendix G: Department of Education Ethics Approval Letter	265
Appendix H: Catholic Education Office Approval Letter	266

LIST OF TABLES AND FIGURES

	Page	
Table 4.1	Characteristics of LBW Preterm Group and Control Group	87
Table 4.2	Descriptive Statistics for Psychometric Test Raw Scores for Each Group	94
Table 4.3	Inferential Statistics and Effect Sizes for the Measures of Cognitive Function across Each Age Group	95
Table 4.4	Raw Correlations between Variables used in Multiple Mediation Analysis	101
Table 4.5	Results of Multiple Mediation Analysis for Each Age Group as Predictors of Fluid Intelligence	102
Table 5.1	A Summary of Recent Computerized Working Memory Training Studies	126
Table 6.1	Final Analysed Sample: Characteristics of Participants by Group	166
Table 6.2	Descriptive Statistics for the Number of Days Between Pre- and Post-Tests	178
Table 6.3	Descriptive Statistics for Psychometric Test Scores of Each Group	180
Table 6.4	Descriptive Statistics for Training Performance in Set Sizes	181
Table 6.5	Summary of Correlations on Cognitive Measures at Pre-test, Differences in Scores at Post-test, and Training Gains of the Intervention Group	190
Figure 3.1	The flow chart of the course of information in Chapter 3	34
Figure 3.2	Structure of WISC-III and WISC-IV	42
Figure 3.3	Cascade Pathway of Birth Status Effects on FSIQ	53
Figure 4.1	Mean Raw Scores on CCFIT for Fluid Intelligence	96
Figure 4.2	Mean Raw Scores on digit span for Working Memory	97
Figure 4.3	Mean Raw Scores on WCST for Cognitive Flexibility	99
Figure 4.4	Multiple Mediation Path Models	101

Figure 5.1	Diagram of the Dual n -back Training Task	120
Figure 5.2	Details of Adapted Experimental Procedure	122
Figure 6.1	Sampling and Flow of Participants	167
Figure 6.2	Digit span Pre-test and Post-test	168
Figure 6.3	Spatial span Pre-test and Post-test	170
Figure 6.4	Study 2: Experimental Procedure	175
Figure 6.5	A Summary of Average Training Levels	183
Figure 6.6	A Summary of Highest Training Levels	184
Figure 6.7	Transfer Effects for RSPM	189

ACKNOWLEDGEMENTS

I would like to thank a number of people, without whom this dissertation would not have been a success. Firstly, I would like to express my deepest appreciation to all my supervisors, Dr Helen Davis, Dr Corinne Reid and Dr Kathy Modecki at Murdoch University. They have given me immense support throughout the development and write-up of my thesis. Their continuous encouragement and patience have given me the strength to accomplish what I have today. Without their valuable time commitment, guidance and thorough feedback, I would not have been able to carry on with what seemed too much of a struggle to accomplish. Thank you!

My sincere appreciation also goes to the Neonatal Intensive Care Unit at King Edward Memorial Hospital. The team includes Dr Noel French, Dr Catherine Campbell and Dr Judy McCallum. Without their outstanding work with children born premature/low birth weight, research on the understanding of these children's development would not be possible.

I would also like to thank the research team from Project Kids Intellectual Development Study (PKIDS), at Murdoch University and at the Neurocognitive Development Unit of the University of Western Australia, for their contribution to data collection. Also, I wish to express many thanks to the principals, teachers, students, parents and volunteers who agreed to participate in this research.

Thank you to my friends, Wen, Dom, Jade, Delphine, Lisa, Hervé, Colleen, Darren, you have created "a home away from home" for me. I also want to thank my family, Mom, Dad, Bonnie and Crystal, for always being there for me unconditionally. Finally, thanks Jimmy, for your patience, support, and love. Thank you all for believing in me.

CHAPTER 1

General Overview

Advancement in obstetric and neonatal care has led to significant reductions in infant mortality and also improvements in the long term survival rates of infants born preterm and low birth weight (LBW; (Saigal & Doyle, 2008). Early intervention programs, such as caffeine for apnea in premature infants (Davis et al., 2010; Schmidt et al., 2007) and antenatal magnesium sulphate therapy (Doyle, Crowther, Middleton, Marret, & Rouse, 2010) have been shown to have short-term neurological benefits. However, there continue to be long-standing concerns regarding academic underachievement and long-term cognitive impairments amongst children born preterm/LBW.

Children born preterm/LBW are at increased risk of academic struggle, with various studies documenting academic achievement to be lower than that of their full term peers (Bowen, Gibson, & Hand, 2002; Johnson, Wolke, Hennessy, & Marlow, 2011; Mulder, Pitchford, & Marlow, 2010; Pritchard, et al., 2009; Roberts, Lim, Doyle, & Anderson, 2011; Taylor et al., 2011). Evidence for academic impairments has been documented not only during childhood but also continuing into adolescence and adulthood (Johnson, et al., 2011; Mathiasen, Hansen, Anderson, & Greisen, 2009; Taylor, et al., 2011). Special educational needs have also been reported to be more prevalent for this population (Lohaugen et al., 2010; Van Baar, Vermaas, Knots, de Kleine, & Soons, 2009).

Given these circumstances, current clinical research needs not only to focus on increasing survival rates but also reducing morbidity through improving life outcomes. Accordingly, this will involve continuation in the work of understanding the mechanisms giving rise to academic difficulties and cognitive disparity, as well as

providing possible solutions aimed at reducing them.

One important line of research is to investigate the precise nature of the cognitive profile of children born preterm/LBW - underachievement can be the result of various cognitive impairments. Previous research on the effect of prematurity/LBW on cognitive ability has largely focused on general intelligence or Spearman's *g* through standard intelligence testing. In the normal population, Full Scale Intelligence Quotient (FSIQ) as measured by Wechsler Intelligence Scale correlates very highly with the *g* factor (Keith, Fine, Taub, Reynolds, & Kranzler, 2006). Binet and Simon (1916) originally designed intelligence tests as practical assessment tools to identify individuals with learning difficulties and predict academic achievement (Nicolas & Levine, 2012). Their predictive power has been verified by existing evidence of high correlations between intelligence and academic achievement ($r = .69 - .81$; (Deary, Strand, Smith, & Fernanades, 2007; Rohde & Thompson, 2007; Tomporowski, Davis, Miller, & Naglieri, 2008; Wechsler, 1991).

There is general consensus that prematurity is associated with lower IQ test performance (Bhutta, et al., 2002; Kerr-Wilson, et al., 2011). However, although evidence points to significant group differences between this clinical group and full-term/normal birth weight (NBW) children, the performance of children born preterm often remains within the normal range (Bhutta, et al., 2002; Kerr-Wilson, et al., 2011). Full term and NBW children with IQs in the 90s generally perform respectably at school, whereas relatively small deficits in IQ often correspond to persisting risk of academic problems amongst children born preterm/LBW (Pritchard, et al., 2009). Perhaps this measurement of intelligence is insufficiently sensitive to answer the question. This suggests that a different theoretical approach may be necessary.

Intelligence tests, such as Wechsler Intelligence Scale for Children (WISC),

provide global and index scores that are designed to maximise practical use. Designers of Wechsler-type IQ tests set out to sample knowledge broadly across many domains, such as working memory and processing speed. Subsequent analyses found a g factor using the measurement model approach. The measurement model approach refers to the process of identifying latent factors of intelligence through factor analyses of normative data representing performance on the instrument. However, this approach results in only a descriptive account of individual differences in performance. It does not, and was not intended to, identify theoretical constructs (Taub, 2002).

In contrast, fluid intelligence (G_f), proposed by Cattell, is a theoretical construct representing one's ability to reason about novel events, which is independent of one's cultural knowledge and experience (Cattell, 1963). Tests that measure G_f , such as Raven's Progressive Matrices (RPM) and the Cattell Culture Fair Intelligence Tests (CCFIT; (Cattell & Cattell, 1949), were developed based on theories of intelligence, using a theoretical model approach, and were designed using narrow non-verbal tests of abstract reasoning. Despite differences in operationalization, g , G_f and FSIQ are sometimes viewed as interchangeable entities (Carpenter, Just, & Shell, 1990; Carroll, 1993, 2003; Keith, et al., 2006). For examples, some authors have suggested that G_f is identical to the g factor based on factor analysis of a range of psychometric cognitive ability tests (Gustafsson, 1984; Keith, et al., 2006; Kvist & Gustafsson, 2008). However, Duncan and his colleagues (1995) suggested FSIQ and G_f part company in some populations. For example, individuals with damage to the prefrontal cortex can evince no change in FSIQ but a deficit in G_f . Possibly, FSIQ represents a historical measure of how well such cases could think and cope in the past, which became embedded knowledge and, therefore, their premorbid G_f would be the best predictor of their current FSIQ, while their current G_f might be a better indicator of how well they

will cope with current and future thinking. Accordingly, tests such as RPM and CCFIT are suitable for theoretically-based investigations of the effect of prematurity/LBW on general intelligence and may provide more insight into the observation that children born preterm/LBW continue to display deficits in academic performance while obtaining normal FSIQ. To understand causal relations in areas of cognitive abilities and provide empirically based suggestions for specific intervention recommendations, we need to incorporate theory and move beyond purely descriptive use of the FSIQ score.

Some researchers have also attempted to explain individual differences in G_f through basic information processing constructs, such as working memory capacity, inhibitory control and speed of processing (A. R. A. Conway, Cowan, Bunting, Theriault, & Minkoff, 2002). These processing parameters are thought to constrain a person's ability to perform complex thinking. Additionally, there has been a long-standing parallel body of research that focuses on the development of G_f and its relationship with the aforementioned information processing constructs (Friedman et al., 2006). Most studies of individuals born preterm/LBW reveal impairments in such information processing constructs, particularly executive function (Aarnoudse-Moens, Smidts, et al., 2009; Luu, Ment, Allan, Schneider, & Vohr, 2011; Nosarti et al., 2007) and working memory (I. S. Baron, Erickson, Ahronovich, Litman, & Brandt, 2010; Luu, et al., 2011).

Executive function (EF) refers to high-order cognitive processing that involves self-regulation of reasoning and planning towards goal-directed behaviour (P. J. Anderson, Howard, & Doyle, 2010; Clark, Pritchard, & Woodward, 2010; Miyake et al., 2000). Working memory (WM) is defined as the mental capacity for temporary storage and active manipulation of information used in a variety of day-to-day activities

(Baddeley & Hitch, 1974; Kane, Conway, Bleckley, & Engle, 2001). However, no investigation has been reported yet on the effects of birth status on G_f and whether it may be attributable to impairments in these information processing parameters.

Researchers investigating individual differences in the normal population have debated the direction of causality between G_f and information processing parameters for a long time (Engel de Abreu, Conway, & Gathercole, 2010; Heitz, Unsworth, & Engle, 2005; Yuan, Steedle, Shavelson, Alonzo, & Oppezzo, 2006). Although controversial, some recent evidence showed that computerized WM training results in increased performance on measures of G_f , which implies a theoretical causal direction, while also suggesting the modifiability of G_f .

Currently, most computerized WM training studies have been directed to individuals with Attention Deficit/Hyperactivity Disorder (ADHD) and university volunteers. Cogmed and dual n -back training programs have captured the most attention and working memory span training studies have just begun to show some success. It is suggested that these interventions may improve trained WM and possibly generalize to non-trained areas of WM and even G_f (Buschkuhl & Jaeggi, 2010; Jaeggi, et al., 2008; Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2009). This evidence represents pioneering success in cognitive intervention. Consequently, success in WM-focused training is potentially much stronger evidence that WM plays a causal role in G_f as it is based on experimental design rather than correlational design. However, current evidence continues to show inconsistent results. These results will be discussed in later chapters. In particular, more studies are required to clarify the utility of WM span training, its effectiveness in typically developing children, as well as its potential for clinical groups such as those born prematurely.

Aims of the Current Studies

The present studies are designed to help further understanding of the underlying mechanisms of the impairments found in preterm/LBW children and test possible casual links between cognitive abilities. These studies go beyond the traditional focus on FSIQ, and instead, seek to assess differences using the theoretically-based construct, G_f . Given strong associations respectively from FSIQ and G_f towards general intelligence, one would expect children born preterm/LBW also to perform less well on measures of G_f , however, no published studies have yet confirmed these differences. Additionally, no studies have investigated the neuropsychological underpinning of g amongst preterm/LBW, using G_f . Thus, the aim of the first study is to test whether any differences in G_f between children born preterm and typically developing children are accounted for by differences in information processing parameters, specifically WM and cognitive flexibility. A mediation modelling approach will be applied to allow the investigation of both mediators concurrently, which has rarely been documented in past research into prematurity and cognition. If WM does indeed serve as a mediator, it would make logical sense that a WM-focused intervention may narrow the gap in G_f between the preterm/LBW population and typically developing children. Even if WM does not account for the between-group difference, WM *training* may still serve as a useful intervention for enhancing G_f . This leads to study two, in which the efficacy of intensive adaptive WM span training in increasing G_f will be tested in typically developing children, with a view to providing preliminary support for the development of an intervention suitable for children born preterm/LBW.

Together, the two studies seek to make both a theoretical contribution to the understanding of the deficits observed in children born preterm/LBW and an applied contribution to clinical practice, by starting the process of developing appropriate

evidence-based interventions. Both of these studies add to the literature that may assist in enhancing cognitive sequelae for preterm birth and possibly ameliorating the practical academic problem found in these children.

Overview of Chapters

A review of the literature on children born preterm/LBW is presented in Chapter 2. In particular, the review will include a comprehensive overview of definitions, incidence and impact of preterm birth/LBW, and academic performance outcomes for such children.

Next, the literature on cognitive outcomes for children born preterm/LBW will be reviewed in Chapter 3. A thorough overview of the preterm/LBW population's cognitive performance in IQ tests and domain-specific abilities will be given.

Following this, the effects of birth status differences in information processing tasks relevant to EF will be described. It will be shown that, although adverse outcomes are observed in academic performance, children born preterm/LBW often fall within the normal IQ range as their same age peers and, therefore, it will be argued that further investigations through the lens of G_f would be useful in understanding the nature of the deficit experienced.

In Chapter 4, Study 1 will be described, which aims to test whether any difference between children born preterm and typically developing children in G_f is accounted for by individual differences in working memory and cognitive flexibility, using a mediation modelling approach.

The literature on computerized working memory training is reviewed in Chapter 5. Recent intervention studies are described, along with descriptions of their variations and treatment outcomes. Differences amongst the training regimes are discussed with concluding comments on the qualities of an ideal working memory training study. It

will be argued that, amongst other training regimes, complex working memory span training involves theoretical and practical components suitable for further investigation with typically developing children and clinical cohorts such as children born preterm/LBW.

Chapter 6 presents Study 2, an evaluation of an intensive computerized WM intervention with typically developing children aged between 7 and 9 years. Participants are assigned to three groups and their performance on working memory, executive attention, speed of processing and G_f are measured and compared following training. The utility of the intervention and the importance of the adaptive nature in the training procedure are discussed.

Finally, Chapter 7 concludes the thesis with an integrated discussion of the findings, clinical implications, and future directions in research for children born preterm/LBW.

CHAPTER 2

Literature Review: Children Born Preterm/Low Birth Weight

Introduction

Previous research has documented less favourable outcomes for children born preterm/LBW than their full-term peers (Aarnoudse-Moens, Weisglas-Kuperus, Van Goudoever, & Oosterlaan, 2009; Bhutta, et al., 2002; Doyle & Anderson, 2010; E. A. Hutchinson et al., 2013; Johnson, et al., 2011). Recent statistics show that the incidence of preterm/LBW status is on the rise due to improvements in modern medical technology, giving children born extremely preterm and extremely low birth weight a higher chance of survival (Wyatt, 2010). However, the differences between this clinical cohort and their peers in terms of cognitive sequelae do not appear to have changed for the better. This chapter will present a literature review to establish the definitions of the basic terminology associated with preterm birth/LBW, current knowledge concerning its incidence, and commonly associated neurological sequelae. Academic outcomes of those born preterm/LBW will also be discussed.

Defining Prematurity

Prematurity can be defined in terms of gestational age (GA) or birth weight (BW). Gestational age can be categorized as preterm, term, or post-term. Preterm birth is defined by the World Health Organization (Blencowe et al., 2012; 2014), as being born at less than 37 weeks of gestation. Those individuals born within 33-36 GA are considered moderately preterm (MPT), those less than 32 weeks GA are considered very preterm (VPT) and those less than 28 weeks are considered extremely preterm (EPT; (Lawn et al., 2010). "Full-term" is defined as being born between 37-41 weeks of gestation, while being born at or beyond 42 weeks of gestation defines post-term birth (Johansson & Cnattingius, 2010). When defined using BW, the World Health

Organization defines low birth weight (LBW) as referring to those born at less than 2,500 g. It can be further categorized as very low birth weight (VLBW) for babies weighing less than 1,500 g and as extremely low birth weight (ELBW) for babies weighing less than 1,000 g (United Nations Children's Fund & World Health Organization, 2004).

When studying prematurity, the inclusion of both BW and GA is important. This is because there are two groups of newborns that are not adequately accounted for when BW is used as the sole criterion for risk. First is the subgroup of LBW infants known as “small for gestational age” (SGA), which means that, although they may be either preterm or born at term, they have a lower than expected gender-specific BW. Their BW is estimated to be below the 10th percentile for its appropriate GA. Another group are affected by “intrauterine growth restriction” (IUGR), frequently documented by a declining rate of fetal growth detected by prenatal ultrasound over the pregnancy period. IUGR is considered as a subset of SGA. The further the BW is below the appropriate GA percentile of a SGA infant, the higher likelihood of the infant being considered IUGR. That said, some IUGR infants may be born at full-term with a LBW, yet their BW may not be low enough to meet criteria for SGA. Conversely, not all SGA infants are small as a result of not reaching their growth potential, and, therefore, would not meet criteria for IUGR. They can be small but otherwise healthy. An infant’s head circumference, aside from birth weight and birth length, has also been often used as an indicator of severity in growth restriction (Campbell et al., 2012; Frisk, Amsel, & Whyte, 2002; Johansson & Cnattingius, 2010; Martin & Dombrowski, 2008; Saleem et al., 2011; Wollmann, 2009). Therefore, sample characteristics on both GA and BW are provided for the participants in Study 1 of Chapter 4.

Children born prematurely can also be categorized according to one of three

typical clinical conditions, namely, spontaneous preterm labour, preterm premature rupture of the membranes (PPROM), and medically induced preterm labour.

Spontaneous preterm birth refers to natural premature contractions before the fetus reaches full-term. This is the most common type of preterm birth and is often associated with maternal factors such as previous preterm deliveries, stress, smoking, substance abuse, as well as low body mass or weight gain during pregnancy. PPRM refers to labour that is initiated by the rupture of membranes. It is often associated with, but not limited to, inflammations, uterine distension, and cervical anomalies. Finally, medically induced preterm labour refers to birth deliveries through Caesarian section or labour induction initiated by the physician to reduce health risk to the fetus or mother. This procedure is often associated with fetal distress, intrauterine growth restriction and maternal complications, such as antepartum bleeding and preeclampsia (Behrman & Butler, 2007; Johansson & Cnattingius, 2010; Moutquin, 2003; Schetter & Glynn, 2011). Although these are informative categorical details, their differentiating effects on preterm/LBW individuals' cognitive outcomes is not the focus of this research and they are, therefore, not included in the descriptions of later studies.

Incidence of Preterm Birth

According to recent statistics, global incidence of preterm birth continues to rise over the years. There was approximately 13 million preterm born babies in 2005, with an approximate 9.6% preterm birth rate (Beck et al., 2010). In 2010, there was approximately 14.9 million preterm born babies, with a preterm birth rate of approximately 11.1% (Blencowe, et al., 2012).

More-developed countries documented 7.5% preterm birth rate as compared to the 12.5% for less-developed countries. A large burden of these figures was accounted for by Africa (31%) and Asia (54%), which documented 85% of the overall preterm

births (11 million). The preterm birth rate for Africa and Asia was documented at 11.9% and 9.1% respectively. North America also had a very high preterm birth rate of 0.5 million, which translates to 10.6% of all births. There was an estimated 0.9 million preterm births in Latin America and the Caribbean and their preterm birth rates were reported at 8.1%. Europe had the lowest preterm birth rate documented at 6.2%, which is approximately 0.5 million preterm births (Beck, et al., 2010).

Comparisons amongst developed countries, such as Australia, Canada, Japan, New Zealand, the United Kingdom and United States of America, on preterm birth rate yield similar figures. These countries also showed a continuous rising trend in preterm birth over the past three decades, with a current rate ranging from 5% to 9% (Lawn, et al., 2010). In line with the trends of these countries, perinatal statistics in Western Australia showed that out of 30,670 births in 2008, 6.7% were of prematurely born (Le & Tran, 2008), followed by more recent statistics showing an increase to an approximate 8.7% (2,708 prematurely born out of 31,264 newborns) in Western Australia (J. A. Hutchinson, 2012).

Commonly Associated Neurological Disorders in Preterm Birth/LBW

Individuals born preterm are vulnerable to complications at birth such as intraventricular haemorrhage (IVH; (Adams-Chapman, 2009; Wyatt, 2010) and white matter injury (Khwaja & Volpe, 2008; Wyatt, 2010), as well as neurological pathology. Two of the most commonly found neurological disorders in preterm/LBW cohorts are Cerebral Palsy (CP; (Krageloh-Mann, 2010; Platt et al., 2007) and Attention Deficit/Hyperactivity Disorder (ADHD; (Amor, Chantal, & Bairam, 2012; Lindstrom, Lindblad, & Hjern, 2011).

CP refers to permanent motor deficiency that causes limitations in developing movements and postures. It stems from non-progressive damage to the developing fetal

or infant brain. Often, this motor disorder spectrum also affects cognitive abilities, perception and sensation, through occurrences of epilepsy (Rosenbaum et al., 2007). The most recent prevalence estimates for CP was 2.11 per 1000 live births (95% CI 1.98-2.25);(Oskoui, Coutinho, Dykeman, Jette, & Pringsheim, 2013). Recent review of 49 studies suggested that the prevalence of CP increased as GA and BW decreased. In particular, low GA individuals were at increased risk for developing CP. Those born EPT showed the highest prevalence (111.80 per 1000 live births; 95% CI 69.53-179.78) whereas those with appropriate GA of more than 36 weeks were of lowest prevalence (1.35 per 1000 live births; 95% CI 1.15-1.59). Patterns were similar when pooled prevalence was reported according to BW. VLBW individuals showed highest prevalence (59.18 per 1000 live births; 95% CI 43.38-73.95) and those with normal BW above 2500 g had the lowest prevalence of CP (1.33 per 1000 live births; 95% CI 1.19-1.49). The prevalence amongst ELBW did not differ significantly from the VLBW group (Oskoui, et al., 2013).

ADHD is a neurodevelopmental and psychiatric disorder that, according to DSM IV-TR, can be categorised into three different subtypes: ADHD/C (Attention-Deficit/Hyperactivity Disorder, Combined Type), ADHD/H (Attention-Deficit/Hyperactivity Disorder, Predominantly Hyperactive-Impulsive Type) and ADHD/I (Attention-Deficit/Hyperactivity Disorder, Predominantly Inattentive Type; (American Psychiatric Association, 2000). Studies on neurodevelopmental trajectories suggest that ADHD stems from a delay in cortical maturation, particularly in the lateral prefrontal cortex (Shaw, Gogtay, & Rapoport, 2010; Shaw et al., 2013). According to earlier reported reviews on children born preterm (Bhutta, et al., 2002) and a recent large-scale study of the Swedish preterm cohort (Lindstrom, et al., 2011), EPT children were at double the risk of full-term school-aged children to be diagnosed with ADHD.

According to Bhutta et al. (2002), being born at lower GA was associated with increased risk of ADHD, reported at a pooled risk ratio (RR) of 2.64 (95% CI 1.85-3.78) compared to controls. A further breakdown in the degree of birth immaturity demonstrated a GA-related gradient. EPT children were at increased risk with an odds ratio (OR) of 2.1 (95% CI 1.4-2.7), while the OR for VPT was 1.6 (95% CI 1.4-1.7). MPT between 33 to 34 weeks and MPT between 35 to 36 weeks were reported with ORs at 1.4 (95% CI 1.2-1.7) and 1.3 (95% CI 1.1-1.4) respectively (Lindstrom, et al., 2011). Recent data gathered from a National Health Interview Survey between 1997 and 2005 in the United States provided similar findings regarding children born with LBW. Prevalence rate was reported highest in children born ELBW (Adjusted odds ratio [AOR] at 2.0; 95% CI 1.3-2.8), then VLBW (AOR 1.8; 95% CI 1.4-2.2) and LBW (AOR 1.4; 95% CI 1.2-1.6). Adjustments included factors such as race, household income, maternal education and other disorders (Boulet, Schieve, & Boyle, 2011).

In addition, studies also suggest that children born preterm/LBW more often fit criteria for ADHD/I rather than ADHD/H (Elgen, Sommerfelt, & Markestad, 2002; B. Hayes & Sharif, 2009). For example, in a recent South German study, VP/VLBW ($n = 281$) were reported to be at increased risk for ADHD/I but not for ADHD/H as compared with full-term controls ($n = 286$). Specially, ORs for ADHD/I were documented at 2.8 ($p < .001$) and 1.7 ($p = .02$) for children at 6 years 3 months and 8 years 5 months respectively. This was in contrast to ORs for ADHD/H which were documented at 1.4 ($p = .396$) and 0.9 ($p = .82$) for children at 6 years 3 months and 8 years 5 months, respectively (Jaekel, Wolke, & Bartmann, 2013).

Prematurity and Academic Performance

Outcomes of Standardized Tests and Teacher Ratings

Academic struggles become noticeable when children born preterm/LBW enter

the school system. According to standardized achievement tests, as well as teacher and parent ratings, these children perform less well than their peers across various academic subjects. In particular, math-related (Bowen, et al., 2002; Pritchard, et al., 2009) and reading-related (Bowen, et al., 2002; Lee, et al., 2011; Roberts, et al., 2011) activities have been found to be profoundly affected. A thorough meta-analysis of 14 studies that compared participants born VPT/VLBW to full-term peers, with ages ranging from five to 18 years, provided combined effect sizes for math at -0.60 and for reading at -0.48 (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). Later studies provided similar evidence with comparable effect sizes. For example, Pritchard et al. (2009) compared children born VPT (≤ 33 weeks GA, $n = 102$) with their full-term peers ($n = 108$) on the Woodcock Johnson Tests of Achievement (WJ-III) subtests at six years of age. Findings revealed that the performance of children born VPT was inferior to their term peers for math fluency ($p < .001$, $d = -0.62$). Consistent with this, significant differences have also been found between children born EPT (< 26 GA, $n = 219$) and their classmates ($n = 153$) at 11 years of age, matched for gender and ethnicity, when assessed using the Wechsler Individual Achievement Test – II (WIAT-II). The children in the clinical group were outperformed by their peers with a 14.8-point difference ($p < .0001$) at a large effect size ($d = -1.0$) for math skills and a 7.7-point difference ($p < .0001$) at a moderate effect size ($d = -0.6$) for reading (Johnson, et al., 2011).

Similar findings have been indicated by studies using ratings from teachers and parents. In a large scale study in Denmark, teachers and parents rated the school performance of children born preterm/LBW as compared to that of their peers at 10 years of age ($n = 5,319$). With all results adjusted for gender, breastfeeding status and parental education level, findings indicated that children with LBW ($< 2,500$ g) showed significantly higher risk of difficulty in both math (AOR 4.46, 95% CI 1.41-15.00) and

reading (AOR 1.85, 95% CI 0.81-4.22) than those with BW between 3,500 g and 3,999 g. However, those with GA in the 33-36 week range (MPT) performed as well as their peers in math, which was inconsistent with other documented ratings in the literature (Kirkegaard, Obel, Hedegaard, & Henriksen, 2006).

Further GA categorization in studies using teacher ratings has suggested difficulties in overall academic achievement among children born preterm. For example, with teachers blinded to students' birth status, research has indicated that children born VPT, at six years of age, were rated as below average or delayed in academic performance two to three times more often than their peers. The most concerning area rated as below average or academically delayed by teachers was math (VPT: 44%, $n = 41$ vs. FT: 14%, $n = 14$). Significant differences from full-term children were discovered in subject areas including math ($p < .001$, $d = -0.67$), reading ($p < .001$, $d = -0.50$), written language ($p < .001$, $d = -0.58$), language comprehension ($p < .01$, $d = -0.43$), handwriting ($p < .001$, $d = -0.61$), spelling ($p < .001$, $d = -0.62$), and physical education ($p < .001$, $d = -0.67$), but not expressive language (Pritchard, et al., 2009).

Similar findings have been reported by Mulder, Pitchford, and Marlow et al. (2010), who compared teacher ratings of 9- to 10-year-old children born VPT (< 31 weeks GA, $n = 48$) to those of matched term controls ($n = 17$). VPT children were rated significantly below average more frequently than their peers, not only for total academic achievement (OR 11.9, 95% CI 1.4-96.9) but also on math (OR 6.5, 95% CI 1.7-25.8) and English (OR 3.8, 95% CI 1.1-13.5).

It is worth noting that, although standardized tests would likely provide less biased reflections, teachers routinely grade their students and thus their ratings are valuable reflections of day-to-day observations. However, these observations may also be affected by general classroom behaviour that may create a negative halo effect, thus

resulting in lower ratings. Even taking these factors into account, studies clearly illustrate that, whether academic abilities are measured by standardized achievement tests or teacher/parent ratings, children born preterm/LBW are at increased risk of academic difficulties compared to their peers.

Special Educational Needs and Academic Attainment

Need for special education and grade retention have also been documented as more prevalent in children born preterm/LBW than their same-age peers. For example, Mulder et al. (2010) reported that teachers rated their VPT group as significantly more likely to need special education (OR 7.2, 95% CI 1.5-35), including one-on-one assistance or small group learning, as well as help from professional care such as an educational psychologist, clinical psychologist or speech pathologist. Interestingly, this does not only appear in VPT children: a study comparing MPT and full-term children at eight years of age also found similar outcomes. The study showed that the group of MPT children were twice as likely to require special education as the general Dutch population. As well, amongst those within mainstream schooling, the rate of grade retention doubled in the MPT group as compared to the term group ($p < .01$; (Van Baar, et al., 2009). This study further demonstrated that, although MPT children appear to be at a lower risk of adverse health outcomes than VPT or EPT, they still require additional educational support and are still highly vulnerable to academic delays.

More special educational needs are also apparent in adults born preterm/LBW. In line with studies on school-aged children born preterm/LBW, studies of adults indicate differences in educational profile. A recent Norwegian study showed that 22% of 19-year-old adults born VLBW ($n = 55$) reported having had special education assistance in the past as compared to 2.5% in the comparison group ($n = 81$; OR 11.02, 95% CI 2.4-51.5, $p < 0.001$; (Lohaugen, et al., 2010).

Furthermore, relatively low academic attainment has been reported in adults born preterm. According to a nation-wide study of 27- to 29-year-old adults from Denmark, educational outcomes were lower in the VPT group ($n = 1,422$) than the full-term group ($n = 192,223$). The VPT group was observed to have completed less tertiary education (OR 0.77, 95% CI 0.69-0.89, $p < 0.001$) than the full-term group. Evidence also showed that the VPT group earned 11.7% ($p < 0.001$) less income, as well as depending more often on social pension (OR 3.27, 95% CI 2.58-4.13, $p < 0.001$) and welfare support (OR 2.14, 95% CI 1.81-2.55, $p < 0.001$) than the full-term group (Mathiasen, et al., 2009).

It is important to point out that significant results reported by the aforementioned studies remained unchanged even after accounting for potential confounding factors. The findings remained after controlling for socio-economic status (SES; (Lohaugen, et al., 2010; Pritchard, et al., 2009) and maternal education (Lohaugen, et al., 2010; Mathiasen, et al., 2009), as well as excluding children with neurocognitive impairments (Johnson, et al., 2011; Pritchard, et al., 2009).

Summary

In summary, with the increasing trend in incidence of preterm/LBW birth and high prevalence of neurodevelopmental disorders such as ADHD and CP associated with this status, it is not surprising that the impact of preterm birth/LBW on later life development has captured a vast interest in the literature. In particular, evidence of impairments in academic performance associated with preterm/LBW have been found in numerous areas of academic skills, particularly in math-related and reading-related activities. These impairments have been shown to persist from early school age to adulthood. These educational outcomes can also have long-lasting effects on an individual's earnings and social welfare dependency (Doyle & Anderson, 2010;

Lohaugen, et al., 2010; Mathiasen, et al., 2009). Chapter 3 will review the extensive research that has been conducted on the possible cognitive causes of these educational outcomes.

CHAPTER 3

Literature Review: Cognitive Outcomes of Preterm Birth

Introduction

Healthcare professionals are not only interested in improving survival rates and neonatal care for children born preterm/LBW but also their long-term developmental outcomes (Ruegger, Hegglin, Adams, & Bucher, 2012; Vohr, 2010). Cognitive functioning is one of the most frequently studied topics amongst children born preterm/LBW. This is because cognitive abilities are highly associated with academic performance ($r = 0.81$; (Deary, et al., 2007), and to a lesser but not insignificant extent with other life outcomes including job status, health and quality of life (Deary & Batty, 2007; Mõttus, Gale, Starr, & Deary, 2012).

In the previous chapter, it was established that children born preterm/LBW, on average, perform less well than their full-term peers in terms of academic outcomes, and these outcomes are likely to persist into adolescence and adulthood. In this chapter, studies comparing performance of children born preterm/LBW and their full-term peers on cognitive abilities will be reviewed. In particular, the focus will be on the seemingly conflicting findings on normal range IQ, observed in a majority of yet not all preterm literature, but poor academic achievement (Bhutta, et al., 2002; Kerr-Wilson, et al., 2011). In an attempt to explain this, along with the majority of past research focusing on traditional IQ, G_f and information processing parameters will be introduced as a potential avenue for theoretically based investigation. The investigation through G_f is also relatively novel in the current literature and therefore worthy of exploration. Given the extensive amount of detail presented in this chapter, Figure 3.1 will assist in understanding the flow of information.

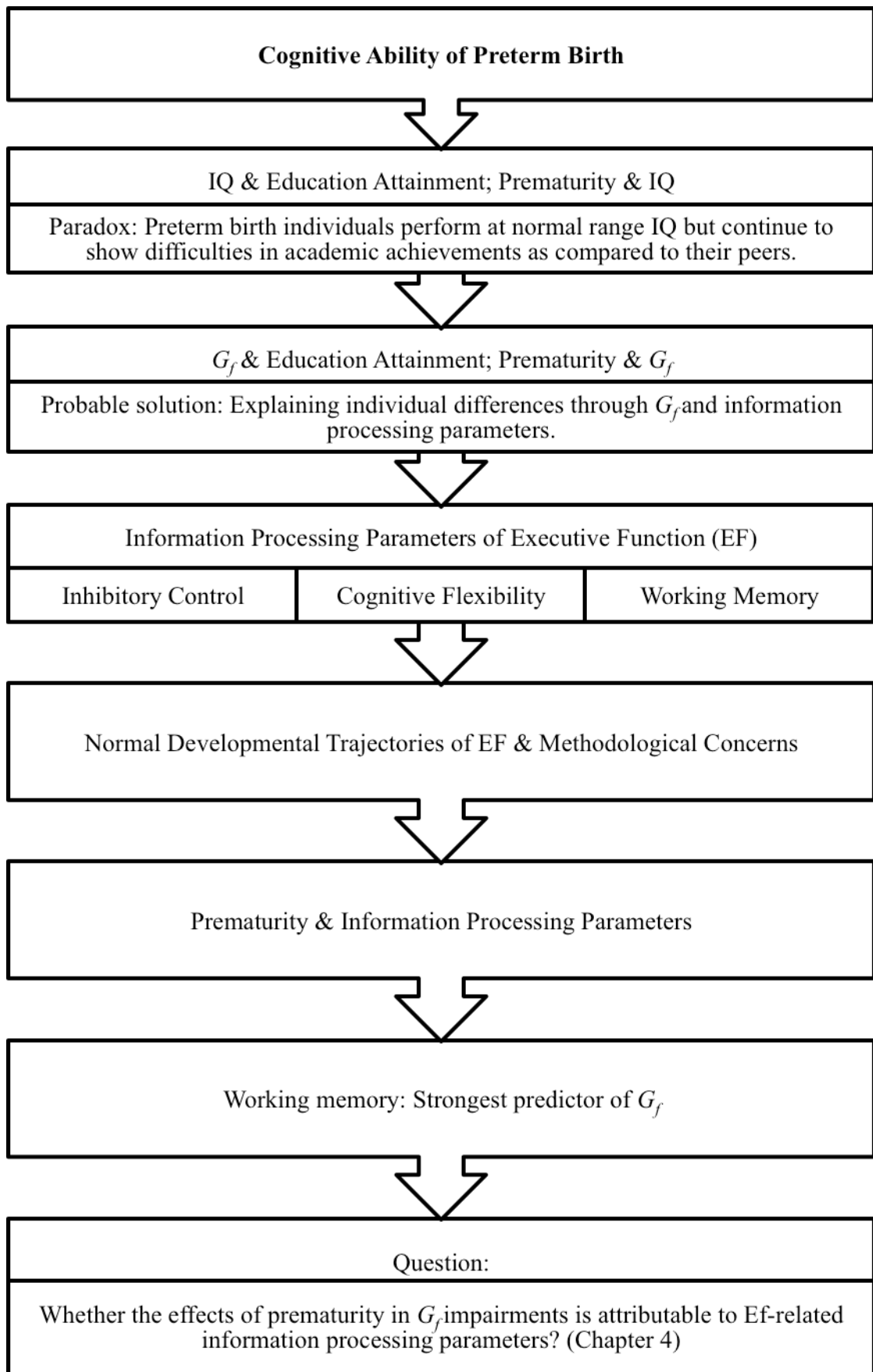


Figure 3.1. The flow chart presents the course of information presented in Chapter 3.

Intelligence Quotient (IQ) and Education Attainment

According to Spearman's (1904) General Cognitive Ability (*g* factor) theory, individual differences in intellectual ability are largely captured by a single, general construct, referred to as *g*. Those seeking to measure this construct generally define *g* as the common variance shared by cognitive ability measures drawn from different content domains (e.g., verbal, spatial, numerical etc.).

IQ tests developed to measure intelligence typically derive a general IQ score from performance on a battery of different cognitive ability measures. Binet and Simon (1916) originally designed such a test with the aim of predicting a person's academic performance (Gottfredson, 1997; Jensen, 1998). The predictive value of psychometric intelligence estimates for individual differences in academic achievement has been extensively researched. It is evident that higher IQ is generally associated with better academic achievement (Deary, et al., 2007; Glutting, Youngstrom, Ward, Ward, & Hale, 1997; Wechsler, 1991). For example, Wechsler (1991) reported high correlations, that ranged from .70 to .81, between its Full Scale IQ score on WISC-III and an achievement composite from the Wechsler Individual Achievement Test (Wechsler, 1992). In corroboration, Deary and his colleagues (2007) analysed data of over 70,000 children from United Kingdom, at age 11, on the Cognitive Abilities Test (CAT) and found that participants' performance positively correlated with all 25 subject scores measured in the national examinations General Certificate of Secondary Education Examination (GCSE) as tested at 16 years old. Effect sizes were noted as ranging from medium to large. In particular, the following categories received the highest correlations, with the CAT's overall intelligence score correlating with overall GCSE point score ($r = .69, p < .001$), GCSE best 8 subject scores ($r = .72, p < .001$), individual's English score ($r = .67, p < .001$), and individual's Math score ($r = .77, p <$

.001).

The use of recognized standardized psychometric instruments and large samples in the aforementioned studies provide strong evidence for the positive correlation between intelligence and later education achievement. Given the strong links between intelligence and academic achievement in the general population, a discussion of IQ test performance in the preterm population is warranted.

Prematurity and IQ

Children born preterm/LBW have been documented to perform less well on IQ tests. Bhutta and colleagues (2002) performed a meta-analysis of 15 studies, with findings involving 1,556 children born preterm and 1,720 controls tested across the age range of five to 14 years. The cognitive scores in their studies were determined from a variety of tests that all used the same mean (100) and standard deviation (15).

Particularly, the WISC was most frequently used, while other assessments included the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), the Kaufmann Assessment Battery of Childhood (KABC), and the British Abilities Scale (BAS). The authors documented in their findings that preterm birth was associated with a 10.9-point difference in intelligence scores (95% CI 9.2-12.5). They also reported that there was a dose-relationship where lower IQ scores were correlated with lower birth weight ($R^2 = 0.51, p < .001$) and gestational weeks at birth ($R^2 = 0.49, p < .001$). However, only two of the studies reported an average IQ of 1 SD below the mean for their preterm birth groups, while all others reported means within the normal range. The mean scores for controls in the meta-analysis were also noted as often scoring above 100. There may be an issue of sampling with lower-scoring full-term children not being well represented.

Although there have been continuous improvements in medical practice and prenatal care over the past four decades, this did not appear to have changed the

neurodevelopmental outcomes of preterm birth/LBW population (I. S. Baron & Rey-Casserly, 2010). Kerr-Wilson, Mackay, Smith and Pell (2011) published an updated meta-analysis reviewing 27 studies that included 3,504 preterm and 3,540 term individuals tested between three and 16 years of age. The authors found consistent results nearly a decade after Bhutta and colleagues' (2002) meta-analysis. The updated analysis included standardized psychometric assessments such as the KABC, the BAS, the McCarthy scale and various versions of the Wechsler IQ scales. Results indicated that there remained an overall IQ score difference of 11.94 points (95% CI 10.47-13.42, $p < 0.001$) favouring those born at term. A significant dose-relationship was found across the GA range (adjusted coefficient: -0.91, 95% CI -1.64, -0.17, $p = 0.018$) where lower GA was associated with lower IQ score. Specifically, it was also indicated that at < 28 weeks GA, a mean difference of 13.9 points was noted (95% CI 11.5-16.2, $p = 0.001$), while an 11.4 point mean difference was found in those between 28-31 weeks GA (95% CI 9.7-13.2, $p = 0.022$) and finally an 8.4 point mean difference for those preterm birth individuals born at ≥ 32 weeks GA (95% CI 6.6-10.2, $p = 0.314$; (Kerr-Wilson, et al., 2011). This recent meta-analysis confirmed not only at least an 8.4-point difference in IQ scores favouring full-term children over children born preterm, but also that a strong dose-response relationship exists between prematurity and IQ. It was documented that IQ dropped steadily with the reduction of each gestation week. Similar to the previous review, a majority of the studies reported preterm groups to have mean IQ scores within the average range, and only four out the 27 the studies found the preterm birth group to have a mean IQ score of 1 SD below the mean. The control groups were better represented in this meta-analysis than Bhutta et al.'s (2002) as the mean scores for the control groups were more evenly distributed.

A strength of the two meta-analyses was that they both excluded studies that

categorized children's birth status according to BW only and also focused on GA as an inclusion criterion due to the possible inclusion of children born full-term but small for gestational age (SGA). A shortcoming for both meta-analyses was that the authors did not report on whether any particular demographic variables played a significant role in the overall differences between children born preterm and their controls.

Nevertheless, more specific observations demonstrate that children born preterm/LBW achieve lower scores on IQ assessments than their peers, even though excluding neurologically impaired participants and adjustments on SES reduced the group differences. For example, in an Australian study that investigated the WISC-III scores of 8-year-old school-aged children born VPT and ELBW ($n = 298$) and their peers weighing more than 2,499 g ($n = 262$), a 9.4-point (95% CI 12.1 to 6.7) difference was found in FSIQ where the control group scored higher than the preterm/LBW group. After excluding children with neurosensory impairments from the analysis, there remained an 8.8 point difference (95% CI 11.6 to 6.1), and a 7.6 point (95% CI 10.3 to 4.9) difference after SES adjustments (P. J. Anderson, Doyle, & the Victorian Infant Collaborative Study Group, 2003).

Similar results were reported in a more recent study with 8-year-old children born EP/ELBW ($n = 189$) and their term born/normal birth weight controls ($n = 173$). Tested on WISC-IV, children in the control group scored higher than the EP/ELBW group with a mean difference of 12.5 points (95% CI 15.5 to 9.5). After SES adjustments, the difference was reduced but continued to be significant (10.2 points, 95% CI 1.7 to 6.6). After SES adjustments and excluding children with neurosensory impairments, the difference was further reduced, yet still significant (8.8 points, 95% CI 12.2 to 5.3; (E. A. Hutchinson, et al., 2013). Accordingly, there is an IQ score difference of approximately half a SD attributable to birth status and not neurosensory

impairments or low SES.

Do Cognitive Differences Measured by IQ Tests Persist into Adulthood?

Studies not only report lower IQ performance for this cohort of children, but cross-sectional data suggest that these deficits persist into adulthood (Allin et al., 2008; Hallin, Hellstrom-Westas, & Stjernqvist, 2010; Pyhala et al., 2011). For instance, a recent study compared 18-year-old participants born EPT/ELBW ($n = 52$), with an average GA of 27 weeks ($SD = 1.0$) and average BW of 1002 g ($SD = 234$ g), to their full term peers ($n = 54$), with an average GA of 40 weeks ($SD = 1.5$) and average BW of 3612 g ($SD = 525$ g). Participants were matched according to gender, age and residential location. Results indicated significant differences between the two groups on measures of IQ, using the WAIS-III, even after adjusting parental educational level (EPT/ELBW FSIQ: $M = 92.8$, $SD = 15.4$ vs Controls FSIQ: $M = 105.7$, $SD = 12.5$, $p < 0.001$) with 95% confidence interval (CI) of the mean difference ranging from -18.27 to -7.45 (Hallin, et al., 2010). The authors did not provide statistical results before SES adjustments were made but did note that the outcome of results were not different. It is also noteworthy that despite differences between the two groups, the mean FSIQ of adults born EPT/ELBW remained within average ranges, similar to outcomes in studies with children. Added to that, the controls are also on the higher side of average.

Pyhala et al. (2011) reported similar persistence when comparing three groups of 21- to 29-year-old young adults (age $M = 25$, $SD = 2.1$), excluding those with neurocognitive impairments, on WAIS-III performance. The first group consisted of VLBW young adults born at term, (Appropriate gestational age [AGA]; <1500 g, $n = 66$), the second group consisted of VLBW young adults born SGA (<1500 g, BW for $GA \leq -2$ SD, $n = 37$), and a final full-term control group (≥ 37 GA, $n = 105$). Results for all IQ indices for all three groups were within normal range. When comparing the whole

VLBW group with the control group, there were significant group differences favouring the control group (FSIQ: OR -0.57, 95% CI -0.83 to -0.31, $p > .001$) with results adjusted for gender and age at the time of testing, as well as after a second adjustment that included gender, age, parental education and head circumference at birth and adulthood (FSIQ: OR -0.38, 95% CI -0.68 to -0.09, $p = .01$). As indicated earlier, confounding variables such as residential location and parental education were measures used to adjust for SES, whereas head circumference was relevant to growth restrictions. Adjustment to these factors narrowed the gap between the clinical cohort and full-term, however, they were not strong enough to explain the differences between the two groups in adulthood.

Outcomes in Domain-Specific Abilities

Aside from investigating global IQ score, domain-specific abilities have been further explored in the literature. In early versions of the Wechsler Scales, FSIQ comprised Verbal IQ (VIQ) and Performance IQ (PIQ). VIQ assessed mainly verbal reasoning using tasks involving verbal questions and responses, while PIQ assessed non-verbal reasoning through tasks with visual stimuli, motor responses, as well as timed tests. Based on factor-analysis, four domain-specific indices could be derived, two of which (Verbal Comprehension Index [VCI] and Freedom from Distractibility Index [FDI]) constituted VIQ, while the other two (Perceptual Organization Index [POI] and Processing Speed Index [PSI]) were combined as a measure of PIQ. Updated versions of Wechsler Scales, such as WISC-IV (Wechsler, 2003), no longer use VIQ and PIQ but now generate four main indices, namely, VCI, Perceptual Reasoning Index (PRI), Working Memory Index (WMI) and PSI, to better reflect their content as measured by their respective groups of subtests (Wechsler, 1991, 2003). Essentially, POI has been renamed PRI with a greater emphasis on fluid abilities, such as including

the Matrix Reasoning and Picture Concept subtests. FDI has been renamed WMI with an addition of the Letter-Number Sequencing subtest to reflect its measured construct more accurately. A visual comparison of the structural differences between WISC-III and WISC-IV is presented below in Figure 3.2.

The comparison of domain-specific abilities amongst individuals born preterm/LBW has suggested a greater magnitude of deficits in PIQ than VIQ. For instance, 6-year-old children born preterm (< 34 weeks GA, $n = 168$) were matched according to gender and BW z score with their full-term controls ($n = 168$) and compared on performance using Wechsler Intelligence Scale for Children – Revised (WISC-R). The BW z score of within 0.1 SD of their term counterparts was considered an acceptable match. Findings revealed that the clinical cohort displayed significantly higher risk of obtaining scores less than 85 on FSIQ (OR 2.17, 95% CI 1.17-4.03, $p < .05$; aOR 2.35, 95% CI 1.20-4.61, $p < .05$) and PIQ (OR 2.01, 95% CI 1.10-3.68, $p < .05$; aOR 2.04, 95% CI 1.09-3.82, $p < .05$) but not VIQ (OR 1.47, 95% CI 0.80-2.72; aOR 1.52, 95% CI 0.78-2.95). Adjustments on confounding variables here included the child's gender, maternal IQ, marital status and residential setting (Talge et al., 2010).

Similar patterns of results were indicated in Pyhala et al.'s (2011) study of young VLBW adults. They found that there were significant group differences favouring the control group over the VLBW group in specific IQ estimates (PIQ: OR -0.68, 95% CI -0.93 to -0.42, $p < .001$; VIQ: OR -0.29, 95% CI -0.57 to -0.02, $p = .03$), with results adjusted for sex and age at the time of testing. However, these differences persisted only in PIQ after a second adjustment that included sex, age, parental education and head circumference at birth and adulthood: PIQ (OR -0.48, 95% CI -0.76 to -0.19, $p = .001$), but VIQ (OR -0.17, 95% CI -0.48 to 0.15, $p = .30$). This suggested that, although the performance in VIQ was lower in the VLBW adults, it was likely

attributable to a lower level of parental education and growth restriction.

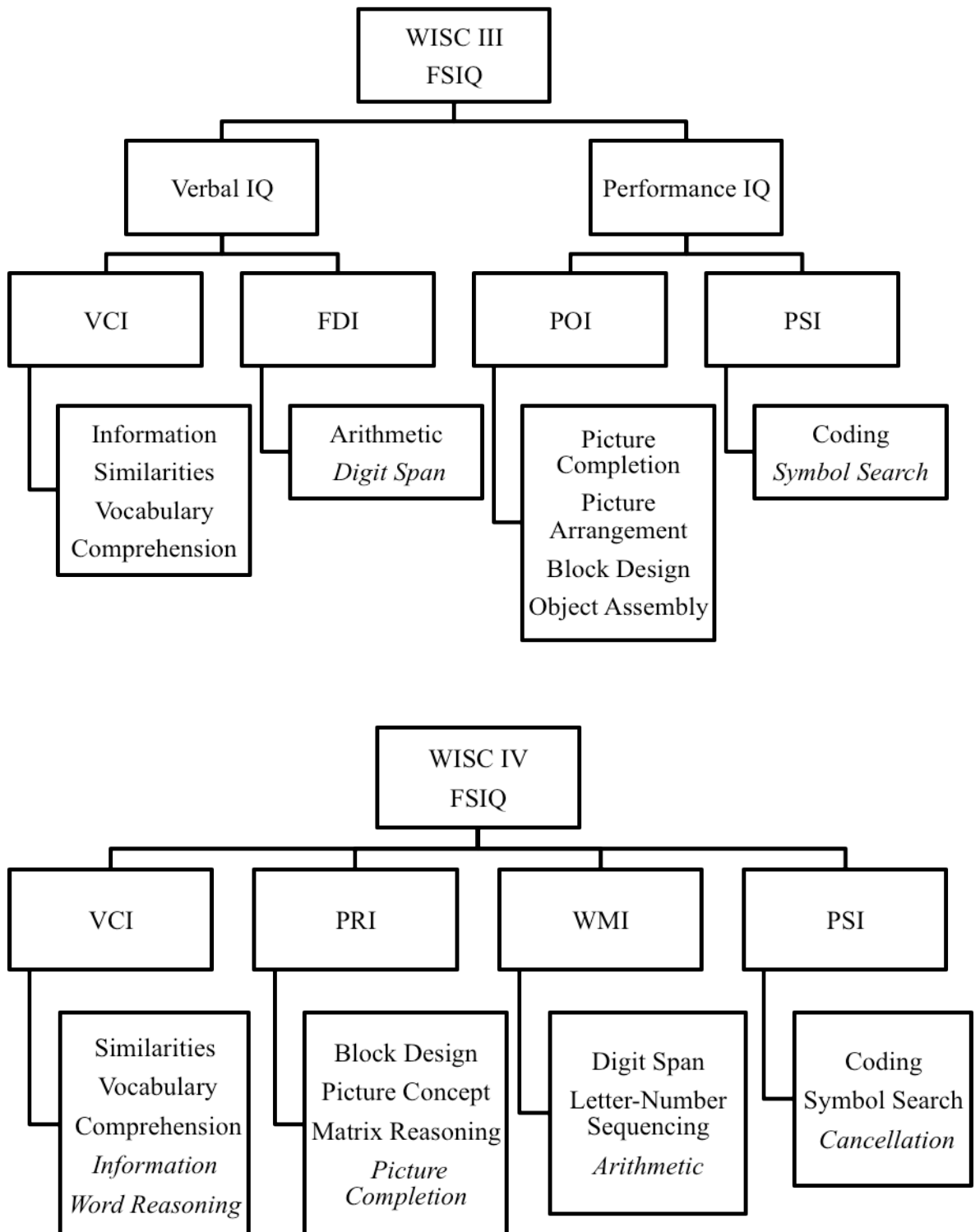


Figure 3.2. The structure of WISC-III and WISC-IV are shown with their respective subtests under each index. Optional subtests are printed in *italics*. 'Mazes' (not shown in the diagram) is another optional subtest under WISC-III PIQ but does not belong to either POI or PSI.

Pyhala et al. (2011) also analysed their results separately for VLBW-AGA (Appropriate for Gestational Age) and VLBW-SGA (Small for Gestational Age), in comparison with the full-term control group. VLBW-AGA performed less well than their control peers for FSIQ and PIQ after the two adjustment analyses respectively: FSIQ (1st adjustment (gender and age): OR -0.50, 95% CI -0.80 to -0.21, $p = .001$; 2nd adjustment (gender, age, parental education and head circumference) OR -0.39, 95% CI -0.71 to -0.07, $p = .02$); PIQ (1st adjustment: OR -0.62, 95% CI -0.91 to -0.33, $p < .001$; 2nd adjustment OR -0.46, 95% CI -0.77 to -0.15, $p = .004$). These findings suggested that adults born at appropriate size for GA continued to show deficiencies in FSIQ and PIQ that were not attributable to a lack of parental education and/or restricted growth. In terms of comparison between the VLBW-SGA and the control group, significant group differences favourable to the control group were found, after both adjustment analyses, only in PIQ (1st adjustment: OR -0.79, 95% CI -1.14 to -0.43, $p < .001$; 2nd adjustment OR -0.53, 95% CI -0.96 to -0.09, $p = .02$). Significant group differences between VLBW-SGA and their peers in FSIQ (OR -0.69, 95% CI -1.06 to -0.33, $p < .001$) and VIQ (OR -0.40, 95% CI -0.77 to -0.02, $p = .04$) were present after the first adjustment but not after the second. Similarly, VLBW adults born SGA continued to show deficiency in PIQ as with VLBW-AGA. However, VLBW adults born SGA showed deficiency in VIQ or FSIQ that was perhaps attributable to parental education and restricted growth. Nonetheless, overall results here indicated that PIQ appears to be affected in young adults with VLBW-AGA and VLBW-SGA most substantially, reflecting difficulties with skills such as non-verbal problem solving, spatial and sequencing tasks.

Other specific analyses revealed more prominent deficits in POI and FDI/WMI. A study documented that 8-year-old VPT/ELBW, as compared to their full-term

controls, displayed significantly lower scores on all indices ($p < .001$), with the greatest mean difference shown in POI at 9.9 points (95% CI -12.7 to -7.2). This was in comparison to a mean difference of 8.2 points in FDI (95% CI -10.8 to -5.5), 6.8 points in VCI (95% CI -9.5 to -4.2), and 6.7 points in PSI (95% CI -9.4 to -4.0). Within-group statistics also showed that VPT/ELBW participants had their lowest average scores on the FDI at 93.1 (14.7). All results remained significant after excluding participants with neurosensory impairments, such as cerebral palsy, blindness and deafness (POI: 8.8 points, 95% CI -11.6 to -6.1; FDI: 8.1 points, 95% CI -10.7 to -5.4; VCI: 7 points, 95% CI -9.6 to -4.3; PSI: 5.9 points, 95% CI -8.6 to -3.3), and adjusting for SES (POI: 8.1 points, 95% CI -11.1 to -5.1; FDI: 7.2 points, 95% CI -10.1 to -4.3; VCI: 5.9 points, 95% CI -8.4 to -2.8; PSI: 5.7 points, 95% CI -8.7 to -2.8). However, the figures also suggested that neurosensory impairment and SES adjustments partially accounted for differences in the cognitive abilities between the two groups (P. J. Anderson, et al., 2003).

Hallin, Hellstrom-Westas and Stjernqvist (2010) demonstrated similar outcomes on domain-specific abilities in their follow-up study comparing adults born EPT ($n = 52$) and matched full-term adults ($n = 54$). Participants were tested at 18 years of age using WAIS-III. They not only showed significant group differences on FSIQ but also across all the four indices. Of particular interest was that, again, the greatest between-group mean difference was found in POI where the full-term group performed better than the EPT group by an average of 16.2 points. Within-group statistics also showed that the preterm birth group performed particularly poorly, on average, on the WMI, (EPT: $M = 88.3$, $SD = 14.6$ vs FT: $M = 96.3$, $SD = 12.1$, $p = .003$) with a mean difference of 95% CI ranging from -13.21 to -2.84, when compared to their own performance on other indices, namely VCI (EPT: $M = 95.0$, $SD = 13.5$ vs FT: $M =$

104.4, $SD = 12.4$, $p < .001$, 95% CI -14.43 to -4.43), POI (EPT: $M = 96.2$, $SD = 18.1$ vs FT: $M = 112.4$, $SD = 14.5$, $p = .001$, 95% CI -22.53 to -9.81), and PSI (EPT: $M = 92.6$, $SD = 14.2$ vs FT: $M = 100.2$, $SD = 12.1$, $p = .004$, 95% CI -12.69 to -2.48). These results were consistent with Anderson et al.'s study (2003), where between-group comparisons showed that the greatest mean differences occurred in perceptual reasoning tasks, favouring the control group and the greatest within-group differences for the preterm/LBW group were working memory related tasks. Since perceptual reasoning tasks were initially grouped under PIQ and working memory related tasks were considered under VIQ (see Figure 3.2), perhaps these results could explain the consistently larger magnitude of deficits in PIQ than VIQ among those born preterm/LBW. Although results here also suggested that children and adults born preterm/LBW are likely to find working memory related tasks most difficult to complete, their perceptual abilities were most markedly different from their same age peers.

Impact of Confounding Factors

Evidence of cognitive impairment is apparent even when various confounding factors have been considered. Studies demonstrate that prematurity continues to be a predictor of lower IQ found in those born preterm/LBW when adjustments have been made (P. J. Anderson, et al., 2003; E. A. Hutchinson, et al., 2013; Pyhala, et al., 2011). SES has been most commonly included for adjustment. This is because SES is associated with adverse birth outcomes including LBW, prematurity and growth restriction. Measures of SES may include parental education, intelligence or occupation, as well as household income or residential location (Kramer et al., 2009). A recent study documented that SES predicted long-term cognitive trajectories found in ELBW, with maternal education being the strongest predictor. In particular, Voss and colleagues

(2012) suggested that children born ELBW had better cognitive developmental outcomes when their mothers were more educated than those who were less educated due to a likelihood of better parenting style and knowledge in cognitive stimulation in the former.

Nonetheless, a recent meta-analysis that covered 27 studies and consisted of just over 7,000 participants (50% preterm and 50% term), with ages ranging from four to 14 years, suggested otherwise. Its findings suggested that the result of significant IQ differences between those born preterm/LBW and their peers did not differ between studies that documented the inclusion of SES adjustments and those that did not ($p = .316$; (Kerr-Wilson, et al., 2011). Given evidence from the meta-analysis, support for the argument that (lack of) maternal education is the strongest predictor of cognitive impairment found in those born preterm/LBW has not been fully justified and would benefit from further investigation. However, it is acknowledged that the risk of preterm birth decreases as SES increases (Thompson, Irgens, Rasmussen, & Daltveit, 2006), and adult SES is moderately associated with IQ (Fergusson, Horwood, & Ridder, 2005; Herrnstein & Murray, 1994; Sternberg, Grigorenko, & Bundy, 2001).

Heritability of maternal IQ may also play a role in the cognitive ability of children born preterm (Talge, et al., 2010). Additionally, some studies document that the chances of preterm birth can be familial (Clausson, Lichtenstein, & Cnattingius, 2000; Pennell et al., 2007; Wilcox, Skjaerven, & Lie, 2008). Studies on twins suggested that heritability for preterm birth could be as high as 40% (Clausson, et al., 2000; Treloar, Macones, Mitchell, & Martin, 2000). Mothers who were born preterm have an increased risk of preterm birth (Porter, Fraser, Hunter, Ward, & Varner, 1997; Wilcox, et al., 2008) and the risk increased as the mother's GA at birth decreased (Porter, et al., 1997). Research on the heritability of preterm birth also suggested that maternal

genetics were more influential to preterm birth deliveries than paternal genetics (Boyd et al., 2009; Svensson et al., 2009; Wilcox, et al., 2008).

Other confounding variables have also been considered. These include neurosensory impairments and head circumference. Studies have shown that prematurity and LBW continue to be associated with lower IQ even after adjusting for neurosensory impairment (P. J. Anderson, et al., 2003) and head circumference (Pyhala, et al., 2011). These confounding variables have been observed to correlate with an increasing degree of prematurity. As discussed in earlier sections, those born premature/LBW have an increased risk of ADHD and CP, which are also commonly associated with neurosensory impairments (Bhutta, et al., 2002; Lindstrom, et al., 2011; Oskoui, et al., 2013; Rosenbaum, et al., 2007). Head circumference has often been used as an indicator for growth restriction found amongst SGA/IGUR (Frisk, et al., 2002) and has been suggested to correlate modestly with IQ in the general population (Jensen, 1994).

Although these aforementioned confounding factors do not appear to fully explain the underlying differences in cognitive development observed between those born premature/LBW and those of typical development/full term, their associations with prematurity and LBW should not be dismissed as they have been shown to narrow the gap of unexplained differences (P. J. Anderson, et al., 2003; E. A. Hutchinson, et al., 2013; Pyhala, et al., 2011; Talge, et al., 2010).

Fluid Intelligence (G_f) and Education Attainment

In order to better understand impaired academic achievement but normal range FSIQ scores found in those born preterm/LBW, and to provide a theoretical basis for an investigation of the effect of prematurity on intelligence, it is proposed that G_f needs to be invoked (Duncan, et al., 1995; Taub, 2002). Traditional use of FSIQ as a measure of

general intelligence stems from factor analysis and, thus, instruments that provide an FSIQ score rely on a measurement model. Moreover, Colom and his colleagues (2002) suggested that FSIQ provides a measure of ‘intelligence in general’ rather than the scientific concept of general intelligence. In contrast, instruments that measure G_f stem from theoretical models of intelligence and are therefore able to assist in formulating and testing causal theories of relationships between cognitive abilities. In reference to the Cattell-Horn theory of intelligence (1963), G_f refers to a theoretical construct that relates to the ability to adapt to novel ideas and situations. It involves complex cognitive processes and relates to one’s ability to reason logically, solve problems and build associations among ideas. Induction and cognition of figural relations are its primary factors, which require non-verbal skills and are usually less culturally influenced. G_f stands in contrast to crystallized intelligence (G_c), which measures one’s accumulation of learned knowledge.

G_f and G_c are perceived as distinct but related constructs (Cattell, 1963; Kaufman, Kaufman, Liu, & Johnson, 2009; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002). For example, McArdle and his colleagues (2002) provided evidence on the $G_f - G_c$ distinction. They analysed data of individuals aged from two to 95 years on G_f and G_c abilities using the Woodcock-Johnson Psycho-Educational Battery – Revised (WJ-R) and found that the two have different growth patterns in the course of lifespan development. They stated that at initial rate of growth, G_f was slower than G_c and that G_f peaked earlier than G_c at approximately 22.8 vs. 35.6 years respectively. G_f also declined faster than G_c after it reaching its peak.

On the other hand, Cattell and Horn (1963) proposed, in their earliest conceptualization of the theory, that G_f facilitates and enhances the ability to acquire G_c which has been referred to as the Investment Theory (Cattell, 1963). Although evidence

for this directional relationship has been scarce, there is evidence showing that they are both associated with educational attainment and essential to academic success. Both G_f and G_c have been documented to correlate with the number of years of schooling ranging from .44-.64 for G_c and .48-.59 for G_f (Kaufman, et al., 2009; Kaufman & Wang, 1992; Schweizer & Koch, 2002; Sharma, Sharma, & Sharma, 2011). Some other authors have also claimed that G_f has a particularly strong influence on academic achievement and academic growth compared to G_c (Ferrer & McArdle, 2004), although other authors have argued otherwise (Postlethwaite, 2011; J. Raven, 1989).

Despite inconsistent views on whether G_f or G_c has a stronger influence, Sharma et al. (2011) demonstrated the two constructs to show distinctive patterns of predictions in academic subjects. In their recent research in India, they tested the performance of Year 11 students ($n = 200$) using Raven's Standard Progressive Matrices (RSPM) as a measure of G_f and the General Mental Ability Test (GMAT) as a measure of G_c and investigated their corresponding correlations with school subjects. Results, irrespective of gender group, revealed that performance on RSPM was more predictive of Math and Science, whereas performance on GMAT was more predictive of Language and Social Science. In particular, RSPM predicted 36% to 58% of variance in Math, and 32% to 35% of variance in Science as compared to only 9% to 13% variance in Math and 8% to 9% variance in Science by GMAT. In contrast, GMAT predicted 37% to 45% of variance in English Language, and 44% to 56% of variance in Social Science as compared to only 10% to 11% variance in English Language and 6% to 13% variance in Science by RSPM (Sharma, et al., 2011).

Prematurity and Fluid Intelligence

Researchers have also proposed that g and G_f are perfectly related or even equivalent constructs, at least in studies within normative cohorts (Gustafsson, 1984;

Keith, 2005; Kvist & Gustafsson, 2008), although others contest this (Blair, 2006; Gignac, 2006; McArdle, et al., 2002). This disagreement stems from clinical evidence in both adults and children. Studies using neuropsychological testing with adults with dorsolateral prefrontal cortex damage indicate that they perform within the normal range when assessed on measures of g , however exhibited poor performance on measure of G_f (Duncan, et al., 1995). Similarly, examinations of children with developmental disorders, particularly those with ADHD and specific learning difficulties, also demonstrated normal-range performance on measures of g , yet compromised cognitive functioning in EF components relating to G_f (Barkley, 1997; Tamm & Juranek, 2012).

Although children born preterm/LBW show performance profiles similar to children with ADHD in terms of normal range IQ, academic difficulties as well as impairments in basic information processing components of EF, which will later be discussed, no published studies have yet explicitly investigated the effects of prematurity on G_f . Nonetheless, an earlier thesis by Davies (2004), investigated G_f performance among children born preterm/LBW as part of an investigation on general intelligence. In one study, he concluded that significant differences existed between children born preterm/LBW ($n = 139$) and their full-term peers ($n = 73$) aged seven to nine years on Cattell Culture Fair Intelligence Test (CCFIT) raw scores ($p < .001$, $\eta^2 = .12$). After controlling for Wechsler FSIQ scores, full-term peers continued to score significantly better than children born preterm/LBW ($p = .01$, $\eta^2 = .03$). His observations also suggested that 9-year-old children in the clinical cohort performed at a level identical to the 7-year-old children in the full-term group. Davies' (2004) study provides preliminary evidence on the effects of preterm/LBW status on G_f . However, replication of such investigations using other G_f measures and age groups is warranted

to confirm his findings.

Examination of other existing empirical evidence supports a likely hypothesis that preterm/LBW will present with lower scores on tests measuring G_f . Firstly (and obviously), factor analysis shows FSIQ and G_f to have strong loadings on Spearman's g (Gustafsson, 1984; Keith, 2005; Keith, et al., 2006; Kvist & Gustafsson, 2008).

Therefore, if preterm/LBW children exhibit lower scores than their peers on FSIQ, then they would also be expected to perform less well on tests of G_f than their peers.

However, this alone would only predict a deficit in G_f proportional to the deficit in FSIQ. Secondly, and more interestingly, domain-specific indices indicate relatively large mean deficits in preterm/LBW children in PIQ (Talge, et al., 2010) and POI (P. J. Anderson, et al., 2003; Hallin, et al., 2010). As these indices tap into non-verbal abilities and include subtests that may be more G_f loaded rather than knowledge based (Wechsler, 1991), we may expect to find a G_f deficit disproportionate to the FSIQ deficit in this population. However, thorough investigation of this possibility remains to be done.

Explaining Individual Differences in Fluid Intelligence through Information Processing Parameters

Information processing parameters are basic cognitive system properties that, when taken together, may explain G_f variance. However, little research has investigated basic cognitive processes in relation to G_f amongst children born preterm/LBW. The most relevant study was published by Rose, Feldman, Jankowski, and Van Rossem (2011) on children born preterm, in which they inspected the effect of prematurity on FSIQ through several basic cognitive processes.

Rose et al. (2011) compared children born preterm ($n = 44$) with full-term peers ($n = 86$), group-matched on SES and gender at 11 years, on measures of working memory, attention, processing speed, representational competence, and FSIQ. The

authors described *representational competence*, a term less commonly documented in the literature, as “the ability to extract commonalities from experiences and represent them symbolically” (Rose, et al., 2011, p. 199). Cognitive abilities were assessed using subtests from the Cambridge Neuropsychological Testing Automated Battery (CANTAB), the Cognitive Abilities Test (CAT), and the Specific Cognitive Abilities Tests (SCA). FSIQ was assessed using WISC-III. Rose and her colleagues (2011) reported that, having excluded those with IQ <70 which constituted five of the children in the preterm group, children born preterm showed pervasive deficits in all four areas. More importantly, taken together, these deficits fully accounted for the preterm/full term differences measured using IQ. A cascade of effects was presented where children born preterm demonstrated impairments on speed and attention tasks, considered to represent elementary information processing, leading to impairments on memory and representational competence tasks, which are more complex, and subsequently resulting in IQ deficits relative to their full-term peers. The cascade model has been adapted in Figure 3.3 for clarity. This represents one of very few studies documenting impairments in core information processing abilities and investigating them as mediating factors in explaining the birth status differences in IQ. However, similar studies using different age groups and a wider range of representative factors are required in the future to further generalise their findings, particularly when factors such as *representative competence* has rarely been documented.

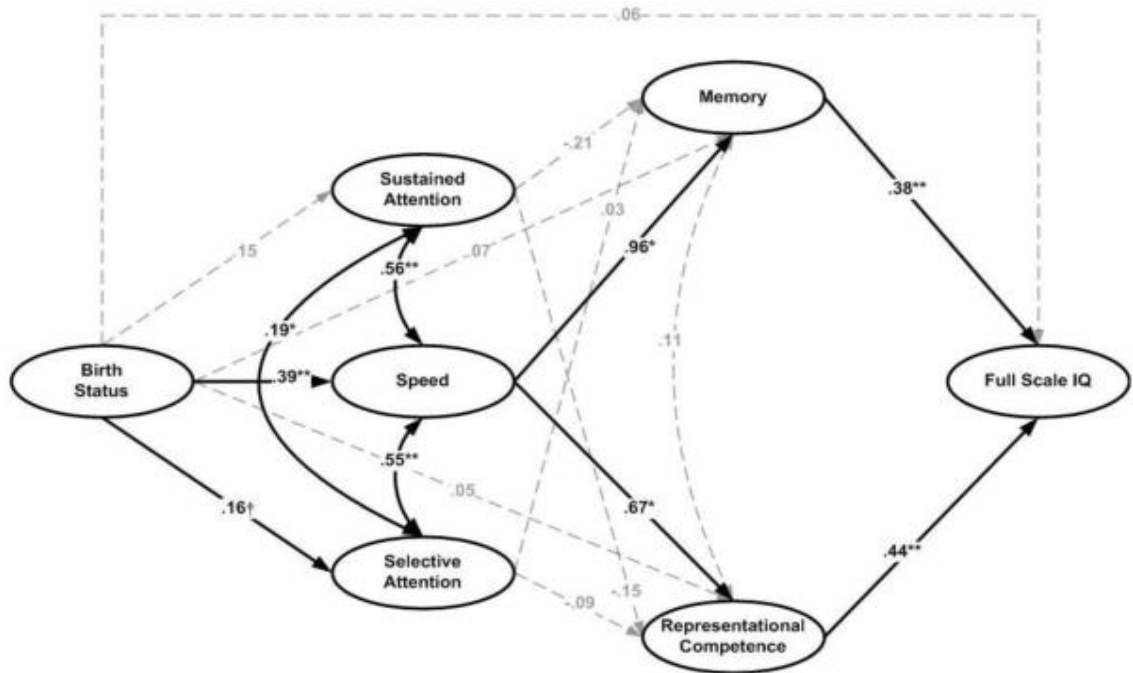


Figure 3.3. The results of standardized estimates of birth status effects on full scale IQ through basic information processing abilities at 11 years old. Significant pathways are indicated in solid lines, † $p < .06$, * $p < .05$, ** $p < .01$. Adapted from “Basic Information Processing Abilities at 11 years Account for Deficits in IQ Associated with Preterm Birth” by S. A. Rose, J. F. Feldman, J. J. Jankowski, and R. Van Rossem, 2011, *Intelligence*, 39(4), p.18.

However, given that children born premature/LBW typically present with average-range IQ, resulting in the previously discussed inability of global IQ to explain this group’s academic struggle as compared to their peers, Rose et al.’s (2011) study may not capture the full scope of cognitive deficits. Nonetheless, their approach may be useful in exploring how core information processing abilities may be substantial mediating factors for possible birth status differences in G_f .

The literature on prematurity/LBW has documented the pivotal roles played by the basic cognitive processes found in executive function, which include aspects such working memory, cognitive flexibility and inhibition. In the next few sections, these processes and the effect of prematurity on each of them will be elaborated.

Executive Function (EF)

EF refers to higher-order cognitive processing that relates to self-regulation of problem solving, reasoning and planning towards goal-directed behaviours (P. J. Anderson, et al., 2010; Clark, et al., 2010; Miyake, et al., 2000). Theories on how to conceptualize EF have been controversial with some researchers arguing for a unitary model (Brydges, Reid, Fox, & Anderson, 2012; Duncan, Emslie, Williams, Johnson, & Freer, 1996; Wiebe, Espy, & Charak, 2008), while others consider EF to comprise related but distinguishable subcomponents (Baddeley, 1996; Miyake, et al., 2000). The factor structure of EF changes developmentally, starting unitary in young children but distinguishable components emerge with development in mid to late childhood. Therefore, subcomponents are more easily identified as individuals age (Best, Miller, & Jones, 2009; Isquith, Gioia, & Espy, 2004).

Factor analyses have led to the proposal of some separate essential subcomponents under the EF umbrella term, including planning, fluency, and attentional control (Mulder, et al., 2009), while Miyake et al. (2000) suggested the distinctive functions of three core abilities. Although the components proposed in Miyake et al.'s (2000) seminal paper on adults may not have conclusively been demonstrated as a comprehensive list, they have been further supported by Lehto, Juujarvi, Kooistra, and Pulkkinen's (2003) evidence in child populations. The proposed factors include inhibitory control, cognitive flexibility and working memory. The authors suggested that these core components were relatively independent of one another, yet also combined to contribute to complex executive tasks.

In daily life, EF takes the form of creating goals, making plans, performing actions and monitoring performance in relation to established goals, while adjusting one's action to maintain effectiveness of a series of goal directed behaviours (Jurado &

Rosselli, 2007).

Inhibitory Control

Inhibitory control refers to the ability to suppress automatic and prepotent responses in order to perform the task at hand. This requires one to focus on the task and refrain from acting on first impulse (Heitz, et al., 2005; Miyake, et al., 2000). Exemplar tasks often used to measure inhibition include the Stroop task (Stroop, 1992), and Go-No-Go task (Cragg, Fox, Nation, Reid, & Anderson, 2009). The Stroop task requires participants to name the colour that the words on a list are printed in. The ink colour may be congruent or incongruent with the meaning of the word, such as seeing the word “RED” that is printed in red ink or seeing the word “RED” but printed in yellow ink, in which the correct response would be *red* and *yellow* respectively. The difference in time between the conditions is used as a measure of inhibition (Stroop, 1992). In one of the many variations of the Go-No-Go task, participants are asked to release a “home” button and press a response button as quickly as possible when a targeted stimulus is presented on the computer screen, referred to as the “go” condition. In the “no-go” condition, where no stimulus or a non-targeted stimulus is presented, participants are to withhold any response and remain pressing on the “home” button. The number of correct response to a “no-go” condition is used as an indicator of inhibition (Cragg, et al., 2009).

Cognitive Flexibility

Cognitive flexibility, also known as shifting, refers to the ability to flexibly switch one’s focus of attention between mental sets and tasks. This means switching from the engagement in one task to commence a new task (Jurado & Rosselli, 2007; Miyake, et al., 2000). From here, cognitive flexibility and shifting are used interchangeably. Examples of tasks considered to measure shifting include Wisconsin

Card Sorting Test (WCST; (Heaton, Chelune, Talley, Kay, & Curtiss, 1993; Miyake, et al., 2000), plus-minus task (Miyake, et al., 2000) as well as the Trail Making Test – Part B (TMT-B; (Reitan, 1971; Reitan & Wolfson, 1992) documented in EF studies (Jurado & Rosselli, 2007; Mulder, et al., 2009). WCST is a sorting task that requires participants to match category cards under one of four stimulus cards. Participants are to switch amongst three principles, namely colour, number, and form, using the feedback provided by the examiner. The raw score on perseverative errors from the task is often used as the outcome measure (Brydges, et al., 2012; Heaton, et al., 1993). The plus-minus task requires participants to solve math problems in three two-digit numbers lists, first in addition and then subtraction, and subsequently switch between addition and subtraction. The measure of shifting is derived from the cost of shifting. This is calculated using the difference between the time in completing the final list - alternating between addition and subtraction - and the average time in completing the first two lists (Miyake, et al., 2000). In the TMT-B task, there are 15 circles on a sheet of paper with a number and a letter in each of them. Participants are required to draw a line to connect the circles alternating between a number and a letter in sequential order (i.e. 1->A->2->B). The measure of shifting is derived from the time used to complete the task (Reitan, 1971).

Working Memory

Working memory (WM) refers to the mental capacity for temporary storage and active manipulation of information used in a variety of everyday activities (Baddeley & Hitch, 1974; Kane, et al., 2001). Thus, WM tasks require the individual not only to store information but also to recode stored information in a way that allows for effective and efficient manipulation of task-related material when needed. Exemplar tasks frequently used include operation span tasks (M. L. Turner & Engle, 1989), digit span tasks

(Wechsler, 1991, 2003), and spatial span tasks (Wechsler, 1991). For all these tasks, the number of correct trials is used as the outcome measures. In an operation span task, participants are presented with a string of words with a distracting task in between each word in the form of a math problem. Participants are to recall the set of words in the order it was presented. The difficulty is increased following successful trials by the increment of another word (M. L. Turner & Engle, 1989). The digit span task requires participants to recall strings of digits verbally presented by the examiner in forward and backward sequences (Wechsler, 1991, 2003). For the spatial span task, the stimulus is presented in a visual-spatial way where the examiner points to square cubes randomly placed on a board. Participants are then asked to repeat where the examiner has pointed in either forward or backward sequence (Wechsler, 1991).

As working memory updating is one of the key processes in both upcoming empirical studies, its theory deserves some further attention. The most commonly referenced models in the conceptualization of WM are the multicomponent model developed by Baddeley and Hitch (1974) and the executive attention model by Kane, Conway, Bleckley, and Engle (2001). Baddeley and Hitch (1974) initially proposed that WM comprises three main components: a domain-general central executive, and two domain-specific subsidiary components - the phonological loop and visuo-spatial sketchpad. The central executive was proposed to be accountable for attentional control, processing information held in the two specific domains and retrieval from long-term memory. The phonological loop was proposed to temporarily store verbal information, and the visuo-spatial sketchpad to temporarily store visual and spatial material. A recent addition to the WM model is a fourth component known as the episodic buffer. It is responsible for binding information coming through from the other WM components (Baddeley, 2000; Baddeley & Logie, 1999).

Kane et al.'s (2001) executive attention model agrees to some extent with that of Baddeley and Hitch (1974), in particular with respect to the existence of a central executive. However, Kane et al. (2001) stressed that the executive attention function, also known as 'controlled attention', is dependent on one's working memory capacity (WMC). According to their definition:

“An executive control capability... is...an ability to effectively maintain stimulus, goal, or context information in an active, easily accessible state in the face of interference, to effectively inhibit goal-irrelevant stimuli or responses, or both.... This attentional control capability allows flexibility in response to environmental demands, whether those demands involve keeping many representations active in some contexts, keeping only one simple goal active in other contexts, or keeping irrelevant representations or responses at bay through inhibition” (Kane, et al., 2001, p. 180).

Engle (2002) explains that in the controlled-attention model, higher WMC is a consequence of higher control attention capability rather than merely the ability to store more items for active processing. Essentially, higher WMC reflects a greater ability to inhibit interference and distraction by means of focused attention.

This concept of WMC was also clearly defined in Heitz, Unsworth and Engle (2005):

“In other words, although we measure WMC quantitatively by the number of items recalled on complex span tasks, the scores on such measures reflect controlled-attention ability rather than the number of “units” of information that can be held in a short-term store. Accordingly, we propose that WMC is an *ability* reflecting the extent to which an individual is able to control attention, particularly in situations involving interference from competing information, activated representations, or task demands” (p. 64).

In support of this model, Kane et al.'s (2001) research showed that there is a positive association between individual differences in working memory span and attentional control. "Higher memory span" adults were more capable than "lower memory span" adults of inhibiting distracters and maintaining attention to relevant information for further processing. They also viewed executive attention as a key connecting component to high level functioning, potentially an essential mechanism of G_f (A.R.A. Conway, Kane, & Engle, 2003; Kane, et al., 2001).

Even though it seems like WM and inhibitory control are both measures of attentional control, a key feature and commonality amongst these multi-system WM models is that there exists a short-term storage (STM) component and an executive attention control component to WM that does not exist in the construct of inhibitory control. STM is usually measured using simple span tasks, for example the digit span forward task that requires an individual to store and recall information directly and immediately. In contrast, the attentional control component is usually measured using complex span tasks, for example operational span tasks and digit span backward tasks that require an individual not only to store information but also to manipulate it in some way to arrive at a correct response (Baddeley, 2000; Baddeley & Hitch, 1974; Cowan, 2000; Engle, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kane, et al., 2001).

Nonetheless, pure measures of storage capacity and attentional control may not be achievable as it is difficult to measure one independently of the other (Heitz, et al., 2005). Presumably, some tasks are more sensitive to capacity than attention and vice versa. As suggested by Unsworth and Engle (2006), overload occurs when the number of items required to be recalled extends beyond one's STM capacity in simple span tasks, therefore a long item length task in simple span tasks also tap executive attention control mechanisms. Similarly, complex span tasks require an individual to utilize both

STM storage and attentional control (Bayliss, Jarrold, Gunn, & Baddeley, 2003).

Therefore, STM and attentional control components of WM are both important in their own varying degrees, with complex span tasks more reliant on attention control in addition to STM storage and vice versa for simple span tasks (Heitz, et al., 2005).

Normal Developmental Trajectories of EF and Methodological Concerns

Literature on the normal developmental trajectories of EF suggests that each component develops differently across the lifespan, but interpretations vary depending on the methodology used to measure the component as well as the complexity of the task (Best, et al., 2009). The study of normal EF development has disproportionately focused on preschool-age children in the past. This is because researchers have been primarily interested in identifying the specific age at which EF, in general or as specific components, emerges (Garon, Bryson, & Smith, 2008). Nevertheless, the study of normal developmental trajectories of EF in typical school-aged children is also important, as changes in EF take place beyond early childhood. New environmental stimuli and new experiences become relevant, ranging across academic, social, and school-related domains (Best, et al., 2009).

According to several authors who support the distinction between the three core components suggested by Miyake et al. (2000), these distinct components are less noticeable and more intertwined in children as compared to adults (Isquith, et al., 2004; Miyake, et al., 2000). However, evidence of component-specific developmental trajectories suggests that components do work separately at different ages. In a comprehensive review of EF development in children after their preschool years, Best and his colleagues (2009) summarized findings on how specific components of EF change developmentally.

Basic inhibitory control abilities emerge as young as one year old.

Developmental changes in inhibition are more noticeable during early childhood/preschool years with rapid development occurring between three and five years of age. Although changes to inhibition continue to occur from school entry level through to adolescence, these changes reflect changes in speed and accuracy rather than fundamental changes in the ability to inhibit prepotent responses, as found in young children. Adult levels of inhibitory control abilities are said to be reached around 14 years of age (Best, et al., 2009; Mulder, et al., 2009).

On the other hand, cognitive flexibility and WM can be detected in preschool age children and evolve in a linear progression beginning in early school age years. At approximately three years of age, the ability to shift in simple tasks with a maximum of two stimulus-response rules emerges. The ability to shift in tasks with more rules can be detected around the age of five years. Between the age of seven to nine years, children demonstrate significant improvement in the ability to switch between multi-dimensional tasks, which then plateaus around the age of 12-years (P. Anderson, 2002; Best, et al., 2009).

A normal developmental trajectory in WM is similar to that of cognitive flexibility. It emerges as early as four years of age. By around seven years of age, children are well capable of performing various WM tasks, particularly those involving visuo-spatial and verbal working memory. Continuous linear increment can be observed until it begins to plateau around 11 years of age and peak around age 20 (Best, et al., 2009; Gathercole, Pickering, Ambridge, & Wearing, 2004; Lehto, et al., 2003).

Nonetheless, no single assessment provides a pure measure of any intended construct (Best, et al., 2009; Senn, Espy, & Kaufmann, 2004). This leads to several methodological concerns, in particular regarding the purity of EF tasks as measures of a single construct and the reliability of EF measures. Task impurity has been a concern

for measures of EF (Burgess, 1997; Rabbitt, 1997). For example, the plus-minus task as documented in Miyake et al.'s (2000) study is used to assess shifting abilities, however, mental math calculations have long been acknowledged to rely heavily on working memory (DeStefano & LeFevre, 2004). Also, an individual may employ other processes that the task was not originally intended to engage (Hughes & Graham, 2002).

The concept of such task impurity also leads us to other methodological issues with EF. Researchers have asserted that it is very difficult, or even impossible, to establish reliable measures of EF (Burgess, 1997; Rabbitt, 1997). This is because the central idea of measuring EF is to assess how well one self-regulates to cope with and solve *new* problems. Thus, the issue of novelty of stimuli arises when a task is administered again after the first time, making the task no longer novel and no longer a measure of EF (Burgess, 1997; Jurado & Rosselli, 2007; Rabbitt, 1997).

In summary, the developmental trajectories of, as well as methodological issues inherent in, EF should be considered when proceeding with both correlational and experimental studies. This is especially true, and perhaps inconvenient, for longitudinal and repeated measure studies that try to establish utility in cognitive interventions.

Prematurity and Information Processing Parameters of EF

Dysfunctions in EF are often discussed in relation to damage to the prefrontal cortex (Eslinger, Flaherty-Craig, & Benton, 2004) as well as other neurological impairments (Narberhaus et al., 2008; Nosarti et al., 2004). Such damage is also commonly found amongst individuals born preterm/LBW. Studies show those children born preterm are vulnerable to prefrontal cortex injuries and thinning of the corpus callosum (Narberhaus, et al., 2008; Nosarti, et al., 2004). These two areas of the brain are interconnected and include a large white matter volume. White matter injuries are also commonly found. These are associated with the occurrence of periventricular

leukomalacia (PVL) and are often present in the development of cerebral palsy. In addition, according to de Kieviet et al. and his colleagues' (de Kieviet, Zoetebier, van Elburg, Vermeulen, & Oosterlaan, 2012) meta-analysis review on 15 studies, children born VPT/ VLBW are found to have an overall reduction in total brain volumes, including the cerebellum, hippocampus and corpus callosum area, as well as volumes in white and grey matter as compared to their term peers. These reductions are also influential towards lower cognitive performance, for example general IQ, EF and memory.

Another frequently noted cerebral injury amongst children born preterm is intraventricular haemorrhage (IVH). Immature blood vessel development at early birth may cause IVH (Edgin et al., 2008; Inder, Lawrence, & Neil, 2010). Given the above, impairments in EF would be considered relatively likely in children born preterm/LBW. Meta-analyses also confirm that EF impairments are present in the group and that degree of prematurity significantly predicts the degree of impairment (Mulder, et al., 2009). The effect of prematurity is also evident across all subcomponents of EF, but particularly WM (Aarnoudse-Moens, Smidts, et al., 2009; Fraello et al., 2011; Mulder, et al., 2009; Nosarti, et al., 2007). The following sub-sections elaborate on the effects of prematurity and LBW on the three basic processes of EF.

Inhibitory Control

Evidence of impairment in inhibitory control has been documented in children born preterm/LBW. Mulder et al.'s (2009) review demonstrated that individuals born preterm, ranging from five to 22 years of age, exhibited impairments in inhibitory control proportional to their GA. A combined moderate effect size ($d = -0.50$) was recorded for inhibition performance with children born preterm (<26 weeks) showing more commission errors on the Go/no-go test than their peers. A much smaller effect

size ($d = -0.16$) was found for individuals born preterm with GA ≥ 26 weeks. A significant relationship was detected between effect size and GA for those born ≥ 26 weeks ($R^2 = .74$; $p = .006$), suggesting that children born more mature showed less impairment on inhibitory control tasks. However, the number of studies included in the review was rather small ($N = 8$).

A recent large sample study with children born VPT aged between four and 12 years provided a more precise estimate of children's inhibitory control ability. A comparison of VPT ($n = 200$) with their full-term peers ($n = 230$) on a stop task that recorded omission errors, commission errors and stop signal reaction time demonstrated standardized mean differences (SMD) of -0.15, -0.41 and -0.43 respectively, after adjusting for age, gender and speed of processing, with comparison group as reference (SMD = 0.0). Children born full-term performed better than the VPT group. Although adjustments for IQ and neurosensory dysfunction were undertaken, the authors did not include any statistical details of the effects of these on the results. Nonetheless, neither of the aforementioned confounds undermined the significant results in favor of the full-term group. This study also demonstrated that although their sample of VPT/VLBW ($M = 95.3$, $SD = 15.8$) performed significantly less well than their full-term controls ($M = 105$, $SD = 13.4$) on IQ ($p < .001$), the VPT/VLBW's performance were within normal range (Aarnoudse-Moens, Duivenvoorden, Weisglas-Kuperus, Van Goudoever, & Oosterlaan, 2012).

Specific age-groups were also reviewed. For example, a study compared age-matched 6-year-old VPT/VLBW (BW: $M = 1042.6g$, $SD = 31.8$, GA: $M = 28.0$, $SD = 1.4$, $n = 50$) with full-term controls ($n = 50$) and found significant group differences after controlling for IQ scores (Aarnoudse-Moens, Smidts, et al., 2009). Findings

showed that the full-term control group performed better than the VPT/VLBW group on both efficiency scores for inhibition from the Go/No-Go (before adjusting for IQ: $p < .01$, $d = -0.83$; after adjusting for IQ: $p = .03$, $d = -0.51$) and Day-Night tasks (before adjusting for IQ: $p < .01$, $d = -1.35$; after adjusting for IQ: $p < .01$, $d = -0.79$). Moreover, GA, maternal education and neurosensory impairments were not significant predictors of inhibition performance. Similar to Aarnoudse-Moens et al. (2012), the two groups differed significantly on IQ measures ($p < .001$) but both VPT/VLBW ($M = 92.5$, $SD = 17.5$) and the full-term group ($M = 109$, $SD = 19.2$) performed within the average range (Aarnoudse-Moens, Smidts, et al., 2009).

Findings presented by Ford et al. (2011) with EPT/ELBW aged seven to nine years provided a more comprehensive picture and included several other confounding factors. It was reported that the clinical cohort performed less well on inhibitory control than their same age full-term controls as measured by the Stroop task ($p = .033$, $\eta^2_p = .005$). Their regression analysis further demonstrated that birth weight was a significant predictor of Stroop performance ($\beta = 0.36$, $p = .038$). Children with a higher birth weight performed better on the Stroop task. Inclusion of SES and neuro-biomedical history, such as respiratory distress syndrome (RDS) and intraventricular hemorrhage, to their regressions suggested evidence of interaction between SES and neuro-biomedical history on Stroop ($\beta = 0.41$, $p = .027$). This suggested that the influence of neuro-biomedical history on Stroop in children born EPT/ELBW was higher in low SES households. Once again, despite these between group differences, participants in the EPT/ELBW group attended mainstream schools, showed no significant neurological disabilities, and displayed average IQ. However, Ford et al.'s (2011) study also demonstrated clearly the significant impact of SES and neuro-biomedical history, despite these factors not fully accounting for group differences.

These inhibitory control impairments are also evident in late teen and adult samples of individuals born preterm/LBW. For example, 16-year-olds born VLBW were compared with a full-term control group on subtests measuring inhibitory control through the Delis-Kaplan Executive Function Scale (D-KEFS). Results indicated that VLBW participants did not perform as well as their peers ($p < 0.05$; (Luu, et al., 2011). Similarly, 19- to 25-year-old VPT adults ($n = 61$) were compared with a peer control group ($n = 64$), where participants were assessed on inhibitory control skills using the Test of Attentional Performance - Incompatibility subtests (TAP/I). Results indicated that the VPT group performed less well than their controls ($p < .05$). Only 18% of VPT adults scored less than 1SD below the mean score of their peers, suggesting that the majority of VPT adults displayed substantial impairments as measured by the TAP/I task (Nosarti, et al., 2007). The aforementioned studies with adolescents and adults all documented normal IQ scores for VPT/VLBW groups.

In general, the literature shows that impairment in inhibitory control is consistent and prevalent in both children and adults born preterm/LBW, despite their average IQ scores. Effect sizes for impairments in inhibition range from moderate to large, which are comparable to those demonstrated in earlier meta-analysis for academic problems found in those born preterm/LBW, specifically in regards to math ($d = -0.60$) and reading ($d = -0.48$; (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). In addition, while Mulder et al. (2009) suggested that greater GA is associated with smaller deficits, Aarnoudse-Moens et al. (2009) found that GA did not significantly predict performance in inhibition tasks. Currently, most studies rely on single measures in their investigations and at different age groups. More studies are needed to understand the precise impact of prematurity on inhibitory control. Specifically, the inclusion of several inhibitory control measures in one study across different age groups

and different subgroups of children born preterm can shed light on their developmental differences in inhibitory control.

Cognitive Flexibility

Deficits in cognitive flexibility have been documented in the preterm/LBW population relative to their full-term peers, but these findings are mixed. Meta-analysis suggested that the measurement tool used to assess shifting has a strong impact on the outcomes of comparison studies. It was documented that studies using the TMT-B to measure shifting performance demonstrated a moderate effect size ($N = 6$, $d = -0.50$, 95% CI 0.36-0.64), with the clinical cohort performing less well than their full-term control peers. This was in contrast to studies that described using sorting tasks, such as WCST, where shifting performance was not significantly different from control peers ($N = 6$, $d = -0.10$, 95% CI -0.06-0.27; (Mulder, et al., 2009).

Findings from another review of 12 EF studies corroborated the aforementioned results using TMT-B as a measure of shifting. The review described the performance of VPT/VLBW individuals, ranging from eight to 22 years of age, as below that of their full-term peers. A reported combined effect size of -0.49 ($p < .001$) was documented, with the TMT-B as measurement. Their reported correlation between BW/GA and shifting performance was not significant (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). This study did not, however, provide details on the impact of SES, neurosensory impairments or IQ on their range of reviewed studies.

A similar effect size to the aforementioned 12-study review was found in a study of 6-year-old, age-matched children, comparing VPT/VLBW and full-term controls. The study used the Object Classification Task for Children (OCTC) as a measure of shifting and found that the control group performed better than the clinical cohort before ($p < .01$, $d = -0.77$), and after controlling for IQ scores ($p = .04$, $d = -0.40$; (Aarnoudse-

Moens, Smidts, et al., 2009). Although it was noticeable that both the *p*-value and the effect size have reduced after controlling for IQ, Friedman et al. (2006) suggested that there is a very low correlation between shifting and IQ in young adults. Further investigations on the impact of SES and neonatal risks did not show predictive variance amongst the 6-year-old sample. However, in contrast to the review, GA was found to explain 12% of variance in the shifting measure. In addition, similar to that documented for inhibitory control, the clinical group continue to perform at normal range IQ despite displaying impairments in cognitive flexibility tasks (Aarnoudse-Moens, Smidts, et al., 2009).

Nonetheless, a majority of recent findings were in agreement with Mulder et al.'s (2009) conclusion regarding the influence of measurement tools. For example, when Aarnoudse-Moens and her colleagues (2012) tested children born VPT/VLBW, between the age of four to 12 years, on a range of EF processes, they did not find any significant between-group differences on shifting using a stimulus-response compatibility task. Also in contrast with findings using TMT-B, Luu et al. (2011) did not find any significant group differences on measures of shifting in children born VPT/VLBW at 16-years of age when assessed using the Delis-Kaplan Executive Function Scale (D-KEFS).

In sum, group differences in cognitive flexibility between the clinical cohort and their full-term peers appear to be inconsistent in both children and adults born preterm/LBW when measurement tools other than TMT-B was used (Aarnoudse-Moens, et al., 2012; Luu, et al., 2011). More evidence on the effect of prematurity/LBW on shifting is required in measures other than TMT-B. This is because given the current evidence, it is unable to ascertain whether those born premature/LBW display deficits pertaining to shifting or an alternate cognitive process to which TMT-B is sensitive.

Ideally, future studies should test the same clinical cohort concurrently with various shifting measures and across age groups.

Working Memory

Psychologists have taken particular interest in WM. Not only because evidence exists for WM's associations with prematurity/LBW, whether it be tested as a general process or in specific forms, but also WM's unique relationship with G_f . Recent meta-analysis of 12 EF studies comparing VPT/VLBW to their full-term peers, ranging from seven to 14-years of age, on measures of digit span detected significant group differences with children born VPT/VLBW displaying lower scores with a combined effect size of -0.36 ($p < .001$; (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). However, the impact of confounding factors such as IQ and SES on the combined effect was not considered in the analysis.

Studies that have considered confounding variables have mixed findings. For example, Ford et al. (2011) documented significant WM deficits in children born EPT/ELBW. They compared children born EPT/ELBW ($n = 45$) with their full-term controls ($n = 45$), aged between seven and nine years, on the combined scores of performance using digit span backwards and spatial span backwards. Inclusion criteria for participants included average IQ, no significant neurological disabilities as well as attendance at mainstream schools. Results indicated group differences with full-term controls performing significantly better ($p = .007$, $\eta^2_p = 0.08$). The interaction between SES and neonatal medical risk significantly predicted the clinical group's performance on WM ($\beta = 0.42$, $p = .022$). It was suggested that EPT/ELBW children growing up in higher SES families had less adverse effects stemming from neonatal medical risk at birth than those in lower SES families.

However, a more recent and larger scale study by Aarnoudse-Moens et al.

(2012) reported otherwise. Children, aged between four and 12-years, born VPT/VLBW ($n=200$) were compared with their full-term peers ($n = 230$). The full-term group performed significantly better than children born VPT/VLBW on the spatial span task ($p = .002$) and the digit span backwards task ($p = .001$). This was indicative of small standard mean differences of -0.34 and -0.32 respectively. These impairments were no longer statistically significant after adjusting for IQ, although all participants performed at average IQ. It was also noted that neurosensory impairments did not have a significant impact on WM tasks.

A closer analysis of existing research demonstrated impairments in specific WM areas, for example, verbal and visuo-spatial WM, amongst children born preterm/LBW. Recently, Clark and Woodward (2010) compared 6-year-old VPT children (GA: $M = 27.9$, $SD = 2.38$; BW: $M=1066$ g, $SD = 316.27$; $n = 103$) with full-term control (GA: $M = 39.51$, $SD = 1.19$; BW: $M=3574.58$ g, $SD = 409.84$; $n = 108$). They tested participants on verbal and visuo-spatial WM performance using the digit span task and Corsi blocks task respectively (including both forward and backward spans). All children displayed average IQ as tested using the WPPSI, although significant between group differences were detected in favour of the full-term group ($p < .001$). Group differences were particularly marked on the backward aspect of each WM task. However, after controlling for SES and neurological impairments, only differences on the Corsi block backward span task remained significant (FT: $M = 9.73$, $SD = 7.29$ vs VPT: $M = 6.35$, $SD = 6.04$, $p = .02$, $d = -0.51$). This study indicated that perhaps the clinical group was more vulnerable to visuo-spatial rather than verbal impairments.

In contrast, another study conducted a comparison between age-matched 6-year-old VPT/VLBW ($n = 50$) and full-term controls ($n = 50$) on verbal WM measured using digit span task. Results indicated that the clinical cohort was out-performed by their

control peers on backward (BWD) spans ($p = .01$, $d = -0.80$), even after controlling for IQ scores ($p = .02$, $d = -0.53$). In addition to that, GA, SES and neurosensory impairments did not significantly predict variances in WM scores. Additionally, children in this sample all performed with average IQ (Aarnoudse-Moens, Smidts, et al., 2009). These findings suggest WM deficits in children born preterm/LBW was not accounted for by other factors.

Fraello et al.'s (2011) investigation was not in total agreement with the above evidence in regards to the specific areas of WM impairments as demonstrated in a recent small sample study comparing preterm/LBW (GA: $M = 28.4$, $SD = 1.8$, BW: $M = 972.7$ g, $SD = 151.4$, $n = 49$) and full-term ($n = 20$) adolescents' performance on WM tasks. Participants were assessed at 12 years of age and matched by age, gender and minority status. The authors used the WISC-III, the Comprehensive Test of Phonological Processing (CTOPP) and the Clinical Evaluation of Language Fundamentals-III (CELF-III) as measures for WM, which included digit span forward and backward tasks, a non-word repetition task, a verbal sentences recall task, as well as a multistep oral concepts and directions recall task. Results revealed no significant differences in digit span forward or backward tasks and non-word repetition, with maternal education ruled out as a potential confound. However, they showed that full-term adolescents performed significantly better than the preterm cohort in tasks that were of higher verbal complexity, such as recalling sentences (FT: $M = 11.1$, $SD = 2.8$ vs PT: $M = 8.5$, $SD = 3.2$, $p = .004$) and recalling concepts and directions (FT: $M = 10.4$, $SD = 3.1$ vs PT: $M = 7.4$, $SD = 2.7$, $p = .001$). It was concluded that as novel verbal information increases in complexity, children born preterm/LBW showed more difficulties in their ability to process and recall this information. Unfortunately, the authors did not provide any effect size statistics for their results. Therefore, further

comparisons could not be made with other studies and generalizations given the small sample size were to be made with caution. Nevertheless, Freallo et al.'s (2011) results here may provide some insight into reading difficulties identified in preterm/LBW and support Kane and Engle's (2001) controlled attention theory. This is because simple WM tasks, such as digit span tasks, require less focused attention, less inhibition of interference and distractions than complex, novel, verbal WM tasks, such as recalling sentences. Focused attention and inhibitory control are essential components of executive attention, which is also positively associated with working memory capacity and levels of cognition. Thus, the reported significant group difference in higher complexity tasks is consistent with a difference in executive attention.

Working memory impairments in individuals born preterm/LBW appeared to also persist into late teens (Luu, et al., 2011) and early adulthood (Hallin, et al., 2010). For example, WM abilities of 16-year old VPT adolescents ($BW \leq 1250$ g, $n = 337$) were compared with their full-term peers ($n = 102$), matched according to age, gender, race and residential area. Performance on the Wechsler Memory Scale (WMS) backward span task was compared between the two groups with covariates adjusting for SES, such as maternal education and minority status. Significant group differences in favour of the full-term group were found in backward span task: mean difference [MD] = -1.7, 95% CI -2.5 to -0.8, $p < .005$). The results were then adjusted for IQ measured by the Peabody Picture Vocabulary Test – Revised (PPVT-R), and results indicated that significant differences continued to be apparent ($MD = -1.0$, 95% CI -1.8 to -0.3, $p < .05$). However, significant differences were no longer detected when further excluding participants with neurosensory impairments (NSI) and those with $IQ < 70$. SES and IQ did not provide a full explanation for the differences in working memory as measured by the backward span task, although the adjustment of IQ did narrow the gap.

Nevertheless, NSI and low IQ participants, together with SES and IQ fully accounted for the differences found in their sample of adolescence. Subsequently, their results for verbal and visual spatial memory after exclusion of participants with NSI and $IQ < 70$ contrasted with those of working memory backward span. When the adolescents were tested using California Verbal Learning Test as a measure of verbal memory and the Rey-Osterrieth Complex Figure Test as visual spatial memory, the three adjustments continued to narrow the discrepancy between the clinical group and the full-term group, but significant differences continue to exist. Results after each adjustment for 1) SES, 2) IQ, and 3) NSI and $IQ < 70$: immediate verbal memory: (1: $MD = -7.1$, 95% CI -9.7 to -4.5, $p < .005$; 2: $MD = -5.2$, 95% CI -7.5 to -2.9, $p < .005$; 3: $MD = -4.2$, 95% CI -6.5 to -1.9, $p < .005$) and delayed verbal memory: (1: $MD = -0.8$, 95% CI -1 to -0.5, $p < .005$; 2: $MD = -0.6$, 95% CI -0.8 to -0.3, $p < .005$; 3: $MD = -0.5$, 95% CI -0.7 to -0.3, $p < .005$), as well as visual spatial memory, again both immediate (1: $MD = -5.6$, 95% CI -7.6 to -3.6, $p < .005$; 2: $MD = -3.9$, 95% CI -5.7 to -2.2, $p < .005$; 3: $MD = -2.8$, 95% CI -4.6 to -1.0, $p < .005$) and delayed (1: $MD = -5.7$, 95% CI -7.6 to -3.7, $p < .005$; 2: $MD = -4.0$, 95% CI -5.7 to -2.3, $p < .005$; 3: $MD = -2.9$, 95% CI -4.7 to -1.1, $p < .005$; (Luu, et al., 2011).

All in all, WM impairments in those born preterm/LBW are consistently evident and these impairments are persistent in both children and adults. Impairments are also apparent in both working memory as measured using STM measures and attentional control measures. Confounding variables, such as SES and NSI played a role in narrowing the gaps of significant differences. Global IQ measures also narrowed the discrepancy, however, similar to findings for inhibitory control and cognitive flexibility, children born preterm/LBW displayed average IQ, which would not lead us to predict WM impairments.

Confounding Variables

Several covariates are often included in the study of EF amongst those born preterm/LBW. These include SES, as measured using residential location, parental education, household income, race and age. The main reason for including these variables is their documented association with EF outcomes (Aarnoudse-Moens, Smidts, et al., 2009; Ardila, Rosselli, Matute, & Guajardo, 2005; Jurado & Rosselli, 2007; Rosselli & Ardila, 2003). The impact of these covariates varies across studies. They can be unique predictors of EF performance in one study (Ford, et al., 2011), but have no predictive value in another (Aarnoudse-Moens, Smidts, et al., 2009). Although evidence shows that these covariates do not account for the entire EF impairment found between those born preterm/LBW and their peers, studies that control for these factors would likely provide a more accurate estimation of the magnitude of differences genuinely attributable to birth status.

Global IQ scores have been included as a potential confounding variable in the EF studies reviewed earlier (Aarnoudse-Moens, Smidts, et al., 2009; Luu, et al., 2011; Nosarti, et al., 2007). The reason to control for IQ was to ensure that any impairment discrepancy found between clinical groups and their peers was not attributable to any pre-existing intelligence differences. As indicated, controlling for IQ does narrow the gaps in the discrepancy of the basic information parameters between the groups (Aarnoudse-Moens, Smidts, et al., 2009). This could be interpreted in several ways. It could be that the EF component in question could explain differences in IQ but still has variance left over. It could also be that controlling for IQ makes no difference because the EF component itself is not related to group differences in IQ. However, researchers have also argued that analysing IQ as a covariate is inappropriate when IQ itself may be meaningfully and intrinsically related to the clinical condition (Dennis et al., 2009;

Miller & Chapman, 2001). For example, Miller and Chapman (2001, p. 45) specified that “IQ would be very likely to be meaningfully related to brain damage, so using IQ as a covariate would disrupt any comparison of brain-damaged and control groups’ performance: IQ differences would almost certainly be *part of* group differences in brain-damage status. As a consequence, removing variance associated with IQ would alter the diagnostic group variable substantively.”

Relationships between Working Memory and Fluid Intelligence

There has been strong parallel research that supports the robust relationship between WM and G_f . Some focus on investigations of WM as part of EF and compare predictive power across the three aforementioned basic components of EF towards G_f (Duan, Wei, Wang, & Shi, 2010; van der Sluis, de Jong, & van der Leij, 2007). Others have provided evidence on specific components within the WM model and their respective relations to G_f (Engel de Abreu, et al., 2010; Tillman, Nyberg, & Bohlin, 2008).

Miyake et al.’s (2000) study of the unity and diversity of EF and their proposed three latent variables of EF (WM, inhibitory control, and shifting), described previously, has driven much of EF-related research. Friedman et al. (2006) replicated Miyake et al.’s (2000) model and tested which of the three EF constructs best predicted G_f . In their study, Raven’s Progressive Matrices (RPM) and WAIS Block Design subtest were used to measure G_f , while a multiple-choice vocabulary subtest and the WAIS Information subtest were used as measures for G_c . Using 234 young adults, aged 16 to 18 years, and applying structural equation modelling (SEM), whereby inter-EF correlations were accounted for, the authors concluded that WM was the only EF construct that significantly ($ps < .001$) accounted for both fluid and crystallized

intelligence measures, with explained variance ranging from 37% to 45%. Results indicated that WM updating predicted both G_f and G_c in similar degrees, with path coefficients of .74 and .79 respectively. The authors also pointed out that there remained unexplained variance of 49% to 57% in intelligence (Ackerman, Beier, & Boyle, 2005; Hornung, Brunner, Reuter, & Martin, 2011). These findings imply that not all EF processes were independently related to intelligence, but WM showed the strongest correlation with intelligence. Furthermore, EF did not fully explain variability in intelligence.

Several published articles have tested how well EF, using Miyake's (2000) model, predicts G_f in children, (Brydges, et al., 2012; Duan, et al., 2010; Lehto, et al., 2003; van der Sluis, et al., 2007). All of these used either Raven's Progressive Matrices (RPM) or the Cattell Culture Fair Intelligence Test (CCFIT) as measures of G_f . However, the use of EF tasks varied extensively. For example, Duan et al. (2010) tested 61 Chinese children with an average age of 11.88 years and confirmed the existence of three distinct, yet correlated, constructs of EF. The measures they used included two 2-back tasks; two Go/no-go tasks; and a local global task as well as digit shifting for measures of WM, inhibitory control and cognitive flexibility, respectively. Their investigation through SEM indicated a significant path coefficient between WM and G_f and a shared variance of approximately 35% ($p < .01$). However, neither inhibitory control nor shifting predicted significant additional variance in G_f and their documented shared variances with G_f were at 19% and 7% respectively. Although this study assisted in depicting a strong relationship between WM and G_f , the reliabilities of their measures were questionable. As reported in their article, all their measures were modified versions of the original task and no reliability figures were provided for reference.

Van der Sluis et al. (2007) were unable to replicate the three constructs proposed

by Miyake (2000) when testing 172 children aged between nine to 12 years. Only two components, WM and cognitive flexibility, were detectable, but not inhibitory control. Nonetheless, they reported that WM accounted for 15.1% of variance in Raven's Standard Progressive Matrices (RSPM). Although WM showed the strongest relation with G_f , the percentage of variance explained was less than those indicated in other EF or WM studies.

Brydges et al. (2012) also tested Miyake's (2000) EF model and sought to replicate Friedman et al.'s (2006) findings in the child population. The authors tested 215 children, between the ages of seven and nine, on measures of inhibitory control, working memory, shifting, as well as G_f and G_c . EF tasks used included the Stroop task, Go/no-go task, and the compatibility reaction time task for inhibition; letter-number sequencing, backward digit span, sentence repetition for WM; and the WCST, verbal fluency, and the letter monitoring task for shifting. The majority of selected tasks were from subtests of the WISC-IV, the British Abilities Scale, and the NEPSY. Results indicated that there was only one general construct to EF rather than the suggested three constructs. In their SEM model, EF was strongly predictive of both G_f and G_c at 80% and 69% variance respectively. Similar to Friedman et al. (2006), there remained unexplained variance of 20% and 31% in G_f and G_c respectively. It was noted that raw scores were used across the whole age range in Brydges et al.'s (2012) study, which could possibly lead to correlations being inflated by the effect of age-related changes. Another possible interpretation of the age-related differences in structural models maybe in terms of the differentiation hypothesis (Deary et al., 1996), also termed 'the law the diminishing returns' by Spearman's (1927). As confirmed by many other authors, the g factor is stronger and accounts for more variance in mental tests when assessed in young children and when cognitive ability levels are low compared to adults

and populations of high cognitive ability (Deary, et al., 1996; Detterman & Daniel, 1989; Jensen, 2003). Nonetheless, Brydges et al.'s (2012) documented invariance testing results that indicated that there were no structural differences in EF and intelligence amongst the age groups. As well, their choice of EF tasks appeared more reliable than those used in previously reviewed papers as all tasks were taken from reliably established assessment batteries. In sum, the aforementioned findings suggest that EF does indeed have a strong association with G_f and, when components of EF are identified, WM and G_f had the most consistent association in children.

Several meta-analyses have been conducted to examine the relationship between WM and G_f in adults. Ackerman, Beier, and Boyle (2005) examined 86 studies and reported a shared variance of 25% ($r = 0.48$) for the two constructs. However, a re-analysis of the published studies in Ackerman et al.'s (2005) using the SEM approach revealed a much higher true correlation ($r = 0.85$) between the latent constructs of WM and G_f (Oberauer, Wilhelm, Schulze, & Suß, 2005), indicating a shared variance of 72%. Several methodological refinements were used in Oberauer et al.'s (2005) study that likely brought about the discrepant results. In particular, Oberauer et al. (2005) commented that Ackerman et al.'s (2005) inclusion criteria for their selection of tasks was unclear and did not adequately represent the construct of interest. Furthermore, Ackerman et al. (2005) used a fixed-effects model to conduct their study, while Oberauer et al. (2005) used a random-effects model. Oberauer et al. (2005) argued that a meta-analysis consists of participants that originated from various different populations, therefore a random-effects model was more appropriate. Their use of analysis method was also different. Oberauer et al. (2005) used SEM instead of aggregation of indicators and correlational analysis with correction for attenuation as in Ackerman et al.'s (2005) study. The relationship between latent variables, excluding error variance, is arguably

more accurate. Subsequently, a re-analysis of 10 latent variable studies, also used in Ackerman et al.'s (2005) meta-analysis, was conducted and yielded a high correlation between WM and G_f ($r = 0.72$) and a shared variance of approximately 50% (Kane, Hambrick, & Conway, 2005).

Given that WM emerges as a distinct construct relatively early in development and contributes to variance in G_f for both adults (Engle, et al., 1999; Shelton, Elliott, Matthews, Hill, & Gouvier, 2010) and children (Engel de Abreu, et al., 2010; Fry & Hale, 2000; Hornung, et al., 2011), the WM- G_f link will now be explored within the multi-component WM theory. According to Engle and Kane (2004) and Conway et al. (2003), WM and G_f are closely related due to their dependence on executive attention control abilities, while others believe that the STM storage capacity, rather than the executive attention control, better explains WM- G_f link (Chuderski, Taraday, Necka, & Smolen, 2012; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Colom, Shih, Flores-Mendoza, & Quiroga, 2006).

A general consensus has been established that WM and G_f , although clearly not isomorphic, are closely associated in children. Earlier reviews have documented correlations between WM and G_f ranging from 0.64 to 0.82 (Fry & Hale, 2000). Further investigations also investigated contributions of specific WM components, such as the STM and executive attention control, in predicting individual differences in G_f in children. For instance, some authors claim that executive control was significant in predicting G_f while STM did not (Bayliss, Jarrold, Baddeley, Gunn, & Leigh, 2005; Engel de Abreu, et al., 2010; Swanson, 2008). In particular, Engel de Abreu et al. (2010) tested 119 children, ranging from kindergarten to Year 2 in school. Measures used in the study included the counting recall and backwards digit recall tasks for complex memory, digit recall and non-word repetition tasks for short-term memory, and

Raven's Coloured Progressive Matrices Test (RCPM) for G_f . The authors concluded that the complex span tasks, as measures of executive attention control, were better predictors of performance on the G_f task than were the STM tasks. They also suggested that complex span tasks rely heavily on executive attention control to effectively complete goal-directed behaviour and inhibit interference to elicit appropriate responses, which was in accordance with Kane et al. (2001).

In divergence from the above, Tillman, Nyberg, and Bohlin (2008) suggested that it was not a question of either/or but that both storage and processing components of WM were significant and influential contributors to the development of higher cognitive reasoning. Nevertheless, more evidence from both children and adults is needed in order to provide a clear understanding on how these components contribute to individual differences in G_f as it is likely that the developmental trajectory across different age groups and variations in the use of measurement tools may influence these findings. Moreover, concurrent comparison of each component in the same sample would afford a more sophisticated opportunity to evaluate relative contribution.

Summary

In conclusion, on global IQ scores, preterm/LBW individuals score lower than their full-term peers by approximately 10-12 points. However, studies not only show that the performance of individuals born preterm/LBW fall within normal range on global IQ score but that they display substantial academic difficulties at school which could continue to later life. It appears that global IQ cannot readily explain the full extent of difficulties of this clinical cohort. Inspection of domain-specific abilities from the traditional IQ test suggests that the children born preterm/LBW display greater deficit in fluid abilities than crystallized abilities. However, no published evidence has been presented to date regarding the investigation of G_f amongst children born

preterm/LBW. The investigation through G_f is relatively novel in the current literature and therefore benefit from further examination.

The literature has also provided ample evidence on the effect of prematurity/LBW on the basic processes within EF. In particular, investigation of the normal developmental trajectories of EF shows that WM and cognitive flexibility are two particularly rapidly changing components in school-aged children. In addition, children born preterm/LBW display relatively large difficulty on tasks measuring WM performance compared to other information processing parameters. While persistent WM impairments are found in school-aged children born preterm/LBW, results are inconsistent for cognitive flexibility. The inconsistency may be due to the use of different measurement tools.

A close relationship has been demonstrated between G_f and EF-related information processing parameters, with WM being viewed as a strong predictor of G_f . However, whether the effect of prematurity in G_f impairments is attributable to these EF-related information processing parameters is still unknown.

CHAPTER 4

Study 1: Effects of preterm birth on fluid intelligence: Investigating working memory and cognitive inflexibility as potential mediators

Introduction

Psychologists have used global IQ measures to examine differences in general intelligence in children born preterm/LBW. Global IQ measures are an obvious starting point because of their ability to predict academic achievement (Deary, et al., 2007; Rohde & Thompson, 2007). From the literature review, children born preterm exhibit at least a 10-12 point decrement in FSIQ, however, their mean scores remain within the average range (Bhutta, et al., 2002; Kerr-Wilson, et al., 2011). Despite this consistent finding, low academic achievement has also continually been documented in children and adults born preterm/LBW (Bowen, et al., 2002; Johnson, et al., 2011; Mulder, et al., 2010; Roberts, et al., 2011; Taylor, et al., 2011). Moreover, neurodevelopmental evidence has shown that children born preterm/LBW perform less well than their peers in the basic information parameters of executive function, including inhibition, WM and, possibly, cognitive flexibility.

Previous research has taken a componential approach, which assumes that information processing components are building blocks of higher cognitive abilities (Sternberg, 1981). For example, evidence indicates that differences between birth groups on global IQ impairments can be fully accounted for by basic information processing parameters (Rose, et al., 2011). However, this evidence does not adequately explain the discrepancy between normal range global IQ scores and consistent reports of poor academic achievement.

Therefore, the examination of theoretically based cognitive constructs, such as G_f , may be fruitful in assisting the understanding of this cognitive discrepancy. Past research has presented us with some success in other contexts. For example, Duncan et

al. (1995) resolved a similar issue in the adult literature where patients with damage to their prefrontal cortex, causing executive function impairments, presented with ‘average’ scores on global IQ measures but exhibited compromised G_f performance. He claimed that the underlying mechanism of the g factor is most appropriately examined through measures of G_f . Research on children with developmental disorders, such as ADHD, reveals a similar pattern (Barkley, 1997; Tamm & Juranek, 2012). Given the pattern of cognitive abilities shown by children born preterm/LBW, perhaps the consideration of G_f will be worthy of investigation.

Factor analyses have documented that G_f and FSIQ have strong loadings on g in the general population, (Gustafsson, 1984; Keith, et al., 2006; Kvist & Gustafsson, 2008). If this is due to their isomorphism, then the clinical cohort is likely to perform lower on G_f measures than their peers, but only to the extent that their FSIQ is lower. However, there has only been preliminary documentation of deficiencies in G_f amongst children born preterm/LBW. Whether the clinical cohort displays developmentally lagged performance on G_f as compared to their peers is also not well documented. If G_f is indeed impaired, it is still unclear whether birth status difference in G_f deficits is attributable to EF in general or specific components or combinations of both.

According to the literature on typical development, the greatest changes affecting children of school age also occur in WM and cognitive flexibility, both of which develop rapidly during early childhood, rather than inhibition (Best, et al., 2009). Studies on how well EF predicts G_f also suggest that WM is a more consistent predictor of G_f than the other two widely identified EF components (Fry & Hale, 2000; Kane, et al., 2005; Kyllonen & Christal, 1990; Shelton, et al., 2010), and that it may possibly play a causal role in G_f (Jaeggi, et al., 2008; Klingberg, et al., 2002; Studer, et al., 2009). In addition, WM has been shown to have strong associations with academic

achievement, particularly in math and reading (Alloway, 2009; Alloway & Alloway, 2010; Alloway, Gathercole, Kirkwood, & Elliot, 2009; Maehler & Schuchardt, 2009), which have been identified as areas of difficulty in children born preterm/LBW (Pritchard, et al., 2009; Roberts, et al., 2011).

Cognitive flexibility also appeared to be impaired in children born preterm/LBW. The ability to switch between stimulus-response rules in simple tasks typically emerges around three to five years of age in normal development. However, preschoolers may struggle in tasks that require switching between more than one set of stimulus-response rules, for example matching with color is switched to matching with shapes within the same task (Espy, 1997). Older children may continue to struggle with complex rules, but a sharp improvement in ability to withhold perseverative behavior has been suggested to take place between seven and nine years of age (P. Anderson, 2002). Provided with this information, the comparison of cognitive flexibility between children born preterm and their same aged peers between seven to nine years old may shed light on developmental trajectory differences.

Studies of cognitive flexibility in preterm/LBW samples have also demonstrated rather mixed results. Variations in measures used, age of sampling group as well as sample size have made it difficult to compare outcomes across studies. Evidence of deficiency in cognitive flexibility in children born preterm/LBW have been frequently documented using TMT-B task (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). However, the TMT-B task requires eye-hand co-ordination and motor skills from using pen and paper to connect testing items. It also uses the time to complete the task as the outcome variable (Reitan & Wolfson, 1992). WCST, on the other hand, is an alternative measure of cognitive flexibility that requires participants to match cards while completing the task without any time constraints (Heaton, et al., 1993). The

administrative difference between the two tasks may impede on test performance for young children and the TMT-B task may underestimate the shifting ability of those children that need more processing time. Therefore, using perseverative errors from WCST, as opposed to the TMT-B task, for investigating cognitive flexibility amongst children born preterm/LBW may benefit exploration of this issue (Mulder, et al., 2009).

Therefore, the current study will investigate the effects of prematurity on G_f and the role of two core information processing parameters, namely working memory and cognitive flexibility, in a group of seven to nine-year-olds. Measures used in the study will include the Cattell Culture Fair Tests for G_f , the digit span tasks for working memory, and the WCST for cognitive flexibility, respectively. Coordinators of the Project Kids Intellectual Development Study (PKIDS) have carefully selected the assessments used as suitable for this age group, representative of these core constructs and with lower construct overlap than other options.

The Current Study

The current study has three aims:

- 1) To test whether children born preterm show G_f impairments as compared to their same age peers;
- 2) To ascertain whether children born preterm show WM and cognitive flexibility deficits relative to their same age peers; and
- 3) To determine whether birth group differences in G_f between preterm and typically developing children are attributable to WM and/or cognitive flexibility.

Hypotheses

- 1) Given similarities in cognitive profile identified in adults with frontal lobe damage and children with ADHD, it is predicted that the preterm cohort will

perform less well than their same age peers in the G_f task.

- 2) If information processing parameters constrain G_f , it is predicted that the current sample of children born preterm will present with deficits in both WM and cognitive flexibility relative to their same age peers.
- 3) If WM is causally related to G_f , and children born preterm show deficits in both, it is predicted that WM will partially mediate birth group differences in G_f . If cognitive flexibility is causally related to G_f , and children born preterm show deficits in both, it is predicted that cognitive flexibility will also partially mediate birth group differences in G_f .

Method

Participants

The archival records of 362 children were considered for analysis in the study. Participants were recruited through PKIDS, which is an ongoing research program based at Murdoch University and previously at the Neuro-cognitive Development Unit (NDU) of the University of Western Australia (UWA). Participants to the program were recruited through primary schools within the metropolitan area of Perth city in Western Australia, within which participants born preterm were recruited through Kind Edward Memorial Hospital. Information packs were distributed to parents and interested parents and participants were contacted upon completion of consent forms.

The data set included typically developing children born between 1988 and 1991, who participated in PKIDS in 1998, and children born preterm born between 1990 and 1992, who participated in PKIDS in 1999. Table 4.1 displays the sample characteristics. One participant was excluded due to missing data for gestation weeks and birth weight, giving a final total of 217 participants (106 females, 111 males) in the preterm group. At the time of assessment, participants in the preterm group ranged from

seven to nine years old ($M = 8.31$, $SD = .70$) and consisted of those with birth weight ranging from 505g to 2760g ($M = 1331.42\text{g}$, $SD = 452.03$) and gestational age ranging between 23-34 weeks ($M = 29.35$, $SD = 2.50$). The comparison control group was made up of 145 typically developing children (63 females, 81 males), group-matched for age ($M = 8.43$, $SD = .92$). The two groups did not differ significantly in age or gender composition as indicated by a Mann-Whitney U test. Within the preterm group, there was a strong positive correlation between birth weight and gestational age, where lower birth weight was associated with shorter gestational age, $r = .75$, $p < .001$.

Table 4.1

Characteristics of LBW Preterm Group and Control Group

	LBW Preterm Group			Control Group		
	7-year olds ($n = 87$)	8-year olds ($n = 69$)	9-years olds ($n = 61$)	7-year olds ($n = 55$)	8-year olds ($n = 47$)	9-years olds ($n = 43$)
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Birth weight (g)	1250.32 (471.26)	1284.20 (385.68)	1500.49 (455.53)			
Gestational age (weeks)	28.66 (2.40)	29.22 (2.51)	30.48 (2.24)			
Age	7.60 (.26)	8.44(.32)	9.18 (.16)	7.45 (.28)	8.54(.27)	9.57 (.30)

A significant difference was found in birth weight between the preterm age groups, $F(2, 214) = 6.34$, $p = .002$. Scheffe's post hoc test showed that the 7-year-olds ($p = .004$, $d = 0.46$) and the 8-year-olds ($p = .022$, $d = 0.38$) were born significantly lighter than the 9-year-olds. However, the birth weight of 7- and 8-year-old children was not significantly different.

A significant difference was also found in gestational weeks between the age groups, $F(2, 214) = 10.53$, $p < .001$. Scheffe's post hoc test showed a similar pattern

where 7-year-olds ($p < .001$, $d = 0.62$) and 8-year olds ($p = .012$, $d = 0.41$) were born significantly earlier than the 9-year olds, with no difference found between the younger two groups.

Measures

Cattell Culture Fair Intelligence Test (CCFIT). The CCFIT (Cattell, 1973) was used to measure participants' G_f . It is a non-verbal assessment that has been commonly used in studies of intelligence (Brydges, et al., 2012; Colom, Abad, Rebollo, & Shih, 2005; Colom, Rebollo, Abad, & Shih, 2006; Dang, Braeken, Ferrer, & Liu, 2012).

The test is intended to measure cognitive ability with little dependence on knowledge, thereby minimizing the influence of cultural factors, educational background and verbal fluency. Individuals are required to identify figural relations through inductive reasoning. PKIDS used Scale 2 Short Form A, which consists of four timed subtests with 46 items to be answered within an allocated time of 12.5 minutes altogether. The item difficulties gradually increase within each subtest (Cattell, 1973). The four subtests are summed to produce a total score, which serves as the outcome variable in the present study.

According to CCFIT's manual (Cattell, 1973), the immediate test re-test correlation for Scale 2 Short Form A, using 650 students, was .73. The internal consistency (Cronbach's alpha) of the test, based on data from 3999 participants was .76. The test also showed criterion validity via a high correlation with another intelligence measure, the WISC, at .70. Form A also shares a variance of .66 attributable to g (Cattell, 1973).

Digit span task. The digit span task is a subtest included in the Wechsler Intelligence Scale for Children – Third Edition (WISC-III) as a measure of WM

(Wechsler, 1991). The subtest consists of a forward task and a backward task, both of which are commonly used in WM literature (Swanson, 2008; Tillman, et al., 2008) and amongst cognitive studies of prematurely born children (Clark & Woodward, 2010; Fraello, et al., 2011).

This task requires participants to verbally repeat sequences of numbers in the same order for digit span forwards and in reverse order for digit span backwards. The number of digits increases as examinees successfully recall two consecutive trials of a given length and testing is discontinued when examinees fail to correctly recall two strings of digits of the same length. The raw scores of digit span forwards and backwards are summed to indicate the number of correct trials for the task. This number is used as the outcome variable for the examinees' WM. Although it is possible to separate out the scores for digit span forwards and backwards respectively, recent studies have demonstrated that the combined score is also adequately suitable in representing WM as a single cognitive ability (Bowden, Petrauskas, Bardenhagen, Meade, & Simpson, 2013).

According to the WISC-III manual (Wechsler, 1991), the reliability coefficient for the digit span subtest, estimated by split half correlations for the specific age groups of 7- and 9-year-olds ranged from .81 to .84 ($n = 200$ for each age group).

Wisconsin Card Sorting Test (WCST). The WCST (Heaton, et al., 1993) measures participants' executive function (EF). In particular, it measures an individual's ability to carry out goal-directed activities together with planning, initiation, and resistance to perseveration.

In this test, participants are presented with four 'key' cards differing in three dimensions: colour, number and form. Participants are required to match each of the 128 stimulus cards to one of the key cards. The matching rules change after every ten

consecutive correct responses, however, the rule is not divulged to the participants.

Thus, the task requires participants to induce the tacit rule by attending to the feedback provided on their responses and change rules appropriately. The examiner provides feedback on each trial by saying ‘Right’ or ‘Wrong’ to inform the participants of their performance. The test terminates once participants complete categorizing on the three dimensions (colour, number, or form) two times, or after sorting all stimulus cards (Heaton, et al., 1993; Romine et al., 2004).

Of the various outcome measures available, raw score on perseverative errors from the WCST was chosen as the outcome measure. A perseverative error is defined as “a failure to shift category after receiving negative feedback from the previous trial” (Barceló & Knight, 2002, p. 353). The task has been known to assess the shifting component of EF. Therefore, a high score in this task reflects a lower level of shifting ability, also referred to as cognitive flexibility (Best, et al., 2009). For the ease of interpretation, this performance is subsequently labelled as ‘cognitive inflexibility’ in this study. Alternate-form reliability from a student sample ($n = 75$) ranged from .25 to .63 (Bowden et al., 1998). Although the test was originally intended for the adult population as a means to identify prefrontal cortical damage, extensive research has used WCST in children to determine its sensitivity and specificity, and has developed further normative data for it to be used in the child population (Bull & Scerif, 2001; Chase-Carmichael, Douglas Ris, Weber, & Schefft, 1999; Heaton, et al., 1993; Romine, et al., 2004).

Procedure

The study was granted ethical approval by the Human Research Ethics Committee of Murdoch University, Western Australia (see Appendix A). It employed a retrospective design where participants’ data were obtained from those who took part at

PKIDS in 1998 and 1999. The program was established in 1995 and runs every year as a two-week school holiday program. Children participating in the present study underwent two consecutive days of psychometric testing in a child-friendly environment. The tests used in this study are a subset of a larger assessment battery. Each testing session lasted no longer than 30 minutes in recognition of the limited attention span of typically and atypically developing primary school aged children. Well-trained post-graduate psychology students administered all the measures in accordance with standardized administration procedures. The WCST and digit span task used in the present study were administered individually, while CCFIT was administered in small groups.

To maintain high engagement levels, standardized psychometric assessment was presented to children as solving ‘puzzles’ in return for tokens. Children were able to exchange their tokens for materials used in building a rocket and a planet in the context of an outer-space alien theme story. During the day, children were also provided with game activities and meals during non-testing times. All children received a small gift and a certificate at the end of their two-day participation with PKIDS. The return rate of participants between Day 1 and 2 is more than 95%, suggesting that children experience the day as fun and engaging. The return rate for longitudinal return after 2 years is more than 85% with the majority of non returners being due to ill-health or families that have moved and could not be contacted.

Results

SPSS for Windows Version 17.0 was used for all data analyses. Results are divided into three sections. The first section details data preparation. The second section provides results for birth group differences on the three cognitive measures.

Developmental trends are also presented. Finally a mediation analysis is used to test

whether effects of birth status on G_f are mediated by measures of working memory and cognitive inflexibility, as well as whether these mediation effects differ according to age group.

Data Preparation

Assumptions were tested before proceeding with the analyses. The assumption of normality was assessed for each cohort and within each age group separately for each measure, using a combination of the Skewness and Kurtosis statistics and visual inspection of histogram and normal Q-Q plots. If inspection of normal Q-Q plots showed observed sample data points clustering tightly along the normal distribution line with only occasional minor deviation, normality was assumed (Field, 2009c; Tabachnick & Fidell, 2001b).

For the digit span task, the assumption of normality was met for the distribution of scores for all ages in both groups. On the CCFIT, no violations of normality were found for the preterm cohort but some were found within the control group, particularly the 8-year-old group was slightly negatively skewed. Although absolute normality could not be assumed, inferential statistics used in the analysis are relatively robust to violations of this assumption (Tabachnick & Fidell, 2001a). The assumption was, however, severely violated on the WCST measure of perseverative errors, with the distributions of both groups showing positive skew. A non-parametric Mann–Whitney U test was then used as a conservative test for birth group differences in WCST performance. Perseverative error scores were then log transformed (first, adding 1 to raw scores). These WCST transformed scores were used in all subsequent analysis. Although minor deviations from normality continued to exist, this would have a negligible impact on robust inferential statistics (Tabachnick & Fidell, 2001a).

The homogeneity of variance assumption, tested using Levene's Test, was met

for the digit span task but was violated for WCST ($p = .049$) and CCFIT ($p = .041$).

With moderate to large sample size and equivalent group size, ANOVA is robust to the violation of this assumption (Allen & Bennett, 2010). Since the current group sizes were not equivalent, as a precaution against this violation, a more stringent p level of .01 was used instead of the usual .05. All tests results reported in the following sections are two-tailed. When post hoc tests are required, the Scheffe's post hoc test is used, as it is most conservative and flexible (Tabachnick & Fidell, 2001c). Effect sizes are provided using Cohen's d and partial η^2 . The effect sizes of Cohen's d are interpreted following Cohen's guidelines where 0.20, 0.50 and 0.80 represent a small, medium, and large effect size, respectively (Cohen, 1992).

Birth Group Differences

Descriptive statistics for the two birth groups' performance on CCFIT, WM, and WCST are provided in Table 4.2.

Table 4.2.

Descriptive Statistics for Psychometric Test Raw Scores for Each Group

Variables	<i>n</i>	Preterm	<i>n</i>	Control
		<i>M (SD)</i>		<i>M (SD)</i>
<i>Gf</i> - CCFIT				
All ages	216	24.14 (6.43)	144	28.94 (5.62)
Age 7	86	21.02 (6.55)	55	26.44 (5.54)
Age 8	69	25.06 (5.45)	47	29.43 (5.86)
Age 9	61	27.51 (5.24)	42	31.67 (3.87)
WM – Digit span				
All ages	215	11.80 (2.92)	144	13.17 (2.76)
Age 7	87	10.68 (2.52)	55	12.18 (2.22)
Age 8	68	12.59 (3.04)	46	13.50 (2.83)
Age 9	60	12.53 (2.83)	43	14.07 (2.97)
Cognitive Inflexibility – WCST				
All ages	213	33.95 (25.77)	137	20.28 (12.94)
Age 7	85	43.69 (29.43)	51	22.69 (12.73)
Age 8	68	29.76 (21.12)	45	18.91 (12.34)
Age 9	60	24.88 (20.13)	41	18.80 (13.69)
Cognitive Inflexibility – WCST (after transformation)				
All ages	213	1.43 (.32)	137	1.25 (.28)
Age 7	85	1.53 (.34)	51	1.31 (.25)
Age 8	68	1.40 (.27)	45	1.21 (.29)
Age 9	60	1.30 (.31)	41	1.20 (.30)

Note. *Gf* = Fluid intelligence; CCFIT = Cattell Culture Fair Intelligence Test; WM = Working memory; EF = Executive function; WCST = Wisconsin Card Sorting Test.

Fluid intelligence: CCFIT.

A 2 x 3 ANOVA was used to test for group differences in the CCFIT raw scores between children born preterm and their typically developing peers in each age group. Analysis indicated that there was a significant main effect of birth group, $F(1,354) = 58.29, p < .001$, partial $\eta^2 = .14$, with the control participants performing better than their preterm born peers. As expected, a significant main effect of age group was also found, $F(2,354) = 32.35, p < .001$, partial $\eta^2 = .16$, indicating that older

participants performed better than younger ones. Post hoc analysis indicated that the 9-year olds scored significantly higher than the 8-year olds ($p = .008$), and the 7-year olds ($p < .001$), and the 8-year olds also scored significantly higher than the 7-year olds ($p < .001$). The birth group x age group interaction effect was not statistically significant. Mean scores for the preterm and control groups across the three age groups are presented in Figure 4.1.

One-way ANOVAs specified that the preterm group performed less well than their matched controls with a medium to large effect size for each separate age group (Cohen, 1992). Refer to Table 4.3 for inferential statistics and effect sizes.

Table 4.3

Inferential Statistics and Effect Sizes for the Measures of Cognitive Function across Each Age Group

Variables	<i>F</i>	η^2	Cohen's <i>d</i>	Contrast
<i>Gf</i> - CCFIT				
All ages	(1,358) = 53.01***	.13	0.77	PT < T
Age 7	(1,139) = 25.80***	.16	0.86	PT < T
Age 8	(1,114) = 16.90***	.13	0.77	PT < T
Age 9	(1,101) = 19.22***	.16	0.87	PT < T
WM – Digit span				
All ages	(1,357) = 19.73***	.05	0.47	PT < T
Age 7	(1,140) = 13.12 ***	.09	0.61	PT < T
Age 8	(1,111) = 2.52	.02	0.30	PT < T
Age 9	(1,102) = 7.28**	.07	0.53	PT < T
Cognitive Inflexibility – WCST (after transformation)				
All ages	(1,348) = 29.00***	.08	0.58	PT > T
Age 7	(1,134) = 16.67***	.11	0.71	PT > T
Age 8	(1,111) = 12.38**	.10	0.67	PT > T
Age 9	(1,99) = 2.83	.03	0.34	PT > T

Note. PT = Preterm birth group; T = Term group.

** $p < .01$, *** $p < .001$.

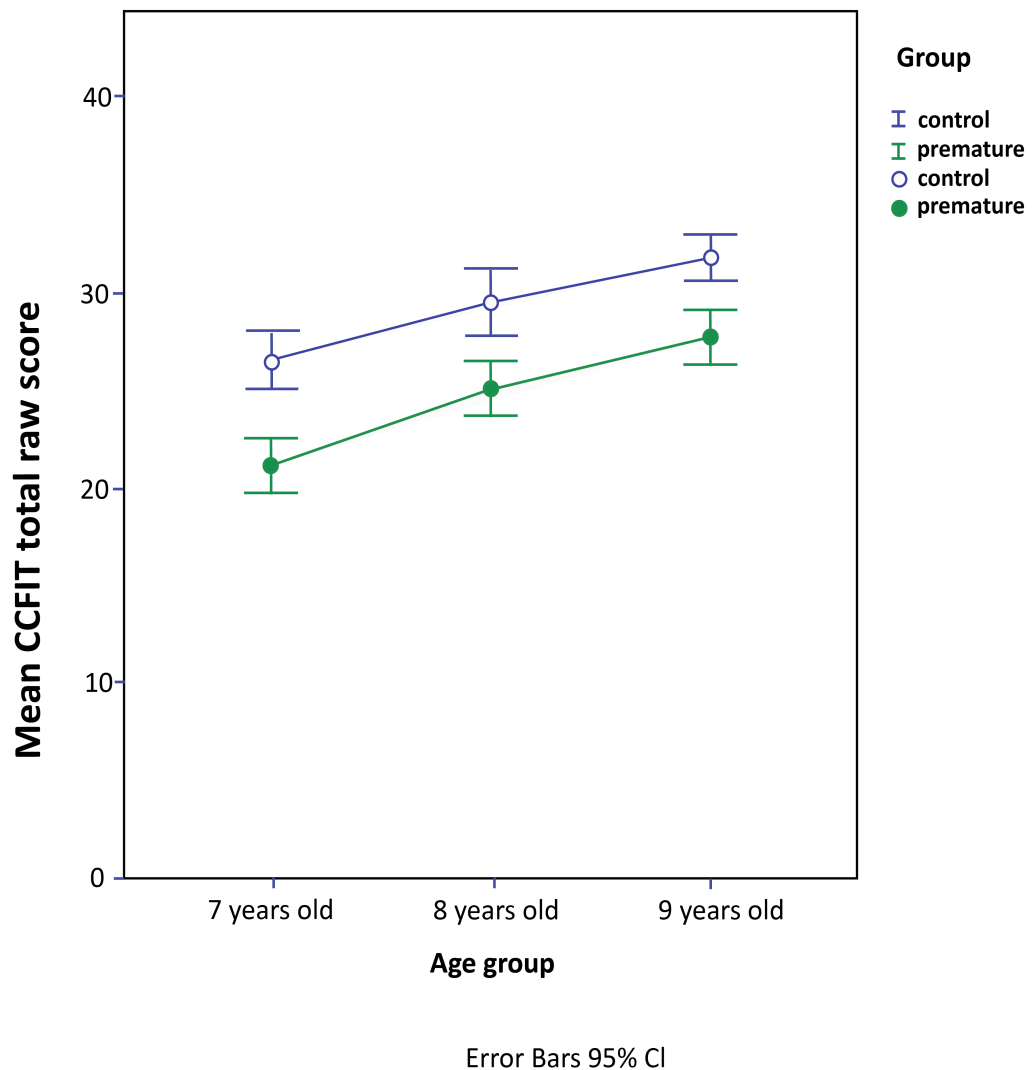


Figure 4.1. Mean CCFIT total raw scores for each of the preterm and their control age group.

Working memory: Digit span task.

A 2 x 3 ANOVA demonstrated that there was a significant effect of birth group on WM, $F(1,353) = 19.72, p < .001$, partial $\eta^2 = .05$, with the control group participants outperforming their preterm born peers. The effect of age group was significant, $F(2,353) = 16.81, p < .001$, partial $\eta^2 = .09$. Post hoc analysis showed that 9-year-old children ($p < .001$) and 8-year-old ($p < .001$) children performed significantly better than 7-year-old children, but the two older groups did not differ significantly from each

other. A follow up simple effects analysis for each birth group showed that the pattern of significance was the same for each age group. The birth group x age group interaction effect was not statistically significant. A visual inspection of the trend can be viewed in Figure 4.2. However, one-way ANOVA showed a significant difference by birth status in 7-year-old and 9-year-old children in their performance on the digit span task in favour of the control group, whereas 8-year-old participants did not show a significant birth group effect. Effect sizes ranged from small to medium (Cohen, 1992). Detailed inferential statistics are presented in Table 4.3.

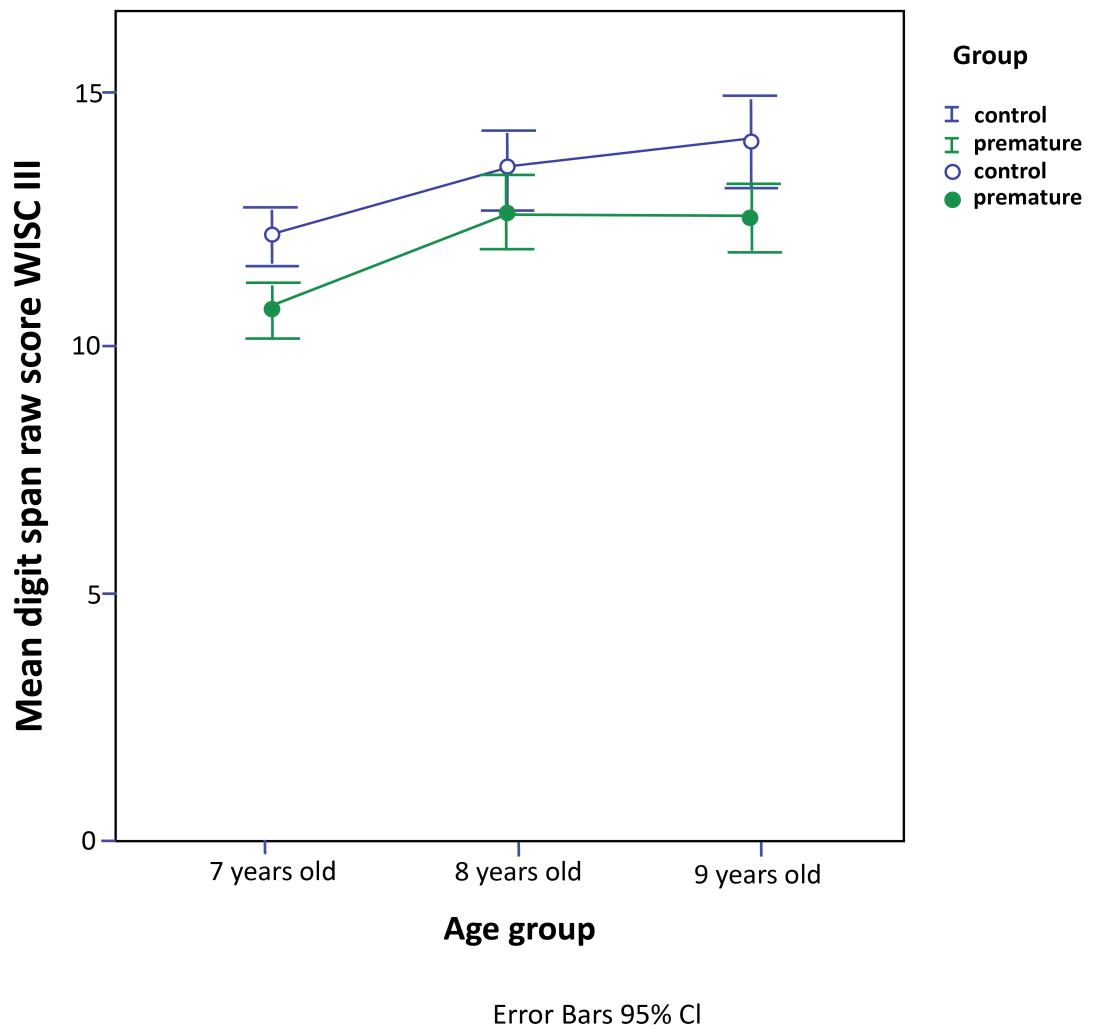


Figure 4.2. Mean digit span total raw scores for each of the preterm and their control age group.

Cognitive Inflexibility: WCST.

A 2 x 3 ANOVA on perseverative errors on the WCST showed a significant effect of birth group, $F(1,344) = 27.29, p < .001$, partial $\eta^2 = .07$, with the preterm group producing consistently more errors than the control group. There was also a significant effect of age group, $F(2,344) = 9.68, p < .001$, partial $\eta^2 = .05$. Post hoc analysis revealed that 9-year-old children ($p < .001$) and 8-year-old ($p = .005$) children performed significantly better than 7-year-old children, but the two older groups were not significantly different from each other. A follow up analysis showed that this pattern of age group significance was present in the preterm group but not in the typically developing children. In children born preterm/LBW children, 7-year olds made more perseverative errors on the WCST than their 8-year-old ($p = .034$) and 9-year-old peers ($p < .001$), but the older two groups did not differ. In the typically developing control group, no significant difference was found in the age group comparisons. Nevertheless, a significant birth group x age group interaction effect was not found. Figure 4.3 presents their different cognitive trajectories.

One-way ANOVA, using transformed WCST scores, analyzed separate age group differences. It indicated that the control group performed better than the clinical cohort among 7- and 8-year-old participants but not the 9-year-old participants, with effect sizes ranging from small to medium (Cohen, 1992). Refer to Table 4.3 for statistics and a comparison of effect sizes with other measures across age groups.

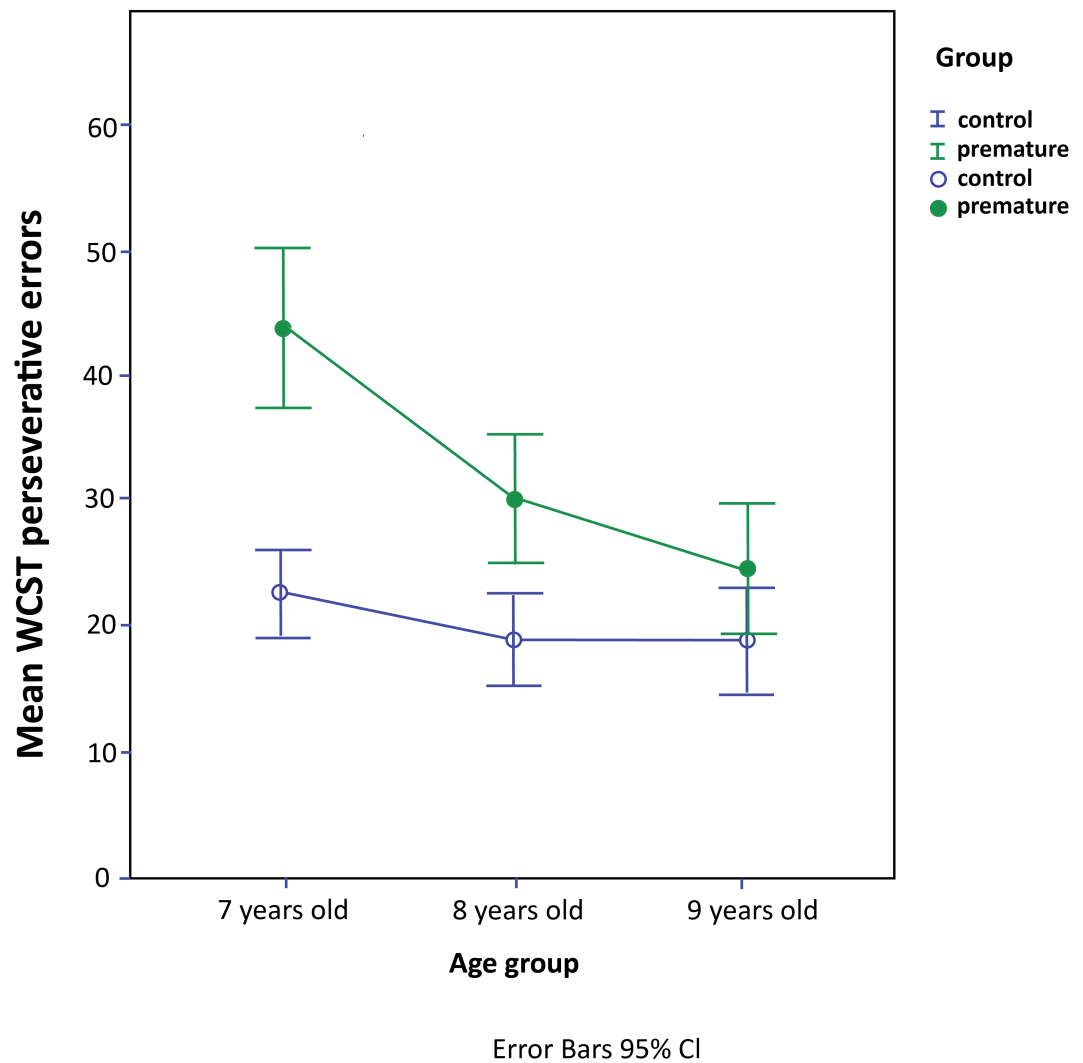


Figure 4.3. Mean WCST perseverative errors for each of the preterm and their control age group.

Mediation Analysis

A multiple mediation analysis was performed to explore the degree to which birth status effects on G_f were mediated by both digit span performance and WCST performance.

Preacher and Hayes' (2004, 2008) SPSS macro script, named PROCESS, on their website: <http://www.afhayes.com/introduction-to-mediation-moderation-and-conditional-process-analysis.html> (2013) provided both Normal Theory Sobel Test and

bootstrapping methods to test direct and specific indirect effects. All results for indirect effects produced by PROCESS generated 95% bias-corrected bootstrap confidence intervals. The number of bootstrap samples was set at 1000. This bootstrapping approach helps alleviate errors in normality assumption that the Sobel test makes when calculating indirect effects. The bootstrapping approach is now considered best practice over the commonly used Baron and Kenny (1986) approach, as the latter has been reported to have relatively low statistical power (A. F. Hayes, 2009; Hoyt, Imel, & Chan, 2008; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Preacher & Hayes, 2008).

In Figure 4.4, the c coefficient depicts the total effect of birth status on G_f before controlling for WM and cognitive inflexibility. The c' coefficient illustrates the direct effect of birth status on G_f after concurrently controlling for WM and cognitive inflexibility. The product of path coefficients $a1$ and $b1$ represent the indirect effect of birth status on G_f through WM. Similarly, the product of path coefficients $a2$ and $b2$ represent the indirect effect of birth status on G_f through cognitive inflexibility. The analyses were first conducted using all participants and then repeated for each age group separately. The data is dummy coded using 1 for the clinical cohort and 0 for typically developing children. Raw correlations and results are presented in Table 4.4 and 4.5 respectively.

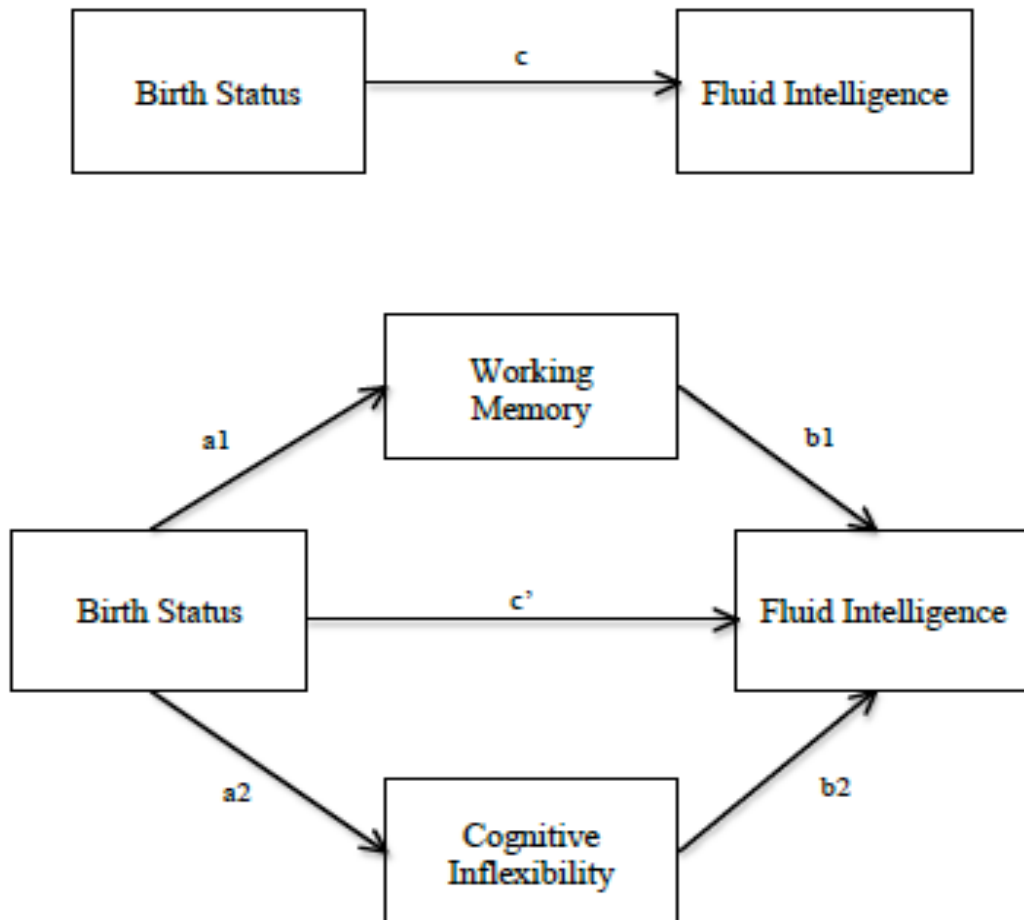


Figure 4.4. This illustrates the multiple mediation analysis path models for the current study. The first model presents a direct pathway and the second model presents indirect/mediated pathways.

Table 4.4

Raw Correlations between Variables used in Multiple Mediation Analysis

	1	2	3	4	5
1. Birth Group	-	-.02	-.36**	-.23**	.28*
2. Age Group		-	.38**	.29**	-.25**
3. CCFIT total score			-	.44**	-.47**
4. Digit Span raw score				-	-.32**
5. WCST (PE)					-

Note. CCFIT = Cattell Culture Fair Intelligence Tests; WCST (PE) = Wisconsin Card Sorting Test (log perseverative errors)

** $p < .01$

Table 4.5

Results of Multiple Mediation Analysis for Each Age Group as Predictors of Fluid Intelligence

Age Group	Birth Status							
	Total effect	Direct effect	Mediation by Working Memory			Mediation by Cognitive Inflexibility		
	c	c'	a1	b1	a1xb1 ^a	a2	b2	a2xb2 ^b
All ages <i>n</i> = 347	-4.90***	-2.85***	-1.30***	.67***	-.87***	.18***	-6.43***	-1.18***
Age 7 <i>n</i> = 136	-5.16***	-3.07**	-1.40**	.35	-.49	.22***	-7.14***	-1.60**
Age 8 <i>n</i> = 111	-4.78***	-3.22**	-.90	.48**	-.43	.20***	-5.78**	-1.14*
Age 9 <i>n</i> = 100	-4.23***	-2.89**	-1.43*	.74***	-1.06*	.10	-2.70*	-.27

Note. Unstandardized coefficients are presented.

^a Specific indirect effect for working memory.

^b Specific indirect effect for cognitive inflexibility.

p* < .05, *p* < .01, ****p* < .001

As presented in Table 4.5, both age-collapsed and separate age group analyses revealed significant total effect and direct effect of birth status towards G_f . The direct effect showed that the preterm cohort performed less well than their peers on CCFIT, after controlling for both WM and cognitive inflexibility performance concurrently. Both specific indirect effects were statistically significant for the combined age groups. According to the Sobel Test, there was a partial mediation effect through WM: $Z = -3.46, p < .001$, with a 95% bias corrected bootstrap confidence interval that did not include 0. Similarly, there was a partial mediation effect through cognitive inflexibility ($Z = -4.19, p < .001$), which was also verified by a 95% bias corrected bootstrap confidence interval that did not include 0. The total indirect effect (unstandardized),

which refers to the sum of both specific indirect effects, was -2.04, with a 95% bias corrected bootstrap confidence interval that did not include 0. Approximately 5% variance in WM is explained by birth status ($R^2 = .0475$, $F = 17.21$, $p < .001$), while 8% variance in cognitive inflexibility is explained by birth status ($R^2 = .0782$, $F = 29.27$, $p < .001$). 13% variance is accounted for from estimating G_f from birth status alone ($R^2 = .1343$, $F = 53.52$, $p < .001$). The two mediators and birth status together accounted for 36% variance in G_f ($R^2 = .3585$, $F = 63.90$, $p < .001$).

When analyses were conducted with each age group separately, results were different. A partial mediation effect of birth status on G_f through WM was not statistically significant for 7-year-olds ($Z = -1.42$, $p = .154$) or 8-year-olds ($Z = -1.31$, $p = .189$), but was statistically significant for 9-year-olds ($Z = -2.17$, $p = .030$). On the other hand, a partial mediation effect of birth status on G_f through cognitive inflexibility was statistically significant in both 7-year-olds ($Z = -2.95$, $p = .003$) and 8-year-olds ($Z = -2.41$, $p = .016$), but was not significant in 9-year-olds ($Z = -1.15$, $p = .248$). The total indirect effect (unstandardized), for the 7-year-old age group was -2.09, for the 8-year-old age group was -1.57, and for 9-year-old age group was -1.33. All results were verified with a 95% bias corrected bootstrap confidence interval that did not include 0.

For the 7-year-old age group, approximately 8% variance in WM is explained by birth status ($R^2 = .0776$, $F = 11.28$, $p = .001$), while 11% variance in cognitive inflexibility is explained by birth status ($R^2 = .1106$, $F = 16.66$, $p < .001$). 14% variance is accounted for from estimating G_f from birth status alone ($R^2 = .1441$, $F = 22.56$, $p < .001$). The two mediators and birth status together accounted for 29% variance in G_f ($R^2 = .2931$, $F = 18.25$, $p < .001$). As for the 8-year-old group, approximately 2% variance in WM is explained by birth status ($R^2 = .0216$, $F = 2.40$, $p = .124$), while 11% variance in cognitive inflexibility is explained by birth status ($R^2 = .1107$, $F = 13.56$, $p < .001$).

16% variance is accounted for from estimating G_f from birth status alone ($R^2 = .1589$, $F = 20.59$, $p < .001$). The two mediator variables and birth status together accounted for 31% variance in G_f ($R^2 = .3119$, $F = 16.17$, $p < .001$). Finally, for the 9-year-old age group, near 6% variance in WM is explained by birth status ($R^2 = .0580$, $F = 6.04$, $p = .016$), while approximately 3% variance in cognitive inflexibility is explained by birth status ($R^2 = .0255$, $F = 2.56$, $p = .113$). 16% variance is accounted for from estimating G_f from birth status alone ($R^2 = .1645$, $F = 19.29$, $p < .001$). The two mediators and birth status together accounted for 39% variance in G_f ($R^2 = .3882$, $F = 20.31$, $p < .001$).

In short, the negative relationship between preterm birth status and G_f , as measured by CCFIT, was partially mediated by both lower WM (digit span) and cognitive inflexibility (perseverative errors) concurrently, when the entire sample was examined. When analysed using separate age groups, again with the other mediator in the model simultaneously controlled, only WM was a significant partial mediator for participants in the 9-year-old age group, whereas only cognitive inflexibility was a significant partial mediator for participants in the 7- and 8-year-old age groups.

Discussion

This study examined the effect of preterm birth status on G_f through two basic information processes, namely WM and cognitive flexibility, in children aged between seven and nine years. There were several major findings. First, children born preterm showed impairments in G_f and in all the basic information processes measured. Second, multiple mediation modelling revealed that differences between the groups in G_f remained after concurrently controlling for both WM and cognitive inflexibility. The modelling also demonstrated that the two basic processes in preterm children partially mediated their deficits in G_f when age-collapsed analysis was performed. Finally, the aforementioned partial mediation results differed when the three different age groups

were analysed separately. These findings are discussed separately below.

Birth Group Differences

Fluid intelligence differences were found between children born preterm and their peers. The preterm group scored significantly lower on the CCFIT than their same age peers as hypothesized. The effect size was large across all age groups as compared with other measures in the study. The performance of both groups also improved with increasing age. Only preliminary findings have previously been documented for the performance of G_f in children born preterm/LBW, which stated that the preterm cohort showed significantly lower scores on the CCFIT (Davies, 2004). The current results corroborated those of Davies (2004) where effect sizes for group differences in G_f ranged from medium to large ($d = -0.50$ to 1.00). The effect sizes for group differences in G_f across children aged 7-to 9-years in the present study ranged from -0.77 to -0.87 . Given that the sample in this study overlapped with that of Davies' (2004) study, with both comprising participants from PKIDS, similar results were not surprising. However, the larger sample in the present study, as compared to Davies' (2004) further confirmed the extent of G_f differences between children born preterm and their same aged peers. In addition, the performance of children born preterm was also consistent with the evidence of strong positive correlations between G_f and the g factor (Gustafsson, 1984; Keith, 2005; Kvist & Gustafsson, 2008) as well as between FSIQ and the g factor (Jensen, 1998; Keith, et al., 2006). Essentially, prior evidence of FSIQ disparity between children born preterm/LBW may, therefore, also be reflected in measures of G_f . The present study's effect size for group differences in G_f is similar to those effect sizes reported in other studies for group difference in FSIQ that ranged from 0.62 to 0.96 (Davies, 2004; Orchinik et al., 2011).

Furthermore, studies that examined domain-specific abilities also confirmed that

both children and adults born preterm/LBW showed more difficulties in nonverbal reasoning tasks (Grunau, Whitfield, & Davis, 2002; Johnson, 2007; Pyhala, et al., 2011; Talge, et al., 2010). They revealed particularly large mean deficits in perceptual reasoning subtests in between group comparisons (P. J. Anderson, et al., 2003; Hallin, et al., 2010). Observations on the current pattern of CCFIT raw scores from Table 4.2 and error bars in Figure 4.1 not only corroborate with the aforementioned findings but also revealed a developmental lag of at least one year. In particular, 9-year old children born preterm in the sample displayed CCFIT raw scores slightly better than 7-year old typically developing children and worse than 8-year-old typically developing children. This developmental lag in G_f also replicated preliminary findings reported by Davies (2004).

Working memory impairments were identified in children born preterm relative to their same-age peers. The current finding was consistent with existing evidence in the literature pertaining to verbal WM, as measured by the digit span task (Aarnoudse-Moens, Smidts, et al., 2009; Ford, et al., 2011). The effect size in the current age-collapsed analysis ($d = -0.47$) was similar to that obtained in the recent meta-analysis ($d = -0.36$) (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). These results were also expected as domain-specific studies have revealed that those born preterm/LBW generally perform particularly poorly on the FDI/WMI in comparison with the remaining three indexes that form FSIQ (P. J. Anderson, et al., 2003; Hallin, et al., 2010). However, their performance at eight years of age did not significantly differ from their peers. These results appear inconsistent with previous research that suggests WM impairments found in young children born preterm/LBW (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009; Ford, et al., 2011) persist to early adulthood (Hallin, et al., 2010). Nonetheless, a comparison for separate age groups showed that the effect size for the 8-

year-old age group ($d = -0.30$) is relatively small (Cohen, 1992) compared to the other age group findings in this study.

A significant difference found between age groups confirmed that there was indeed a relatively large improvement in WM for both birth status groups from age seven to eight. Children after the age of seven begin to adopt more sub-vocal rehearsal strategies (Gathercole & Hitch, 1993; Pickering, Gathercole, & Peaker, 1998), where verbal rehearsal of auditory information can assist in short term memory storage. Digit span in the study is reflective of verbal working memory ability and the results may be related to the emergence of such strategic processes thus leading to a developmental improvement in task performance for both groups. Perhaps typically developing children continue to develop after this initial strategy switch while children born preterm require more time to consolidate the switch and move on to further improvements. As results further indicated, the performance on the digit span task of the 9-year-old children born preterm was virtually identical to 8-year-old children born preterm. This explains the re-emergence of a significant between-group difference on the digit span task at nine years. However, while this explanation has theoretical integrity, it could also be an issue of random sampling.

Nonetheless, the pattern of digit span raw scores from Table 4.2 and overlapping error bars in Figure 4.2 are consistent with a developmental lag of at least one year, similar to the performance differences in CCFIT. The 8-year old children born preterm display digit span mean scores comparable to those of 7-year old typically developing children indicating substantial delay.

Cognitive inflexibility was found in the preterm group as they demonstrated significantly more perseverative errors on WCST as compared to their control peers. The effect size in the current age-collapsed analysis ($d = -0.58$) was stronger than that

obtained in the recent meta-analysis ($d = -0.49$) (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009), even though this study used a sorting task and Aarnoudse-Moens et al.'s (2009) meta-analysis only included studies using TMT-B, thought to be more sensitive to effects of preterm birth, as the measure of cognitive flexibility.

The current findings corroborate some earlier documented results (Aarnoudse-Moens, Smidts, et al., 2009; Nosarti, et al., 2007), but conflict with others (Aarnoudse-Moens, et al., 2012; Luu, et al., 2011). Of particular interest was that the current results diverged from the conclusions drawn in Mulder et al.'s (2009) meta-analysis, which documented significant differences between the preterm cohort and their peers only when TMT-B was used as a measurement tool, as opposed to the use of WCST. However, separate age group analyses showed a decrease in the number of perseverative errors as children age, which indicates an improvement in performance on the WCST task. Group differences were no longer significant for this study's sample of 9-year-old preterm children. It appears that these results demonstrated a trend of increase in performance and that significant group differences diminished as the children aged. These results did not support the notion that impairments in the cognitive flexibility process of EF persist (Nosarti, et al., 2007), but rather the interpretation of a "catching up" trend. A maturational lag of approximately two years is observed in the pattern of WCST perseverative errors transformed raw scores from Table 4.2 and error bar overlaps in Figure 4.3. The 9-year-old children born preterm displayed mean perseverative errors comparable to those observed in 7-year-old typically developing children. The certainty of any "catching up" trend would have to be confirmed from longitudinal studies in the future.

Mediation Analysis

Very few studies have attempted to explain the general intelligence disparity

between preterm/LBW and their full-term peers through basic processes. One of the first was performed by Rose and her colleagues (2011), in which they suggested that information processing variables, including speed, attention, memory and representational competence, completely mediated IQ differences between their preterm/LBW and full-term samples of 11 years of age. The current multiple mediation analysis provided an alternative investigation by examining G_f and its concurrent relationships with basic information process in a cohort of children born preterm. The close relationship between G_f and WM has been well-established and it might seem obvious that if WM predicts G_f , and both WM and G_f impairments are found in children born preterm then the reason for G_f deficiency is poor WM. However, this is not a necessary logical consequence because the causes of between group differences can be different from causes of within-group differences.

From the current analysis, where all other mediators in the model were controlled simultaneously, it was concluded that there was a significant direct effect of birth status on G_f whether data were analysed combining age groups or treating them separately. Added to that, both WM and cognitive inflexibility were significant partial mediators of the path from birth status to G_f in the age-collapsed analysis. Further analyses by separate age groups indicated that WM was a significant partial mediator only for participants in the 9-year-old age group and cognitive inflexibility was a significant partial mediator only for participants in the 7- and 8-year-old age groups. As children move from less to more differentiation of EF, the impact of each component on G_f changes.

Although the present study's approach was compatible with Rose et al.'s (2011) study in which basic parameters are theorised to serve as building blocks for higher cognitive abilities closely related to g , measured in FSIQ or G_f tasks, direct comparison

of their study and the current study was difficult as the use of measures and age groups of participants were different. While the current study adopted simultaneous mediation modelling, Rose et al.'s (2011) study developed a cascade model through an SEM approach. The divergence of the present results also stems from the less comprehensive coverage of basic information processing parameters than that of Rose et al. (2011). Aside from the difference in coverage of measuring tasks used, the two studies also had no overlap in the use of measures. Another divergence was the inclusion of three consecutive age groups rather than one, as in Rose et al. (2011). Since measures of basic information parameters are often affected by developmental trajectories at different age (Best, et al., 2009), the current study was able to provide comparisons on how birth status effects on G_f were accounted for by WM and cognitive inflexibility differently at each age group. The current analysis is the first to demonstrate WM and cognitive inflexibility as significant partial mediators of birth status effect on G_f by means of multiple mediation analysis. The results of the mediation analysis established that G_f impairments in children born preterm should not be neglected, as significant differences in G_f remained present throughout the three age groups over and above the significant mediation effects detected.

Furthermore, the current analysis of significant mediation effects for WM and cognitive inflexibility simultaneously indicated that, in these age groups, the two components were not equivalent and were working separately. Although developmental research has suggested that EF may be viewed as a unitary construct for typically developing children ranging from 7- to 9-year-olds (Brydges, et al., 2012), the current results on children born preterm were in agreement with typical developmental trajectory evidence of separate EF components in children (Lehto, et al., 2003). Whether EF is viewed as a unitary construct or as factor components still seems worthy

of debate (Duan, et al., 2010; Miyake, et al., 2000; Wiebe, et al., 2008).

Limitations

A number of limitations should be considered when interpreting the current findings. First of all, the use of a single measurement tool to assess one information processing component might not be the ideal approach given that task impurity and reliability of EF measures have been raised as methodological concerns (Best, et al., 2009; Burgess, 1997; Rabbitt, 1997). In regards to task impurity, for example, the WM measure in this study is a combination of verbal storage and executive control. This fits well with some theories where short-term storage capacity and attentional control are conceptually linked (Engle, 2002; Heitz, et al., 2005; Kane, et al., 2001) but less well with others that claim WM to have distinct multiple facets (Baddeley & Hitch, 1974). There continues to be debate over which aspects of WM are most important in predicting G_f (Engel de Abreu, et al., 2010; Hornung, et al., 2011). Nevertheless, studies have pointed out that even simple span tasks, such as digit span, are capable of tapping executive attention control (Unsworth & Engle, 2006). The present study used the combined score of forward digit span and backward digit span that perhaps is more sensitive to short-term storage capacity, but since no task is a pure measure of either short-term storage or attentional control, conclusions can only be drawn about participants' ability in general WM but not whether short-term storage or attentional control was more predictive of birth status effects on G_f . A possible methodological refinement would be to include more than one measure for each cognitive construct, as well as including measures that are distinctively sensitive to short-term storage and attentional control respectively. The use of a latent-variable approach would also increase the strength of the study. Despite the aforementioned concerns, previous empirical researchers (Clark, et al., 2010; Friedman, et al., 2006; Mulder, et al., 2009)

have commonly treated the measures utilized in this study as valid and reliable assessments tools for measuring the constructs in question.

Secondly, the present study was unable to control for confounding factors, such as SES and pre-existing differences in FSIQ, which may influence the strength of the study. Recruitment of participants for PKIDS occurs across a wide range of areas around Perth, Western Australia, to cover children in a variety of SES settings. Given that SES is associated with adverse birth outcomes (Kramer, et al., 2009), there could be a chance that participants with more financial resources and more favourable developmental background volunteered for research. It is also worth considering that attendance at PKIDS involves two full days of free activities for the children during the school holiday period and this is often appealing to lower SES families. Since the present study did not account for confounding variables, birth group differences may be at least partly attributable to SES differences. Unfortunately the results here cannot argue against that entirely, however, evidence in previous meta-analysis of 7000 participants suggested that SES was not enough to account for significant results on cognitive outcomes and explain all the differences between the birth groups (Kerr-Wilson, et al., 2011). Aside from the aforementioned factors, the impact of perinatal characteristics, such as infections and IVH, for the clinical population was not available. Therefore, the results of the current study were unable to demonstrate how perinatal characteristics play a role in the cognitive development in children born preterm.

One other limitation is that, even though the sample size was 362 for the age-collapsed analyses, when the sample was broken down by age group, the cell size became relatively small. There is a possibility that the different mediation pattern for different age groups stems from a lack of power in smaller groups, causing mediators to drop out. However, this does not explain the pattern of reversal in significant partial

mediation found from WM to cognitive inflexibility and the presence of reasonable effect size in most age groups.

Finally, the birth weight and GA of the controls were not included in the archival data therefore specific characteristics for the control sample could not be provided. However, the PKIDS program has a tick box on the enrolment form for parents to indicate whether their typical child was LBW or preterm. Therefore, the PKIDS program would know that they are *not* preterm, but information on their birth weight was not included.

Clinical Implications and Future Directions

The current findings are amongst the first to confirm deficits in G_f among children born preterm, and that both WM and cognitive inflexibility serve as significant partial mediators at different age groups. Although there has yet to be consistent evidence on early intervention for the remediation of WM or G_f , the current results can assist in designing an appropriate educational environment for children born preterm.

Although this research does not present strong enough evidence to claim causality and provide precise recommendation for the development of practical early intervention, it has identified specific areas of impairment that would likely benefit from intervention. In particular, educational plans that avoid the overload of cognitive systems involving WM and cognitive flexibility would seem to be beneficial *cognitive-accommodation* strategies to support the learning progress of children born preterm. For example, when developing reading comprehension exercises, the teacher can try to keep the WM load low by allowing children to refer back to their books more frequently and avoid frequent changes in the rules involved in the exercises. However, *cognitive intervention* rather than the treatment-as-usual model of *accommodation* of deficits may have better chances of generating generalizability or transfer of abilities.

Recent progress in WM training programs suggests that this may be a fruitful area for further exploration as a neuro-remediation option for children born preterm. There is still debate about whether WM training generalizes to other untrained WM tasks, executive functions, or even G_f , but some promising results have been documented (Chein & Morrison, 2010; Jaeggi, et al., 2008; Klingberg et al., 2005; Redick et al., 2012). In addition to WM training, cognitive flexibility training may also be appropriate since the current study has demonstrated that the effects of birth status on G_f is also partly attributable to cognitive inflexibility in children born preterm. No research has yet explored this approach, but some studies have been done on the effects of alternating thinking strategies and aerobic exercise (Diamond & Lee, 2011), which may warrant further investigation. Given that cognitive flexibility is also a crucial component of EF and correlated with WM and inhibitory control, training may perhaps have an effect on other EF components and G_f . However, the current sample demonstrated a ‘catching up’ trend in cognitive inflexibility that no longer mediated the effects of birth status on G_f amongst the eldest age group, and given previously documented weaker associations between cognitive flexibility and G_f than between WM and G_f (van der Sluis, et al., 2007), perhaps training in WM should be the priority in research exploration.

Several possible improvements to methodology should be considered in future replications and exploratory research amongst children born preterm. For example, the inclusion of both verbal and visuo-spatial WM tasks would be more comprehensive in providing measures of both short-term memory and executive attention control as they might also be sensible in reflecting components commonly addressed in WM theories (Baddeley & Logie, 1999; Engle, 2002; Kane, et al., 2001). These additional measures would also provide more understanding of which specific WM component better

predicts G_f amongst children born preterm: executive attentional control or STM storage. The inclusion of another commonly addressed EF component, inhibitory control, would also increase the strength of the study as it would then fully address all key components documented in Miyake's (2000) seminal paper on EF structure.

Also, as suggested by MacKinnon, Fairchild, and Fritz (2007), proper use of a mediation framework allows researchers to identify which mediating variable is likely to be causally related to the outcome, which then allows researchers to target that mediating variable during intervention in order to change the outcome. Therefore, future research can investigate how additional combinations of cognitive processes may play a role in currently identified differences in G_f amongst children born preterm using multiple mediation modelling or more advanced SEM approaches. Such research would work towards identifying specific causal relations as a remedy to some of the academic delays identified amongst children born preterm.

Finally, given that attention deficits are prominent in children born preterm, it is difficult to ascertain that the results of the study and particularly the performance on assessments for children born preterm were not affected by attention deficits. Perhaps children born preterm are capable of managing complex information when they are fully focused, however, whether they have given their full attention on assessments is unknown. It is therefore important for future studies to ensure that trained and skilled researchers administer the psychological assessments and to maintain good rapport and motivational effort throughout testing. As noted in previous sections, well-trained post-graduate psychology students administered all assessments in the PKIDS program.

Summary

In conclusion, these findings indicated that prematurity is associated with impairment in G_f and that this impairment may be partly attributable to WM and

cognitive flexibility deficits. If indeed deficits in these core cognitive constructs underpin the academic struggles of children born preterm, then recent cognitive training research that demonstrates an ability to enhance WM, executive function and even G_f (Jaeggi, et al., 2008; Jaeggi, Buschkuhl, Jonides, & Shah, 2011) is of great practical interest to researchers, parents and teachers of the preterm birth population.

CHAPTER 5

Literature Review: Computerized Working Memory Training

Introduction

Considering the fact that children born preterm/LBW display, on average, noticeably lower scores on tasks measuring WM and G_f , what, if anything, can be done to help these children catch up with their peers? Given the positive associations between these cognitive abilities and academic outcomes, it appears desirable for researchers to look into ways of improving them.

Past attempts to improve cognitive functioning have ranged from pharmacological means (Elliott et al., 1997; D. C. Turner, Blackwell, Dowson, McLean, & Sahakian, 2005) to interventions through physical exercises (Tomprowski, et al., 2008) and musical training (Hetland, 2000; Moreno et al., 2011; Schellenberg, 2004). In recent decades, researchers have begun to suggest the use of intensive computerized WM training as a plausible way of enhancing WM, and some have even detected generalization effects to higher-order cognitive abilities, such as G_f (Jaeggi, et al., 2008; Klingberg, et al., 2002).

Cattell (1987) noted that G_f increases until around the age of 15 then gradually decreases as one gets older. A similar curve was also demonstrated in McArdle et al.'s (2002) study, where G_f reaches its peak age around 23 years old before it gradually decreases. The expected rate of increase in G_f between the ages of two and 19 years were shown to be much faster than between the ages of 20 and 75 years. The authors suggested that when individuals have reached G_f peak age, measurements of cognitive abilities taken at the time should show little to no improvement. Nonetheless, recent WM interventions have suggested otherwise as their findings point to the ability to enhance G_f performance after it has peaked, which further refutes the necessary stability

of intelligence. Thus, a review of the literature on computerized WM intervention is presented in this chapter. This includes understanding the different types of training and their variations, exploring the potential modifiability of both WM and G_f , and understanding the importance of adaptive training involved in the learning process.

Types of Computerized Working Memory Training

The most extensively described intensive computerized WM training interventions are Cogmed training, n -back training, and complex working memory (CWM) span training. Klingberg, Forssberg and Westerberg (2002) were among the first to publish their work on the success of WM training and subsequently developed the Cogmed Working Memory Training Group (Pearson Inc., 2011). The overall theoretical assumption behind their training is that WM is one of the most essential processing functions that underlie other cognitive abilities, such as logical reasoning (Hornung, et al., 2011; Tillman, et al., 2008). Furthermore, WM and EF are closely related to prefrontal cortex functioning and impairments in this area have been prominent in individuals with ADHD (Castellanos, Sonuga-Barke, Milham, & Tannock, 2006; Stuss, 2011). Therefore, much of the Cogmed group's early research focused on children with ADHD.

There are now three versions of Cogmed training available, specifically targeted at different age groups: Cogmed JM for pre-schoolers, Cogmed RM for school-age children, and Cogmed QM for adults (Pearson Inc., 2011). Lohaugen et al. (2011) provide a comprehensive description of the computer tasks. In brief, the training includes a variety of WM tasks, allowing participants to train in areas of short-term storage, processing and manipulation of information, as well as immediate recall. The presented stimuli are either in visuo-spatial or verbal formats, or both (J. Holmes, Gathercole, & Dunning, 2009). A coach is also included in the training process to

provide the family and participant with weekly feedback regarding performance and motivational support (Lohaugen, et al., 2011).

The Cogmed program includes a battery of adaptive training tasks that tap overlapping mental processes. This diversity of tasks is likely to increase the chance that overall training will lead to generalized training-related cognitive benefits. This overlap has been justified through a number of studies that suggest WM in general, rather than specific domains, is the most influential cognitive function that predicts higher-order logical reasoning. However, such overlap could affect the determination of which particular task or process is key for generalization (Morrison & Chein, 2011). It is important then, theoretically, to consider other computerized WM training aside from Cogmed. In recent decades, many other researchers have embraced the rising popularity of cognitive training and proposed other training regimens, such as dual n -back and CWM.

Dual n -back training was initially tested in Jaeggi, Buschkuhl, Jonides, and Perrig's (2008) pioneering study. They investigated the effects of adaptive dual n -back training on measures of G_f . The authors proposed that n -back training integrated executive processes and attentional control similar to those tapped by reasoning tasks, therefore leading to successful transfer effects. The neural mechanisms, specifically activity in the lateral prefrontal cortex and parietal cortex, have been shown to mediate the relationship between demanding WM tasks, such as dual n -back task, and G_f (Gray, Chabris, & Braver, 2003).

In the dual n -back task, individuals are asked to respond to visual and auditory stimuli simultaneously. The stimuli are presented as a cue at a spatial location on the computer screen and an audio presentation of a letter of the alphabet. The goal is to identify whether one of the current stimuli is identical to one of the stimuli n trials back.

Similar to Cogmed training, an adaptive feature is included. The difficulty is adjusted according to individuals' performance on the task, where n increases or decreases with successful and unsuccessful attempts, respectively. The task requires individuals to be trained on this highly demanding WM task at a difficulty level tailored to them. Studies have also used a single n -back task, where individuals are only presented with and required to respond to either a visual or auditory stimulus. Figure 5.1 illustrates an example of the dual 2-back condition (Jaeggi, et al., 2008).

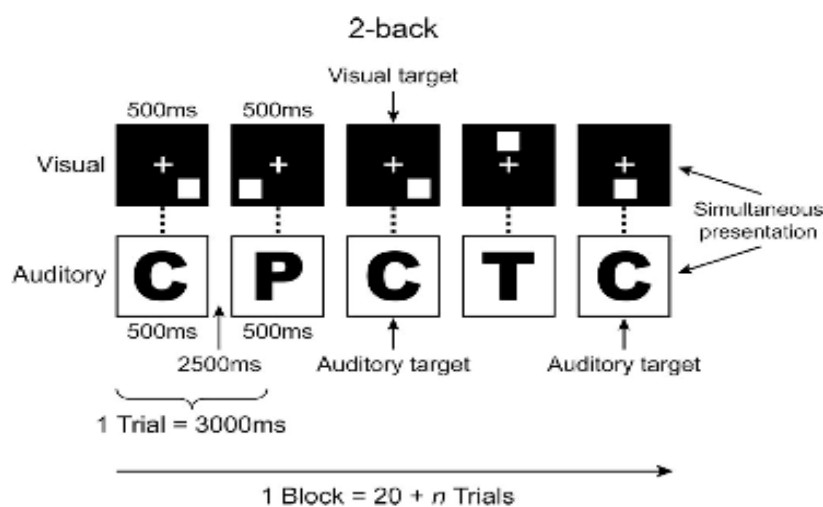


Figure 5.1. A visual example of the dual n -back training task, 2-back condition. Adapted from “Improving fluid intelligence with training on working memory” by S. M. Jaeggi, M. Buschkuhl, J. Jonides, and W. J. Perrig, 2008, *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), p. 6830.

More recently, additional published studies have investigated the effect of CWM span on enhancing WM and G_f . Studies have directed attention to the reliability and validity of CWM tasks as measures of working memory capacity and executive attentional control, both of which have strong associations with G_f . Therefore, individual differences in performance on CWM tasks have been viewed as a promising predictor of complex high-order cognition. (A.R.A. Conway et al., 2005; Unsworth, Redick, Heitz, Broadway, & Engle, 2009). CWM span tasks may take many forms, but

they must consist of two essential elements. One is the “to-be remembered (TBR)” (Unsworth, et al., 2009, p. 636) item and the other is some type of cognitive processing activity. TBR items are typically recalled in sequential order but range across various kinds of presentation stimuli, such as spatial locations, numbers, letters, words, and shapes. TBR items are interspersed through the cognitive processing task, which is unrelated to the TBR items and also takes varying forms, such as counting, solving mathematical problems, and judging whether the given stimuli are upside down or not (Unsworth, et al., 2009). Loosli, Buschkuehl, Perrig, and Jaeggi (2012) describe an example of such a task in their study with typically developing children. Refer to Figure 5.2 for an illustration of the task adapted from their original article.

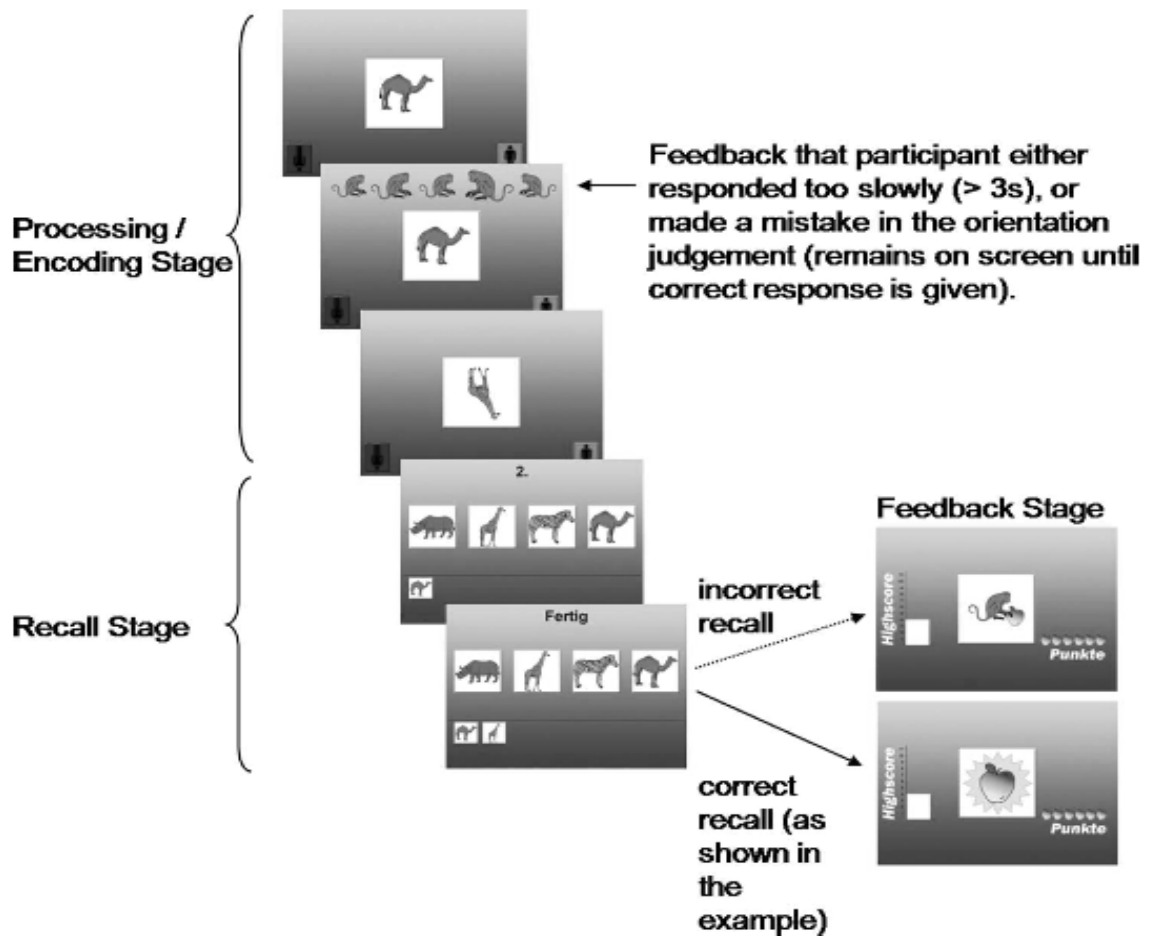


Figure 5.2. A visual example of the animal adaptive span task. Each trial consists of two stages. The participant is to encode whether the animal is right way up or upside down during the Processing/Encoding Stage. During the Recall Stage, the participant is asked to recall the animals presented in the Processing/Encoding Stage in sequential order. The number of animals presented for each trial is adaptive to the participant's performance. Adapted from "Working memory training improves reading processes in typically developing children" by S. V. Loosli, M. Buschkuehl, W. J. Perrig, and S. M. Jaeggi, 2011, *Child Neuropsychology*, 18(1), p. 6.

As seen in Figure 5.2, during Part 1: the processing/encoding stage, participants are required to identify whether each animal presented in a sequence in the centre of a computer screen was right way up or upside down. Participants are to click on the bottom right/green icon or the bottom left/red icon, respectively. Simultaneously, participants are required to remember the sequence of animals in their order of

appearance. An error prompt appears on the screen when participants exceeds a time limit to respond or provide an incorrect answer. In Part 2: the recall stage, participants are provided with the four choices of animals and are required to recall their appearance in sequential order. The task has an adaptive feature where the difficulty of the task matches each participant's performance. This is determined by using the number of animals required to be remembered in each trial, termed "set size". Moreover, the maximum number of animals is indicated, labelled as "Highscore", on the Feedback screen page (Loosli, et al., 2012).

Variations in Training Regime

Aside from the different types of computerized WM training, there are also variations in the training regimes that researchers have used within their cognitive training studies. These include variations in training schedule, dosage, and the inclusion of passive and/or active controls. Table 5.1 summarizes the variations in these computerized WM training studies.

Training schedules varied slightly across different cognitive training studies. Typically, Cogmed training research describes a training schedule of at least 20 sessions in 4-6 weeks with the number of sessions ranging from 20-25 (Dahlin, 2011; J. Holmes, et al., 2009; J. Holmes et al., 2010; Klingberg, et al., 2005; Klingberg, et al., 2002; Lohaugen, et al., 2011; Lundqvist, Grundstrom, Samuelsson, & Ronnberg, 2010; Westerberg et al., 2007). Similarly, studies evaluating *n*-back training (Chooi & Thompson, 2012; Jaeggi, et al., 2011; Redick, et al., 2012; Studer, et al., 2009) and CWM span training report using a training schedule of 20 sessions within 4-6 weeks (Chein & Morrison, 2010), with the exception of a few. In particular, Jaeggi et al.'s (2008) pilot dual *n*-back study tested participants across 8, 12, 17, and 19 days and Loosli et al.'s (2012) CWM training study tested participants rather briefly over 10

sessions across 2 weeks.

The duration of each session varies more extensively than the overall training schedule. They range from as brief as 12 minutes (Loosli, et al., 2012) to as long as a full hour (Lundqvist, et al., 2010). There does not appear to be a consistent timeframe for the duration of training each day across reviewed computerized WM studies. For example, two studies that tested children with ADHD report using 20-25 minutes of training per day (J. Holmes, et al., 2010; Klingberg, et al., 2002), while another reports at least a 40-minute training session for the same clinical population (Klingberg, et al., 2005). Studies on children with low WM, special needs, and of adolescents born ELBW report a training dosage that ranges from 30-40 minutes per session (Dahlin, 2011; J. Holmes, et al., 2009; Lohaugen, et al., 2011).

The inclusion of both a passive no-contact control group and an active non-adaptive training control group in addition to the adaptive training group are important elements in evaluating cognitive interventions. A passive no-contact control group enables a study to control for any maturational changes and possible practice effects. However, participants in a no-contact control group may be aware that they are not participating in any intervention, and thus may have less motivation in performing at their best on their cognitive assessments because they expect no change in their performance scores. The inclusion of an active, yet non-adaptive training control group, where participants train on a low dosage WM task, acts as an alternative treatment group. It controls for training adherence through a non-challenging but related training task. It also assists in controlling for expectancy effect, also known as the Hawthorne effect (Shipstead, Redick, & Engle, 2010).

According to current literature, none of the Cogmed training evaluations included both types of control groups except for those performed on children at

preschool age (Nutley et al., 2011; Thorell, et al., 2009). Thorell et al. (2009) included an inhibition training group along with a non-adaptive active control group, whereas Nutley et al. (2011) included a non-verbal training group and a combined training group, which was trained on both WM and non-verbal skills. In contrast, three studies evaluating *n*-back training included both passive and active control groups (Chooi & Thompson, 2012; Redick, et al., 2012; Studer, et al., 2009). None of the currently reviewed CWM span training studies included both passive and active control groups in their investigations (Chein & Morrison, 2010; Loosli, et al., 2012), which limits their strength in controlling for the aforementioned confounding factors.

In summary, training studies generally evaluate WM training over at least 20 sessions within a 4-6 week training period. However, the wide variation in daily training time across different studies makes it difficult to compare results and determine an optimal training duration for future studies. Nonetheless, the inclusion of both a passive and an active control group is deemed more appropriate, given the aforementioned methodological benefit.

Near Transfer and Far Transfer Effects

Studies have investigated whether WM training enhances abilities that closely resemble processes involved in the trained task, referred to as near transfer effects, as well as whether WM training extends to improved performance in different processing domains, referred to as far transfer effects (Barnett & Stephen, 2002; Chooi, 2012). A summary of recent computerized working memory training results is presented in Table 5.1. Results for follow-up assessments are described in later sections.

Table 5.1 A Summary of Recent Computerized Working Memory Training Studies

Author (year)	Population (age in years)	Adaptive Training (n = ?)	Training Dosage	Time (mins)	Control group (n = ?)	Near Transfer to WM	Far transfer to cognitive tasks	Far transfer to Gf	Follow up
Klingberg et al. (2002)	ADHD (7-15)	Cogmed (n = 7)	~ 24 sessions in 5-6 weeks	25	Active (n = 7)	Span board*	Stroop*	RCPM*	n/a
Klingberg et al. (2002)	Healthy volunteers (7-15)	Cogmed (n = 4)	26 sessions in 5 weeks	25	none	Span board*	Stroop* CRT	RCPM*	n/a
Klingberg et al. (2005)	ADHD (7-12)	Cogmed (n = 24)	20 sessions in 20 days	40	Active (n = 20)	Span board* Digit span*	Stroop*	RCPM*	3 mths
Westerberg et al. (2007)	Stroke patients (34-65)	Cogmed (n = 9)	25 sessions in 5 weeks	40	Passive (n = 9)	Digit span* Span board* Delayed recall Claeson-Dahl RUFF*	Stroop PASAT*	RSPM	n/a
Thorell et al. (2009)	Preschool (4-5)	Cogmed (n = 17)	~ 25 sessions in 5 weeks	15	Passive (n = 16) Active 1 (n = 18) Active 2 (n = 14)	Span board* Word span*	Day-Night Stroop Go/no-go CE Go/no-go OE* Nepsy* Block Design Go/no-go RT	n/a	n/a
Holmes et al. (2009)	Low WM (8-11)	Cogmed (n = 22)	20 session in 5-7 weeks	35	Active (n = 20)	Complex span* Following-instructions*	WASI WORD WOND	n/a	6 mths

Prematurity, Cognitive Abilities & Intervention

Author (year)	Population (age in years)	Adaptive Training ($n = ?$)	Training Dosage	Time (mins)	Control group ($n = ?$)	Near Transfer to WM	Far transfer to cognitive tasks	Far transfer to Gf	Follow up
Holmes et al. (2009)	Low WM (8-11)	Cogmed ($n = 22$)	20 session in 5-7 weeks	35	Active ($n = 20$)	Complex span* Following-instructions*	WASI WORD WOND	n/a	6 mths
Holmes et al. (2010)	ADHD (8-11)	Cogmed ($n = 25$)	20-25 sessions in 6-10 weeks	20-25	While on medication at pre-test	AWMA*	WASI	n/a	6 mths
Lundqvist et al. (2010)	Brain-injured patients (27-63)	Cogmed ($n = 10$)	21-25 sessions in 5 weeks	45-60	Passive ($n = 11$)	Block span forward/backwards* Listening span Picture span	PASAT* Color-Word Interference Test*	n/a	20 wks
Nutley et al. (2011)	Preschool (4-4.5)	Cogmed ($n = 24$)	25 sessions in 5-7 weeks	15	Passive ($n = 25$) Active 1 ($n = 25$) Active 2 ($n = 27$)	Visuo-spatial grid* AWMA* Word span	Repeated Patterns Sequential Orders Classifications Block Design	RCPM	n/a
Lohaugen et al. (2011)	ELBW (14-15)	Cogmed ($n = 16$)	25 sessions in 5 weeks	30-40	Passive full-term ($n = 11$) Active full-term ($n = 19$)	Digit span* LNS* Spatial span* WMS (Verbal learning)*	n/a	n/a	6 mths

Author (year)	Population	Adaptive Training	Training Dosage	Time (mins)	Control group	Near Transfer to WM	Far transfer to cognitive tasks	Far transfer to Gf	Follow up
Gruewaldt et al. 2013	VLBW (5-6)	Cogmed JM (<i>n</i> = 20)	25 sessions in 5 weeks	10-15	n/a	Digit span (Forward) (Backward) Spatial span (Forward) (Backward)*	Nepsy: Visual attention total time Auditory attention and response set* Phonological processing* Comprehension of instructions Repetition of nonsense words* Memory for faces* Narrative memory* Sentence repetition*	n/a	n/a
Gruewaldt et al. 2015	VLBW (5-6)	Cogmed JM (<i>n</i> = 20)	25 sessions in 5 weeks	10-15	Passive VLBW age-matched (<i>n</i> = 17)	Digit span (Forward) (Backward) Spatial span (Forward) (Backward)*	Nepsy: Visual attention total time Auditory attention and response set* Phonological processing* Comprehension of instructions Repetition of nonsense words* Memory for faces* Narrative memory* Sentence repetition*	n/a	7 mths

Prematurity, Cognitive Abilities & Intervention

Author (year)	Population	Adaptive Training	Training Dosage	Time (mins)	Control group	Near Transfer to WM	Far transfer to cognitive tasks	Far transfer to Gf	Follow up
Jaeggi et al. (2008)	University students (26)	Dual <i>n</i> -back (<i>n</i> = 35)	12-19 days	25	Passive (<i>n</i> = 35)	Digit span* Reading span	n/a	BOMAT*	n/a
Studer et al. (2009)	University students (mean age 19)	Dual <i>n</i> -back (<i>n</i> = 21) Single <i>n</i> -back (<i>n</i> = 25)	20 sessions in 4 weeks	15-20	Passive (<i>n</i> = 43)	n/a	n/a	RAPM* BOMAT*	n/a
Jaeggi et al. (2011)	Healthy children (7-10)	Single <i>n</i> -back (<i>n</i> = 32)	20 sessions in 4-6 weeks	15	Active (<i>n</i> = 30)	n/a	n/a	RSPM* TONI*	3 mths
Chooi and Thompson (2012)	University students (mean age 19)	Dual <i>n</i> -back (<i>n</i> = 22)	8 days / 20 days	30	Passive (<i>n</i> = 45) Active (<i>n</i> = 26)	Complex span	Vocabulary Tests Perceptual speed tasks	RAPM	n/a
Redick et al. (2012)	University students (18-30)	Dual <i>n</i> -back (<i>n</i> = 24)	20 sessions in 4-5 weeks	30-40	Passive (<i>n</i> = 20) Active (<i>n</i> = 29)	Symmetry span Running Letter span	SynWin Control Tower ATClab Vocabulary General Knowledge Letter Comparison Number Comparison	RAPM RSPM CCFIT Paper Folding Letter Sets Number Series Inferences Analogies	n/a

Author (year)	Population	Adaptive Training	Training Dosage	Time (mins)	Control group	Near Transfer to WM	Far transfer to cognitive tasks	Far transfer to Gf	Follow up
Jaeggi et al. (2008)	University students (26)	Dual <i>n</i> -back (<i>n</i> = 35)	12-19 days	25	Passive (<i>n</i> = 35)	Digit span* Reading span	n/a	BOMAT*	n/a
Studer et al. (2009)	University students (mean age 19)	Dual <i>n</i> -back (<i>n</i> = 21) Single <i>n</i> -back (<i>n</i> = 25)	20 sessions in 4 weeks	15-20	Passive (<i>n</i> = 43)	n/a	n/a	RAPM* BOMAT*	n/a
Jaeggi et al. (2011)	Healthy children (7-10)	Single <i>n</i> -back (<i>n</i> = 32)	20 sessions in 4-6 weeks	15	Active (<i>n</i> = 30)	n/a	n/a	RSPM* TONI*	3 mths
Chooi and Thompson (2012)	University students (mean age 19)	Dual <i>n</i> -back (<i>n</i> = 22)	8 days / 20 days	30	Passive (<i>n</i> = 45) Active (<i>n</i> = 26)	Complex span	Vocabulary Tests Perceptual speed tasks	RAPM	n/a
Redick et al. (2012)	University students (18-30)	Dual <i>n</i> -back (<i>n</i> = 24)	20 sessions in 4-5 weeks	30-40	Passive (<i>n</i> = 20) Active (<i>n</i> = 29)	Symmetry span Running Letter span	SynWin Control Tower ATClab Vocabulary General Knowledge Letter Comparison Number Comparison	RAPM RSPM CCFIT Paper Folding Letter Sets Number Series Inferences Analogies	n/a

Author (year)	Population	Adaptive Training	Training Dosage	Time (mins)	Control group	Near Transfer to WM	Far transfer to cognitive tasks	Far transfer to Gf	Follow up
Chein and Morrison (2010)	University students (mean age 20)	CWM span ($n = 19$)	20 sessions in 4 weeks	30-45	Passive ($n = 21$)	Complex span*	Stroop* NDRT* Verbal reasoning Spatial reasoning	RAPM	n/a
Loosli et al. (2012)	Healthy children (9-11)	CWM ($n = 20$)	10 sessions in 2 weeks	12	Passive ($n = 20$)	n/a	SLT*	TONI	n/a

Note. RCPM = Ravens Colored Progressive Matrices; CRT = Choice Reaction Time; PASAT = Paced Auditory Serial Attention Test; RSPM = Raven's Standard Progressive Matrices; Go/no-go CE = Go/no-go Commission Errors; Go/no-go OE = Go/no-go Omission Errors; Go/no-go RT = Go/no-go Reaction Time; WASI = Wechsler Abbreviated Scale of Intelligence; WORD = Wechsler Objective Reading Dimensions; WOND = Wechsler Objective Number Dimensions; AWMA = Automated Working Memory Assessment; LNS = Letter-Number Sequencing; WMS = Wechsler Memory Scale; NEPSY = A Developmental Neuropsychological Assessment; BOMAT = Bochumer Matrizen Test; RAPM = Raven's Advanced Progressive Matrices;; TONI = Test of Nonverbal Intelligence; CCFIT = Cattell Culture-Fair Intelligence Test; NDRT = Nelson-Denny Reading Test; SLT = Salzburger Lesetest for reading process.

*Significant change compared to control group at post-assessment.

Near Transfer to Working Memory

Near transfer to WM tasks has often been measured by digit span tasks, span board tasks and CWM tasks. An explanation for near transfer to occur is that the adaptive WM training is effective in improving WM abilities that do not entail task-specific strategies. If the training task truly involved temporary storage and active manipulation of information strategies, then subsequently those who participated in training should be able to generalize their abilities to other tasks that utilize the same set of skills.

Examples of significant near transfer to WM have been found across several studies after Cogmed training. It was difficult to compare results given the differences in use of measures in each study, however, out of the total 10 reviewed studies that used Cogmed training, all 10 reported one or more measures with significant near transfer effects. Several studies that tested children with ADHD found improvements in visuo-spatial working memory (J. Holmes, et al., 2010; Klingberg, et al., 2005; Klingberg, et al., 2002) and verbal working memory (J. Holmes, et al., 2010; Klingberg, et al., 2005). According to Holmes, Gathercole and Dunning (2009), significant near transfer effects occurred in a group of children with low working memory. Results indicated that there were significant training gains for the adaptive training group on all measures of the Automated Working Memory Assessment, including tasks of verbal short-term memory ($p = .01$, $d = 0.62$), visuo-spatial short-term memory ($p < .01$, $d = 1.2$), verbal memory ($p < .01$, $d = 1.55$), and visuo-spatial memory ($p < .01$, $d = 1.03$). In addition, the authors also found significant gains in a following-instructions task ($p < .01$, $d = 0.83$; (Gathercole, Durling, Evans, Jeffcock, & Stone, 2008).

Aside from the groups with ADHD or low WM, the effectiveness of Cogmed training has also been examined across other clinical populations. For example, adults

with stroke (Westerberg, et al., 2007) and brain injuries (Lundqvist, et al., 2010), children with special needs (Dahlin, 2011), as well as children of ELBW (Lohaugen, et al., 2011) have all demonstrated successful near transfer to working memory tasks.

Only two out of the 10 studies gathered data from typically developing participants, specifically, preschool children aged four to five years old, and they too found significant near transfer to WM measures (Nutley, et al., 2011; Thorell, et al., 2009). No other studies have investigated typically developing participants of other school age.

In further elaboration, two of the studies on adults with brain injuries and stroke documented a significant near transfer effect to non-trained WM tasks with a large effect size. For example, Westerberg et al. (2007) documented significant differences after training in favour of the training group in the digit span task ($p = .005$, $d = 1.58$) and span board task ($p = .05$, $d = 0.83$). Similarly, Lundqvist et al. (2010) observed significant improvements in all WM measures ($p < 0.01$) except for span board forward at 4-week post-assessment.

As for children with clinical backgrounds, Dahlin (2011) compared the effects of Cogmed WM training in a no treatment control group to those children with special needs. Evidence pointed to significant near transfer effects on WM related non-trained tasks, including measures of digit span forward ($p < .05$, $d = 0.66$) and backward ($p < .05$, $d = 0.67$), span board forward ($p < .01$, $d = 0.98$), but no near transfer on the span board backward task. However, the author's findings here should be interpreted cautiously because the training group received considerably more attention and reinforcement than the no treatment control group, which received only the usual special education support. This could stimulate additional positive impact through motivation on performance of outcome measures. There also appeared to be a large

discrepancy between the numbers of participants in the training group as compared to the control group, which may lead to plausible bias. Refer to Table 5.1 for sampling characteristics.

To date, only three computerized WM training studies document their utility in individuals of LBW. Lohaugen et al. (2011) were the first to demonstrate the effectiveness of Cogmed training in ELBW adolescents. Participants, ranging from 14 to 15 years old, were age-matched across three groups: an ELBW group ($n = 16$) and a term-born control group ($n = 19$), both of which participated in training, and a third full-term control group that did not participate in any training ($n = 11$). Results indicated that there were significant improvements in non-trained tasks for both groups in training. Both ELBW and full-term groups in training displayed significant improvements in digit span score (ELBW: $p < .01$; FT: $p < .001$), spatial span score (ELBW: $p < .001$; FT: $p < .001$), and letter-number sequencing (LNS; ELBW: $p < .05$; FT: $p < .001$) at post-intervention. Both groups also improved on performance of verbal learning tasks except for the subtest that measured delayed memory in word lists. A comparison of results with the non-training group of full-term participants indicated no practice effect on the outcome measures. The authors claimed that the training program was as effective for ELBW adolescents as it was for the controls in enhancing WM. However, the authors' claim could be strengthened if an active control ELBW group was also present. The inclusion of an ELBW clinical peer group would allow further comparison of the ELBW cohort and the full-term cohort, and the extent of their improvements from training with reference to their own cohort as controls.

Grunewaldt and her colleagues (Grunewaldt, Lohaugen, Austeng, Brubakk, & Skranes, 2013; Grunewaldt, Skranes, Brubakk, & Lohaugen, 2015) recently published two other studies from Norway that investigated the effectiveness of the Cogmed

program in 5 to 6 years old VLBW children. In the earlier study, Grunewaldt et al. (2013) used a Stepped Wedge design. Both groups (Group 1: $n = 9$; Group 2: $n = 11$) were tested before the training and four months after the training. Outcome measures included the digit span and spatial span tasks as well as subtests from the NEPSY. The authors combined the results of the two groups in analysing the training effects because of the small sample size. Findings revealed that, after five weeks of training and four weeks of treatment as usual, the participants showed significant improvements in spatial span backwards ($p = .01$), auditory attention ($p = .012$), phonological processing ($p = .004$), repetition of nonsense words ($p = .017$), memory for faces ($p = .001$), narrative memory ($p = .003$) and sentence repetition ($p = .005$). Grunewaldt et al.'s (2015) subsequent study improved slightly in methodology and included a control group. The authors compared the data they gathered from the previous study (Grunewaldt, et al., 2013) and compared the VLBW participants' ($n = 20$) training results with a VLBW age-matched control group ($n = 17$). The same outcome measures were used. Assessments were administered to both groups four months after the intervention group completed training and then again at 7-months follow-up. Results indicated that only spatial span backwards ($p = .003$), memory for faces ($p = .012$), and narrative memory ($p = .002$) remained significant at follow-up. Despite significant results, Grunewaldt et al.'s study (2013; 2015) did not provide an immediate post-training assessment. Without such information, the significant improvements found four weeks after training can be a result of other activities that the participants were involved in rather than a genuine effect of the computerized intervention.

Studies using n -back training have not been as successful in achieving near transfer in non-trained WM as with those observed in Cogmed training studies. Only one out of the five reviewed studies documented a significant near transfer to WM.

Jaeggi et al.'s (2008) pilot study on healthy university students was one rare example. In Jaeggi et al.'s (2008) experiment, 70 healthy university student participants (mean age = 25.6, SD = 3.3) were placed into either the intervention group ($n = 35$) or the control group ($n = 35$). The intervention group was divided into four different training groups, and matched on days of training with four control groups that received only the pre- and post-test assessments within the corresponding time intervals. The main difference between the four pair of groups was the time duration of training, specifically across 8 days, 12 days, 17 days and 19 days with $n = 16, 22, 16,$ and 15 participants respectively. Outcome measures for WM included the digit span task, tested in all groups, and the reading span task, tested in all except for the 8-day training group. Results revealed that participants in training not only improved on the training task but also improved in performance on the digit span task. Despite relatively less successful reports of near transfer after dual n -back training, Jaeggi et al.'s (2008) finding on group \times session interaction effects, as opposed to the pre-post effects documented in most Cogmed training studies, provided evidence on a training dosage response. A group \times session interaction was detected indicating a dose response in near transfer to the digit span task, where more training led to a greater near transfer effect ($p < 0.001$; $\eta p^2 = .17$).

Amongst other n -back training studies that have included non-trained WM measures, no significant near transfer effects have been found (Chooi & Thompson, 2012; Redick, et al., 2012). Chooi and Thompson (2012) argued against the success of the approach after their attempt to replicate Jaeggi et al.'s (2008) dual n -back study showed non-significant results for all transfer effects. Chooi and Thompson (2012) assigned their participants to three groups and, within each group, the participants were further divided into 8-day or 20-day training schedules: intervention (8-days: $n = 9$; 20-

days: $n = 13$), active control (8-days: $n = 15$; 20-days: $n = 11$) and passive control (8-days: $n = 22$; 20-days: $n = 23$). Participants in the intervention group were trained in the adaptive dual n -back version of the training task for 30 minutes, four days per week, whereas the active controls were trained in a non-adaptive dual 1-back version of the task for 20 trials each session. There was no mention of the total time differences spent on the computer across the two groups. The passive control group did not participate in any training. Despite their findings, the described replication appeared to have rather small cell sizes of less than 20 participants in both intervention and active control group as compared to other dual n -back studies. This would have reduced the statistical power of their analysis.

Similarly, near transfer to non-trained WM has not been documented in CWM span training thus far (Chein & Morrison, 2010; Loosli, et al., 2012). Even though Chein and Morrison reported significant improvements in WM ($p < .005$, $d = 1.42$) for their training group after CWM training, their assessment tasks for WM were spatial and verbal memory tasks that resembled their CWM training task. Therefore, this does not constitute a valid near transfer to non-trained WM.

In general, there is consistency in successful near transfer effects after Cogmed training but not other methods of training. Effect sizes for near transfer after Cogmed training ranged from medium ($d = 0.62$) to large ($d = 1.58$). The relative number the participants in the studies across the three training types did not appear to affect the results of near transfer, most studies had 10 to 20 participants in each comparison group. According to successful Cogmed findings, the training schedule varied across a time range from 15 to 60 minutes. Unfortunately, there is no study yet to suggest whether shorter or longer duration of training per day led to more successful results. However, it was observed that younger children were assigned to less training time per

day and school students were trained with more. Nonetheless, overall training remains the same across age groups, and training is commonly documented from 20-25 sessions over a span of 5-7 weeks.

Far transfer to Other Cognitive Tasks

In regards to far transfer to other cognitive tasks, studies typically included measures of executive function (EF), while some also included processing speed and reading-related measures. Across the various intervention studies, it was noted that Stroop was one of the most commonly used measures amongst others, as listed in Table 5.1.

Far transfer to measures of EF may be explained by the relationship between WM and components of EF. As reviewed earlier, WM is an essential part of EF and findings have also demonstrated that both WM and EF rely on prefrontal cortex functioning. Therefore, training on WM would tap similar neurological systems, thus leading to possible transfer effects (Castellanos, et al., 2006; Miyake, et al., 2000; Stuss, 2011).

Several Cogmed training studies have found far transfer to cognitive tasks other than WM. In particular, statistically significant improvements as measured by the Stroop task were found after Cogmed training in two studies of 7- to 15-year-old children with ADHD (Klingberg, et al., 2005; Klingberg, et al., 2002), as well as children with special needs ($p < .01$; (Dahlin, 2011) and one recent study in adults with brain injury ($p < .01$; (Lundqvist, et al., 2010). However, the effect sizes for these significant results were quite small, ranging from a Cohen's d of 0.06 to 0.34 (Dahlin, 2011; Klingberg, et al., 2005). Other studies of Cogmed with stroke patient adults (Westerberg, et al., 2007) and typically developing pre-schoolers (Thorell, et al., 2009) documented no significant improvements on the Stroop task.

Processing speed measured by choice reaction time tasks and Go/no-go task were included in a small study with healthy adults (Klingberg, et al., 2002) and preschoolers (Thorell, et al., 2009), respectively, however, no improvements were recorded after Cogmed training. The generalizability of some of these studies was very limited due to the recruitment of a very small sample, especially those with fewer than 10 participants in each group (Klingberg, et al., 2002; Westerberg, et al., 2007).

Some studies have included reading and math-related measures, as well as global intelligence assessments. Only one study of children with special needs has demonstrated significant improvements in reading comprehension tasks after Cogmed training as measured by the Progress in International Reading Literacy Study and IEA Reading Literacy ($p > .001$, $d = 0.88$; (Dahlin, 2011). Whereas Holmes et al. (2009) found no significant differences in his participants, after training with Cogmed, in terms of VIQ and PIQ performance from the Wechsler Abbreviated Scales of Intelligence (WASI), or the reading and mathematical subtests from the Wechsler Objective Reading Dimensions (WORD) and Wechsler Objective Number Dimensions (WOND), respectively.

Contrary to Cogmed training, very few dual n -back training studies included far transfer measures to capture changes in cognitive areas other than G_f . Two recent studies have attempted to explore the presence of transfer on verbal and perceptual tasks, for example, by including vocabulary and perceptual speed tasks, aside from the usual reasoning tasks. As expected, no far transfer to these tasks was found. In addition, neither of the studies found any far transfer effects in any of the measures included (Chooi & Thompson, 2012; Redick, et al., 2012). These authors claim that dual n -back training does not lead to any near or far transfer effects.

Nonetheless, far transfer to both EF tasks and reading-related tasks has been

documented after CWM span training. Chein and Morrison's (2010) study of university students reported significant improvements on between-group comparisons for the Stroop task ($p = .039$, $d = 0.57$) and the Nelson-Denny Reading Test ($p = .04$, $d = 0.58$) after CWM span training that included verbal and spatial CWM tasks. However, on repeated measures analyses, group x session interaction showed no effect ($p = .13$, $\eta p^2 = .056$), whereas using only successfully trained participants to test for interaction showed marginal significance effect ($p = .053$, $\eta p^2 = .103$). They claimed that they were able to replicate findings of enhanced performance in cognitive control shown in Klingberg et al.'s (2005) study. Similar interaction patterns for the reading task were demonstrated, with initially no significant group x session interaction ($p = .077$, $\eta p^2 = .08$) followed by subsequent significance restricting the sample to only successfully trained participants ($p = .013$, $\eta p^2 = .174$). Successfully trained participants were defined as those showing statistically significant improvement in the spatial CWM task. Nevertheless, this was the first published study to show improvements in reading comprehension using CWM training.

Following Chein and Morrison (2010), Loosli et al. (2012) also found a significant far transfer effect to reading processes using Salzburger Lesetest (SLT) after only 10 sessions of CWM training in healthy children aged nine to 11 years. In particular, they found significant transfer effects for word reading ($p < .01$, $d = 1.03$) and text reading ($p < .05$, $d = 0.72$) from the SLT. Despite Loosli et al.'s (2012) findings, the study included no measures of untrained WM tasks and thus raised the question of whether significant improvements in reading were due to the CWM span task used in the study specifically or a WM process that was transferable. Nonetheless, the authors' findings indicated that the effects of short duration training should not be underestimated since they did find significant outcomes despite all previous findings

pointing to at least 20 sessions over 4-6 weeks of training to achieve any training gains. Perhaps this calls for future investigation into shortened duration training, specifically, and the optimal training schedule to achieve successful gains, generally.

As described, both CWM training studies found consistent results in reading tasks. This appeared in line with literature that suggests a strong link between complex span tasks and reading comprehension (Daneman & Merikle, 1996). More consideration in the future of whether this type of training task included reading-related activity and whether cross-task stimulus similarity is necessary for transfer effects is important.

In general, more consistent evidence is required to conclude that WM training shows far transfer to other cognitive tasks, including EF, processing speed and reading-related measures. Current literature suggests only a consistently small effect size found using the Stroop task as a measurement outcome after Cogmed training. Questions remain regarding other WM training regimes and their utility in far transfer. Evidence of transfer, or of lack of transfer, to other cognitive functions besides WM and G_f could also shed light on what mechanisms underlie successful generalization.

Far Transfer to Fluid Intelligence

Another area of interest in the literature is far transfer to G_f after adaptive computerized WM training. Experimental studies have demonstrated successful far transfer results, however, not all studies have replicated successful findings. One possible explanation for successful transfer stems from the close but distinctive relationship between WM and G_f . As elaborated in previous chapters, studies have documented a high correlation between WM and G_f ranging from 0.72 to 0.85 in adults (Kane, et al., 2005; Oberauer, et al., 2005) and a correlation ranging from 0.64 to 0.82 in children (Fry & Hale, 2000).

Performance on Raven's Progressive Matrices has been most commonly used as

the outcome measure. Of those studies that included such observations, improvement in the Raven's task has been reported after Cogmed training in children with ADHD (Klingberg, et al., 2005; Klingberg, et al., 2002), but not healthy pre-schoolers (Nutley, et al., 2011) or adult patients with stroke (Westerberg, et al., 2007). To elaborate on underlying mechanisms, the Cogmed training regimen appears to fit with Baddeley and Hitch's (1974) model of WM as it utilizes training exercises that focus mainly on visuo-spatial components. There has been evidence in support of this STM component of WM and its link with G_f (Chuderski, et al., 2012; Hornung, et al., 2011), which has also been reviewed in Chapter 4.

On the other hand, the use of n -back training and complex span training appears to fit the argument that training on any form of adaptive, challenging WM task may lead to successful transfer, especially to measures of G_f . Evidence from n -back and complex span training may relate to Kane et al. (2001) and Engle's (2002) view on WM and WMC. Kane et al. (2001) and Engle (2002) emphasized the importance of controlled attention. In particular, higher WMC is seen as a consequence of higher control attention capability and it is also suggested as a key component of higher-order cognitive functions. The n -back task and complex span task both require executive attention and improvements in these tasks may be indicative of improvements in controlled attention. Thus, gains in training through challenging WM tasks may transfer to improvements in G_f via their influence on controlled attention.

Most of the improvements in G_f have been found after training with the dual n -back task. The first success in far transfer to G_f was documented in Jaeggi et al.'s (2008) pilot study of healthy university students. Outcome measures for G_f included Raven's Advanced Progressive Matrices (RAPM) for the group trained for eight days, while the other three groups (12, 17, and 19 days of training) were tested on the Bochumer

Matrizen Test (BOMAT). Significant improvements were found in BOMAT for the three groups but not in the RAPM. These documented gains in G_f remained significant after controlling for performance differences in digit span ($p < 0.001$), and reading span ($p < 0.01$) as well as averaged n -back level achieved during the last session ($p < 0.001$). Therefore, the authors concluded that this far transfer effect to G_f was attributable to the dosage of dual n -back training in which more training led to more improvement in G_f ($p < 0.001$, partial $\eta^2 = 0.48$). Further observation on the relationship between inter-individual differences in G_f and subsequent training gain demonstrated a main effect for performance group (low- G_f and high- G_f group), which indicated that those participants with a lower G_f to begin with showed larger training gains (Jaeggi, et al., 2008).

However, interpretation of this successful far transfer to G_f as a result of n -back training should be made with caution. It was unexplained, in their dual n -back study (Jaeggi, et al., 2008), why RAPM was used only in the group trained for eight days while other groups were tested on the BOMAT. A possible explanation may stem from potential ceiling effects with RAPM in university students, which would restrict the ability to detect change with WM training. Whereas BOMAT, as claimed in Jaeggi et al.'s (2010) later study, had less ceiling effect issues and left more scope for scores to improve. Indeed, there was no certainty as to whether these factors could have contributed to the significant findings for groups trained longer than eight days, but the consistency in measures used would highly improve the strength of the comparison. Furthermore, they described shortening the testing time on G_f measures (RAPM and BOMAT) without clear justification other than "to keep the pre- and post-test sessions short enough" (Jaeggi, et al., 2008, p. 6832). Whether this change affected participants' performance on the tests was unknown and could have important impact to the proposed findings.

Nevertheless, later studies testing the effectiveness of the single n -back task have also successfully demonstrated far transfer to G_f . Students from the National Taiwan Normal University were trained in either the dual n -back task or the single n -back task with performance compared with a no-training group. Similar to the earlier study, the RAPM and the BOMAT were used to measure far transfer effects, but were administered to all groups. The groups were matched according to their age, gender, and pre-training test performance. Both training groups showed greater improvements on measures of G_f than the control group. Comparing the two training groups, the single n -back group appeared to have improved significantly more than the dual n -back group on their trained task ($p < .05$). Evidence also indicated that the single n -back task was as effective as the dual n -back task in enhancing G_f (RAPM: $p < .01$; BOMAT: $p < .05$; (Studer, et al., 2009). However, no comparisons were made with an active control group of either non-adaptive nature or non-WM nature.

Furthermore, evidence of far transfer to G_f has been demonstrated in typically developing children using n -back tasks. Jaeggi, Buschkuhl, Jonides, and Shah (2011) recently tested the adaptive single n -back task on children where they modified and presented task stimuli in a child-friendly manner similar to a video game. The study included 62 typically developing 7- to 10-year-old students, of whom 32 participated in the intervention group and 30 in the active control group. This time, the authors included an active control group, which was trained on knowledge and vocabulary-based tasks, but no waitlist control. The duration of training was controlled between groups. All participants took part in the same pre- and post-assessments on measures of G_f using Raven's Standard Progressive Matrices (RSPM) and the Test of Nonverbal Intelligence (TONI). A composite score was derived from the two measures for analysis. A 10-question self-report questionnaire was also administered to participants

at post-test as a measure of engagement in their training. They were able to replicate findings showing a significant far transfer effect to G_f from their n -back training only when they split the intervention group into two, namely a large training gain and small training gain group: significant group differences were only found when comparing the large training gain group with other groups (large training group vs small training group: effect size $d = 0.80$; large training group vs active control group: effect size $d = 0.55$). The results of the post-training questionnaire were in line with their other results, showing that participants in the small training gain group gave higher ratings of difficulty during training. Moreover, rather than an emphasis on dose of training time, as mentioned in their first dual n -back study (Jaeggi, et al., 2008), the findings for this study indicated that the amount of training gain on the n -back task was the essential element to achieve far transfer to G_f (Jaeggi, et al., 2011). The authors' conclusion here supported the theory that higher WMC, which required better ability in executive attentional control, was the essential underlying mechanism behind G_f improvements in training (Engle, 2002; Kane, et al., 2001). However, this calls into question the direction of causality. If participants with the potential to do well at G_f measures also seem to learn to do the n -back task more effectively than those with less potential to do well, then it is no longer a true experimental effect because participants are not effectively getting the same treatment. Nonetheless, participants in Jaeggi et al.'s (2011) study reported that the initially high- G_f group displayed smaller training gains than the initially low- G_f group and also stated that the initially high- G_f and initially low- G_f groups did not differ significantly in the amount of transfer displayed.

Jaeggi et al.'s (2011) study not only provided evidence to extend the effectiveness of n -back training to typically developing children but also described findings on their self-report questionnaire. This assessment of participants' engagement

level was analysed and linked back to their findings, which other computerized WM studies have rarely included. It appears that the anxiety associated with the increase in the cognitive demands of the task played a role in individual performance when evaluating training effects. Although the Cogmed online training program is now available to the public as a paid training program and offers regular coaching by trained staff, very few published studies have incorporated analyses regarding the coaching component and how it may be linked with their findings and individual performance. Some may have mentioned its existence briefly (Lohaugen, et al., 2011; Nutley, et al., 2011) and some have omitted it altogether (J. Holmes, et al., 2009). Whether or not participants' self-reported experience contributes to the utility of an effective cognitive training program continues to be an area that requires more investigation in future studies.

Moreover, despite described success in far transfer effects to G_f , both of Jaeggi et al.'s (2008; 2011) studies included only one control group, for example, either a passive control group or an active control group, for comparison. The inclusion of both would lead to stronger comparison analyses, as noted earlier, enabling the studies to account for possible practice effects and expectancy effects, respectively. Each of these studies being the first to link dual n -back and single n -back training, respectively, with improved G_f have room for methodological refinement.

In contrast to these successes, Redick et al. (2012) as well as Chooi and Thompson (2012) were unable to find transfer to G_f in their dual n -back training studies on university students. The two studies included both passive and active control groups, with Redick et al.'s (2012) active control group participating in an adaptive visual search task, which is a similarly challenging yet different task, and Chooi and Thompson's (2012) active control group participating in a lower dose training schedule.

These two recent studies made improvements in their methodology in determining the effectiveness of dual n -back training as they also included multiple outcome measures covering both near transfer and far transfer. Indeed, the results of these two studies highlighted the need to view earlier findings with more caution as their improved methodology, as compared to that of Jaeggi et al. (2008; 2011), did not support earlier successful transfer (Slagter, 2012; Sternberg, 2008). However, it should also be noted that both of these studies used a relatively long daily training duration than previously documented dual n -back studies. A prolonged training period may lead to a possible reduction in engagement level and the repetition from a long duration of dual n -back task can also lead to boredom and decrease in motivation. These factors together may have significant impacts to successful learning (Green & Bavelier, 2008; Vygotsky, 1978/1997; Yerkes & Dodson, 1908).

In terms of CWM training studies, none has documented successful far transfer to G_f (Chein & Morrison, 2010; Loosli, et al., 2012). A closer look at their training schedules (see Table 5.1), may suggest a possible explanation for their failed attempts. Previously documented successful far transfer to G_f has utilized a schedule with a minimum training time of 15 minutes and no longer than 25 minutes each day over a 4-6 weeks period. However, Chein and Morrison (2010) requested their participants to train for 30-45 minutes per session. As stated earlier, prolonged training may have an impact motivation and learning (Endler, Rey, & Butz, 2012; Vygotsky, 1978/1997; Yerkes & Dodson, 1908). As for Loosli et al. (2012), the case was an opposite one where participants were requested to train for only 12 minutes per session over 2 weeks. Their training schedule is much shorter than previously documented successful attempts in the literature and thus it comes to no surprise when no far transfer to G_f was detected.

In summary, the literature on far transfer to G_f has shown that dual n -back

training gained more success than other training regimes. Successful results stem from a training time that ranged from 15-25 minutes each day over 20 sessions from 4-6 weeks. Those that reported using training time longer than 30 minutes showed no transfer. Available effect sizes suggested an improvement on G_f ranging from small ($d = 0.32$) to medium ($d = 0.65$) when intervention groups were compared with other control groups. Although evidence appeared to be consistent at this point, it is apparent that many of the successful transfer effects to G_f stem from research by Jaeggi (2008) and her colleagues as well as from the same university (Jaeggi, et al., 2010; Studer, et al., 2009). More research from different research parties would enhance the reliability and decrease possible researchers' bias of these current outcomes. Moreover, given the theoretical explanation behind the utility of CWM span and dual n -back training, CWM span training should also demonstrate transfer to higher-cognitive functions such as those tested in RSPM tasks. Perhaps the use of previously successful training schedules is needed in the replication of CWM training studies.

Does Training Gain Persist?

The next question of concern is whether the described near transfer and far transfer effects detected immediately post intervention persist and for how long they are sustained. Only six out of 17 of the reviewed studies included follow-up assessments to determine whether improvements in training persisted. Of those that did, follow-up assessments were performed from three to six months post intervention (J. Holmes, et al., 2009; J. Holmes, et al., 2010; Jaeggi, et al., 2011; Klingberg, et al., 2005; Lohaugen, et al., 2011; Lundqvist, et al., 2010). Refer to Table 5.1 for details regarding studies with follow-up assessments.

More follow-up assessments were found amongst Cogmed training studies than other types of training. For example, Klingberg et al.'s (2005) study on children with

ADHD reported that, as compared with baseline, their participants' gain on the span board task ($p < .001$, $d = 0.93$) and digit span task ($p < .01$, $d = 0.59$) remained significant at 3 months follow-up assessment (span board task: $p < .002$, $d = 0.92$; digit span task: $p < .03$, $d = 0.57$). Their effect sizes were comparable with those recorded immediately post intervention. However, performance gains on Stroop and RCPM were no longer significant at follow-up assessment (Stroop: post assessment: $p < .025$, $d = 0.34$; follow-up: $p = .07$, $d = 0.25$; RCPM: post assessment: $p < .01$, $d = 0.45$; follow-up: $p < .12$, $d = 0.30$). Lohaugen et al. (2011) also reported that their ELBW cohort's performance improvement on digit span ($p < .01$) and spatial span ($p < .001$) remained significant at six months follow-up assessment when results were compared with baseline performance (follow-up assessment for digit span: $p < .05$; spatial span: $p < .01$) but performance on the Letter-Number Sequencing task was no longer significantly higher than baseline. From the results of these follow-up assessment, near transfer was more sustainable than far transfer with larger effect sizes holding up slightly better than smaller effect sizes.

Similar results were documented for Cogmed training studies on children with low working memory (J. Holmes, et al., 2009) and ADHD (J. Holmes, et al., 2010) at their 6-month follow-up assessments. When comparing performance between baseline and follow-up, measures on visuo-spatial STM, verbal WM, and visuo-spatial WM remained significant (J. Holmes, et al., 2009; J. Holmes, et al., 2010). However, both studies documented that their participants' performance on measures of verbal STM were not sustained, with results of a significant decline between post-intervention assessment and follow-up assessment ($p < .05$, $d = 0.47$) for the ADHD cohort (J. Holmes, et al., 2010) as well as non-significant differences between baseline and follow-up (post assessment: $p < .01$, $d = 0.62$; follow-up: $p < .13$, $d = 0.44$) for the low

WM cohort (J. Holmes, et al., 2009). Holmes et al. (2009) further documented a significant result at 6-month follow-up for their low WM cohort, compared with baseline performance, on measures of math using WOND ($p < .01$, $d = 0.49$), which had not previously been significant at post-assessment ($p < .31$, $d = 0.11$). The authors suggested that this was expected because any cognitive gain that transfers to learning required some time before significant results could be reflected in standardized tests. However, this was the only study to document a significant result at follow-up but not at post-test.

The only dual n -back training study with an inclusion of a follow-up assessment was published by Jaeggi et al. (2011). The study reported that significant gains on their RSPM and TONI composite score remained significant compared to baseline at both post assessment ($p < .05$, $d = 0.80$) and 3 month follow-up ($p < .05$, $d = 0.92$). No studies from CWM training included follow-up assessments. Clearly, more studies that include follow-up procedures are necessary in order to compare maintenance amongst different training regimes.

The Importance of the *Adaptive* Nature of Training

The importance of the adaptive nature of WM training is discussed here. Adaptive WM training refers to providing participants with an individual training experience whereby the training taxes WM with an optimally challenging dose. This means that the difficulty of the training task is automatically adjusted for each trial to match the individual's WM performance. This adjustment is implemented in small increments or decrements after the individual has demonstrated mastery or inadequacy, respectively, of their current level of difficulty. For example, as the participants' performance improves, the difficulty of the training task increases in order for it to maintain a suitable cognitively challenging level, whereas if the participants'

performance deteriorated, the difficulty decreases (Green & Bavelier, 2008; Pearson Inc., 2011).

Adaptiveness is considered an important feature in the currently discussed computerized WM training literature. Concepts stemming from the Zone of Proximal Development Theory (ZPD; (Vygotsky, 1978/1997) and Yerkes-Dodson Law (Yerkes & Dodson, 1908) may assist in understanding its importance as they have been influential to the literature of motivation, learning, and task performance. According to the ZPD theory, Vygotsky (1978/1997) proposes that as the proximity between the difficulty level of a task and one's capability increases, the chance of successful learning also increases. In contrast, discrepancy between the two decreases the chance of successful learning. Similarly, the Yerkes-Dodson Law (Yerkes & Dodson, 1908) proposes that optimal arousal from difficulty level of a task maximizes performance gain. This law states that, in order for one to actively learn or seek knowledge, a certain degree of arousal and motivation, which provides intrinsic energy for one to move towards directed behavior, is necessary. In simple tasks, high arousal leads to optimal task performance. However, simple tasks provide little inherent arousal, though this can be optimized by conditions that increase arousal, for example external reward or punishment. This arousal forms an inverted U-shape with performance on task. Difficult tasks, on the other hand, induce too much arousal and the anxiety associated with the increase in task difficulty and arousal after exceeding the optimal level tends to block learning rather than activate learning. Therefore, tasks well matched to a person's ability optimize arousal.

Endler, Rey and Butz (2012) investigated the impact of adaptive features of a small group of participants ($n = 37$), ranging from 16 to 48 years of age, using an adaptive e-learning program. The authors concluded that their results fit well with the

ZPD ((Vygotsky, 1978/1997) and Yerkes-Dodson Law (Yerkes & Dodson, 1908) with key areas, such as interest, challenge, and anxiety, being influential to motivation and learning. Accordingly, by adapting the difficulty level of the WM computer training task to one's ability on a trial-to-trial basis, it is likely to optimise individuals' arousal for learning as well as maintain a suitable level of motivation for successful learning (Endler, et al., 2012; Teigen, 1994; Vygotsky, 1978/1997; Yerkes & Dodson, 1908).

Other Factors Affecting Learning and Possible Transfer of Learning

There are other determinants of successful learning and transfer of learning aside from the importance of an adaptive task with emphasis on motivation and arousal discussed above. Factors such as feedback and reward are also likely to play an important role. Additionally, variability in the task and context can increase the chance of transfer.

Feedback and rewards are considered a learning signal to allow participants and experimenters to distinguish whether answers to questions are correct or not. With correct interpretation of feedback information, algorithms can be created to tailor adaptive training programs. As well, feedback allows participants to assess their own accomplishments or lack thereof (Green & Bavelier, 2008) and adjust their effort accordingly. On the other hand, the use of extrinsic reinforcement and reward has often been found to be useful for learning and is indeed used by academic teachers in schools for typically developing children as well as those with learning disabilities (Flora, 2004). The usefulness of a reward also has important implications. According to Green and Bavelier (2008), the value and worth of an identical reward is different for every person and is referred to as "the relative desirability of a reward" (p.11). Accordingly, feedback and rewards also impact motivation to a different extent for different individuals.

The possibility of transfer has been related to variability. Transfer between tasks is proportional to the overlap in stimulus-response pairings. If two tasks use completely different stimuli, or require completely different responses (or both), then typically no transfer can be found. Variability refers to extending the range of stimulus-response pairings in training, thus the second task will be more likely to be in common with it. If the second task has nothing in common, no transfer is expected, regardless of the variability of the trained task. Both the variability of the task, where learning is undertaken, and the context in which stimuli are presented can affect transfer success. When variability is high, it enhances the flexibility of the learning and thereby reduces the chance of dependence on learning only strategies or information specific to a specific form of the task. In other words, higher variability enables higher generalization probabilities (Green & Bavelier, 2008). The transfer of learning is the exception rather than the rule, therefore when transfer occurs it captures scientists' attention. Thus, some have considered the current evidence of transfer effects following WM training as mere build up of familiarity with Raven's like matrices and stimuli (Moody, 2009; Shipstead, et al., 2010).

Summary of the Various Training Regimes

Cogmed Training. Some consistent evidence has shown that Cogmed training improves non-trained WM demanding tasks, but conclusions about their transfer effects to higher-order cognition, such as G_f , cannot be drawn from existing evidence. Aside from Klingberg et al.'s (2005; 2002) earlier studies of children with ADHD and four healthy adults, no other Cogmed training have shown both near and far transfer to non-trained WM and G_f respectively. Given existing strong evidence on the predictability of G_f from WM (Engel de Abreu, et al., 2010; Fry & Hale, 2000; Hornung, et al., 2011), these results subsequently raise concerns about the generalizability of Cogmed training.

Does Cogmed training work within the non-clinical population? Klingberg et al. (2002) suggested that initial impairments in cognitive function or WM were not required in order to see improvements in WM through Cogmed training. They claimed this because their four healthy participants demonstrated significant improvements in the span board task, Stroop task, and RCPM task after training. However, this conclusion seems rather overreaching given the small number of participants. Although there have been recent published articles on the efficacy of Cogmed training with healthy pre-schoolers (Nutley, et al., 2011; Thorell, et al., 2009), these studies continue to show successful transfer to non-trained WM measures and yet no transfer to G_f measures. Thus, whether the receptiveness to benefits from WM training is dependent on initial WM impairments or not, it appears inconclusive due to limited studies with healthy participants across different age groups, particularly in regards to transfer effects to G_f . In addition, on one level, if WM training can assist or “cure” individuals with cognitive impairments, then training for non-clinical cohorts do not provide much important information. However, at a theoretical level, if training enables researchers to distinguish between WM training benefits and WM training benefits that apply only to those who show deficiency in what is being trained, then it is of much more importance as it also distinguishes underlying differences between normal individuals and those with deficiency.

Accordingly, a subsequent concern is the use of clinical populations in the experimental research with Cogmed training. Many studies investing the utility of Cogmed training invited clinical populations, including ADHD, patients with stroke and brain damage, individuals born premature and low birth weight. The reliance on difficult-to-access populations with various cognitive and motor limitations constrains the full experimental investigation of what makes the training task effective. It also

limits studies' power to identify significant effects.

Dual *n*-back Training. Research on dual *n*-back training has provided some astonishing outcomes with their demonstration of success in far transfer to G_f after training. However, the conclusion about the theoretical interpretation of these findings is unclear and the utility of dual *n*-back training is still debated for several reasons. First of all, evidence on successful transfer to G_f is based on studies that did not include an active control group (Jaeggi, et al., 2008; Studer, et al., 2009). When comparisons were made only between the training group and a passive group, results of far transfer could be explained by various factors other than training, for example test-retest practice effects as well as motivational and engagement level. Similarly, significant transfer results from comparisons between the intervention group and the active control group (Jaeggi, et al., 2011) without a passive/wait-list control group fails to control for placebo effects. In dual *n*-back studies where all three groups were included, transfer to G_f could not be detected (Chooi & Thompson, 2012; Redick, et al., 2012). This methodological concern raises doubts of the utility of this training regime.

Secondly, the aforementioned studies showing successful transfer to G_f differ from those that did not find successful transfer in terms of training schedule. The former group detected transfer in daily training that ranged from 15-25 minutes whereas, paradoxically, the latter group, which did not detect transfer, asked their participants to train for 30-40 minutes each day. Could shorter training lead to more success than longer training? Conclusions about the utility of dual *n*-back training cannot be made until replication of successful transfer using the same training regime and schedule is presented.

Moreover, dual *n*-back training can be very challenging as it requires participants to respond to visual and auditory stimuli simultaneously. It has been tested

primarily amongst university students (Chooi & Thompson, 2012; Jaeggi, et al., 2008; Redick, et al., 2012; Studer, et al., 2009). The suitability of the task for children is doubtful.

Complex Working Memory Span Training. Kane et al. (2001) explained that individuals with higher memory span were more proficient in inhibiting distracters and focusing on pertinent stimuli required for further processing, where a high WM capacity did not only refer to the ability to store more items for active processing but also included a higher attentional control capability. CWM training aims to increase WM capacity and has been suggested to predict higher-order cognitive functions, such as those measured by G_f tasks (A.R.A. Conway, et al., 2005; Engle, 2002; Kane, et al., 2001; Unsworth, et al., 2009). Therefore, this approach provides a firm theoretical background in explaining the underlying mechanism of any successful near and far transfers.

However, research using CWM training remains rather new as compared to the use of Cogmed and dual n -back. Although CWM training has been tested in both children (Loosli, et al., 2012) and adults (Chein & Morrison, 2010), their training results were only compared with a passive control group. Again, this raises the problem of Hawthorne effect and their results of successful transfer may only be practice effects. Furthermore, the training schedules in these studies varied from research that previously documented successful transfer and, therefore, proper comparisons are not possible. Therefore, there continue to be many gaps to fill in determining the effectiveness of CWM training.

Ideal Characteristics of Working Memory Training Studies

Provided with the current understanding of WM training and recommendations from recent reviews, the ideal working memory training study should meet certain

methodological requirements. First, participants should be randomly assigned to the different groups. This would ensure that the effects of training are not a result of pre-existing group differences. Second, the intervention group should be compared to both an active control group and a passive control group. This is to account for Hawthorne and practice effects, respectively. Third, ideally there should be multiple measures as indicators of any specific construct. This would enhance the likelihood that any transfer effects are a result of training and not a result of learning task-specific strategies.

In addition, motivational factors and issues with participants' compliance should be given attention, since intensive WM training requires participants to train continuously every day on an adaptively challenging task, repetitively. Motivation and engagement in the training task affects learning success (Green & Bavelier, 2008; Vygotsky, 1978/1997; Yerkes & Dodson, 1908). This holds true particularly in evaluating WM training with children (Jaeggi, et al., 2011; Prins, DAVIS, Ponsioen, Brink, & van der Oord, 2011) because continual modification in WM training tasks to provide motivational support, such as including game-like elements, may enhance learning abilities (Harlen & Crick, 2003; Prins, et al., 2011). Existing studies have often overlooked the importance of this issue and seldom report their procedures in coping with this factor. Some studies have included questionnaires after training to learn about participants' engagement level or difficulties that they experience (Lohaugen, et al., 2011; Nutley, et al., 2011), but these findings have not been reported quantitatively so that comparison across studies can be made.

Participants' compliance is another issue that has seldom been addressed. Given that daily training can be repetitive and boring, it can be difficult for researchers to ensure participants comply with pre-assigned training schedules. Only Cogmed studies have described the use of a coach to ensure that participants comply with training

routines (Lohaugen, et al., 2011). Without properly addressing the possibility of non-compliance, the study could face a high dropout rate in participants, thus reducing the power of any experimental study.

Summary

In conclusion, while more evidence to confirm the effectiveness of Cogmed, *n*-back, and CWM training is still required, the aforementioned studies provided valuable evidence that adds to the computerized WM training literature. The understanding of the underlying mechanisms behind WM training and whether or not it provides enough justification to refute the notion of intelligence being a fixed trait continue to be controversial. However, reviewed findings are amongst the first to challenge the consensus that an individual's WMC (Z. Chen & Cowan, 2008; Cowan, 2000; Miller, 1956; Rickers, AuBuchon, & Cowan, 2010) and G_f (Cattell, 1987; McArdle, et al., 2002) are fixed. They propose the possibility that intensive WM training may improve performance on trained tasks as well as generalizing to non-trained abilities.

Given the discussion of cognitive impairments amongst children born preterm/LBW in earlier chapters and that the sequelae of impairments stretch into adolescence and adulthood, early intervention for this group seems desirable. However, there are still many gaps to be filled in the literature of cognitive training in preparation for extensive research amongst the children born preterm/LBW. As seen to date, only one Cogmed study (Lohaugen, et al., 2010) has focused on this population and its results remained within the scope of near transfer. On the other hand, dual *n*-back appears too difficult and challenging for children born preterm/LBW. Although training on their single *n*-back version with children (Jaeggi, et al., 2011) has demonstrated far transfer, the theoretical explanation of this effect remains uncertain due to a lack of concurrent near transfer evidence. In contrast, the theoretical underpinning for CWM

adaptive span training is relatively clear. Although CWM training has only been tested in undergraduates (Chein & Morrison, 2010) and adolescents (Loosli, et al., 2012), more research using modified game-like and friendly versions may disclose their suitability for children born preterm/LBW. However, given that the CWM training regime is still at its early stage of investigations, typically developing children should be tested before experimenting with more vulnerable clinical cohorts.

CHAPTER 6

Study 2: Adaptive working memory span task training to increase fluid intelligence in typically developing children

Introduction

Many recent studies have suggested that computerized WM training may enhance WM performance, as well as G_f (Jaeggi, et al., 2008; Jaeggi, et al., 2011; Klingberg, et al., 2005; Klingberg, et al., 2002). Computerized WM training may be beneficial for children born preterm/LBW as well. This is because previously discussed outcomes clearly demonstrate that prematurity at birth is associated with WM impairments (Pritchard, et al., 2009) and as demonstrated in study one, not only did children born preterm display impairments in G_f but WM was also a significant partial mediator of differences in G_f . In addition, WM and academic achievement are highly correlated, particular in respect to math and reading underachievement pertinent in this clinical cohort (Alloway, 2009; Alloway & Alloway, 2010).

Reported findings of significant improvement in performance of the WM training task itself as well as transfer to non-trained cognitive tasks were amongst the first to cast serious doubt on the permanence of individual differences in WM capacity (Cowan, 2000; Rickers, et al., 2010) and G_f (Cattell, 1987; McArdle, et al., 2002). These findings are of great interest because past attempts to increase intelligence have not had much success (Moreno, et al., 2011; Tomporowski, et al., 2008; D. C. Turner, et al., 2005). Currently published studies have mostly experimented with university students (Chein & Morrison, 2010; Chooi & Thompson, 2012; Jaeggi, et al., 2008; Redick, et al., 2012; Studer, et al., 2009), and clinical groups including brain injured patients (Lundqvist, et al., 2010; Westerberg, et al., 2007), children with ADHD (J. Holmes, et al., 2010; Klingberg, et al., 2005; Klingberg, et al., 2002) and low working memory (J. Holmes, et al., 2009). Only one study assessed the outcome of WM training

in children of LBW, which reported significant near transfer to several measures of untrained WM following Cogmed training. However, no far transfer measures were included in the study (Lohaugen, et al., 2011). As described above, several WM training studies have been done with clinical groups. The frequent reliance of many previous studies on difficult-to-access populations with various cognitive and motor limitations has restricted the full experimental investigation of what makes the training task effective, as well as limiting studies' power to detect significant effects if they exist. Typically developing school-aged children have rarely been involved, despite the obvious relevance of such research to a school-aged population.

Amongst the growing number of computerized WM training studies in the field, the chosen task for the current study – the complex working memory (CWM) training – will be one adopted from Loosli et al.'s (2012) study. The CWM training used a classic and specific complex span task that allowed clear depiction of the association between any improvements in training and the processes involved in complex span task. CWM training on a narrow task stands in contrast to Cogmed WM training studies that used a combination of tasks in their training program. CWM training would not be expected to show transfer given the lack of overlapping stimulus-response demands between tasks. Therefore, if transfer effects are present, the only theoretical explanation for it is Kane et al.'s (2001) theory on executive attention. As well, stepping away from Cogmed's resource demanding methods for engaging children in order to keep them practising and dual *n*-back task's challenging simultaneous auditory/visual requirements, it becomes important to see whether CWM tasks can be adapted to make them sufficiently enjoyable and optimally challenged to persist with.

Therefore, this current study's main goal is to assess the utility of an adaptive computerized CWM span task in typically developing children to attain preliminary

support, or otherwise, for the potential use of this training in children born preterm/LBW. If CWM training is successful in achieving a better score on measures of G_f that is due to genuine increase in G_f through enhancing executive attention, then it is likely to transfer to real life and may assist in further remediation of academic difficulties found in children born preterm/ LBW. In contrast to Loosli et al.'s (2012) original study, the present study will extend the training from their 10 sessions to the more commonly adopted 20-day regime. Other methodological changes are also discussed below.

Previously published papers on CWM training have not documented the inclusion of an active control training group to control for practice effects and possible Hawthorne effects (Chein & Morrison, 2010; Loosli, et al., 2012). Therefore, this study will include a non-adaptive training group that practices the same span training at a lower dosage without challenging WM capacities as opposed to only a non-training passive control group. The total time spent in front of the computer for the non-adaptive version of training will also be controlled, making it identical to the adaptive intervention group (Shipstead, et al., 2010). Essentially, if the adaptive training group shows more transfer improvements than the non-adaptive active control group and the non-training passive control group, then there is evidence to confirm the significant contribution of the adaptive complex WM span training. However, if the performance on the measures of the two training groups demonstrates no significant differences, then the results may be interpreted as a Hawthorne effect.

The particular age groups, between seven and eight years, are examined. This is because it has been suggested that cognitive development transpires quickly at this early school age and that critical periods of change begin around the age of seven years (Diamond, 2002). This age group is also chosen in preference to the 9-year-old age

group because of the findings in Study 1. Given that the 9-year-old age group showed birth status differences in G_f that were partially mediated by WM and that the preterm cohort appeared to demonstrate at least one year of developmental lag in WM and G_f , training that serve as early intervention and learning programs that enhance the overall educational outcomes of children (Blair, 2002; Ramey et al., 2000) would need to happen before the age of nine.

This study will also include more diverse cognitive measures than Loosli et al.'s (2012) study, allowing it to test the breadth of generalization and to establish its limits. Outcome measures will include measures of untrained WM: digit span and spatial span tasks; a measure of executive attention: the Stroop task; a measure of processing speed: the choice reaction time task; and a measure of G_f : Ravens Progressive Matrices. It is expected that if improvements are observed in the intervention group on the adaptive span training task, demonstrating effortful and motivated training, then untrained WM tasks, executive function tasks and the G_f task will also improve if the training effects are due to changes in underlying working memory capacity and executive attentional control, and these are causally related to G_f . In contrast, a significant transfer effect is not expected to show in speed of processing tasks following adaptive span training because speed is only moderately correlated to working memory, and, according to cascade models (Fry & Hale, 1996) is causally “upstream” from WM.

The Current Study

This second study examines the utility of computerized adaptive animal span WM training for typically developing children, as a preliminary step to developing future interventions for children born preterm/LBW. The aim is to investigate whether near and far transfer occur to a greater degree in the intervention group than the active and passive control groups. To this end, the occurrence of the following phenomena in the

intervention group after adaptive training will be assessed:

- 1) Whether specific training effects on the working memory span task are observed after training.
- 2) Whether near transfer effects to untrained working memory tasks are present.
- 3) Whether a far transfer effects to executive function is present.
- 4) Whether a far transfer effect to speed of processing is present.
- 5) Whether a far transfer effect to fluid intelligence is present.

Hypotheses

To the extent that training changes the underlying construct of WM, rather than merely familiarising participants with superficial aspects of the training task, it is hypothesized that:

- 1) There will be a significant improvement in the training task.
- 2) There will be significant near transfer effects to the digit span and spatial span tasks.
- 3) There will be a significant far transfer effect to the Stroop task.
- 4) There will not be a significant far transfer effect to a reaction time task.
- 5) There will be a significant far transfer effect to the RSPM task.

Method

Participants

Initial recruitment included 92 Year 2 students from six public primary schools, three Catholic primary schools and four referrals recruited through a snowballing method from the suburbs of Perth, Western Australia. Consent forms were signed by each school's principal to allow for distribution of invitation letters. Year 2 teachers of participating schools distributed invitation letters containing information about the study, together with consent forms to their students and parents. Interested participants

replied by completing the consent forms and returning them to their class teachers. A total of 638 information sheets and consent forms packages were distributed to the participating schools with a return rate of 13.8%. Copies of invitation letters are presented in Appendices B through E.

Next, screening questions were asked via telephone to ensure that participants met inclusion and exclusion criteria. Inclusion criteria were that participants were, at the time of testing, (1) attending Year 2 at school and (2) had a computer with Internet access at home. Exclusion criteria included participants who had (1) hearing or visual disabilities that would influence pre/post assessments and the use of the online computer intervention program, (2) current diagnosis of clinically significant mental health illness requiring therapy, (3) current diagnosis of intellectual disabilities, or (4) current medical illness that required ongoing medical attention.

After the initial screening process conducted via telephone, one student was excluded for not meeting the screening criteria and 10 decided not to proceed with the study. A total of 81 students proceeded through pre-test assessment and were allocated to one of three groups. However, 17 students did not complete 20 days of training as required and thus were excluded from the current analyses. Another student was excluded from the analyses due to high distractibility at testing, which led to a discontinuation of post-test assessment. A participant flow chart through recruitment and testing processes is presented in Figure 6.1.

Sixty-three students were included in the final analyses. Participants ranged from 6.72 - 8.51 years old ($M = 7.40$, $SD = .39$) at the time of initial testing. Participants were pseudo-randomly assigned to three groups: the intervention group ($n = 21$) the active control group ($n = 19$) and the passive control group ($n = 23$). Age and gender proportions are shown in Table 6.1. The attrition rate between pre and post-test were

25%, 32.14% and 4.16% for the intervention, active control and passive control group respectively.

Participants' age did not differ significantly across the three groups, $F(2, 60) = .49, p = .614$. Pearson's chi-square also indicated no significant differences in gender composition across the three groups, $\chi^2(2) = .290, p = .865$.

Table 6.1

Final Analysed Sample: Characteristics of Participants by Group

	Intervention group	Active control group	Passive control group
Mean Age (SD)	7.46 (.44)	7.41 (.38)	7.34 (.38)
Gender proportion (M:F)	13:8	13:6	14:9

Note. SD = Standard deviation; M = Male; F = Female

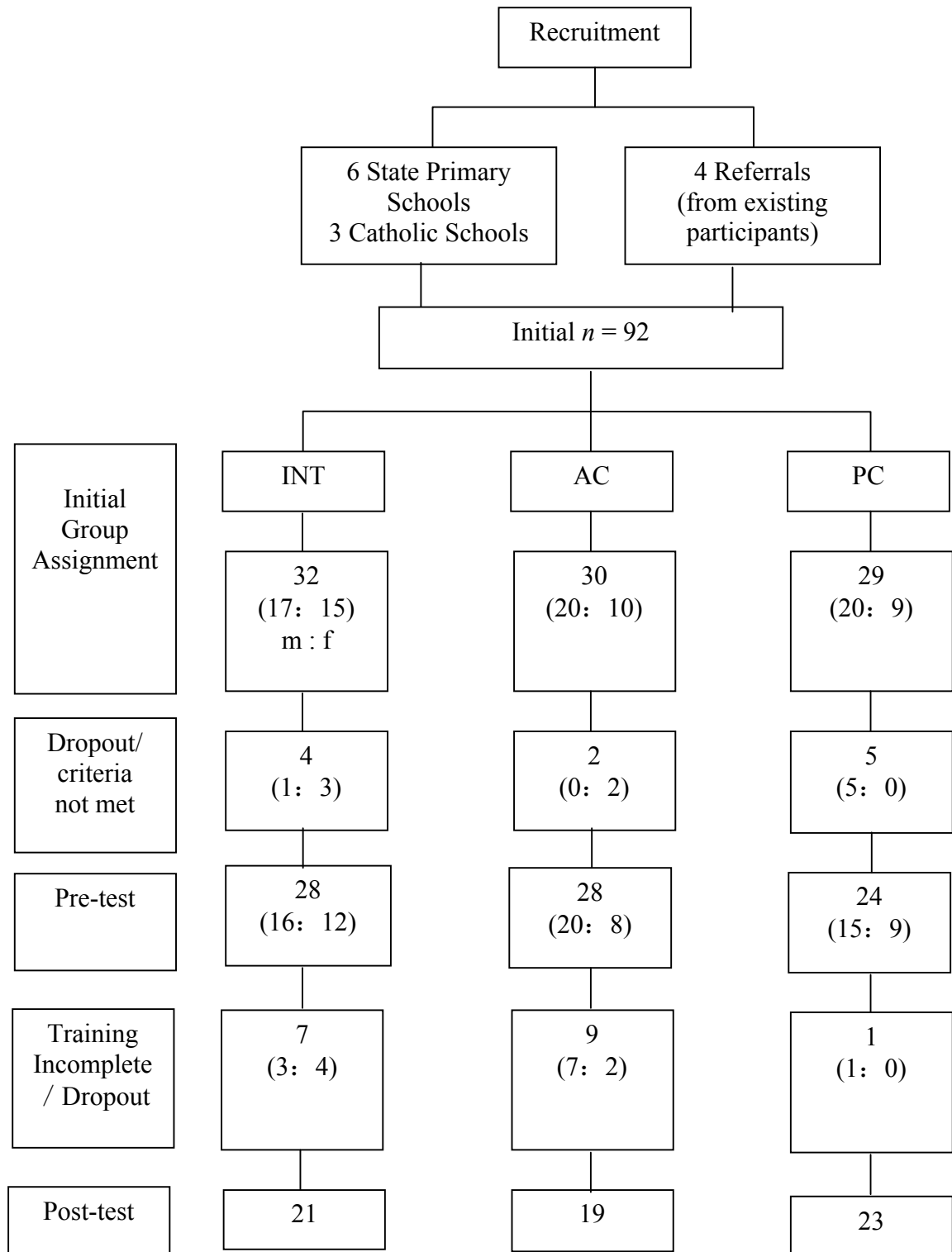


Figure 6.1. Recruitment, response rate and the flow of participants in the current adaptive working memory span training experiment. INT = Intervention group; AC = Active control group; PC = Passive control group. The ratio of males to females are included and presented as (number of males : number of females) in the figure. One student from the PC group did not meet criteria.

Measures

Digit span Task. This task measured children’s verbal WM and was one of the subtests from the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV). The task has been commonly used in the literature of WM (Swanson, 2008; Tillman, et al., 2008) and cognitive studies in clinical cohorts, for example prematurely born children (Clark & Woodward, 2010; Fraello, et al., 2011).

As described in study 1, the task required the participant to verbally recall sequences of digits, first, in sequential order for the forward span and then in reverse order for the backwards span. The combined raw score of total correct trials in both spans was used as the outcome measure. Internal consistency reliability coefficient for digit span subtest by split half correlations is .87. The test-retest reliability coefficient from 243 children over an interval of 32 days, is .81 (P. E. Williams, Weiss, & Rolhus, 2003).

A parallel version of the task was used at post assessment. The parallel version was devised by reversing the digits within each trial that consisted of the same number of digits. Examples of a 2-digit and a 3-digit trial can be found in Figure 6.2.

<u>Digit Pre-test</u>		<u>Digit Span Post-test</u>	
<u>Forward</u>	<u>Backward</u>	<u>Forward</u>	<u>Backward</u>
Example of a 2-digit trial			
2-9	2-1	6-4	3-1
4-6	1-3	9-2	1-2
Example of a 3-digit trial			
3-8-6	5-7-4	2-1-6	9-5-2
6-1-2	2-5-9	6-8-3	4-7-5

Figure 6.2. Examples of the Digit Span pre-test and its corresponding post-test used in the current study.

Spatial span board task. This task measured participants’ visual WM. The task

used in the current study is one of the optional subtests from the Wechsler Memory Scale – Third Edition (WMS-III; (Wechsler, 1997). This has also been referred to as the Corsi block task. Various versions are commonly used in studies of WM (Colom, et al., 2005) and cognitive intervention research (Klingberg, et al., 2005; Lohaugen, et al., 2011).

The test consisted of 10 randomly located cubes on a plastic flat surface with numbers 1-10 printed on the sides facing the examiner. The examiner tapped on the blocks at a rate of one block per second in a specific sequence starting with two blocks. Examinees were tested on their ability to hold the locations of a visual-spatial sequence in memory and replicate them. In particular, they were required to replicate the tapping sequence, using their finger, in the appropriate order for forward and backward span, similarly to the digit span task. As with digit span task, the length of the sequence was increased after every second trial. The test was discontinued when the child failed to replicate both trials within a level of difficulty. The outcome variable for participants' visual WM was the number of correct trials. According to the testing manual, the reliability coefficient for test-retest intervals ranging from 2 to 12 weeks for the spatial span task was documented at .72, while average internal consistency was documented at .79 (Wechsler, 1997).

Parallel versions were also used at post-test. It was devised in a similar manner to that described in the digit span task where the order of the numbers has been switched. Figure 6.3 shows examples of a 2-block and a 3-block trial.

<u>Spatial Span Pre-test</u>		<u>Spatial Span Post-test</u>	
<u>Forward</u>	<u>Backward</u>	<u>Forward</u>	<u>Backward</u>
Example of a 2-block trial			
3-10	7-4	10-3	4-7
7-4	3-10	4-7	10-3
Example of a 3-block trial			
1-9-3	8-2-7	3-9-1	7-2-8
8-2-7	1-9-3	7-2-8	3-9-1

Figure 6.3. Examples of the Spatial Span pre-test and its corresponding post-test used in the current study.

The Stroop Test. This task was used to measure participants' inhibitory ability. The test included four conditions, each presented on a separate sheet of paper, with 30 items in each condition. In the first 'Word' condition, participants were required to read the listed colour words ('red', 'blue', or 'green') aloud. All the words were printed in black ink. In the second 'Colour' condition, participants were required to name the colours that strings of Xs are printed in. The items were listed as 'XXXX' in colours of red, blue, or green. In the third 'Congruent' condition, participants were required to name the colour of the words 'red', 'blue', or 'green' printed in their matching colours. Finally, in the 'Incongruent' condition, participants were to name the printed colour of the words 'red', 'blue', or 'green', printed in colours that did not correspond to the printed word. Other than the 'Word' condition, their correct response was always the ink colour of the presented item (J.-Y. Chen, 1996; Homack & Riccio, 2004; MacLeod, 1991).

Participants were timed on each of the testing conditions. The difference in time between the 'Incongruent' condition and the 'Colour' condition was used as outcome measure. This same method of scoring has been performed in other studies as a measure of inhibition in children (Brydges, et al., 2012). Temporal reliability for each separate

condition has been documented as good ($r > .80$; (Homack & Riccio, 2004). The same version of the Stroop task was used at both pre- and post-test.

Reaction Time Task. A 4-choice reaction time task was used to measure children's speed of processing. Various versions of reaction time tasks have been used to measure processing speed in the cognitive literature (Fry & Hale, 2000; Rose, et al., 2011) and computerized WM training studies (Klingberg, et al., 2002).

A DOS-based computer program was created and used for the display of stimuli to participants. Four different coloured circles (blue, green, red, white) were used as stimuli and appeared in random order in the centre of the screen, one at a time. The circle appeared in the centre of the 15-inch screen computer with a black background. Each colour was presented 32 times within a total of 128 trials. The task required participants to press a corresponding colour key indicated on the keyboard. Children began by completing eight practice trials before being tested on 128 trials within a single block. Participants' mean reaction time was used as the outcome variable for subsequent data analysis. The instructions were presented on the computer screen and read aloud to the participants as follows:

“Press the button that matches the colour on the screen. Go as quickly as you can without making mistakes.”

Raven's Standard Progressive Matrices (RSPM; (J. C. Raven, 1958). This test was used to measure children's G_f . Empirical studies of intelligence and cognitive intervention have also used this as their measure of G_f (Dang, et al., 2012; Jaeggi, et al., 2011; Klingberg, et al., 2005). RSPM contains 60 items with 12 items in each of the five sets (A,B,C,D,E) provided in a booklet. The difficulty progressively increases within each section and between sections. In the test, participants are required to identify the missing part of a targeted geometrical design from amongst six to eight

possible choices, Sets A and B provide six choices and Sets C, D, and E provide eight choices. Burke (1972) documented the odd even split-half reliability coefficient at .96 ($n = 567$). A test-retest correlation coefficient, from a smaller sample ($n = 11$), was recently documented at .83 (J. E. Williams & McCord, 2006).

In the current study, the items were divided into two parallel versions with 30 items each. The level of difficulty for items in each set was matched using existing results of Rasch analysis from 600 Year 1 and 2 students tested by the cognitive developmental psychology laboratory at Murdoch University. The raw score of the test was used as the outcome variable.

Working Memory Training task. The intervention group was trained using an adaptive animal span WM task. The task was an adapted version of the BrainTwister program provided by the University of Bern (Buschkuehl, et al., 2008). The program automatically terminated after participants had attended to the task for 15 minutes on each training day. The program consisted of two stages and used four stimulus animal pictures throughout the training: cow, dog, horse, and pig. In the first stage, the Processing/Encoding Stage, animals appeared in the middle of the screen beginning with one animal at level one and increasing by one animal for each subsequent level. The number of animals is referred to as the 'set size'. Participants were asked to indicate whether the animals on the screen were upside down or the right way up by clicking on one of the up/down arrows that were located on the left and right of the computer screen, respectively, with a computer mouse. These animals appeared in different random orders for each trial. Crocodiles appeared at the top of the screen to indicate an error if participants clicked on the wrong arrow or exceeded the 3-second time limit without responding. This was to ensure that participants did not make random selections in the process of the task.

In the second stage, the Recall Stage, participants were required to indicate the sequence of the animals that had been presented in the first stage of the task by clicking on the animals in the correct order of appearance. The four animals reappeared on the screen for participants to choose. The recalled animals then appeared underneath a red line on the screen in the sequence of the participant's choice. This was done by clicking on each animal in order, at which the animal appears at the bottom of the screen as confirmation. No changes can be made once a choice has been confirmed.

After the completion of the two stages, a screen appeared with a histogram on the side and a basket in the middle. The histogram indicated the highest set size that the participant had reached and the basket held accumulated candies, representing the number of trials to which the participant had correctly responded. The next trial began after the Feedback Stage.

The training task had an adaptive feature whereby the set size was determined according to three rules: Rule 1 – If no prompt was required in Part 1 and recollection was correct in Part 2, then the next set size is increased by one animal; Rule 2 – If a prompt was required in Part 1 but participants reproduced a correct sequence in Part 2, then the next set size remains unchanged; Rule 3 – Regardless of the outcome in Part 1, if participants provided an incorrect sequence in Part 2, then the next set size decreases by one animal. Feedback is provided after each trial through a point system represented by a candy. Participants earn one candy for every correct attempt as described in Rule 1. No candies are earned or taken away in any other circumstances. These rules allowed the program to adjust its difficulty for each individual participant. The three rules were identical to the ones used in Loosli et al.'s (2012) study on CWM training with children.

The active control group completed a non-adaptive version of the same task. In the non-adaptive version, the number of animal stimuli presented remained constant at

two. This low dosage yet similar activity allowed participants to have the same degree of contact with the researcher, thus making the adaptive feature the key difference between the groups (Shipstead, et al., 2010). Both the adaptive and non-adaptive versions of the training tasks were located online and participants used an assigned login number and password to access it. Children and parents could also click on an 'Instructions' link on the screen to access the game instructions and the contact details of the researcher at any time. Participants' task performance data were stored online, accessible only by the researcher and supervisors of the present study. An example of the flow of a trial is captured from the actual task that the participants view online and is presented in Figure 6.4.

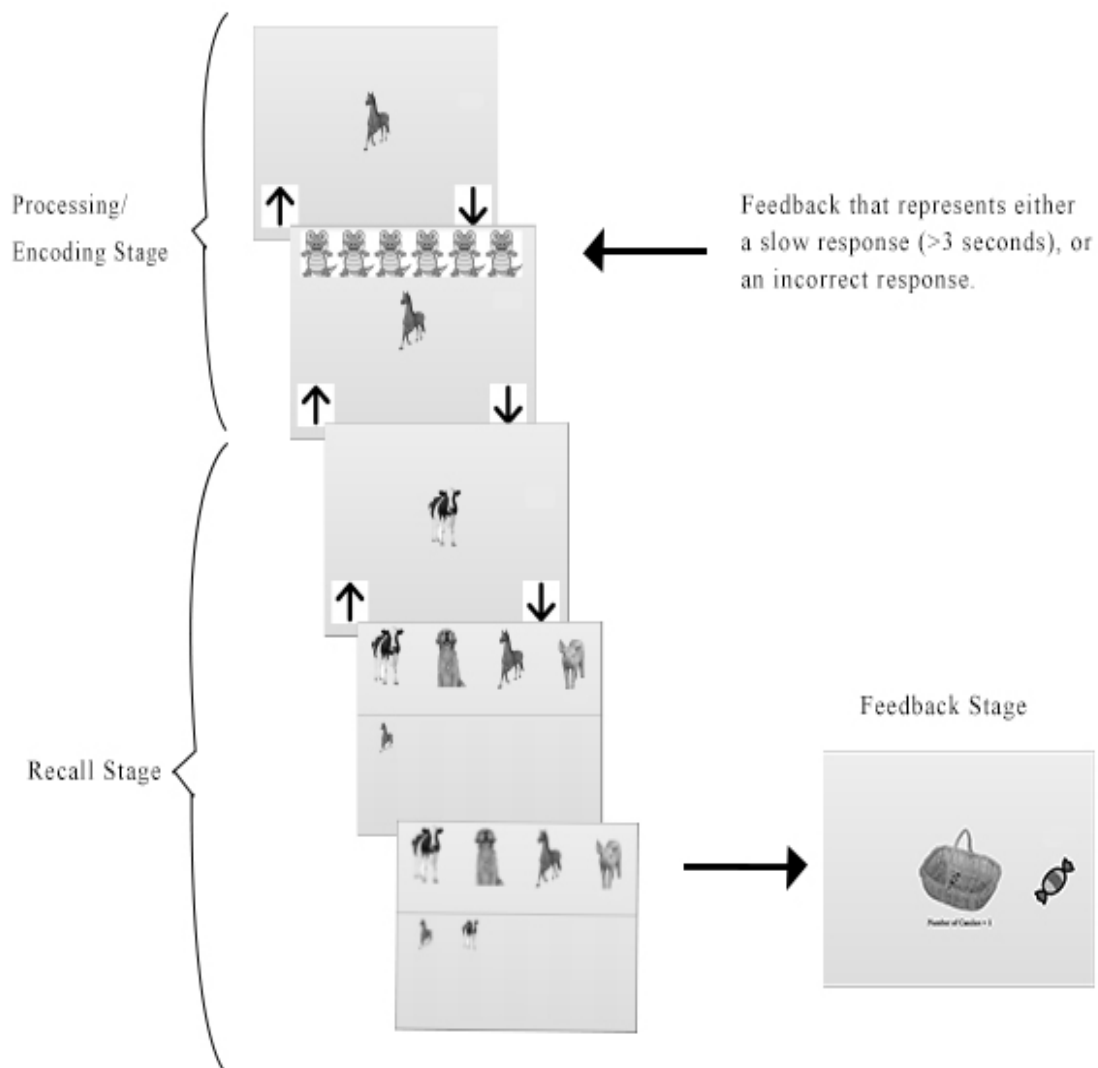


Figure 6.4. Experimental Procedure of the Adaptive Working Memory Span Training. This demonstrates a set size of two animals in sequential order as seen on the computer screen.

Procedure

This study was granted ethical approval by the Human Research Ethics Committees of Murdoch University in Western Australia (see Appendix F), the Department of Education of Western Australia (see Appendix G), and the Catholic Education Office of Western Australia (see Appendix H).

Two major periods of data collection were conducted over two school terms and holidays, with participants joining the research at different times from different schools throughout the data gathering process. Participants were assigned to one of the three groups pseudo-randomly by continuously matching gender and the number of participants per group prior to baseline assessment. Pseudo-random assignment has been used in other cognitive training studies (Jaeggi, et al., 2011). Subsequently, participants were assigned a number in one of the groups according to a pre-set sequential order (#1 - INT, #2 - AC, #3 - PC, #4 - AC, #5 - INT, #6 - PC), and so on. This pattern was used in order to assist later categorization of the online training task, where intervention group participants would key in their personal code (which was always odd) and an odd number activated the adaptive version of training, while an even number activated a non-adaptive version for the active control group. Only participants who completed 20 days of training within six weeks from pre-assessment were included in the data analyses. The researcher, also the author of this thesis, completed all data collection and assessments following verbatim test instructions for this unblinded study.

For participants who met criteria for the study, a time convenient to the participant and a parent was arranged for individual pre-testing in the child's home. All participants took pre- and post-tests on the five outcome measures prior to and after training. The intervention group and the active control group completed 4-6 weeks of online WM training in the adaptive and non-adaptive versions of the training task, respectively. These participants accessed the task online each day and completed 15 minutes of training for 20 days, 5 days a week and excluding weekends. However, if participants missed a day of training during the weekdays, they were given the option to make up for the day of missed training on the weekends. The passive control group did

not engage in any online training.

For those assigned to the intervention group and active control group, once pre-testing was completed, the researcher demonstrated to both the participant and the parent how to use the online program as well as a reinforcement schedule. To ensure that participants in the intervention and active control group were reinforced to continue throughout the online training, participants received a star sticker on a progress chart each day upon completion of the task. Parents of participating children were asked to assist with monitoring task completion at home while the researcher provided all the required materials. When participants received five stars consecutively on their progress chart, which indicated that they have trained for 5 days, they received a small gift. Participants assigned to the passive control group also received a small gift following pre- and post-testing. The passive control group was informed that they would have an opportunity to use the adaptive version of the online program upon completion of the research project. All participants were notified that they could withdraw from the study at any time if they did not wish to continue.

The Kruskal-Wallis Test was used to examine whether there was a significant difference between the number of days in between pre- and post-test for each group. This test was used because the assumption of normality regarding the number of days in between training was violated for all three training groups. Refer to Table 6.2 for relevant descriptive statistics. Results indicate that there were no significant differences in the number of days between pre- and post-testing for the intervention group (Mean Rank = 33.07), the active control group (Mean Rank = 28.39), and the passive control group (Mean Rank = 34.00), $\chi^2 = 1.09$, $df = 2$, $N = 63$, $p = .579$, $\eta^2 = .018$, $d = 0.27$.

Table 6.2

Descriptive Statistics for the Number of Days Between Pre- and Post-Tests

	Intervention group	Active Control group	Passive Control group
Lower Bound <i>M</i>	33.69	32.72	34.45
Upper Bound <i>M</i>	38.59	37.38	38.68
Median	38	34	36

Results

SPSS for Windows Version 17.0 was used for all data analysis. Results are divided into four sections. The first section will describe general data preparation. Next, the assessment of the specific training effects for the intervention group will be presented. Then the examination of pre-test differences, followed by investigations of transfer effects will be presented.

Data Preparation

The specific training effects for the intervention group were analysed using paired sample *t*- tests and, where normality assumptions were violated, the results were analysed concurrently using Wilcoxon Signed Rank Test to ensure accurate interpretations. Descriptive statistics for all cognitive measures at pre-test and post-test of all three groups are provided in Table 6.3, as well as their respective pre- and post-test comparisons. Subsequent correlational analysis was used to determine whether there were significant associations between inter-individual differences and training performance.

A series of one-way between groups ANOVAs was used to perform between group comparisons at pre-test amongst the three groups. The assumption of normality was tested using Shapiro-Wilk statistics ($p < .05$) and Q-Q Plots were inspected. The

homogeneity of variances assumption was also tested using a Levene's test ($p < .05$). Welch's F was used when this assumption was violated (Field, 2009b). Although ANOVA is robust to its assumption violations, a non-parametric Kruskal-Wallis one-way ANOVA was concurrently used to ensure the validity of the test results when violations were detected.

Subsequently, following the recommendation of Bonate (2000), a series of ANCOVAs, which have the advantages of accounting for regression to the mean and good statistical power compared to other methods when assumptions are met, were used to analyze post-test differences in reflection of transfer effects on each measure. Since ANCOVA has been suggested to be robust to moderate violations of the assumption of normality (Bonate, 2000, p. 96), unless data indicated severe violations to the normality assumption, current data were not transformed. Upon detection of significant effects, a planned contrast was performed comparing the intervention group with both active control group and passive control group respectively (Field, 2009a; Tabachnick & Fidell, 2001a).

Table 6.3

Descriptive Statistics for Psychometric Test Scores of Each Group

Measures	Pretest			Posttest		<i>p</i>	<i>d</i>
	<i>n</i>	M	SD	M	SD		
Digit Span ^a							
INT	21	13.14	2.37	13.52	2.56	.329	0.15
AC	19	12.11	1.49	12.32	1.64	.578	0.14
PC	23	11.74	1.82	12.48	1.81	.038*	0.41
Spatial Span ^a							
INT	21	11.00	2.98	12.00	3.10	.161	0.33
AC	19	10.74	3.05	10.95	2.70	.734	0.07
PC	23	10.13	2.49	11.35	2.50	.016*	0.49
Stroop ^b							
INT	21	32.36	12.10	29.71	13.39	.224	0.21
AC	19	38.60	12.30	25.18	8.68	.001**	1.28
PC	23	35.64	16.33	28.96	15.17	.089	0.42
Reaction Time ^c							
INT	21	1178.71	162.22	1146.00	112.61	.202	0.24
AC	19	1316.95	223.76	1222.53	198.84	.001*	0.76 ^d
PC	23	1205.65	180.88	1153.78	151.13	.027*	0.46 ^d
RSPM ^e							
INT	21	13.86	3.81	17.48	4.09	.000***	0.92
AC	19	13.42	3.61	14.89	4.07	.037*	0.38
PC	23	12.13	4.28	13.83	4.13	.031*	0.40

Note. Comparisons between pre- and post-test are presented using paired sample *t*-test. Cohen's *d* is provided for effect size.

^aTotal points scored.

^bTime difference between Incongruent condition and Neutral condition (secs).

^cAverage time (ms).

^dThe Wilcoxon Signed Test is used for Reaction Time task in both Active and Passive control group because of violations of normality assumption, with effect size provided in *r*.

^eTotal trials correct.

p* < .05, *p* < .01, ****p* < .001

Specific Training Effects: Intervention Group

To analyse whether performance on the training task changed after intervention, several measures of baseline performance and performance after training were recorded

and compared. Two measures of baseline performance were included and defined:

Baseline A – the highest set size achieved in the first session; Baseline B – the average set size achieved in the first session. As the highest set size was not always achieved in the final training session, and is a less reliable indicator of performance than average set size, both performance measures are shown in Table 6.4. Figure 6.5 and Figure 6.6 show a summary of the participants' training progress using their highest and average set size performance in each session, respectively.

Table 6.4

Descriptive Statistics for Training Performance in Set Sizes			
	First Session	Throughout Training (1 st to 20 th session)	Final Session (20 th session)
	Mean (SD)	Mean (SD)	Mean (SD)
Highest set size achieved	4.86 (1.01) ^a	7.67 (1.35)	6.10 (1.58)
Average set size achieved	2.81 (0.93) ^b	5.62 (1.43)	4.19 (1.25)

Note. ^a Baseline A: highest set size achieved on the first training day.

^b Baseline B: overall average set size achieved on the first training day.

Two separate analyses were performed comparing Baseline A and Baseline B, respectively, with participants' training progress. Firstly, Baseline A was compared with both the highest set size achieved on the final session and the highest set size achieved throughout the 20-days of training.

When comparing Baseline A with highest set size achieved in the participants' final training session, violations to normality assumption were detected, therefore a paired sample *t*-test was conducted concurrently with a Wilcoxon Signed Ranks Test. Both tests provided similar significant findings with participants performing better on their final training session than on Baseline A, $t(20) = -3.67, p = .002, d = 0.96$. The Wilcoxon Signed Rank Test showed that 16 students improved on the training task,

three students did not perform as well as they did at baseline, and two remained the same, $T = 25$, $z = -2.86$ (corrected for ties), $p = .004$, $r = -.67$.

Similar procedures were carried out using Baseline A and highest set size achieved throughout the 20 sessions. Both paired sample t -test and the Wilcoxon Signed Rank Test's results showed a significant result. The paired sample t -test indicated participants who completed the adaptive animal span training improved significantly on task performance, $t(20) = -10.65$, $p < .001$, $d = 2.4$, when Baseline A span was compared with the highest set size achieved throughout training. Similarly, the Wilcoxon Signed Rank Test showed that students had an overall significant improvement on the training task, $T = 0$, $z = -4.00$ (corrected for ties), $p < .001$, $r = -.92$. Twenty students showed improvement after training and one showed no difference in performance after training, no student's performance became worse.

Secondly, Baseline B was compared with the average set size achieved on the final session and the average set size obtained throughout the 20-days of training, both of which demonstrated results consistent with comparisons using Baseline A. A comparison between Baseline B and participants' average set size achieved on their final training session using a paired sample t -test indicated significant improvement between students' first and last training session, $t(20) = 6.50$, $p < .001$, $d = 1.27$. Again, the Wilcoxon Signed Ranks Tests were used due to violations of the normality assumption. Results indicated the same significant findings, $T = 5$, $z = -3.70$ (corrected for ties), $p < .001$, $r = -.85$. Eighteen students showed improvement, one did not perform as well during the final training session as their first training session, while two students showed no differences on their average level performance.

Students' average performance in their first session as indicated by Baseline B was also compared with their average set size reached throughout training. Paired

sample t -test results indicated a significant increase with training, $t(20) = -11.46$, $p < .001$, $d = 2.38$. However, the assumptions of normality for both variables were violated. Again a Wilcoxon Signed Ranks Test was used to confirm the analysis. Similar results were found, $T = 0$, $z = -4.05$ (corrected for ties), $p < .001$, $r = -.88$. In this comparison, all 21 students showed an increase in their average level achievement with training.

To provide a percentage of improvement, the following formula was used, as documented in Chooi and Thompson (2012):

$$\text{Improvement \%} = \frac{\text{Average highest training score} - \text{Average first training score}}{\text{Average highest training score}} \times 100$$

Participants in the intervention group yielded a mean improvement of 48.93%.

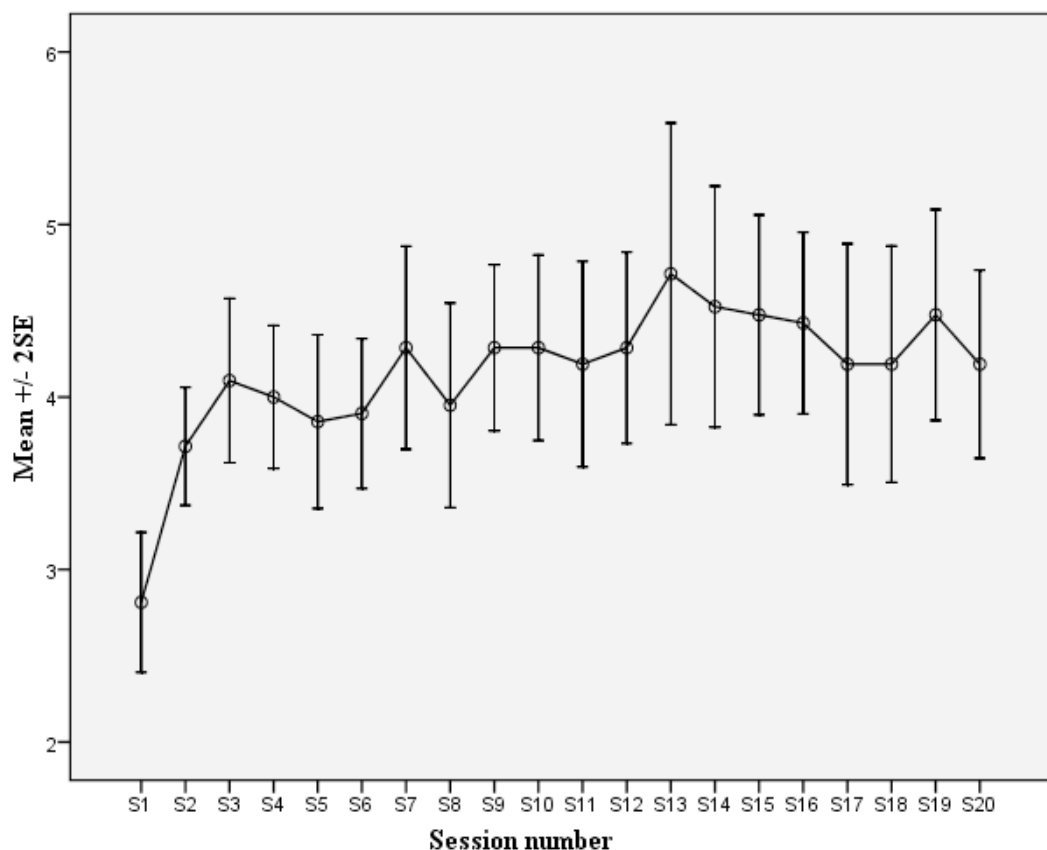


Figure 6.5. The mean training set sizes achieved and their standard errors for each training session across the 20 days of training are presented for the intervention group.

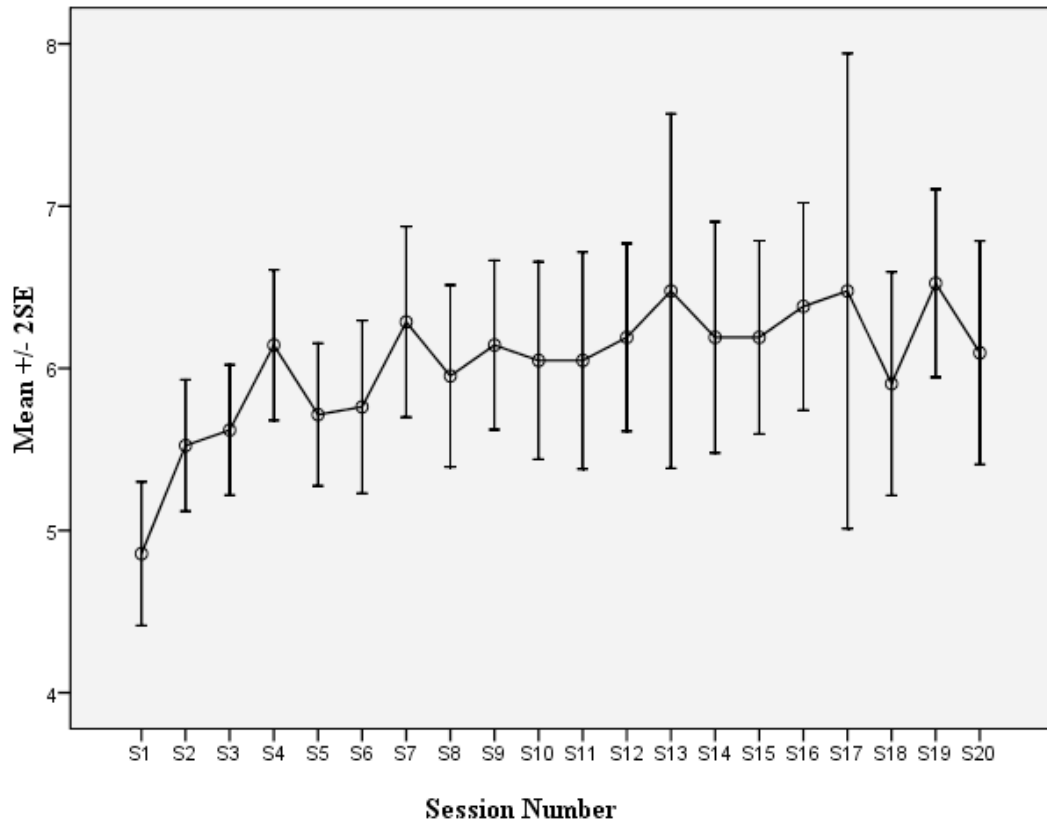


Figure 6.6. The highest training set sizes achieved and their standard errors for each training session across the 20 days of training are presented for the intervention group.

Individual Differences and Training Performance

To further investigate whether training performance was related to individual differences within the intervention group, correlational analysis were performed using cognitive performance at pre-test, differences in scores between pre and post assessment, as well as two measures of training gain. Training gain was calculated using (a) the difference between Baseline A and highest set size achieved throughout training, (b) the difference between Baseline B and the highest session average achieved throughout training. Pearson's correlational analysis indicated no significant associations between cognitive performance as measured at pre-test and the two

measures of training gain. Similarly, no significant correlations were found between changes in cognitive performance and the two types of training gains. The normality assumptions for data on digit span pre-test and training gain (a) were violated, however, an inspection of their Spearman's rho correlation with the respective variables did not change the results. Refer to Table 6.5 for a summary of the correlation analysis.

Between Group Comparisons at Pre-Test

Groups' performance was analysed to determine whether significant pre-test differences were present. The Shapiro-Wilk statistics showed that the assumption of normality was violated for some of the tasks, however further examination of Q-Q plots did not suggest severe deviations from normal distribution. Kruskal-Wallis one-way ANOVA was used to confirm findings. The assumption of homogeneity of variances was met for all the tasks.

Digit span. The assumption of normality was violated for the intervention group ($p = .019$), and the active control group ($p = .048$), but no extreme scores were detected within each group. Results of ANOVA indicated no significant differences at pre-test between groups, $F(2, 60) = 3.05$, $p = .06$, $\eta^2 = .09$, $d = 0.64$. A non-parametric Kruskal-Wallis one-way ANOVA confirmed that there were no statistically significant differences between the pre-test performance in the intervention group (*Mean Rank* = 37.88), the active control group (*Mean Rank* = 30.32), and the passive control group (*Mean Rank* = 28.02), $\chi^2 = 3.51$, $df = 2$, $N = 63$, $p = .173$, $\eta^2 = .06$, $d = 0.49$.

Spatial span. The assumption of normality was violated for both the active control group ($p = .036$) and passive control group ($p = .024$). Again, no extreme scores were observed within each group. Results of ANOVA showed no significant differences at pre-test between groups, $F(2, 60) = .547$, $p = .582$, $\eta^2 = .02$, $d = 0.27$. A Kruskal-Wallis one-way ANOVA indicated no statistically significant differences in the pre-test

performances on spatial span task between the intervention group (*Mean Rank* = 34.93), the active control group (*Mean Rank* = 32.18), and the passive control group (*Mean Rank* = 29.17), $\chi^2 = 1.11$, $df = 2$, $N = 63$, $p = .575$, $\eta^2 = .02$, $d = 0.27$.

Stroop. All assumptions were met for Stroop scores. Results of the ANOVA indicated that there were no statistically significant differences among the three groups' performance on the Stroop task at pre-test, $F(2, 60) = 1.01$, $p = .369$, $\eta^2 = .03$, $d = 0.37$.

Reaction time task. The assumption of normality was violated for the active control group ($p = .012$) and passive control group ($p = .003$). One extreme score was observed in the passive control group in terms of average time required for completing the task. However, this score remained within the range of scores observed in the active control group, thus it is interpreted as a true reflection of the students' ability at the time of testing. Results of ANOVA suggest no significant differences at pre-test between groups, $F(2, 60) = 2.97$, $p = .059$, $\eta^2 = .09$, $d = 0.63$. A Kruskal-Wallis one-way ANOVA indicated no statistically significant differences among the pre-test reaction time performances of the intervention group (*Mean Rank* = 28.02), the active control group (*Mean Rank* = 39.55), and the passive control group (*Mean Rank* = 29.39), $\chi^2 = 4.68$, $df = 2$, $N = 63$, $p = .096$, $\eta^2 = .075$, $d = 0.57$.

RSPM. All assumptions were met for RSPM scores. Results of the ANOVA indicated that there was no statistically significant difference among the three groups' performance on the Ravens task at pre-test, $F(2, 60) = 1.15$, $p = .323$, $\eta^2 = .04$, $d = 0.392$.

Transfer Effects: Between Group Comparisons at Post-Test

A series of one-way ANCOVAs were used to test for differences at post-test on

each measure, with group (intervention, active control, or passive control) as the independent variable fixed factor, post-test scores as the dependent variable and pre-test scores as a covariate.

Digit span. The Shapiro-Wilk test showed minor violations of the assumption of normality for the Passive Control group ($p = .05$). Linearity was evident in scatterplots. Both assumptions of homogeneity of regression slope, $F(2, 57) = .71, p = .499$, and homogeneity of variances, $F(2, 60) = .28, p = .758$, were met. Results of the final ANCOVA indicated that there were no statistically significant differences between the three groups' performance on digit span total at post-test, $F(2, 59) = .56, p = .574, \eta p^2 = .02$.

Spatial span. The Shapiro-Wilk test showed that the assumptions of normality were supported for all three groups. Linearity was evident in scatterplots. Assumption of homogeneity of regression slope, $F(2, 57) = .10, p = .908$, and homogeneity of variances, $F(2, 60) = 2.58, p = .084$, were met. Again, results indicated no statistically significant difference between the three groups' performance on spatial total at post-test, $F(2, 59) = .83, p = .442, \eta p^2 = .03$.

Stroop task. All groups satisfied the assumption of normality and linearity. Both assumptions of homogeneity of regression slope, $F(2, 57) = .302, p = .057$, and homogeneity of variances, $F(2, 60) = .1351, p = .267$, were met. Final ANCOVA results showed no statistically significant difference between the three groups' performance on the Stroop task at post-test, $F(2, 59) = .1708, p = .190, \eta p^2 = .055$.

Reaction time task. The assumption of normality was violated for both control groups at pre and post-test but not for the intervention group. Due to severe violations, data were transformed. Log transformation best achieved normal distributions for all groups at pre- and post-test. ANCOVA was carried out due its robustness and the

evidence of covariate scores achieving a normal distribution. Assumptions of linearity, along with assumption of homogeneity of regression slope, $F(2, 57) = 2.08, p = .135$, and homogeneity of variances, $F(2, 60) = .23, p = .792$, were met. The three groups did not differ significantly at post-test performance on the reaction time task, $F(2, 59) = .24, p = .785, \eta p^2 = .01$.

RSPM. The Shapiro-Wilk test showed that the data met assumptions of normality for all three groups. Scatterplots indicated linearity. Assumptions of homogeneity of regression slope, $F(2, 57) = .44, p = .645$ and homogeneity of variances, $F(2, 60) = 1.31, p = .276$, were met. There was evidence of a statistically significant difference between the three groups' performance on the Ravens task at post-test, $F(2, 59) = 4.29, p = .018, \eta p^2 = .13$. Planned contrasts revealed that the intervention group performed significantly better than both the active control group ($p = .019$), and the passive control group ($p = .011$), while the active control group and the passive control group were not significantly different ($p = .891$). Adjusted group means were 16.92 ($SD = .649$) for the intervention group, 14.657 ($SD = .678$) for the active control group, and 14.530 ($SD = .623$) for the passive control group. Refer to Figure 6.7 for a visual representation of the significant transfer effect on RSPM.

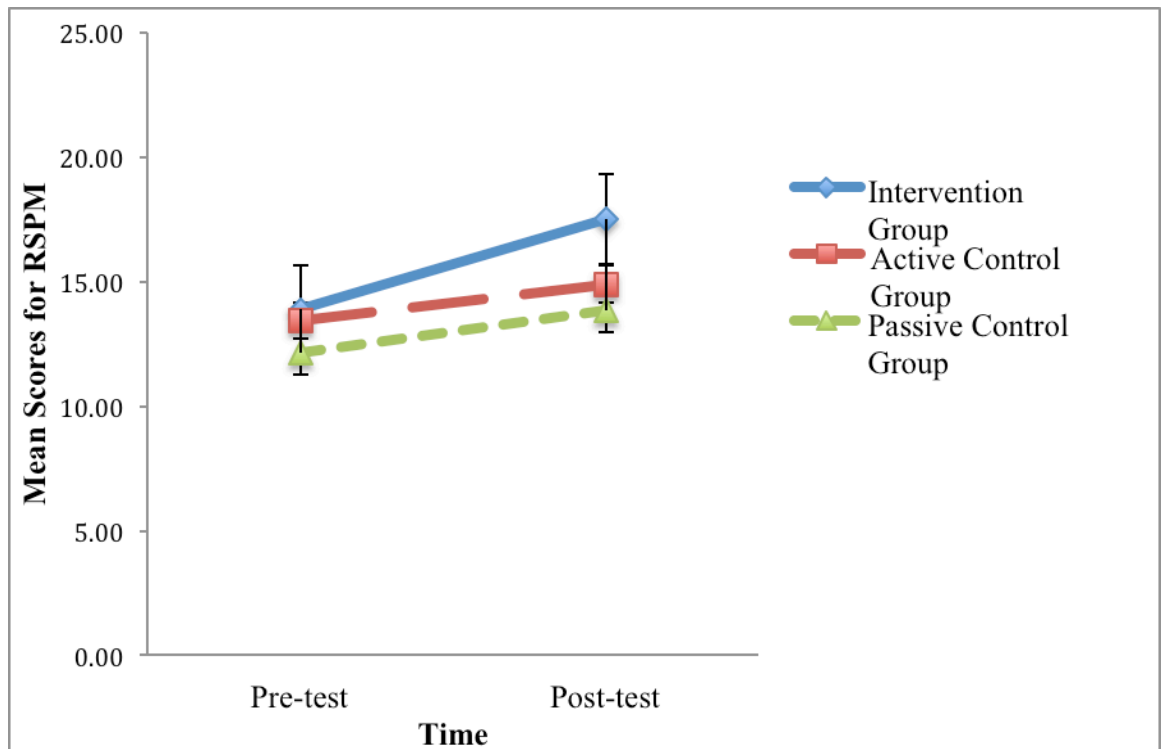


Figure 6.7. Transfer effects on mean scores and their standard errors for the Raven's Standard Progressive Matrices assessment at Pre-test and Post-test for each group.

Table 6.5

Summary of Correlations on Cognitive Measures at Pre-test, Differences in Scores at Post-test, and Training Gains of the Intervention Group

	1	2	3	4	5	6	7	8	9	10	11	12
1. DS pre-test	-	-.26	.20	.00	-.02	-.37	-.07	.19	.33	.06	.27	.07
2. DS difference		-	-.02	-.13	.00	-.25	.10	-.04	-.14	.13	-.32	-.29
3. SS pre-test			-	-.49*	-.42	-.10	-.26	-.01	.66**	.25	.06	.09
4. SS difference				-	.33	-.19	.10	.16	-.35	.27	.09	-.04
5. Stroop pre-test					-	-.26	.40	-.03	-.54*	.20	-.20	-.10
6. Stroop difference						-	.16	-.18	-.04	-.42	.22	.18
7. CRT pre-test							-	-.72**	-.09	-.15	.01	.11
8. CRT difference								-	-.19	.20	-.04	-.31
9. RSPM pre-test									-	-.27	.19	.24
10. RSPM difference										-	-.01	-.02
11. Training gain (a)											-	.78**
12. Training gain (b)												-

Note. DS = Digit Span task; SS = Spatial Span task; CRT = Choice Reaction Time task; RSPM = Raven's Standard Progressive Matrices; Training gain (a) = the difference between Baseline A and the highest set size achieved throughout training; Training gain (b) = the difference between Baseline B and the highest average set size achieved throughout training. ** $p < .01$, * $p < .05$

Discussion

The goal of this study was to investigate whether computerized WM training, specifically in the form of CWM span task, could demonstrate near and far transfer effects. It was hypothesized that participants in the intervention group would demonstrate significant improvements in the training task and that both near transfer and far transfer would be detected with the exception of performance on the speed of processing task. Results indicated significant improvements in the training task, which supported the initial hypothesis. However, results concerning transfer of training to untrained tasks only partially supported the hypothesis. No significant near transfer to measures of non-trained WM, or far transfer to measures of executive attention, was detected. However, significant far transfer to G_f was detected. The implications of these findings are discussed.

Implications of Training Performance for the Intervention Group

Analyses of intervention participants' change in performance on the training task indicated statistically significant improvements throughout the training period. Their performance curves, using both highest and average training set size for each session across the 20 training days, revealed an improvement in performance from the first training session up to approximately session 13 with performance levelling off thereafter (refer to Figure 6.5 and 6.6). This training trend was contrary to Studer et al. (2009) and Jaeggi et al.'s (2008) *n*-back training studies where continuous improvement was reported on the trained task throughout training. One possible explanation is that children in the current experiment had reached a point where they were no longer engaged and motivated to achieve further. Perhaps a more refined and personalised reinforcement schedule would keep the children engaged (Green & Bavelier, 2008). An alternative explanation could be that the children in the study, in contrast to adults in the

aforementioned studies, had simply reached their maximum capacity. Children of school age may have limited scope for devising strategies to keep improving when faced with relatively complex and challenging tasks, resulting in children plateauing after 13 sessions.

Nonetheless, the participants in this experiment improved in the training task as much as those documented in other WM training studies. The present calculation of participants' percentage of improvement in the training group was 48.93%. No direct comparison could be made because the only CWM training study providing an improvement percentage was one that trained their participants for 10 days and reported a 23% increase in specific trained performance (Loosli, et al., 2012). However, the current improvement percentage was slightly higher than the documented improvements in Chooi and Thompson's (2012) 20-day dual n -back study (44%), although they found no transfer effects. The current participants' improvements were very similar to Jaeggi et al.'s (2008) 19-days dual n -back training (47%) that also documented far transfer effects in G_f . Although direct comparisons cannot be made due to the difference in training regime, these similar percentages of improvements could be a good indication of comparable commitment and effort demonstrated amongst the training participants.

Transfer to Working Memory

Near transfer to untrained WM measures after adaptive WM training would logically have been predicted, given that this CWM training involved elements shared with the untrained WM measures. However, this prediction was not upheld. The current results did not indicate significant transfer to untrained WM tasks, namely digit span and spatial span. Evidence for near transfer to these measures after training in CWM has been scarce. Chein and Morrison's (2010) CWM training experiment found significant between group differences and a significant group x session interaction demonstrating

near transfer, however their claims were based on performance improvements in short term memory measures that highly resembled the CWM training task, (i.e., both used letters and locations as stimuli). Moreover, their results were only compared with a passive control group, which as mentioned in the previous chapter, lacked the control of expectancy effect. Therefore, their significant findings could merely be a result of practice effect that entailed training in specific strategies.

Nonetheless, the present results showing no near transfer to measures of WM can be compared with other computerized WM interventions. The findings were consistent with Redick et al. (2012) as well as Chooi and Thompson (2012), where both studies' participants were trained using dual *n*-back tasks. On closer observation, the effect size for near transfer effect at post-test on the digit span task ($p = .56$, $\eta p^2 = .02$) and the spatial span task ($p = .83$, $\eta p^2 = .03$) in current study were similar to those detected in Redick et al.'s (2012) study as measured by symmetry span ($p = .59$, $\eta p^2 = .02$) and running letter span ($p = .82$, $\eta p^2 = .011$).

One possible explanation for unsuccessful near transfer from the present study's adaptive animal span training could be that simple span tasks are more reliant to short-term memory storage (Heitz, et al., 2005). Therefore the current adaptive span training, which was a complex span task, overlapped less with the skills involved in simple span assessments. Although evidence suggests that the use of combined scores for digit span and spatial span tasks are adequate in representing WM (Bowden, et al., 2013), perhaps separating forward and backward span scores may increase the sensitivity to this complex span training and provide scores towards respective WM components, namely short-term memory capacity and executive attentional control.

Transfer to Other Cognitive Tests

In regards to far transfer effects to other cognitive measures, no significant

group differences were detected at post-test on the Stroop. The results partially corroborated Chein and Morrison's (2010) study that reported no training group x session interaction effect of CWM training on performance of Stroop initially, but subsequently claimed a marginally significant interaction effect on the Stroop task performed on comparisons between the subgroup who demonstrated training success and their control group. Although the present study and theirs both used a CWM training regimen, a discrepancy in intervention lies in the duration of training time. Chein and Morrison (2010) trained their participants for 30-45 minutes each session whereas the participants in the current study was only trained for 15 minutes. The current results were also consistent with Thorell et al. (2009) and Lundqvist et al.'s (2010) studies who used Cogmed training and found no significant transfer to the Stroop task. However, the empirical evidence for successful transfer to Stroop was presented in the initial Cogmed studies reported by Klingberg (2005; 2002).

The current literature on whether WM training is transferable to attentional control tasks appears mixed and the need for replication to confirm the absence or presence of transfer is necessary. Thus far, most of the existing evidence has relied on identifying transfer to the Stroop task (Chein & Morrison, 2010; Klingberg, et al., 2005; Klingberg, et al., 2002; Thorell, et al., 2009; Westerberg, et al., 2007). The use of other attentional control tasks would provide a more comprehensive picture. Given that evidence of far transfer to attentional control is absent in the current study but present in Chein and Morrison's (2010) study, perhaps considerably longer training duration may be necessary for transfer, as measured by Stroop task, to occur.

In contrast to attentional control, speed of processing has not been well studied within computerized WM training. The initial reason to include a processing speed task was to allow the testing and potential falsification of models of WM training that claim

that benefits to G_f occur through WM improvement. If the present results showed transfer to a broad range of tasks, including measures of other suggested causes of individual differences in G_f , then it would cast doubt on the specific WM interpretation. The current findings showed no significant transfer effect at post-test on the choice reaction time task. As hypothesized, WM training had no significant impact on measures of speed of processing. Although no direct comparisons to studies using CWM training could be made, WM training studies using other training tasks also documented non-significant results (Chooi & Thompson, 2012; Klingberg, et al., 2002). Chooi and Thompson (2012) even recorded worsening of perceptual speed performance after their dual n -back training in both their 8-day ($d = -0.62$) and 20-day training groups ($d = -1.56$). As suggested by Fry and Hale (1996), the relationship between speed of processing, WM and G_f is a directional one where “differences in speed have a direct effect on working memory capacity, and these individual differences in memory are a direct determination of fluid intelligence” (p. 241). From this perspective, changes in WM to cause upstream changes in speed would not necessarily be expected.

Transfer to G_f

This is the first study reporting a significant far transfer effect to the untrained RSPM task following adaptive CWM training. The type of training task, as well as the choice of training schedule, for school-aged children appears to be factors influential to this success. The current experiment used a classic CWM adaptive span training: it contained ‘TBR’ items (i.e., animals) interspersed with a type of cognitive processing activity (i.e., judging whether the animals were upside down or right side up (Engle, 2002; Kane, et al., 2001; Unsworth, et al., 2009). According to Kane et al.’s (2001) executive attention model, attentional control underlies performance on G_f tasks, and complex memory span are good measures of this. Therefore, training effects on CWM

tasks should show far transfer, and it did. Although there is a possibility that the score increment found in RSPM may be interpreted as the intervention group becoming better at performing the task, the intervention group did performed significantly better than the two control groups and thereby addressing practice effects. In addition, the features of the RSPM task were also not in any way similar to the CWM training, thereby suggesting genuine transfer that did not rely of strategy-specific learning.

Why, then, is far transfer to G_f present in the current study and not in other studies and, particularly, other CWM training experiments with university students (Chein and Morrison, 2010) and healthy school-aged children (Loosli, et al., 2012)? The training schedule appears to be an important factor to G_f transfer on top of the choice of training task. The present study's results were consistent with the earlier dual n -back (Jaeggi, et al., 2008; Studer, et al., 2009) studies with healthy undergraduate volunteers and the single n -back study (Jaeggi, et al., 2011) with typically developing children, both of which described using the same duration of approximately 20-days training and 15 minutes of training time per day to achieve far transfer effects to G_f tasks. Evidence of far transfer effects were also found in Klingberg et al.'s (2002) study of children with ADHD, although they described using a slightly longer duration and dosage of 25 minutes of training each day that lasted for 24 to 26 days. However, most of the studies that did not detect far transfer to G_f adopted a relatively long daily training time of at least 30 to 40 minutes (Chein & Morrison, 2010; Chooi & Thompson, 2012; Redick, et al., 2012; Westerberg, et al., 2007), and one especially brief intervention of only 12 minutes per training day (Loosli, et al., 2012). These include training using Cogmed (Nutley, et al., 2011; Westerberg, et al., 2007), dual n -back (Chooi & Thompson, 2012; Redick, et al., 2012), and CWM tasks (Chein & Morrison, 2010; Loosli, et al., 2012).

Observations on adaptive training schedules and the occurrence of far transfer to

G_f in the current study and past interventions suggest that the optimal timing lies between approximately 15-25 minutes each day for at least 20 days across four to six weeks. This appears to be a safe and promising schedule to adopt in WM training with healthy school-aged participants.

The results of the present study are also more readily interpretable as genuine far transfer to G_f as it included several important methodological strengths. Unlike studies that rely only on a passive control group where documentations of significant transfer may only be a result of expectancy effects despite the presence of significant G_f transfer (Jaeggi, et al., 2008; Studer, et al., 2009), the present study included both active and passive control groups, which ruled out both Hawthorne and practice effects, respectively. The two other CWM studies (Chein & Morrison, 2010; Loosli, et al., 2012), despite documenting an absence of far transfer to G_f , also lacked the inclusion of active control groups.

Although, far transfer to the G_f task was detected, which was arguably occurred through core information processing changes, it appears puzzling that no far transfer effects were found to the Stroop task that most closely represented attention control ability. Although no precise conclusion can be drawn that explains this pattern of results, a comparison of the current adaptive animal CWM span task to the two respective tasks - the Stroop task and RSPM task - may shed light on this phenomenon. Both the CWM task and RSPM used accuracy across a broad time limit as their dependent variable, whereas the Stroop task used time of completion. In order to perform well in the CWM task and RSPM task, participants had to spread their attention across multiple stimuli with little time constraint. However, in order to perform well in the Stroop task, participants were required to narrow their focus of attention, ignore irrelevant stimuli and perform the task as quickly as possible. Moreover, the Stroop task

may be more sensitive to reading abilities, while neither the CWM nor RSPM tasks were reading-dependent. The above comparisons implied that perhaps there is more than one type of attentional control ability. Stroop and RSPM tasks tapped different abilities, with the current CWM training task more relevant to the RSPM task than the Stroop task. However, this explanation does not appear to fit well with the marginally significant interaction effect found on the Stroop task followed by a longer duration of CWM training as described in Chein and Morrison (2010), as compared to the current shorter duration of CWM training. Therefore, further investigation is warranted as to why there is a presence of far transfer to G_f without the presence of transfer to attention control. Whether the explanation stems from a matter of task sensitivity, the difference in training duration, or something entirely different is unclear.

Parallel Versions of Assessments

The intention with pre-post measurements was to identify changes that occur after training and make comparisons across the three groups. Requesting participants to take part in the same assessment twice within a short period of time may create practice/retest effects. Although this study's inclusion of both active and passive control groups already assisted in addressing practice effects, the addition of parallel forms used in the study further minimizes retest effects. Some other studies have used parallel forms in their studies (Chein & Morrison, 2010; Jaeggi, et al., 2008; Jaeggi, et al., 2011; Loosli, et al., 2012), often by separating odd and even numbers items; while some did not use parallel forms (Klingberg, et al., 2005; Klingberg, et al., 2002; Nutley, et al., 2011).

The type of parallel forms used may also contribute to the possible reasons behind the lack of transfer found in untrained WM tasks and the attentional control Stroop task. In particular, it may stem from the use of post-test that lacked novelty compared to pre-test. Specifically, the WM measures used were semi-parallel forms

where the order of items was changed, but not the stimuli themselves. Similar to WM measures, the trials within the choice reaction time task were randomised at pre- and post-test. However, the Stroop task was identical at pre- and post-test. Only RSPM were split into two parallel forms, with different items matched for difficulty, and represent truly novel items for participants at the two testing times. Given that the most valid measures of EF should assess how well one self-regulates to solve new problems (Burgess, 1997; Jurado & Rosselli, 2007; Rabbitt, 1997), the pre- and post-test used for WM and attentional control may have lacked adequate validity as measures of EF at post-test, due to their lack of novelty. Thus, it could be that if CWM training enhances EF, this is only detectable when a novel, yet parallel version of assessment is used at pre- and post-test. Therefore, RSPM may have a higher validity than other measures in the study to adequately reflect on true changes before and after training. If the tasks do lose novelty, participants, in general, are also expected to perform better on the tasks. Overall practice effects from the active and passive control group appear to support this line of argument. At post-test, participants in both control groups performed significantly better in the reaction time task and the RSPM task, while the passive control group also performed significantly better at post-test in the digit span task, spatial span task. They also performed somewhat better on the Stroop task but results were not significant. Practice effects on RSPM are usually greater for those with low scores on first testing (J. C. Raven, 1958). Given that the intervention group performed slightly better than the other groups at pre-test, some practice effects on the RSPM task from the passive control group can be expected (Refer to Table 6.3).

Limitations

There are several limitations to the present study. First, the use of single measures to represent specific domains may have an impact on the validity of the results.

No measure is 'process pure' and able to capture a specific cognitive domain without systematic and unsystematic measurement error. Changes in single test scores may be driven by ability or interest or other random factors, and thus cannot provide definitive evidence to support that underlying change stems solely from CWM training (Shipstead, Hicks, & Engle, 2012). In order to demonstrate that WM training instigates real change in any one cognitive domain, tasks that represent the same domain should all demonstrate significant improvement, not just one type of domain-specific task. For example, if CWM training is successful in improving G_f , then tasks that measure G_f such as both RSPM and Cattell Culture Fair Intelligence Tests should also show improvements. The use of more than one measure for each cognitive domain would assist in ruling out task-specific skills or strategies inherent in a single measurement outcome. That said, the tests in the current study have been widely used in the literature and have sound psychometric properties in their relevant domain literature. For example, RSPM has been considered the best measure of G_f (Carpenter, et al., 1990) and used in the majority of WM intervention studies (Jaeggi, et al., 2008; Klingberg, et al., 2002; Redick, et al., 2012).

A second limitation concerns the administration of reinforcement. Although materials were provided for use as reinforcement for the continuation of participants' daily participation, this factor was not captured in any quantitative or qualitative terms. Parents of participants in the intervention group and the active control group were required to provide their children with stickers and weekly prizes when corresponding training was completed, but it was not clear whether parents complied with these instructions, or whether the rewards had the desired effect on motivation. The lack of a desired effect on motivation could explain the high number of drop-outs/ incompletion training within the intervention group. Perhaps a formal questionnaire requesting

participants to rate their interest in or boredom with the task and how difficult they found it, completed after each training session, could provide more informative and quantitative data on motivation and engagement levels. Real time online data regarding time breaks that the participants may have taken in between trials during each training session could also give a good indication of motivation behaviour. In the absence of such information, we cannot be certain whether participants' performance levelling off after the thirteenth session can be explained by reaching their maximum capability or by losing motivation.

Motivational factors that impact on the performance or compliance level of the active control group is also a concern, since the repetition of their task through a presentation of stimuli of a fixed difficulty in training may easily result in boredom and learned helplessness as described in recent critiques for intensive computerized WM training (Morrison & Chein, 2012; Shipstead, et al., 2012). An indication of this in the present study was that the drop-out rate for the active control group was considerably higher than the two other groups. Other cognitive studies have not documented specific drop-out rates, thus specific comparisons were unavailable. A possible way to amend the impact of boredom could be to widen the range of animals used as stimulus in the task for both active control group and intervention group. This would allow for more variety in stimulus context to counterbalance some of the boredom associated with the task as well as decreasing stimulus familiarity. According to Dempster (1981), the more familiar the stimulus presented in STM tasks, the easier the task becomes due to the ease in chunking. Nonetheless, the use of low dosage non-adaptive control groups has occurred in several other cognitive intervention studies (J. Holmes, et al., 2009; Klingberg, et al., 2005; Klingberg, et al., 2002) and the choice of non-adaptive training task is in line with recently suggested methodologies (Redick, et al., 2012; Shipstead, et

al., 2012). Despite concerns of demotivation, the descriptive statistics and corresponding effect sizes (Table 6.3) for pre and post-test scores of both active and passive control groups did not appear to demonstrate worsening performance. On the contrary, several significant differences at post-test appeared in the active control group, including the Stroop task, reaction time task as well as the RSPM in the current study.

There are, however, several alternative choices of training task for the active control group that may also be suitable. First, the use of an alternative adaptive training that does not tax on WM abilities, such as reading comprehension. This would provide participants in the active control group with similar arousal to learning and the Hawthorne effect can be controlled with a better match on expectation between the groups (Shipstead, et al., 2012). Another type of active control is to present trials in a range of difficulties at random rather than systematically matched to participants' current performance. This may perhaps provide better control in terms of investigating whether adaptiveness is the key to successful learning.

In addition to the above, all parts of the research were conducted by one researcher, including recruitment, allocation of participants, explanation of training procedures to the participants and family, as well as pre- and post-assessments. The researcher, thus, was not blind to the conditions to which the participants had been allocated. There could be a possibility of researcher bias. However, as the current analysis showed no initial pre-test differences amongst the three groups and the researcher closely followed verbatim written instructions for all testing procedures, the current results are not readily attributable to bias alone.

A final limitation was that this study did not include follow-up analysis to assess further changes, if any, in following weeks or months in comparison to the two control

groups. This was due to the limitations of resources and funding. Very few current studies on WM training have included follow-up studies. Those that did also documented mixed findings. Cogmed training showed that near transfer was more sustainable than far transfer with large effect sizes being better maintained than small effect sizes. For example, Klingberg et al. (2005) documented transfer to digit span, spatial span and Stroop tasks that was maintained after three months while transfer to RCPM was not maintained. Jaeggi et al. (2011) reported the only dual n -back study showing significant improvements in G_f maintained at 3-months follow-up. Amongst the scarce follow-up analyses, results tend to reveal a reduction in effects over time.

Clinical Implications

The key clinical implication is that the findings of successful far transfer to G_f provide preliminary support for the use of CWM training in other clinical populations that demonstrate impairments in G_f , particularly children born preterm/LBW. From the results of Study 1, evidence shows that children born preterm performed at lower G_f than their same age peers and that the effect was partially mediated by low WM and cognitive inflexibility.

There are also obstacles to successful training that subsequently leads to transfer of G_f . These may include participating in WM training tasks that are too easy, having difficulty self-regulating to keep training in a repetitive task, as well as the availability of feedback and rewards. The current adaptive animal complex WM span task is able to address the aforementioned obstacles. This training is suitable for 7-year-old typical children with success in increasing G_f , therefore it might also be expected to be suitable for 8- to 9-year-old children born preterm, given evidence from Study 1 of at least one year's developmental delay in the preterm cohort's performance as compared with their peers. In addition, this form of training does not appear to be overly demanding on

either parents or children as it can be completed at home on their personal computer that is connected to the internet, without special laboratory equipment or continuous guidance from expert personnel. Parents are also provided with reinforcement materials to assist in maintaining children's motivation, while the program provides trial-to-trial feedback to enhance arousal for continuous training. The training is not overly demanding for the participants but its adaptive feature allows each participant to train at a challenging yet individualized level. This adaptive feature has been argued to be an essential element to successful learning and maintaining a close proximity between the task difficulty level and the individual's capability should maximize performance gain (Vygotsky, 1978/1997; Yerkes & Dodson, 1908). Concrete evidence has been provided from the results that adaptive training is useful relative to the control groups. The comparatively larger effect size on improvement scores in G_f in the intervention group ($p < .001$, $d = 0.92$) than the active control ($p = .037$, $d = 0.38$) and passive control group ($p = .031$, $d = 0.40$), as well as successful learning in specific trained performance as reflected in a nearly 50% improvement after training, appear to support the utility of the adaptive feature.

Future Directions

These promising findings are just one step on the way to the development and validation of early intervention neuro-rehabilitation programs for children born preterm and other clinical groups. Future studies have many methodological and pragmatic issues to address.

As a priority, future studies should consider the practical application of gains from WM training. What remains unknown is whether transfer is detectable in behaviours of daily living, such as educational achievement (Sternberg, 2008).

Achievements tests and teacher ratings can be included in future research to examine

transfer effects outside psychometric tests. If training gains remain evident only on psychometric tests, then it is of little practical significance to real life.

Another obvious future question is whether the current findings replicate in children born preterm/LBW. Now that the current study provides evidence for genuine increase in G_f following CWM training in school-aged typically developing children, we can be more confident that children born preterm/LBW will not be exposed to unnecessary risks. Moreover, if school-aged children born preterm/LBW do perform better at G_f tasks after CWM training, then achievement tests and teacher ratings can also be included in future research to examine transfer effects beyond psychometric tests. Researchers should also investigate impacts for different age groups as cognitive impairments have been suggested to persist into adolescent and adulthood.

Furthermore, a comparison of the different training regimes and the variations of training programs would be useful. A study that includes all three types of training regimes, namely the Cogmed training, the dual n -back training and the complex span training, with appropriate control groups and a standardized training schedule would shed light on which training regime is most useful or whether they have different effects. Currently, there is no documentation on the ideal training schedule such as the duration of training time per day or the duration of overall training. With the extensive variations in training schedules in existing studies, direct comparisons of training utilities are difficult.

Dose-relationships are also important to explore given the motivational challenges of sustaining repetitive tasks, especially for children. The drop out rate in this study is likely to be evidence of how challenging this program was. Fine tuning minimum requirements for change is critical in making the program accessible to children.

Future studies should also be mindful not only to include more than one measure for each cognitive domain in question as to rectify concerns of task impurity (Burgess, 1997; Rabbitt, 1997), but also to include measures of SES (Aarnoudse-Moens, Smidts, et al., 2009; Ardila, et al., 2005) and a motivational questionnaire (Lohaugen, et al., 2011; Nutley, et al., 2011) as discussed earlier. These would help to more accurately account for the strength of training gains and transfer.

Summary

In conclusion, the results of Study 2 demonstrated that intensive training using a focused adaptive CWM span training task could lead to transfer to an untrained G_f task. Unlike other studies, the presence of far transfer effect to G_f was not attributable to Hawthorne effects. Given that the study used a classic CWM training regime with the inclusion of both active and passive control group, as well as reliable parallel versions of assessment reflecting the features of the RSPM task that were not in any way similar to the CWM training, the significant differences in G_f gains were unlikely to be a consequence of practice or general familiarity effects but, rather, a consequence of the WM training. This success provides preliminary support on the utility of adaptive CWM training to increasing G_f and, therefore, may be beneficial to other clinical groups, such as children born preterm/LBW, who also demonstrated impairments in G_f .

The use of brief WM training to increase intelligence or specific cognitive abilities is still relatively new. Despite the growing number of published studies, the literature still lacks concrete evidence on the underlying mechanisms of successful transfer effects. Therefore, refinement and replications of the investigations of WM training utility across different laboratories and age groups, such as the current study, are important. More replications of successful far transfer effects are required to confirm the utility of CWM training and these replications should be completed with

Prematurity, Cognitive Abilities & Intervention

careful thinking on the use of methodology.

CHAPTER 7

Summary and General Discussion

This chapter consists of three main sections. The first section provides a summary of the thesis, including the main goals of the proposed research. The second section provides summaries of the two studies with their major findings and their integrated implications for children born preterm/LBW. Finally, it ends with recommendations for future research and a conclusion of the thesis.

Summary of Current Research

The main purpose of this thesis was to extend the current understanding of the underlying differences in cognitive and learning outcomes for children born preterm/LBW, with the anticipation that this would direct the development of future research. Children born preterm/LBW often display *average* IQ in the current literature; yet frequent academic difficulties and persistent EF-related impairments. Children born preterm/LBW often do not evince detectable impairments, particularly in global measures, but continue to struggle at school or lag behind their peers in more basic cognitive processes (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009; Kerr-Wilson, et al., 2011; Pritchard, et al., 2009) such as WM and cognitive flexibility. In addition, the fact that IQ and G_f could part company has been demonstrated in children with ADHD, who also displayed EF impairments against an average IQ (Barkley, 1997; Duncan, et al., 1995; Tamm & Juranek, 2012). Given that the majority of current research focused on traditional IQ, investigations through G_f is a novel approach to the literature. G_f is also closely related to working memory and possibly other EF constructs, thus investigations through G_f may also assist traditional IQ in explaining the discrepancy of lower performance in EF tasks and academic results in those children born preterm with normal IQ. Therefore, the current choice of approach to gaining a better understanding

of the observed discrepancy was through investigations in their development of G_f and subsequently, ways of increasing G_f .

Summary of Findings

This section is divided into three sub-sections. The first is a summary of the findings in Study 1 and implications for developmental understanding of the impact of children born preterm. The second sub-section is a summary of the findings in Study 2 and its implications for theories of intelligence. The final sub-section discusses the clinical implications of these findings for children born preterm/LBW.

Summary of Study 1

Study 1 aimed to investigate whether children born preterm, from seven to nine years of age, exhibit G_f impairments. Further, whether birth status effects in G_f differences could be explained by differences in basic information processing parameters, namely WM and cognitive inflexibility. Results indicated that children born preterm did indeed show impairments in G_f and that these differences were partially mediated by WM and cognitive inflexibility. WM and cognitive inflexibility were strong predictors of G_f , each of them played a separate mediating role and there remained unexplained variance in the pathway model. Remaining unexplained variance may stem from other cognitive elements such as processing speed and attention, as suggested by Rose et al. (2011). On further analysis, children born preterm appeared to show developmental delay of at least one year compared to their same age peers, rather than a permanent deficit. This has important implications for the potential modifiability of these functions, and together with the understanding of their similarities and differences in developmental trajectories, may shed light on some of the neurodevelopmental challenges for children born preterm, whilst also pointing to areas that require further exploration.

Similar Developmental Trajectories in G_f

The two groups of participants in Study 1 appeared to have the same developmental trajectory, with children born preterm performing at least one year behind (refer to Figure 4.1). In the typically developing population, G_f changes rapidly in a linear manner in early to middle childhood. The changes of fluid ability in children over one year is of comparable magnitude to changes in adulthood that spans across approximately 11 years (Kaufman, et al., 2009; McArdle, et al., 2002; Schweizer & Koch, 2002; Sharma, et al., 2011). This study is amongst the first to demonstrate that there was a similar pattern of linear progression in mean scores on the G_f task between children born preterm and their peers across the three age groups. There was evidence of a developmental delay rather than a deficit. Given these similarities, remediation programs that serve to enhance G_f appears fruitful for children born preterm as a way of catching up to their same age peers.

Different Developmental Trajectories in WM

In the current study, the children born preterm and the typically developing children appeared to have different developmental trajectories in WM. Studies on the normal developmental trajectory of WM suggested that WM emerges as young as four years of age. By six to seven years, children are fully equipped with various WM components, such as short-term memory storage, phonological loop and visuo-spatial sketchpad. The development of WM increases in a linear fashion and peaks around age 20 (Best, et al., 2009; Gathercole, et al., 2004; Lehto, et al., 2003). According to existing literature, children born preterm/LBW displayed consistent deficiencies in the performance of WM tasks as compared to their same age peers (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009; Ford et al., 2011). Given that different measures and levels of task complexity, as well as gestational age and birth weight, can all influence

the outcomes of group comparison, a distinct developmental trend has not been proposed by existing studies. However, studies have detected a significant group difference with a combined effect size of -0.36 between children born preterm/LBW and their full-term peers, over the age range from seven to 14 years of age (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009).

From the results of the present study, the sample of typically developing children demonstrated the linear progression described in the literature. However, the sample of preterm in the current study deviated from this linear progression. Children born preterm struggled with storing and manipulating information more than their same age peers. There was a very similar pattern of mean scores indicating a developmental delay between ages seven and eight, but not between eight and nine. The effect size detected in the present study was also relatively high compared to that recorded in earlier reviews ($d = 0.30 - 0.61$). This suggested that children born preterm are likely to have their own developmental trajectory, which is perhaps much more delayed after developing the basic WM abilities at eight years of age (Refer to Figure 4.2). If this deceleration in WM development generalises to the preterm/LBW population, then it may be that children aged at seven could benefit from early intervention programs and those between eight and nine could benefit from remedial programmes, both of which target WM abilities. Furthermore, given the adaptive nature of the discussed WM training and close monitoring of training progress, the differences in WM developmental trajectory between children born preterm and those born at term in the older age groups is likely to have minimal affect on the generalization of the positive WM training effects in term born peers towards children born preterm. Nonetheless, the signs of different trajectories found in the present study could also be the result of random sampling, since this is a between-groups (cross-sectional) study and therefore

within-groups (longitudinal) data would help to disambiguate the current observation.

Different Developmental Trajectories in Cognitive Flexibility

In the current study, the children born preterm and typically developing children also showed different developmental trajectories in cognitive flexibility. The developmental trajectory of cognitive flexibility in the normal population is similar to that of WM in the normal population. According to the existing literature, the ability to shift between tasks and rules emerges as young as three years of age in typically developing children. Children at this age can shift between simple stimulus-response rules within a task measuring cognitive flexibility. These simple tasks should involve no more than two distinct rules. Significant improvements emerge around the age of five to six years where individuals are capable of shifting between more rules in more complex tasks. Between the ages of seven and nine years, children demonstrate significant improvement in the ability to switch between multi-dimensional tasks, but this plateaus at around the age of 12 years (Best et al., 2009; P. Anderson, 2002).

As compared to typically developing children, research on the developmental trajectory of cognitive flexibility in children born preterm/LBW is not well documented. Consistent evidence of deficit has been reported from studies using TMT-B as an assessment tool. Documentation of significant differences between VPT/VLBW and full-term groups, ranging from eight to 22 years old, was presented with a moderate effect size (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009). Evidence has been mixed for assessment tasks other than the TMT-B, where between-group differences in cognitive flexibility are not reliably found in children born VPT from 4-12 years old (Aarnoudse-Moens, et al., 2012) or at 16 years old (Luu et al., 2011).

Study 1 did not use the TMT-B, however, it did find significant group differences of a moderate effect size in 7- and 8-year-old children. The younger children

born preterm exhibited greater difficulty in shifting their focus of attention between tasks than their same age peers and made more errors on the WCST task. However, rather than a developmental lag, children born preterm in the present study demonstrated a “catching up” trend (Refer to Figure 4.3). Contrary to the normal developmental trajectory research, the current findings showed significant improvements in the preterm sample but not in the normal developing children across each age group. Given this finding, perhaps parents and teachers need not worry about the cognitive flexibility of children born preterm as, perhaps, it is only a matter of time when these children catch up to their peers. However, the degree of cumulative effect during the period of delay is unclear.

Summary of Study 2

The aim of the second study was to evaluate the suitability and the utility of a child-friendly version of adaptive CWM span training in improving G_f . Given that WM is highly predictive of G_f and that recent studies have suggested the possibility of improving WM and G_f after WM cognitive training, testing the effectiveness of WM training in typically developing children at seven years of age was expected to shed light on whether it is also suitable as early intervention and remedial programs for children born preterm/LBW between seven and nine years of age. Following a classic CWM span training of 20 days’ duration, 15 minutes each day, results indicated that training led to successful learning in the intervention training group. Results also showed that the adaptive WM span training task did not lead to improvements in non-trained measures, including WM, controlled attention, and speed of processing, however, training did demonstrate successful far transfer to the assessment of untrained G_f . The results of having far transfer in G_f in the absence of near transfer to WM and attentional control task were unexpected. A precise explanation cannot be drawn to this

pattern of findings at this point and further investigation is warranted.

Results with Higher Validity than Existing Studies

Many of the previous studies that documented successful generalization from training gains have used only passive/no-contact control groups. This raised concerns over the validity of these studies as participants in the training group of these studies may have performed better due to expectation associated with being actually in training and more contact or coaching from the researchers in the experiment (Redick, et al., 2012; Shipstead, et al., 2010). In response to these concerns of validity, the current study has demonstrated successful far transfer to G_f after adaptive CWM training with the inclusion of both active and passive control groups. Therefore, the results of adaptive CWM training can be said to have resulted in genuine changes in G_f over and above those seen in both control conditions.

Theoretical Contributions

The results of Study 2 also make other theoretical contributions. First, the current findings on successful far transfer effects to G_f add supporting evidence to the literature regarding a theoretical causal, directional, relationship between WM and G_f . Such experimental evidence is considered much stronger confirmation of the causal role of WM towards G_f , compared to correlational studies that documented naturally occurring associations between WM and G_f ranging from 0.64 to 0.82 (Fry & Hale, 2000).

Second, the current investigation with typically developing children supports Klingberg et al.'s (2002) notion that initial impairment in WM is not a pre-requisite for seeing cognitive transfer after computerized WM training. Consequently, this supports the idea that G_f can be modified after all (Jaeggi, et al., 2008; Sternberg, 2008) and that intelligence may not be as resistant to change as proposed by some researchers (Jensen,

1998; Redick, et al., 2012).

Integrated Practical Implications for Children Born Preterm/LBW

The two studies presented in this thesis make contributions to clinical practice with children born preterm. Study 1 provided information that may assist in targeting cognitive abilities for remedial programmes and promoting educational support for possible ways to increase WM and G_f amongst children born preterm.

To elaborate, the results of Study 1 provided evidence of developmental delay in G_f and also found a relatively large sized deficit in G_f ($d = -0.77$) in the present study compared even to academic difficulties found in math in previous studies ($d = -0.60$; (Aarnoudse-Moens, Weisglas-Kuperus, et al., 2009), a subject that has been suggested to be affected most profoundly amongst the clinical cohort (Bowen, et al., 2002; Pritchard, et al., 2009). Given that math has been known to rely heavily on WM (DeStefano & LeFevre, 2004) and G_f is predictive of math performance (Sharma, et al., 2011), it is of priority and great educational interest to investigate further the relationship surrounding these constructs. With this information, together with evidence on the modifiability of G_f , the present results provide hope for children in the clinical cohort that perhaps computerized WM training may increase G_f and subsequently lead to better academic outcomes.

In order to provide recommendations for possible treatment, causal directions need to be understood and the current thesis has attempted to achieve this through experimental manipulation in Study 2. Several factors suggested that recent success in computerized WM training to enhance WM and possibly G_f may be fruitful for children born preterm/LBW. First, WM has also been found to partially mediate birth status effects on G_f in Study 1. Second, WM has been viewed as underpinning academic difficulties in children born preterm/LBW (Alloway, 2009) and finally, parallel research

suggests a strong correlation between WM and G_f (Fry & Hale, 2000).

The detection of successful far transfer on G_f following adaptive CWM training in the sample of typically developing children in Study 2 provides support for its use with children born preterm/LBW. This clinical population may now be invited for experimental procedures without risking unnecessary stressful events on top of existing medical interventions that they may have already experienced. Furthermore, given that the findings in Study 2 provide evidence of positive effects of adaptive CWM training for normal developing children aged seven to nine years old, not attributable to practice or Hawthorne effects, it would be predicted that the same training regime accompanied with the same training schedule and dosage would be beneficial to children born preterm/LBW.

Future Directions

Continuous research is needed to fully understand the underpinning of cognitive differences between preterm/LBW and their control peers as well as determining appropriate remedial programs to enhance their cognitive abilities and academic outcomes. It is recommended that future research to not only refine and replicate the present research as documented in limitations in earlier sections but also to move forward in methodologies and pose practical research questions.

Future research should seek to verify the importance of basic information processes in explaining G_f differences between preterm/LBW and their full-term peers through multiple mediation modelling or SEM analyses as this line of approach is relatively sparse when compared against studies that demonstrate group differences. The use of a latent variable approach in future research, for both theoretical understanding of cognitive abilities in children born preterm/LBW and practical intervention experiments, may also enhance methodological power.

A subsequent step forward would be bridging research to connect the understanding of birth status effects on G_f through EF-related components to their impact on the academic performance in children born preterm/LBW. This would provide us with answers as whether developmental delay found in G_f adequately explains academic difficulties amongst the clinical cohort. Questions remain as to whether deficits in G_f explain discrepancies in academic achievement more adequately than the G_f/G_c conglomerate measured by FSIQ. In addition, the birth status effects on G_c and its relationship with academic achievement amongst children born preterm/LBW have yet to be thoroughly investigated. Concurrently, clarification of these issues shall further assist in the development of educational remedial programs to improve the sequelae of children born preterm/LBW.

Another area deemed fruitful to explore is whether this generalized gain in G_f also transfers to practical, real life behaviours, because if they do not, then they are of little practical significance (Sternberg, 2008). Therefore, aside from examining whether adaptive CWM training can increase performance on untrained G_f amongst children born preterm/LBW, research can also target whether training can generalize to achievement and behaviour assessments.

Conclusions

The incidence of children born preterm/LBW remains high and will continue to rise given the continuous improvements in medical procedures that lead to the increase in survival rates. Therefore, there is an ongoing need not only to monitor the changes in cognitive abilities between this clinical cohort relative to full-term typically developing children, but also to understand the underlying mechanisms of identified differences in intelligence and the role of basic information processes.

The understanding of developmental trajectories and cognitive abilities

underlying G_f deficits in children born preterm/LBW, together with the implication that intensive adaptive CWM training can increase G_f , will better prepare psychologists in offering suggestions to this clinical cohort with appropriate educational and remedial programmes in the future.

References

- Aarnoudse-Moens, C. S. H., Duivenvoorden, H. J., Weisglas-Kuperus, N., Van Goudoever, J. B., & Oosterlaan, J. (2012). The profile of executive function in very preterm children at 4 to 12 years. *Developmental Medicine & Child Neurology*, *54*, 247-253. doi: 10.1111/j.1469-8749.2011.04150.x
- Aarnoudse-Moens, C. S. H., Smidts, D. P., Oosterlaan, J., Duivenvoorden, H. J., & Weisglas-Kuperus, N. (2009). Executive function in very preterm children at early school age. *Journal of Abnormal Child Psychology*, *37*, 981-993. doi: 10.1007/s10802-009-9327-z
- 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*, *124*(2), 717-728. doi: 10.1542/peds.2008-2816
- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*(1), 30-60. doi: 10.1037/0033-2909.131.1.30
- Adams-Chapman, I. (2009). Insults to the developing brain and impact on neurodevelopmental outcome. *Journal of Communication Disorders*, *42*, 256-262. doi: 10.1016/j.jcomdis.2009.03.010
- Allen, P., & Bennett, K. (2010). *PASW statistics by SPSS: A practical guide, version 18.0*. South Melbourne, Vic: Cengage Learning Australia.
- Allin, M., Walshe, M., Fern, A., Nosarti, C., Cuddy, M., Rifkin, L., . . . Wyatt, J. (2008). Cognitive maturation in preterm and term born adolescents. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*, 381-386. doi: 10.1136/jnnp.2006.110858

- Alloway, T. P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Assessment, 25*(2), 1-7. doi: 10.1027/1015-5759.25.2.xxx
- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106*, 20-29. doi: 10.1016/j.jecp.2009.11.003
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliot, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*(2), 606-621.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th Text Revision ed.). Washington, DC: American Psychiatric Association.
- Amor, L. B., Chantal, S., & Bairam, A. (2012). Relationship between late preterm birth and expression of attention-deficit hyperactivity disorder in school-aged children: Clinical, neuropsychological, and neurobiochemical outcomes. *Research and Reports in Neonatology, 2*, 77-83. doi: 10.2147/RRN.S34674
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 8*(2), 71-82. doi: 10.1076/chin.8.2.71.8724
- Anderson, P. J., Doyle, L. W., & the Victorian Infant Collaborative Study Group. (2003). Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. *The Journal of the American Medical Association, 289*(24), 3264-3272.
- Anderson, P. J., Howard, K., & Doyle, L. W. (2010). Executive function development

- in preterm children. In C. Nosarti, R. M. Murray & M. Hack (Eds.), *Neurodevelopment Outcomes of Preterm Birth: From Childhood to Adult Life*. New York: Cambridge University Press.
- Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology*, 28(1), 539-560. doi: 10.1207/s15326942dn2801_5
- Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, 49A(1), 5-28. doi: 10.1080/713755608
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-422.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). New York: Academic.
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). New York: Cambridge University Press.
- Barceló, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia*, 40(3), 349-356. doi: 10.1016/S0028-3932(01)00110-5
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65-94. doi: 10.1037/0033-2909.121.1.65
- Barnett, S. M., & Stephen, J. C. (2002). When and where do we apply what we learn? A

- taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612-637. doi: 10.1037//0033-2909.128.4.612
- Baron, I. S., Erickson, K., Ahronovich, M. D., Litman, F. R., & Brandt, J. (2010). Spatial location memory discriminates children born at extremely low birth weight and late-preterm at age three. *Neuropsychology*, 24(6), 787-794. doi: 10.1037/a0020382
- Baron, I. S., & Rey-Casserly, C. (2010). Extremely preterm birth outcome: A review of four decades of cognitive research. *Neuropsychology Review*, 20(4), 430-452. doi: 10.1007/s11065-010-9132-z
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Bayliss, D. M., Jarrold, C., Baddeley, A., Gunn, D. M., & Leigh, E. (2005). Mapping the developmental constraints on working memory span performance. *Developmental Psychology*, 41(4), 579-597. doi: 10.1037/0012-1649.41.4.579
- Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Child Psychology: General*, 132(1), 71-92. doi: 10.1037/0096-3445.132.1.71
- Beck, S., Wojdyla, D., Say, L., Betran, A. P., Merialdi, M., Requejo, J. H., . . . Van Look, P. F. (2010). The worldwide incidence of preterm birth: A systematic review of maternal mortality and morbidity. *Bulletin of the World Health Organization*, 88, 31-38. doi: 10.2471/BLT.08.062554
- Behrman, R. E., & Butler, A. S. (2007). Biological pathways leading to preterm birth *Preterm Birth: Causes, Consequences, and Prevention* (pp. 169-206).

Washington DC: The National Academies Press.

- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29*(3), 180-200. doi: 10.1016/j.dr.2009.05.002
- Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *Journal of American Medical Association, 288*(6), 728-737.
- Binet, A., & Simon, T. (1916). The development of intelligence in children: The Binet-Simon Scale. *Williams & Wilkins, 11*.
- Blair, C. (2002). School Readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *The American Psychologist, 57*(2), 111-127. doi: 10.1037/0003-066X.57.2.111
- Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Science, 29*, 109-160.
- Blencowe, H., Cousens, S., Z., O. M., Chou, D., Moller, A.-B., Narwal, R., . . . Lawn, J. E. (2012). National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet, 379*(9832), 2162-2172. doi: 10.1016/S0140-6736(12)60820-4
- Bonate, P. L. (2000). *Analysis of pretest-posttest designs*. Boca Raton, FL: Chapman & Hall.
- Boulet, S. L., Schieve, L. A., & Boyle, C. A. (2011). Birth Weight and Health and Developmental Outcomes in US children, 1997-2005. *Maternal and Child*

Health Journal, 15(7), 836-844. doi: 10.1007/s10995-009-0538-2

Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The reliability and internal validity of the Wisconsin Card Sorting Test. *Neuropsychological Rehabilitation: An International Journal*, 8(3), 243-254. doi: 10.1080/713755573

Bowden, S. C., Petrauskas, V. M., Bardenhagen, F. J., Meade, C. E., & Simpson, L. C. (2013). Exploring the dimensionality of digit span. *Assessment*, 20(2), 188-198. doi: 10.1177/1073191112457016

Bowen, J. R., Gibson, F. L., & Hand, P. J. (2002). Educational outcome at 8 years of children who were born extremely prematurely: A controlled study. *Journal of Paediatrics and Child Health*, 38, 438-444. doi: 10.1046/j.1440-1754.2002.00039x

Boyd, H. A., Poulsen, G., Wohlfahrt, J., Murray, J. C., Feenstra, B., & Melbye, M. (2009). Maternal contributions to preterm delivery. *American Journal of Epidemiology*, 170(11), 1358-1364. doi: 10.1093/aje/kwp324

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence*, 40(5), 458-469. doi: 10.1016/j.intell.2012.05.006

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, 19(3), 273-293. doi: 10.1207/S15326942DN1903_3

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of Frontal and Executive Function* (pp. 79-113). Hove: Psychology Press Ltd.

Burke. (1972). Raven's progressive matrices: Validity, reliability, and norms. *The*

- Journal of Psychology*, 82(2), 253-257. doi: 10.1080/00223980.1972.9923815
- Buschkuehl, M., Jaeggi, S., Kobel, A., & Perrig, W. J. (2008). *BrainTwister: A collection of cognitive training tasks*: The University of Bern, Department of Psychology, Division of Experimental Psychology and Neuropsychology.
- Buschkuehl, M., & Jaeggi, S. M. (2010). Improving intelligence: a literature review. *Swiss Medical Weekly*, 140(19-20), 266-272.
- Campbell, M. K., Cartier, S., Xie, B., Kouniakakis, G., Huang, W., & Han, V. (2012). Determinants of small for gestational age birth at term. *Paediatric and Perinatal Epidemiology*, 26, 525-533. doi: 10.1111/j.1365-3016.2012.01319.x
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97(3), 404-431.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.
- Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports g and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5-21): Elsevier/Pergamon.
- Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *Trends in Cognitive Sciences*, 10(3), 117-123. doi: 10.1016/j.tics.2006.01.011
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22.
- Cattell, R. B. (1973). *Measuring intelligence with the Culture Fair Tests*. Champaign, Illinois: Institute for Personality and Ability Testing.

- Cattell, R. B. (1987). The natural history of ability: Distribution and relation to sex and age. In R. B. Cattell (Ed.), *Intelligence: Its structure, growth and action* (pp. 153-212). Amsterdam: Elsevier Science Publishers B. V.
- Cattell, R. B., & Cattell, A. K. B. (1949). *Culture Fair Intelligence Test*. Champaign, Illinois: Institute for Personality and Ability Testing.
- Chase-Carmichael, C. A., Douglas Ris, M., Weber, A. M., & Schefft, B. K. (1999). Neurologic validity of the Wisconsin Card Sorting Test with a pediatric population. *The Clinical Neuropsychologist*, *13*(4), 405-413. doi: 10.1076/1385-4046(199911)13:04;1-Y;FT405
- Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review*, *17*(2), 193-199. doi: 10.3758/PBR.17.2.193
- Chen, J.-Y. (1996). A problem in measuring the stroop facilitation and interference effects: Implications for measuring performance change in general. *Perceptual and Motor Skills*, *83*, 1059-1070.
- Chen, Z., & Cowan, N. (2008). Core verbal working-memory capacity: The limit in words retained without covert articulation. *The Quarterly Journal of Experimental Psychology* *62*(7), 1420-1429. doi: 10.1080/17470210802453977
- Chooi, W.-T. (2012). Working memory and intelligence: A brief review. *Journal of Educational and Developmental Psychology*, *2*(2), 42-50. doi: 10.5539/jedp.v2n2p42
- Chooi, W.-T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, *40*, 531-542. doi: 10.1016/j.intell.2012.07.004
- Chuderski, A., Taraday, M., Necka, E., & Smolen, T. (2012). Storage capacity explains

- fluid intelligence but executive control does not. *Intelligence*, 40, 278-295. doi: 10.1016/j.intell.2012.02.010
- Clark, C. A. C., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology*, 46(5), 1176-1191. doi: 10.1037/a0019672
- Clark, C. A. C., & Woodward, L. J. (2010). Neonatal cerebral abnormalities and later verbal and visuospatial working memory abilities of children born very preterm. *Developmental Neuropsychology*, 35(6), 622-642. doi: 10.1080/87565641.2010.508669
- Clausson, B., Lichtenstein, P., & Cnattingius, S. (2000). Genetic influence on birthweight and gestational length determined by studies in offspring of twins. *British Journal of Obstetrics and Gynaecology*, 107, 375-381.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. doi: 10.1037/0033-2909.112.1.155
- Colom, R., Abad, F. J., Garcia, L. F., & Juan-Espinosa, M. (2002). Education, Wechsler's Full Scale IQ, and g. *Intelligence*, 30(5), 449-462. doi: 10.1016/S0160-2896(02)00122-8
- Colom, R., Abad, F. J., Rebollo, I., & Shih, P. C. (2005). Memory span and general intelligence: A latent-variable approach. *Intelligence*, 33, 623-642. doi: 10.1016/j.intell.2005.05.006
- Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: a reanalysis of key studies. *Memory & Cognition*, 34(1), 158-171.
- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, 32, 277-296.

doi: 10.1016/j.intell.2003.12.002

Colom, R., Shih, P. C., Flores-Mendoza, C., & Quiroga, M. A. (2006). The real relationship between short-term memory and working memory. *Memory, 14*(7), 804-813. doi: 10.1080/09658210600680020

Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*(2), 163-183. doi: 10.1016/S0160-2896(01)00096-4

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span task: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547-552. doi: 10.1016/j.tics.2003.10.005

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Science 24*, 87-185. doi: 10.1017/S0140525X01003922

Cragg, L., Fox, A., Nation, K., Reid, C., & Anderson, M. (2009). Neural correlates of successful and partial inhibitions in children: An ERP study. *Developmental Psychobiology, 51*(7), 533-543. doi: 10.1002/dev.20391

Dahlin, K. I. E. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing, 24*(4), 479-491. doi: 10.1007/s11145-010-9238-y

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422-

433. doi: 10.3758/BF03214546

Dang, C.-P., Braeken, J., Ferrer, E., & Liu, C. (2012). Unitary or non-unitary nature of working memory? Evidence from its relation to general fluid and crystallized intelligence. *Intelligence*, *40*, 499-508. doi: 10.1016/j.intell.2012.05.002

Davies, S. R. (2004). *An investigation of the cognitive nature of general intelligence*. Doctor of Philosophy/Master of Psychology (Clinical), The University of Western Australia. Retrieved from http://repository.uwa.edu.au:80/R/?func=dbin-jump-full&object_id=8211&silo_library=GEN01

Davis, P. G., Schmidt, B., Roberts, R. S., Doyle, L. W., Asztalos, E., Haslam, R., . . . Tin, W. (2010). Caffeine for apnea of prematurity trial: Benefits may vary in subgroups. *The Journal of Pediatrics*, *156*(3), 382-287. doi: 10.1016/j.jpeds.2009.09.069

de Kieviet, J. F., Zoetebier, L., van Elburg, R. M., Vermeulen, R. J., & Oosterlaan, J. (2012). Brain development of very preterm and very low-birthweight children in childhood and adolescence: a meta-analysis review. *Developmental Medicine & Child Neurology*, *54*(4), 313-323. doi: 10.1111/j.1469-8749.2011.0426.x

Deary, I. J., & Batty, G. D. (2007). Cognitive epidemiology. *Journal of Epidemiology and Community Health*, *61*(5), 378-384. doi: 10.1136/jech.2005.039206

Deary, I. J., Egan, V., Gibson, G. J., Austin, E. J., Brand, C. R., & Kellaghan, T. (1996). Intelligence and the differentiation hypothesis. *Intelligence*, *23*(2), 105-132. doi: 10.1016/S0160-2896(96)90008-2

Deary, I. J., Strand, S., Smith, P., & Fernanades, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21. doi: 10.1016/j.intell.2006.02.001

Dempster, F. N. (1981). Memory span: Sources of individual and development differences. *Psychological Bulletin*, *89*(1), 63-100. doi: 10.1037/0033-

2909.89.1.63

- Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of International Neuropsychological Society, 15*(3), 331-343. doi: 10.1017/S1355617709090481
- DeStefano, D., & LeFevre, J. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353-386. doi: 10.1080/09541440244000328
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low-IQ groups. *Intelligence, 13*(4), 349-359. doi: 10.1016/S0160-2896(89)80007-8
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466-503). New York: Oxford University Press.
- Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4-12 years old. *Science, 333*(6045), 959-964. doi: 10.1126/science.1204529
- Doyle, L. W., & Anderson, P. J. (2010). Adult outcome of extremely preterm infants. *Pediatrics, 126*, 342-351. doi: 10.1542/peds.2010.0710
- Doyle, L. W., Crowther, C. A., Middleton, P., Marret, S., & Rouse, D. (2010). Magnesium sulphate for women at risk of preterm birth for neuroprotection of the fetus. *Cochrane Database of Systematic Reviews 2009*(1), 1-96. doi: 10.1002/14651858.CD004661.pub3
- Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive

- functions and intelligence on 11- to 12-year-old children. *Psychological Test and Assessment Modeling*, 52(4), 419-431.
- Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, 33(3), 261-268.
- Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognitive Psychology*, 30(3), 257-303. doi: 10.1006/cogp.1996.0008
- Edgin, J. O., Inder, T. E., Anderson, P. J., Hood, K. M., Clark, C. A. C., & Woodward, L. J. (2008). Executive functioning in preschool children born very preterm: Relationship with early white matter pathology. *Journal of International Neuropsychological Society*, 14, 90-101. doi: 10.1017/S1355617708080053
- Elgen, I., Sommerfelt, K., & Markestad, T. (2002). Population based, controlled study of behavioural problems and psychiatric disorders in low birthweight children at 11 years of age. *Archives of Disease in Childhood*, 87(2), F128-F132. doi: 10.1136/fn.87.2.F128
- Elliott, R., Sahakian, B. J., Matthews, K., Bannerjea, A., Rimmer, J., & Robbins, T. W. (1997). Effects of methylphenidate on spatial working memory and planning in healthy young adults. *Psychopharmacology*, 131(2), 196-206.
- Endler, A., Rey, G. D., & Butz, M. V. (2012). Towards motivation-based adaptation of difficulty in e-learning programs. *Australasian Journal of Educational Technology*, 28(7), 1119-1135.
- Engel de Abreu, P. M. J., Conway, A. R. A., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence*, 38, 552-561. doi: 10.1016/j.intell.2010.07.003
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current*

Directions in Psychological Science, 11(1), 19-23.

- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 44, pp. 144-199). New York, NJ: Elsevier.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331. doi: 10.1037/0096-3445.128.3.309
- Eslinger, P. J., Flaherty-Craig, C. V., & Benton, A. L. (2004). Developmental outcomes after early prefrontal cortex damage. *Brain and Cognition*, 55(1), 84-103. doi: 10.1016/S0278-2626(03)00281-1
- Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology*, 13(4), 495-499. doi: 10.1080/87565649709540690
- Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven II: childhood intelligence and later outcomes in adolescence and young adulthood. *Journal of Child Psychology and Psychiatry*, 46(8), 850-858. doi: 10.1111/j.1469-7610.2005.01472.x
- Ferrer, E., & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, 40(6), 935-952. doi: 10.1037/0012-1649.40.6.935
- Field, A. (2009a). Analysis of covariance, ANCOVA (GLM 2) *Discovering Statistics Using SPSS* (3rd ed., pp. 395-420). London: SAGE Publications Ltd.
- Field, A. (2009b). Comparing several means: ANOVA (GLM 1) *Discovering Statistics Using SPSS* (pp. 347-394). London: SAGE Publications Ltd.

- Field, A. (2009c). Exploring assumptions *Discovering Statistics Using SPSS* (3rd ed., pp. 131-165). London: SAGE Publications Ltd.
- Flora, S. R. (2004). The reality of extrinsic rewards and intrinsic motivation *The Power of Reinforcement* (pp. 61-73). New York: State University of New York Press.
- Ford, R. M., Neulinger, K., O'Callaghan, M., Mohay, H., Gray, P., & Shum, D. (2011). Executive function in 7-9-year-old children born extremely preterm or with extremely low birth weight: Effects of biomedical history, age at assessment, and socioeconomic status. *Archives of Clinical Neuropsychology*, *26*, 632-644. doi: 10.1093/arclin/acr061
- Fraello, D., Maller-Kesselman, J., Vohr, B., Katz, K., Kesler, S., Scheider, K., . . . Spann, M. (2011). Consequence of preterm birth in early adolescence: The role of language on auditory short-term memory. *Journal of Child Neurology*, *26*(6), 738-742. doi: 10.1177/0883073810391904
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172-179. doi: 10.1111/j.1467-9280.2006.01681.x
- Frisk, V., Amsel, R., & Whyte, H. E. A. (2002). The importance of head growth patterns in predicting the cognitive abilities and literacy skills of small-for-gestational-age children. *Developmental Neuropsychology*, *22*(3), 565-593. doi: 10.1207/S15326942DN2203_2
- Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*(4), 237-241.
- Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*, 1-34.

- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrated framework. *Psychological Bulletin*, *134*(1), 31-60. doi: 10.1037/0033-2909/134.1.31
- Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, *22*, 1019-1037. doi: 10.1002/acp.1407
- Gathercole, S. E., & Hitch, G. J. (1993). Developmental changes in short-term memory: A revised working memory perspective. In A. Collins, S. E. Gathercole, M. A. Conway & P. E. Morris (Eds.), *Theories of memory* (pp. 189-210). Hove, England: Erlbaum.
- Gathercole, S. E., Pickering, S. L., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177-190. doi: 10.1037/0012-1649.40.2.177
- Gignac, G. E. (2006). Evaluating subtest 'g' saturation levels via the single trait-correlated uniqueness (STCU) SEM approach: Evidence in favor of crystallized subtests as the best indicators of 'g'. *Intelligence*, *34*(1), 29-46. doi: 10.1016/j.intell.2005.06.001
- Glutting, J. J., Youngstrom, E. A., Ward, T., Ward, S., & Hale, R. L. (1997). Incremental efficacy of WISC-III factor scores in predicting achievement: What do they tell us? *Psychological Assessment*, *9*(3), 295-301.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*(1), 79-132. doi: 10.1016/S0160-2896(97)90014-3
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*(3), 316-322. doi: 10.1038/nn1014

- Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging, 23*(4), 692-701. doi: 10.1037/a0014345
- Grunau, R. E., Whitfield, M. F., & Davis, C. (2002). Pattern of learning disabilities in children with extremely low birth weight and broadly average intelligence. *Archives of Pediatrics & Adolescent Medicine, 156*(6), 615-620. doi: 10.1001/archpedi.156.6.615
- Grunewaldt, K. H., Lohaugen, G. C. C., Austeng, D., Brubakk, A.-M., & Skranes, J. (2013). Working memory training improves cognitive function in VLBW preschoolers. *Pediatrics 131*(3), e747-e754. doi: 10.1542/peds.2012-1965
- Grunewaldt, K. H., Skranes, J., Brubakk, A.-M., & Lahaugen, G. C. C. (2015). Computerized working memory training has positive long-term effect in very low birthweight preschool children. *Developmental Medicine & Child Neurology, 3*(51), 1-7. doi: 10.1111/dmcn.12841
- Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*(3), 179-203. doi: 10.1016/0160-2896(84)90008-4
- Hallin, A., Hellstrom-Westas, L., & Stjernqvist, K. (2010). Follow-up of adolescents born extremely preterm: Cognitive function and health at 18 years of age. *Acta Paediatrica, 99*, 1401-1406. doi: 10.1111/j.1651-2227.2010.01850.x
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education, 10*(2), 169-207. doi: 10.1080/0969594032000121270
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408-420. doi: 10.1080/03637750903310360
- Hayes, A. F. (2013). Introduction to mediation, moderation, and conditional process

- analysis: A regression-based approach Retrieved May 10, 2012, from <http://www.afhayes.com/introduction-to-mediation-moderation-and-conditional-process-analysis.html>
- Hayes, B., & Sharif, F. (2009). Behavioural and emotional outcome of very low birth weight infants - literature review. *Journal of Maternal-Fetal and Neonatal Medicine*, 22(10), 849-856. doi: 10.1080/14767050902994507
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.
- Heitz, R. P., Unsworth, N., & Engle, R. W. (2005). Working memory capacity, attention control, and fluid intelligence. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of Understanding and Measuring Intelligence* (pp. 61-77). Thousand Oaks, CA: Sage Publications Inc.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.
- Hetland, L. (2000). Listening to music enhances spatial-temporal reasoning: Evidence for the "Mozart Effect". *Journal of Aesthetic Education*, 34(3/4), 105-148.
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children *Developmental Science*, 12(4), F9-F15. doi: 10.1111/j.1467-7687.2009.00848.x
- Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology*, 24, 827-836. doi: 10.1002/acp.1589
- Homack, S., & Riccio, C. A. (2004). A meta-analysis of the sensitivity and specificity

- of the Stroop Color and Word Test with children. *Archives of Clinical Neuropsychology*, *19*, 725-743.
- Hornung, C., Brunner, M., Reuter, R. A. P., & Martin, R. (2011). Children's working memory: Its structure and relationship to fluid intelligence. *Intelligence*, *39*, 210-221. doi: 10.1016/j.intell.2011.03.002
- Hoyt, W. T., Imel, Z. E., & Chan, F. (2008). Multiple regression and correlation techniques: Recent controversies and best practices. *Rehabilitation Psychology*, *53*(3), 321-339. doi: 10.1037/a0013021
- Hughes, C., & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions? *Child and Adolescent Mental Health*, *7*(3), 131-142. doi: 10.1111/1475-3588.00024
- Hutchinson, E. A., De Luca, C. R., Doyle, L. W., Roberts, G., Anderson, P. J., & The Victorian Infant Collaborative Study Group. (2013). School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics*, *131*(4), e1053-e1061. doi: 10.1542/peds.2012-2311
- Hutchinson, J. A. (2012). *Western Australia's mothers and babies 2010: Twenty-eighth annual report of the Western Australian midwives' notification system*. Department of Health, Western Australia.
- Inder, T. E., Lawrence, R. K., & Neil, J. J. (2010). Neuroimaging: Imaging the preterm brain. In C. Nosarti, R. M. Murray & M. Hack (Eds.), *Neurodevelopmental Outcomes of Preterm Birth: From Childhood to Adult Life*. New York: Cambridge University Press.
- Isquith, P. K., Gioia, G., & Espy, K. A. (2004). Executive function in preschool children: Examining through everyday behavior. *Developmental Neuropsychology*, *26*(1), 403-422. doi: 10.1207/s15326942dn2601_3

- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(19), 6829-6833. doi: 10.1073/pnas.0801268105
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(25), 10081-10086. doi: 10.1073/pnas.1103228108
- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning - implication for training and transfer. *Intelligence*, *38*(6), 625-635. doi: 10.1016/j.intell.2010.09.001
- Jaekel, J., Wolke, D., & Bartmann, P. (2013). Poor attention rather than hyperactivity/impulsivity predicts academic achievement in very preterm and full-term adolescents. *Psychological Medicine*, *43*(1), 183-196. doi: 10.1017/S0033291712001031
- Jensen, A. R. (1994). Psychometric g related to differences in head size. *Personality and Individual Differences*, *17*(5), 597-606. doi: 10.1016/0191-8869(94)90132-5
- Jensen, A. R. (1998). *The g Factor: The Science of Mental Ability*. Westport, CT: Praeger.
- Jensen, A. R. (2003). Regularities in Spearman's Law of Diminishing Returns. *Intelligence*, *31*(2), 95-105. doi: 10.1016/S0160-2896(01)00094-0
- Johansson, S., & Cnattingius, S. (2010). Epidemiology of preterm birth. In C. Nosarti, R. M. Murray & M. Hack (Eds.), *Neurodevelopmental Outcomes of Preterm Birth: From Childhood to Adult Life* (pp. 1-16). New York: Cambridge

University Press.

- Johnson, S. (2007). Cognitive and behavioural outcomes following very preterm birth. *Seminars in Fetal & Neonatal Medicine*, *12*, 363-373. doi: 10.1016/j.siny.2007.05.004
- Johnson, S., Wolke, D., Hennessy, E., & Marlow, N. (2011). Educational outcomes in extremely preterm children: Neuropsychological correlates and predictors of attainment. *Developmental Neuropsychology*, *36*(1), 74-95. doi: 10.1080/87565641.2011.540541
- Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*, 213-233. doi: 10.1007/s11065-007-9040-z
- Kane, M. J., Conway, A. R. A., Bleckley, M. K., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169-183. doi: 10.1037//0096-3445.130.2.169
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66-71. doi: 10.1037/0033-2909.131.1.66
- Kaufman, A. S., Kaufman, J. C., Liu, X., & Johnson, C. K. (2009). How do educational attainment and gender relate to fluid intelligence, crystallized intelligence, and academic skills at 22-90 years? *Archives of Clinical Neuropsychology*, *24*, 153-163. doi: 10.1093/arclin/acp015
- Kaufman, A. S., & Wang, J.-J. (1992). Gender, race, and education differences on the K-Bit at ages 4 to 90 years. *Journal of Psychoeducational Assessment*, *10*(3), 219-229. doi: 10.1177/073428299201000302

- Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests and Issues* (pp. 581-614). New York: Guilford Press.
- Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Fourth Edition: What does it measure? *School Psychology Review, 35*(1), 108-127.
- Kerr-Wilson, C. O., Mackay, D. F., Smith, G. C. S., & Pell, J. P. (2011). Meta-analysis of the association between preterm delivery and intelligence. *Journal of Public Health, 34*(2), 209-216. doi: 10.1093/pubmed/fdr024
- Khwaja, O., & Volpe, J. J. (2008). Pathogenesis of cerebral white matter injury of prematurity. *Archives of Disease in Childhood Fetal & Neonatal Edition, 93*(2), F153-F161. doi: 10.1136/adc.2006.108837
- Kirkegaard, I., Obel, C., Hedegaard, M., & Henriksen, T. B. (2006). Gestational age and birth weight in relation to school performance of 10-year-old children: A follow-up study of children born after 32 completed weeks. *Pediatrics, 118*(4), 1600-1606. doi: 10.1542/peds.2005-2700
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstorm, K., . . . Westerberg, H. (2005). Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 177-186.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*(6), 781-791.

- Krageloh-Mann, I. (2010). Clinical outcome: neurological sequelae following preterm birth. In C. Nosarti, R. M. Murray & M. Hack (Eds.), *Neurodevelopmental Outcomes of Preterm Birth: From Childhood to Adult Life*. New York: Cambridge University Press.
- Kramer, M. S., Wilkins, R., Goulet, L., Seguin, L., Lydon, J., Kahn, S. R., . . . Platt, R. W. (2009). Investigating socio-economic disparities in preterm birth: Evidence for selective study participation and selection bias. *Paediatric and Perinatal Epidemiology*, 23(4), 301-309. doi: 10.1111/j.1365-3016.2009.01042.x
- Kvist, A. V., & Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's Investment theory. *Intelligence*, 36, 422-436. doi: 10.1016/j.intell.2007.08.004
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability (little more than) working-memory capacity?! *Intelligence*, 14, 389-433. doi: 10.1016/S0160-2896(05)80012-1
- Lawn, J. E., Gravett, M. G., Nunes, T. M., Rubens, C. E., Stanton, C., & the GAPPS Review Group. (2010). Global report on preterm birth and stillbirth (1 or 7): Definitions, description of the burden and opportunities to improve data. *BMC Pregnancy and Childbirth*, 10(Suppl 1), S1.
- Le, M., & Tran, B. (2008). *Perinatal Statistics in Western Australia, 2008: Twenty-sixth Annual Report of the Western Australian Midwives' Notification System* Department of Health, Western Australia.
- Lee, E. S., Yeatman, J. D., Luna, B., & Feldman, H. M. (2011). Specific language and reading skills in school-aged children and adolescents are associated with prematurity after controlling for IQ. *Neuropsychologia*, 49, 906-913. doi: 10.1016/j.neuropsychologia.2010.12.038

- Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*(1), 59-80. doi: 10.1348/026151003321164627
- Lindstrom, K., Lindblad, F., & Hjern, A. (2011). Preterm birth and Attention-Deficit/Hyperactivity Disorder in schoolchildren. *Pediatrics, 127*(5), 858-865. doi: 10.1542/peds.2010-1279
- Lohaugen, G. C. C., Antonsen, I., Haberg, A., Gramstad, A., Vik, T., Brubakk, A., & Skranes, J. (2011). Computerized working memory training improves function in adolescents born at extremely low birth weight. *Journal of Pediatrics, 158*(4), 555-561. doi: 10.1016/j.jpeds.2010.09.060
- Lohaugen, G. C. C., Gramstad, A., Evensen, K. A., Martinussen, M., Lindqvist, S., Indredavik, M., . . . Skranes, J. (2010). Cognitive profile in young adults born preterm at very low birthweight. *Developmental Medicine & Child Neurology, 52*, 1133-1138. doi: 10.1111/j.1469-8749.2010.03743.x
- Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18*(1), 62-78. doi: 10.1080/09297049.2011.575772
- Lundqvist, A., Grundstrom, K., Samuelsson, K., & Ronnberg, J. (2010). Computerized training of working memory in a group of patients suffering from acquired brain injury. *Brain Injury, 24*(10), 1173-1183. doi: 10.3109/02699052.2010.498007
- Luu, T. M., Ment, L., Allan, W., Schneider, K., & Vohr, B. R. (2011). Executive and memory function in adolescents born very preterm. *Pediatrics, 127*(3), e639-e646. doi: 10.1542/peds.2010-1421
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology, 58*, 593-614. doi:

10.1146/annurev.psych.58.110405.085542

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002).

A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7(1), 83.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163-203.

Maehler, C., & Schuchardt, K. (2009). Working memory functioning in children with learning disabilities: Does intelligence make a difference? *Journal of Intellectual Disability Research*, 53, 3-10. doi: 10.1111/j.1365-2788.2008.01105.x

Martin, R. P., & Dombrowski, S. C. (2008). Low birth weight and prematurity: effects on child developmental outcomes *Prenatal exposures: Psychological and educational consequences for children* (pp. 29-45): Springer.

Mathiasen, R., Hansen, B. M., Anderson, A.-M. N., & Greisen, G. (2009). Socio-economic achievements of individuals born very preterm at the age of 27 to 29 years: a nationwide cohort study. *Developmental Medicine & Child Neurology*, 51, 901-908. doi: 10.1111/j.1469-8749.2009.03331.x

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology*, 38(1), 115-142. doi: 10.1037//0012-1649.38.1.115

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *The Psychological Review*, 63(2), 81-97. doi: 10.1037/h0043158

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 40-48. doi: 10.1037//0021-

843X110.1.40

- 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, 49-100. doi: 10.1006/cogp.1999.0734
- Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence*, 37, 327-328. doi: 10.1016/j.intell.2009.04.005
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, 22(11), 1425-1433. doi: 10.1177/0956797611416999
- Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, 18, 46-60. doi: 10.3758/s13423-010-0034-0
- Morrison, A. B., & Chein, J. M. (2012). The controversy over Cogmed. *Journal of Applied Research in Memory and Cognition*, 1, 208-210. doi: 10.1016/j.jarmac.2012.07.005
- Möttus, R., Gale, C. R., Starr, J. M., & Deary, I. J. (2012). 'On the street where you live': Neighbourhood deprivation and quality of life among community-dwelling older people in Edinburgh, Scotland. *Social Science & Medicine*, 74(9), 1368-1374. doi: 10.1016/j.socscimed.2011.12.050
- Moutquin, J.-M. (2003). Classification and heterogeneity of preterm birth. *BJOG: an International Journal of Obstetrics and Gynaecology*, 110(20), 30-33. doi: 10.1046/j.1471-0528.2003.00021.x
- Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Developmental of

- executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology*, 34(4), 393-421. doi: 10.1080/87565640902964524
- Mulder, H., Pitchford, N. J., & Marlow, N. (2010). Processing speed and working memory underlie academic attainment in very preterm children. *Archives of Disease in Childhood Fetal & Neonatal Edition*, 95(4), F267-F272. doi: 10.1136/adc.2009.167965
- Narberhaus, A., Segarra, D., Caldu, X., Gimenez, M., Pueyo, R., & Botet, F. (2008). Corpus callosum and prefrontal functions in adolescents with history of very preterm birth. *Neuropsychologia*, 46(1), 111-116. doi: 10.1016/j.neuropsychologia.2007.08.004
- Nicolas, S., & Levine, Z. (2012). Beyond intelligence testig: Remembering Afred Binet after a century. *European Psychologist*, 17(4), 320-325. doi: 10.1027/1016-9040/a000117
- 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, 571-581. doi: 10.1017/S1355617707070725
- Nosarti, C., Rushe, T. M., Woodruff, P. W. R., Stewart, A. L., Rifkin, L., & Murray, R. M. (2004). Corpus callosum size and very preterm birth: relationship to neuropsychological outcome. *Brain*, 127, 2080-2089. doi: 10.1093/brain/awh230
- Nutley, S. B., Soderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science*, 14(3),

591-601. doi: 10.1111/j.1467-7687.2010.01022.x

Oberauer, K., Wilhelm, O., Schulze, R., & Suß, H.-M. (2005). Working memory and intelligence - Their correlation and their relation: Comment on Ackerman, Beier, and Boyle *Psychological Bulletin*, *131*(1), 61-65. doi: 10.1037/0033-2909.131.1.61

Orchinik, L. J., Taylor, H. G., Espy, K. A., Minich, N., Klein, N., Sheffield, T., & Hack, M. (2011). Cognitive outcomes for extremely preterm/ extremely low birth weight children in kindergarten. *Journal of International Neuropsychological Society*, *17*(6), 1067-1079. doi: 10.1017/S135561771100107X

Oskoui, M., Coutinho, F., Dykeman, J., Jette, N., & Pringsheim, T. (2013). An update on the prevalence of cerebral palsy: a systematic review and meta-analysis. *Developmental Medicine & Child Neurology*, 1-11. doi: 10.1111/dmcn.12080

Pearson Inc. (2011). Cogmed working memory training. Retrieved October 30, 2012, from <http://www.cogmed.com>

Pennell, C. E., Jacobsson, B., Williams, S. M., Buus, R. M., Muglia, L. J., Dolan, S. M., . . . Relton, C. (2007). Genetic epidemiologic studies of preterm birth: guidelines for research. *American Journal of Obstetrics and Gynecology*, *196*, 107-118. doi: 10.1016/j.ajog.2006.03.109

Pickering, S. J., Gathercole, S. E., & Peaker, S. M. (1998). Verbal and visuospatial short-term memory in children: Evidence for common and distinct mechanisms. *Memory & Cognition*, *26*(6), 1117-1130. doi: 10.3758/BF03201189

Platt, M. J., Cans, C., Johnson, A., Surman, G., Topp, M., Torrioli, G., & Krageloh-Mann, I. (2007). Trends in cerebral palsy among infants of very low birthweight (<1500 g) or born prematurely (>32 weeks) in 16 European centres: a database study. *Lancet*, *369*(9555), 43-50. doi: 10.1016/S0140-6736(07)60030-0

- Porter, T. F., Fraser, A. M., Hunter, C. Y., Ward, R. H., & Varner, M. W. (1997). The risk of preterm birth across generations. *Obstetrics and Gynecology*, *90*(1), 63-67.
- Postlethwaite, B. E. (2011). *Fluid ability, crystallized ability, and performance across multiple domains: A meta-analysis*. University of Iowa. Retrieved from <http://ir.uiowa.edu/cgi/viewcontent.cgi?article=2639&context=etd>
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717-731. doi: 10.3758/BF03206553
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879-891. doi: 10.3758/BRM.40.3.879
- Prins, P. J. M., DAVIS, S., Ponsioen, A., Brink, E. T., & van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychology, Behavior, and Social Networking*, *14*(3), 115-122. doi: 10.1089/cyber.2009.0206
- Pritchard, V. E., Clark, C. A. C., Liberty, K., Champion, P. R., Wilson, K., & Woodward, L. J. (2009). Early school-based learning difficulties in children born very preterm. *Early Human Development*, *85*, 215-224. doi: 10.1016/j.earlhumdev.2008.10.004
- Pyhala, R., Lahti, J., Heinonen, K., Pesonen, A.-K., Strang-Karlsson, S., Hovi, P., . . . Raikkonen, K. (2011). Neurocognitive abilities in young adults with very low birth weight. *Neurology*, *77*(23), 2052-2060. doi: 10.1212/WNL.0b013e31823b473e
- Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive

- function. In P. Rabbitt (Ed.), *Methodology of Frontal and Executive Function* (pp. 1-37). Hove: Psychology Press Ltd.
- Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science, 4*(1), 2-14. doi: 10.1207/S1532480XADS0401_1
- Raven, J. (1989). The Raven Progressive Matrices: A review of national norming studies and ethnic and socioeconomic variation within the United States. *Journal of Educational Measurement, 26*(1), 1-16. doi: 10.1111/j.1745-3984.1989.tb00314.x
- Raven, J. C. (1958). *Standard progressive matrices*. London: H.K. Lewis.
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K., Fried, D. E., Hambrick, D. Z., . . . Engle, R. W. (2012). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 141*, 1-21. doi: 10.1037/a0029082
- Reitan, R. M. (1971). Trail making test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33*, 575-581. doi: 10.2466/pms.1971.33.2.575
- Reitan, R. M., & Wolfson, D. (1992). A short screen examination for impaired brain functions in early school-age children. *Clinical Neuropsychologist, 6*(287-294).
- Rickers, T. J., AuBuchon, A. M., & Cowan, N. (2010). Working memory. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 573-585. doi: 10.1002/wcs.50
- Roberts, G., Lim, J., Doyle, L. W., & Anderson, P. J. (2011). High rates of school readiness difficulties at 5 years of age in very preterm infants compared with term controls. *Journal of developmental and behavioral pediatrics, 32*(2), 117-

124. doi: 10.1097/DBP.0b013e318206d5c9

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*, 83-92. doi: 10.1016/j.intell.2006.05.004

Romine, C. B., Lee, D., Wolfe, M. E., Homack, S., George, C., & Riccio, C. A. (2004). Wisconsin Card Sorting Test with children: A meta-analytic study of sensitivity and specificity. *Archives of Clinical Neuropsychology*, *19*, 1027-1041. doi: 10.1016/j.acn.2003.12.009

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2011). Basic information processing abilities at 11 years account for deficits in IQ associated with preterm birth. *Intelligence*, *39*, 198-209. doi: 10.1016/j.intell.2011.03.003

Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., . . . Jacobsson, B. (2007). A report: the definition and classification of cerebral palsy April 2006. *Developmental Medicine & Child Neurology*, *109*, 8-14. doi: 10.1111/j.1469-8749.2007.tb12610.x

Rosselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition*, *52*(2003), 326-333. doi: 10.1016/S0278-2626(03)00170-2

Ruegger, C., Hegglin, M., Adams, M., & Bucher, H. U. (2012). Population based trends in mortality, morbidity and treatment for very preterm- and very low birth weight infants over 12 years. *BMC Pediatrics*, *12*(17), 1-12. doi: 10.1186/1471-2431-12-17

Saigal, S., & Doyle, L. W. (2008). An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet*, *371*(9608), 261-269. doi: 10.1016/S0140-6736(08)60136-1

Saleem, T., Sajjad, N., Fatima, S., Habib, N., Ali, S. R., & Qadir, M. (2011).

- Intrauterine growth retardation - small events, big consequences. *Italian Journal of Pediatrics*, 37(1), 1-4. doi: 10.1186/1824-7288-37-41
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, 15(8), 511-514. doi: 10.1111/j.0956-7976.2004.00711.x
- Schetter, C. D., & Glynn, L. M. (2011). Stress in pregnancy: Empirical evidence and theoretical issues to guide interdisciplinary research. In R. Contrada & A. Baum (Eds.), *The Handbook of Stress Science: Biology, Psychology, and Health* (pp. 321-343). New York: Springer Publishing Company.
- Schmidt, B., Roberts, R. S., Davis, P. G., Doyle, L. W., Barrington, K. J., Ohlsson, A., . . . Tin, W. (2007). Long-term effects of caffeine therapy for apnea of prematurity. *The New England Journal of Medicine*, 357(19), 1893-1902. doi: 10.1056/NEJMoa073679
- Schweizer, K., & Koch, W. (2002). A revision of Cattell's Investment Theory: Cognitive properties influencing learning. *Learning and Individual Differences*, 13(1), 57-82. doi: 10.1016/S1041-6080(02)00062-6
- Senn, T. E., Espy, K. A., & Kaufmann, P. M. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology*, 26(1), 445-464. doi: 10.1027/s15326942dn2601_5
- Sharma, A., Sharma, P., & Sharma, D. (2011). Cognitive correlates of different academic subjects in school setting. *International Journal of Psychological Studies*, 3(2), 142-148. doi: 10.15539/ijps.v3n2p142
- Shaw, P., Gogtay, N., & Rapoport, J. (2010). Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories *Human Brain Mapping*, 31, 917-925. doi: 10.1002/hbm.21018
- Shaw, P., Malek, M., Watson, B., Greenstein, D., de Rossi, P., & Sharp, W. (2013).

- Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 74(8), 599-606. doi: 10.1016/j.biopsych.2013.04.007
- Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B. D., & Gouvier, W. D. (2010). The relationships of working memory, secondary memory, and general fluid intelligence: Working memory is special *Journal of Experimental Psychology: Learning Memory and Cognition*, 36(3), 813-820. doi: 10.1037/a0019046.
- Shipstead, Z., Hicks, K., & Engle, R. W. (2012). Cogmed working memory training: Does the evidence support the claims? *Journal of Applied Research in Memory and Cognition*, 1, 185-193. doi: 10.1016/j.jarmac.2012.06.003
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, 50(3-4), 245-276.
- Slagter, H. (2012). Conventional working memory training may not improve intelligence. *Trends in Cognitive Sciences*, 1-2. doi: 10.1016/j.tics.2012.10.001
- Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-292.
- Spearman, C. (1927). Amount of G in different kinds of education *The abilities of man: their nature and measurement* (pp. 199-221). Oxford, England: Macmillan.
- Sternberg, R. J. (1981). Testing and cognitive psychology. *The American Psychologist*, 36(10), 1181-1189. doi: 10.1037/0003-066X.36.10.1181
- Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), 6791-6792. doi: 10.1073/pnas.0803396105
- Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly* 47(1), 1-41. doi: 10.1353/mpq.2001.0005

- Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, *121*(1), 15-23. doi: 10.1037/0096-3445.121.1.15
- Studer, B. E., Jaeggi, S. M., Buschkuhl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2009). *Improving fluid intelligence - Single N-back is as effective as dual N-back*. Paper presented at the 50th Annual Meeting of the Psychonomic Society, Boston, MA.
- Stuss, D. T. (2011). Traumatic brain injury: relation to executive dysfunction and the frontal lobes. *Current Opinion in Neurology*, *24*(6), 584-589. doi: 10.1097/WCO.0b013e32834c7eb9
- Svensson, A. C., Sandin, S., Cnattingius, S., Reilly, M., Pawitan, Y., Hultman, C. M., & Lichtenstein, P. (2009). Maternal effects for preterm birth: A genetic epidemiologic study of 630,000 families. *American Journal of Epidemiology*, *170*(11), 1365-1372. doi: 10.1093/aje/kwp328
- Swanson, H. L. (2008). Working memory and intelligence in children: What develops? *Journal of Educational Psychology*, *100*(3), 581-602. doi: 10.1037/0022-0663.100.3.581
- Tabachnick, B. G., & Fidell, L. S. (2001a). *Analysis of Covariance Using Multivariate Statistics* (pp. 195-242). Boston: Pearson Education, Inc.
- Tabachnick, B. G., & Fidell, L. S. (2001b). Cleaning up your act: Screening data prior to analysis *Using Multivariate Statistics* (pp. 60-116). Boston: Pearson Education, Inc.
- Tabachnick, B. G., & Fidell, L. S. (2001c). Review of univariate and bivariate statistics *Using Multivariate Statistics* (pp. 33-59). Boston: Pearson Education, Inc.
- Talge, N. M., Holzman, C., Wang, J., Lucia, V., Gardiner, J., & Breslau, N. (2010).

- Late-preterm birth and its association with cognitive and socioemotional outcomes at 6 years of age. *Pediatrics*, *126*(6), 1124-1131. doi: 10.1542/peds.2010-1536
- Tamm, L., & Juranek, J. (2012). Fluid reasoning deficits in children with ADHD: Evidence from fMRI. *Brain Research*, *1465*, 48-56. doi: 10.1016/j.brainres.2012.05.021
- Taub, G. E. (2002). Moving beyond g: Linking theory, assessment, and interpretation in the measurement of intelligence. *The International Journal of Sociology and Social Policy*, *22*(11/12), 132-148. doi: 10.1108/01443330210790229
- Taylor, H. G., Klein, N., Anselmo, M. G., Minich, N., Espy, K. A., & Hack, M. (2011). Learning problems in kindergarten students with extremely preterm birth. *Archives of Pediatrics & Adolescent Medicine*, *165*(9), 819-825. doi: 10.1001/archpediatrics.2011.137
- Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. *Theory Psychology*, *4*(4), 525-547. doi: 10.1177/0959354394044004
- Thompson, J. M. D., Irgens, L. M., Rasmussen, S., & Daltveit, A. K. (2006). Secular trends in socio-economic status and the implications for preterm birth. *Paediatric and Perinatal Epidemiology*, *20*(3), 182-187. doi: 10.1111/j.1365-3016.2006.00711.x
- Thorell, L. B., Lindqvist, S., Bergman, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106-113. doi: 10.1111/j.1467-7687-2008.00745.x
- Tillman, C. M., Nyberg, L., & Bohlin, G. (2008). Working memory components and intelligence in children. *Intelligence*, *36*. doi: 10.1016/j.intell.2007.10.001
- Tomporowski, P. D., Davis, C. L., Miller, P. H., & Naglieri, J. A. (2008). Exercise and

- children's intelligence, cognition, and academic achievement. *Educational Research Review*, 20(2), 111-131. doi: 10.1007/s10648-007-9057-0
- Treloar, S. A., Macones, G. A., Mitchell, L. E., & Martin, N. G. (2000). Genetic influences on premature parturition in an Australian twin sample. *Twin Research*, 3, 80-82.
- Turner, D. C., Blackwell, A. D., Dowson, J., McLean, A., & Sahakian, B. J. (2005). Neurocognitive effects of methylphenidate in adult attention-deficit/hyperactivity disorder. *Psychopharmacology*, 178(2-3), 286-295. doi: 10.1007/s00213-004-1993-5
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28(2), 127-154. doi: 10.1016/0749-596X(89)90040-5
- United Nations Children's Fund, & World Health Organization. (2004). *Low birthweight: Country, regional and global estimates*. New York, NY: UNICEF.
- Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, 54, 68-80. doi: 10.1016/j.jml.2005.06.003
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17(6), 635-654. doi: 10.1080/09658210902998047
- Van Baar, A. L., Vermaas, J., Knots, E., de Kleine, M. J. K., & Soons, P. (2009). Functioning at school age of moderately preterm children born at 32 to 36 weeks' gestational age. *Pediatrics*, 124(1), 251-257. doi: 10.1542/peds.2008-2315

- van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, 35, 427-449. doi: 10.1016/j.intell.2006.09.001
- Vohr, B. R. (2010). Neuropsychological outcome: Cognitive and functional outcomes of children born preterm. In C. Nosarti, R. M. Murray & M. Hack (Eds.), *Neurodevelopment Outcomes of Preterm Birth: From Childhood to Adult Life*. New York: Cambridge University Press.
- Voss, W., Jungmann, T., Wachtendorf, M., & Neubauer, A. P. (2012). Long-term cognitive outcomes of extremely low-birth-weight infants: the influence of the maternal education background. *Acta Paediatrica*, 101(6), 569-573. doi: 10.1111/j.1651-2227.2012.02601.x
- Vygotsky, L. S. (1978/1997). Interaction between learning and development. In M. Gauvain & M. Cole (Eds.), *Readings on the Development of Children* (pp. 29-36). New York: W. H. Freeman and Company. (Reprinted from: *Mind and Society*, pp. 79-91, by Vygotsky, L., 1978, Cambridge, MA: Harvard University Press).
- Wechsler, D. (1991). *Manual for Wechsler Intelligence Scale for Children - Third Edition (WISC-III)*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1992). *Wechsler Individual Achievement Test (WIAT)*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Memory Scale - III: Administration and scoring manual*. San Antonio: The Psychological Corporation.
- Wechsler, D. (2003). *Wechsler intelligence scale for children - Fourth Edition (WISC-IV)*. San Antonio, TX: The Psychological Corporation.
- Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Ostensson, M. L., Bartfai,

- A., & Klingberg, T. (2007). Computerized working memory training after stroke - A pilot study. *Brain Injury, 21*(1), 21-29. doi: 10.1080/02699050601148726
- Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. latent structure. *Developmental Psychology, 44*(2), 575-587. doi: 10.1037/0012-1649.44.2.575
- Wilcox, A. J., Skjaerven, R., & Lie, R. T. (2008). Familial patterns of preterm delivery: Maternal and fetal contributions. *American Journal of Epidemiology, 167*(4), 474-479. doi: 10.1093/aje/kwm319
- Williams, J. E., & McCord, D. M. (2006). Equivalence of standard and computerized versions of the Raven Progressive Matrices Test. *Computers in Human Behavior, 22*, 791-800. doi: 10.1016/j.chb.2004.03.005
- Williams, P. E., Weiss, L. G., & Rolhus, E. L. (2003). WISC-IV technical report #2 psychometric properties *WISC-IV Technical Manual #2: The Psychological Corporation, a Harcourt Assessment Company.*
- Wollmann, H. A. (2009). Children born small for gestational age: Definitions and etiology. In W. Keiss, S. D. Chernausk & A. Hokken-Koelega (Eds.), *Small for Gestational Age: Causes and Consequences*, (Vol. 13, pp. 1-10). Switzerland: Karger.
- World Health Organization. (2014, November 2014). Preterm birth, from <http://www.who.int/mediacentre/factsheets/fs363/en/>
- Wyatt, J. (2010). The changing face of intensive care for preterm newborns. In C. Nosarti, R. M. Murray & M. Hack (Eds.), *Neurodevelopmental Outcomes of Preterm Birth: From Childhood to Adult Life* (pp. 17-29). New York: Cambridge University Press.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity

of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459-482.

Yuan, K., Steedle, J., Shavelson, R., Alonzo, A., & Oppezzo, M. (2006). Working memory, fluid intelligence, and science learning. *Educational Research Review*, *1*, 83-98. doi: 10.1016/j.edurev.2006.08.005

Appendix A



Research Ethics Office
Division of Research and Development

www.murdoch.edu.au

Thursday, 25 August 2011

Dr Helen Davis
Murdoch University

Dear Helen,

Project No.	2009/131
Project Title	Project K.I.D.S (Kids Intellectual Development Study): an umbrella application for Murdoch's proposed program of research associated with UWA's longitudinal research on the cognitive, social and emotional development of children

On behalf of the Murdoch University Human Research Ethics Committee, I certify that this project is renewed until **31 August 2012**, subject to any conditions listed below. This approval is effective ONLY with respect to the project as described in the original application and any subsequent amendments that have received approval.

As a condition of the approval of your human research ethics application you are required to report immediately anything, which might affect ethical acceptance of your project's protocols, including:

- Adverse effects on subjects
- Proposed changes in the protocols
- Unforeseen events that might affect continued ethical acceptability of the project.

Kind Regards,

Dr. Erich von Dietze
Manager of Research Ethics

cc: Angela Corts del Castillo, Jessicka Pokuncinski, Kai Van Der Linden, Helen Hoi Lam KO and Trish Kennedy

HREC Renewal Receipt 100111

CRICOS Provider Code: 001253
ABN 61 616 369 313

Appendix B

Participating School: Letter of Approval**Research title:****Can we improve children's fluid intelligence through working memory training?**

I, _____ (Principal's name) of

_____ (School's name) have read the research information sheet. Any questions about the research have been answered.

I agree to distribute information sheets and consent forms to students of Year 2 in our school. However, participation in your research will be the decision of the children and their parents.

We understand that all information is treated as confidential and will not be released by the investigator unless required to do so by law, where such circumstances will be explained.

We agree that research data gathered for this study may be published, provided participating children's names or other information that might identify the children, are not used.

My signature below indicates that I have understood the information provided.

School Principal's signature:

Date:

Chief Investigator:

Helen Ko (Student researcher)

Dr Helen Davis (Principal supervisor)

Dr Corinne Reid (Co-supervisor)

This study has been approved by the Murdoch University Human Research Ethics Committee (Approval 2011/079). If you have any reservation or complaint about the ethical conduct of this research, and wish to talk with an independent person, you may contact Murdoch University's Research Ethics Office (Tel. 08 9360 6677 or e-mail ethics@murdoch.edu.au). Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Appendix C

Participant Invitation Letter**Research title:****Can we improve children's problem solving through working memory training?**

Dear Participant and Parents,

We invite you to participate in our research study looking at the effects of computer memory training on children's memory and intelligence. This study is part of my degree in Doctor of Psychology at Murdoch University, supervised by Dr Helen Davis and Dr Corinne Reid.

Nature and Purpose of Study

Research has shown that some kinds of memory training may not only improve people's memory, but also their ability to solve unfamiliar problems. We would like to know whether this training has any benefits for children in Year 2.

What would I be asked to do?

All children who participate will be asked to complete some tasks and puzzles involving memory and thinking. These tasks will take about 30-40 minutes and children will complete them individually at home, at a time convenient to you. Children will then be put into one of two groups. One group of children will be asked to play an animal game on computer for 15 minutes each day for 20 days (excluding Saturdays and Sundays). The game will be set up online for children to access from their home computer, and we will arrange initial training of the program to include the parents and participating child. Parents will be asked to assist in monitoring the child's progress. The second group of children will not be given the game to play immediately. After 20 days, all children will complete again the tasks and puzzles involving memory and thinking. At the end of the study, participating children who did not play the game will be given the opportunity to play.

Do I have to take part?

No. Parents and children are completely free to say yes or no. The research team will respect your decision whichever choice you make, and will not question it.

What if I wanted to change my mind?

If you say yes, but then you or your child want to stop participating, that's OK. Just let me know and you can stop at any time.

What will happen to the information I give - is it private and confidential?

The answers given on the tasks and scores achieved on the animal game will remain strictly confidential and will be seen only by the investigator and supervisor. We will record names so that we can keep track of different children's progress during the study, but as soon as a child completes the study, his or her name will be removed from all data and no one will be able to identify them. Any other contact information provided to us will also be shredded as soon as the child completes the study.

At the end of the study in December 2012, a summary of the results will be sent to you and the school. This will also be made available on the School of Psychology Research Results website:

http://www.psychology.murdoch.edu.au/researchresults/research_results.html

No individual children will be identified.

Is this research approved?

The research has been approved by Murdoch University Human Research Ethics Committee.

Who do I contact if I wish to talk about the project further?

If you would like to talk about the project further or have any concerns and questions, you are welcome to contact us either by email or by phone (see below).

OK – so how do I become involved?

If you **do** want to be a part of this project, then please do not hesitate to contact me by email: hko.research@gmail.com or phone: 0406143208. We will then arrange a time convenient for you and your child to begin assessment.

This letter is for you to keep.

Yours sincerely,

Helen Ko
(Student researcher)
Email: hko.research@gmail.com
mbl: 0406143208

Dr Helen Davis
(Principal supervisor)
h.davis@murdoch.edu.au
tel: 9360 2859

Dr Corinne Reid
(Co-supervisor)
corinne.reid@uwa.edu.au

This study has been approved by the Murdoch University Human Research Ethics Committee (Approval 2011/079). If you have any reservation or complaint about the ethical conduct of this research, and wish to talk with an independent person, you may contact Murdoch University's Research Ethics Office (Tel. 08 9360 6677 or e-mail ethics@murdoch.edu.au). Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Appendix D

PARENT CONSENT FORM

Research title: Can we improve children's problem solving through working memory training?

I _____ (insert name of parent/guardian) and

_____ (insert child's name) have read the research information sheet. Any questions we have about the research have been answered.

We agree to participate in the study and understand that it will include completing memory and thinking tasks and may also include 20 days of playing an online computer game. However, we know that we may change our minds at any time and withdraw from the research project without having to provide further explanations.

We understand that all information is treated as confidential and will not be released by the investigator unless required to do so by law, where such circumstances will be explained to me.

We agree that research data gathered for this study may be published, provided my child's name or other information that might identify my child, is not used.

My signature below indicates that we have understood the information provided.

Parent /guardian's signature:

Date:

Phone Number (H): _____ (M): _____

Email:

Chief Investigator:

Helen Ko (Student researcher)

Dr Helen Davis (Principal supervisor)

Dr Corinne Reid (Co-supervisor)

This study has been approved by the Murdoch University Human Research Ethics Committee (Approval 2011/079). If you have any reservation or complaint about the ethical conduct of this research, and wish to talk with an independent person, you may contact Murdoch University's Research Ethics Office (Tel. 08 9360 6677 or e-mail

Appendix E

School-aged Participants' Consent Form**Research title: Can we improve children's problem solving through working memory training?**

My name is _____ (write your own name).

My parents have explained the study to me and all my questions have been answered.

I would like to take part in this memory game study.

I am happy to do some puzzles before and after I start playing the online game for the next 20 days (not Saturdays and Sundays).

I am happy to play the game on my own and will try to give it my best shot.

I know that I can choose not to answer your questions if I do not want to.

I know that I can stop doing any of the puzzles or playing the game if I do not want to.

Child's Signature

Chief Investigators:

Helen Ko (Student researcher)

Dr Helen Davis (Principal supervisor)

Dr Corinne Reid (Co-supervisor)

This study has been approved by the Murdoch University Human Research Ethics Committee (Approval 2011/079). If you have any reservation or complaint about the ethical conduct of this research, and wish to talk with an independent person, you may contact Murdoch University's Research Ethics Office (Tel. 08 9360 6677 or e-mail ethics@murdoch.edu.au). Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Appendix F



www.murdoch.edu.au

Research Ethics Office
Division of Research and Development

Wednesday, 20 July 2011

Dr Helen Davis
School of Psychology
Murdoch University

Chancellery Building
South Street
MURDOCH WA 6150
Telephone: 9360 6677
Facsimile: 9360 6686
human.ethics@murdoch.edu.au
www.research.murdoch.edu.au/ethics

Dear Helen,

Project No. 2011/079
Project Title Can we improve children's fluid intelligence through working memory training?

Your response in support of the above project was reviewed by the Murdoch University Research Ethics Office and was;

APPROVED

Approval is granted on the understanding that research will be conducted according the standards of the *National Statement on Ethical Conduct in Human Research (2007)*, the *Australian Code for the Responsible Conduct of Research (2007)* and Murdoch University policies at all times. You must also abide by the Human Research Ethics Committee's standard conditions of approval (see attached). All reporting forms are available on the Research Ethics web-site.

I wish you every success for your research.

Please quote your ethics project number in all correspondence.

Kind Regards,

A handwritten signature in black ink, appearing to read "E. von Dietze".

Dr. Erich von Dietze
Manager of Research Ethics

cc: Helen Ko

HREC Outright Approval Letter 100111

CRICOS Provider Code: 00125J
ABN 61 616 369 313

Appendix G



Government of Western Australia
Department of Education

Your ref :
Our ref : D11/0374096
Enquiries :

Ms Helen Ko
Flat 56 Room 1
Murdoch University Village
CANNING VALE WA 6155

Dear Ms Ko

Thank you for your completed application received 7 June 2011 to conduct research on Department of Education sites.

The focus and outcomes of your research project, *Can we improve children's fluid intelligence through working memory training?*, are of interest to the Department. I give permission for you to approach site managers to invite their participation in the project as outlined in your application. It is a condition of approval, however, that upon conclusion the results of this study are forwarded to the Department at the email address below.

Consistent with Department policy, participation in your research project will be the decision of the schools invited to participate, individual staff members, the children in those schools and their parents. A copy of this letter must be provided to site managers when requesting their participation in the research. Researchers are required to sign a confidential declaration and provide a current Working with Children Check upon arrival at the Department of Education site.

Responsibility for quality control of ethics and methodology of the proposed research resides with the institution supervising the research. The Department notes a copy of a letter confirming that you have received ethical approval of your research protocol from the Murdoch University Research Ethics Office.

Any proposed changes to the research project will need to be submitted for Department approval prior to implementation.

Please contact Ms Allison McLaren, R/Evaluation Officer, on (08) 9264 5512 or researchandpolicy@det.wa.edu.au if you have further enquiries.

Very best wishes for the successful completion of your project.

Yours sincerely

ALAN DODSON
DIRECTOR
EVALUATION AND ACCOUNTABILITY

22 June 2011

Appendix H



CATHOLIC EDUCATION

OFFICE OF WESTERN AUSTRALIA

DIRECTOR OF CATHOLIC EDUCATION

29 July 2011

Miss Helen Ko
School of Psychology
Murdoch University
90 South Street
MURDOCH WA 6150

Dear Helen

**RE: CAN WE IMPROVE CHILDREN'S FLUID INTELLIGENCE
THROUGH WORKING MEMORY TRAINING?**

Thank you for your completed application received 25 July 2011, whereby this PhD will examine whether training children with an 'animal adaptive span' working memory task will improve children's working memory, and if this training will transfer to fluid intelligence.

I give in principle support for the selected Catholic schools in Western Australia to participate in this valuable study. However, consistent with CEOWA policy, participation in your research project will be the decision of the individual principal and staff members.

Responsibility for quality control of ethics and methodology of the proposed research resides with the institution supervising the research. The CEOWA notes that Murdoch University Human Research Ethics Committee has granted permission for the duration of this research project (Project Number: 2011/079).

Any changes to the proposed methodology will need to be submitted for CEOWA approval prior to implementation.

The focus and outcomes of your research project are of interest to the CEOWA. It is therefore a condition of approval that the research findings of this study are forwarded to the CEOWA.

Further enquiries may be directed to Tanya Davies at davies.tanya@ceo.wa.edu.au or (08) 6380 5379.

I wish you all the best with your research.

Yours Sincerely

Ron Dullard