Executive Function Development: A Comparison of Monolingual and Bilingual Children in Ireland

Claire Stephens BSc

Supervisors: Dr. Judith Wylie & Prof. Carol McGuinness

A thesis submitted to Queen's University, Belfast, in fulfilment of the thesis requirement for the degree of Doctor of Philosophy (PhD)

School of Psychology September 2013

Abstract

Researchers continue to deliberate on the areas of cognition affected by the bilingual experience. Recent studies have focused on the executive function (EF), an area showing enhanced skills for younger and older bilinguals. However, research has yet to decipher how EF advantages develop through middle childhood. Furthermore, few studies have examined the impact of successful forms of bilingual education (BE), such as immersion education (IE) on children's EF development. In most current studies, socioeconomic status (SES) has been controlled for and few have investigated its effects on disadvantaged or low-SES bilinguals.

To further explore these issues, two studies were conducted as part of this thesis. The first study employed a longitudinal design to examine the effects of bilingualism on mid- and low-SES childrens' EF development. A number of EF tasks suggested to tap children's specific and unified functions were assessed over a 3-year period. The study had three testing phases (each approximately 1 year apart) and tested children between 8 and 12 years old. At Time 1 of the longitudinal study, 147 participants were recruited and the retention rate was 96% across time. Results indicated that children's EF development improved as a function of age and that the IE experience was successful in enhancing certain EFs such as the unified EF component but not others, such as inhibitory control (IC). While SES played an important role in children's EF development, the aspects most affected were non-verbal IQ and English language skills. Developmental findings demonstrated the benefits of longitudinal research and the potential for EF improvements in bilinguals with increased exposure and proficiency in the second language (L2).

The second study examined how language proficiency (LP) and experience with the L2 (Irish) affected children's EF skills by comparing 19 children from Gaeltacht or Irish speaking areas of Ireland with monolingual and IE participants at Time 3. Results indicated that careful consideration must be taken when comparing children from unique linguistic environments such as the Irish Gaeltacht as socio-political factors may impact on children's language experience and subsequently their EF performance. Despite these issues, children from the mid-SES Gaeltacht group outperformed all other groups on the unified EF task. Implications of findings are discussed in relation to bilingualism, EF and Irish IE research.

Acknowledgments

First and foremost I must thank an Chomhairle um Oideachas Gaeltachta & Gaelscolaíochta for funding this research as well as all the children who volunteered to take part in the studies, along with their parents, teachers and principals for kindly letting me come back each year and who made the process so enjoyable.

To my supervisor, Dr. Judith Wylie who has guided me through the process and who has challenged many of my assumptions, thank you. To Dr. Gerry Mulhern and Professor Carol McGuinness: I believe your ideas and suggestions greatly improved this research.

To all the staff in the school of psychology, without whom this research could not have been carried out or completed, particularly Martin and Paul and all the hard-working staff in the workshop. To Holly, Lynne, Lynda, Judith, Bernie and everyone in administration who helped out whenever I was stuck and who even kept me fed during the final stages.

I would like to thank a number of people who helped me to remain grounded in my research as well as reminding me of why I began this process and what the future can hold: (Dr.-to-be) Anita Durso, Dr. Jenny Maguire and Paul Quinn. Speaking to each of you maintained my motivation and passion for psychology and I am so grateful for the experiences that each of you has given me, thank you.

To the most amazing group of friends, who encouraged me to laugh, ‘keep swimming’ and not to lose sight of the light at the end of the tunnel. To the Dublin ladies, there are too many of you to name but you know who you are, the go-to people when you need to laugh. To all in the Queen’s Surf Club who helped maintain my sanity by encouraging me to get in the water and who never gave up asking when I was coming out. To my psychology girls: Anita, Carmel, Kristen and Stacey, no words for you guys but we all know we’re in it for the long-haul. To the Taters crew, I fully believe that making it through our year together set us up for anything and let’s hope there are more adventures to come. Finally, to a group of people who were constant supports over the last few years and who understood like no others could. I am so grateful that we all stuck it out together and I will never forget your kindness: Ali, Aoibheann, Eimear, and Steph, thank you.

To my extended family, the Carrolls, Eogans, Hopkins', Morrisons' and Stephens' whose words of wisdom and support were always welcome and especially to Dr. John Morrison who inspired this journey in the first place and who has demonstrated the power of good research. To the Lowry family who took me in, fed and watered me during the final stages as well as providing some of the home comforts that I missed along the way…here's to many more nights by the Chimera to come…

Finally, to the four people who supported me through to the final full stop. To Tom, who managed to put up with the dramas and late-night panics even from 10, 553 miles away (I counted) and who failed to be deterred by any of it, thank you and thank you for coming back (to Clom 2014). To Annette, Brian and Laura who reminded me every day of what is important and who have been editors, PAs, friends, secretaries and bank managers all in one. ☺ I will never be able to express my appreciation to you four in particular but I'm determined to find a way.

Grá mór xxx

TABLE OF CONTENTS

CHAPTER ONE – BACKGROUND TO BILINGUALISM AND IMMERSION EDUCATION

CHAPTER TWO - EXECUTIVE FUNCTION AND BILINGUALISM

CHAPTER THREE - BILINGUALISM AND COGNITIVE DEVELOPMENT

CHAPTER FOUR - METHODOLOGY FOR THESIS STUDIES

CHAPTER FIVE - THE IMPACT OF BILINGUALISM AND SOCIOECONOMIC STATUS ON CHILDREN'S EXECUTIVE FUNCTION DEVELOPMENT

APPENDICES

List of Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AIC	Akaike Information Criterion
ASD	Autism Spectrum Disorder
AWMA	Automated Working Memory Assessment
BE	Bilingual Education
BIC	Bayesian Information Criterion
CC	Creature Count
CE	Central Executive
CFA	Confirmatory Factor Analysis
EFA	Exploratory Factor Analysis
ELS	Extra Learning Support
ERPs	Event-Related Brain Potentials
EF	Executive Function
IC	Inhibitory Control
IGC	Individual Growth Curve
IE	Immersion Education
L1	First Language
L2	Second Language
LG	Language Group
LP	Language Proficiency
LSB	Low-SES Bilingual Group
LSM	Low-SES Monolingual Group
LSN	Low-SES Native Group
LTM	Long-Term Memory
M	Mean
MSB	Mid-SES Bilingual Group
MSM	Mid-SES Monolingual Group
MSN	Mid-SES Native Group
OW	Opposite Worlds
PE	Practice Effects
PFC	Prefrontal Cortex
PPVT	Peabody Picture Vocabulary Scale

RSPM	Raven’s Standard Progressive Matrices
RT	Reaction Time
SD	Standard Deviation
SES	Socioeconomic Status
SEM	Structural Equation Modelling
SIP	Speed of Information Processing
STM	Short-Term Memory
SW	Same Worlds
TEA-Ch	Test of Everyday Attention for Children
TH	Threshold Hypothesis
ToH	Tower of Hanoi
TMT	Trail Making Test
VS	Visuospatial
WCST	Wisconsin Card Sorting Test Working
WM	Working Memory
WMTB-C	Working Memory Test Battery for Children

List of Tables

List of Figures

OUTLINE OF THESIS AND SUMMARY OF CHAPTERS

The research component of this thesis is the culmination of two experimental chapters, one a longitudinal and one a cross-sectional design. As a number of issues were covered within this thesis including the developmental nature of executive function (EF), bilingualism, immersion education (IE) and socioeconomic status (SES), the first three chapters provide an overview of the current literature in some of these areas. An outline of the chapters within the thesis is provided below.

Chapter One provides a background to the literature and issues within bilingualism and IE research. Topics including definitions of bilingualism and the success of IE programmes are discussed. A brief overview of the Irish immersion and the Gaeltacht cases in particular are outlined in this chapter.

Chapter Two introduces the EF, an aspect of cognitive development which has received a growing level of interest as an indicator of children's academic and lifelong success. A number of different models have been proposed to conceptualise the EF and the models chosen for this thesis are outlined in this chapter. Issues surrounding EF measurement and definitions are discussed as well as the developmental trends of specific and unified functions through middle childhood.

Chapter Three brings together topics from Chapter One and Chapter Two by outlining the areas of cognitive development thought to be hindered and advanced by the bilingual experience. Theories for why bilingualism should affect cognitive development and the EF in particular are outlined.

As this thesis used mostly standardised assessments to assess children's EF development, Chapter Four outlines the methodology for each task as well as the procedures and designs for the longitudinal and cross-sectional studies.

Chapter Five outlines the first study in this thesis: a longitudinal study with monolinguals and bilinguals from low and mid-SES backgrounds. In order to examine the impact of bilingualism on specific as well as unified EFs, the study looked at children's EF development over a three year period. Issues of SES in particular are discussed in this chapter as well as findings and implications for future research.

The second study in this thesis is outlined in Chapter Six and looks more closely at issues of language proficiency (LP). In this study, bilinguals in IE were categorised according to their levels of Irish productive vocabulary to examine whether these categorisations had any effect on EF outcomes at Time 3. A third language group was recruited in this study to further assess the impact of language demographics on EF skills by comparing children from an Irish speaking, Gaeltacht area of Ireland with monolinguals and IE bilinguals at Time 3. Findings are discussed.

Finally, Chapter Seven discusses the findings from Chapter Five and Chapter Six as well the implications of this research for bilingualism and IE research. Limitations and the potential for future research are also discussed.

CHAPTER ONE
BACKGROUND TO BILINGUALISM AND IMMERSION EDUCATION

1.1. INTRODUCTION

This chapter aims to provide a background and context to the experimental chapters of this thesis (Chapter 5 and 6). A discussion of the issues facing historic and modern bilingualism research is provided in sections 1.2. and 1.3, particularly surrounding the on-going debate on how best to define and assess bilingual groups. As this thesis recruited participants from immersion education (IE) primary schools in Ireland, a background to IE and a brief discussion of how previous sections relate to the Irish context is covered in section 1.4.

1.2. DEFINING BILINGUALISM

One of the first challenges when studying bilingualism is deciding how to define language groups (LGs). Defining *who* may be bilingual and deciding *the degree of fluency* in both the first (L1) and second (L2) languages necessary to be considered bilingual is part of this challenge. Although many studies assign participants to monolingual or bilingual groups, language is not simply a dichotomous variable and there can be great variations in levels of fluency within both the L1 and L2 (Carlson & Meltzoff, 2008; Kroll & Bialystok, 2013). From speaking to listening, from reading to writing, these variations in language production (encoding) and reception (decoding) may be fundamental when assessing LG differences (Baker, 2011; Dopke, 1992, Weinreich, 1953). Equally, it is very difficult to find an individual who has not been influenced by a language beyond their mother tongue (Bialystok, 2001). Ellen Bialystok has discussed these points eloquently in her book, *Bilingualism in Development* (2001) where she comments that, (p. I): "…*the idea of an "uncontaminated" monolingual is probably a fiction*".

Despite evidence suggesting that over half of the world's population are now bilingual (Grosjean, 1999), traditional opinion has plagued the scientific literature on bilingualism. Mackey (1967, p. 11) writes: "…*bilingualism, far from being exceptional, is a problem which affects the majority of the world's population".*

Here, bilingualism is described as a ‘problem’, an issue to be dealt with and understood rather than a societal norm. Although being monolingual is portrayed as normative, worldwide, bilingual and multilingual speakers outnumber monolingual speakers (Adesope, Lavin, Thompson, & Ungerleider, 2010; Tucker, 1998). Furthermore, it has been reported that two-thirds of the world’s children grow up within some form of bilingual environment (Crystal, 1997) although each of these children will differ in how they have experienced and interacted with either language. Such differences may foster unique attitudes towards the L1 and L2 and, more crucially, different levels of competence in each of their languages (Baker, 2011; Bialystok, 2001; Cummins, 2000). Furthermore, any changes in how a child experiences each of their languages may significantly alter the nature of their bilingualism as language context can strongly influence language ability (Cummins, Baker, & Hornberger, 1991, Butler & Hakuta, 2004). For instance, Baker (2003) suggested that a lack of exposure to the L2 outside of the academic environment (e.g. bilingual education) can affect children’s L2 development, as the language becomes associated with the school curriculum but not children’s peer culture (see also Ó Duibhir, 2009).

Consequently, the process of determining whether to define an individual as ‘bilingual’ or as a ‘second language learner’ is critical to any study of bilingualism (Genesee, Paradis, & Crago, 2004) and continues to be debated within current research. This debate has generated a wide-variety of linguistic cut-offs with a range of definitions proposed by researchers through the years (Baker, 2011; Bialystok, 2001; Romaine, 1995) although definitions tend to fit within two categories of classification: *strict* (section 1.2.1) *and lenient* (section 1.2.2).
Bloomfield’s early definition (1933) for bilingualism is an example of a ‘strict definition’. He maintained that to be classified as bilingual an individual must be fully fluent in two languages. However, deciding what it means to be ‘fully fluent’ is unclear due to the range of linguistic skills that can be differentiated, e.g. receptive versus productive competence. In contrast, Macnamara (1996a, 1967) gave a very ‘lenient definition’ for bilinguals, describing them as people who possess at least one of the four language skills (speaking, writing, listening, reading) in an L2 to a minimal degree. Under his definition, a person who is brought up as a native (L1) English speaker but can also read a little in French (L2) would be classified as

bilingual. Understanding how researchers choose to define their bilingual groups is crucial when drawing conclusions from their results and can limit the generalizability of their findings.

1.2.1. Strict definitions for bilingualism

There is a clear division between strict and the lenient researchers who attempt to define bilingualism. Bloomfield's (1933) suggestion that individuals need to have acquired native-like control in two languages is an example of a more stringent view. Subsequently, much of the literature has adopted a term to describe those who are equally proficient in two languages to a presumed native degree, known as '*balanced bilinguals*' (Albert & Obler, 1978) and researchers have striven to recruit such participants in their studies. However this definition can be extremely limiting as it is challenging to recruit bilinguals who display complete linguistic balance or equal proficiency in both their L1 and L2 as most, if not all will have a preference or *dominance* in one or other of their languages and will display larger productive or receptive vocabularies in that language (Argyri & Sorace, 2007; Baker, 2011; Dornic, 1980). Furthermore, these strict definitions often lack clarity. What is required for 'native-like' abilities and how may someone be classified as having native-like proficiency in one language or another? After all, there can be varying levels of linguistic skill and competence even within monolingual populations (Bialystok, 2001).

1.2.2. Lenient definitions for bilingualism

To compensate for rigid definitions, some researchers have chosen to adopt broader conceptualisations for their bilingual groups. Grosjean (1989, 2010) described a bilingual as someone who has the ability to function in each language according to his or her given needs, arguing that the emphasis of definitions should be on the regular use of both languages rather than the degree of fluency. Similarly, Haugen (1953) defined bilingualism as the point at which a speaker in one language can produce complete meaningful utterance in the other language and Abutalebi and colleagues (2007) defined bilinguals as people who use their two languages (a native first language or L1 and a learned second language or L2) in their everyday lives. A lenient definition was provided by Diebold (1964), who developed the term ‘incipient bilingualism’ to describe the initial state of contact between two languages. Under this terminology, any individual with minimal competence in an L2 could be classified under the bilingual umbrella.

Two of the most highly regarded linguists of the 20th century, Mackey and Weinreich proposed definitions that were deliberately vague and consequently raised as many questions as those they attempted to avoid (see Baetens Beardsmore, 1991). Mackey described bilingualism as, "…*the alternate use of two or more languages by the same individual*" (1967, p. 57). He remarked (1968) that the point at which the speaker of an L2 becomes bilingual is either arbitrary or impossible to determine. Similarly, Weinreich described bilingualism as, "…*the practice of alternatively using two languages*…" (1953, p. 5). Such generalised definitions provide no constraints on how well both languages must be comprehended and spoken or whether the level of bilingualism can be graded or compared according to common linguistic skills. Baker (2011) noted that the substitution of highly constrained definitions with ones, which are overly lenient, is not a satisfactory method for resolving this conceptual debate. Indeed lenient definitions may only create further ambiguity and imprecision when attempting to define bilingualism. However, if researchers do choose to impose excessively strict criteria for their language groups then individuals who nonetheless have been affected by their linguistic experience may be excluded from important research and valuable data may be lost (Bialystok, 2001).

One prevailing view is that bilingualism should be described as a matter of degree rather than a categorical variable (Grosjean, 1996; Butler & Hakuta, 2004; Kroll & Bialystok, 2013). However, there are still no accepted standards for classifying children on the basis of an objective bilingualism scale (Carlson & Meltzoff, 2008) and it is unclear at what point along such a scale can an individual be classified as bilingual. In other words, assigning children to either a monolingual or bilingual group may undermine the complexity of bilingualism and diminish the intricacies of children’s linguistic experience and skill (Kroll & Bialystok, 2013). Furthermore, a range of environmental and linguistic factors may impact on the classification of bilingual groups. Some of these factors are discussed below.

1.2.3. Factors Influencing Definitions

Nowadays, it is recognised that bilingualism is a multidimensional construct, reflected by the multidisciplinary nature of researchers interested in this area (Diaz, 1983). Therefore, Baker (2011) suggested that bilingual samples should be defined according to a number of dimensions rather than a single classification. Some of these dimensions are discussed briefly in the following sections.

1.2.3.1. Degree of bilingualism

There is yet to be definitive evidence regarding the minimum amount of exposure necessary for a bilingual to acquire competence in the L2 similar to a monolingual's L1 skill (Hamers & Blanc, 2000). However, language proficiency (LP) can depend on how an individual performs in each of the four language skills: listening, speaking, reading and writing. While some may be able to listen to and understand the written language (receptive bilinguals) others may go on to develop the ability to write and speak in their L2 (productive bilinguals; Baker, 2011). Genesee (1994) noted that studies often report students' comprehension skills as more advanced than productive skills following IE programmes. An example of the contrast between productive and receptive bilingualism can be seen from the Gaelic/English bilingual communities of Scotland. Dorian (1982) noted that speakers often had minimal control of Scottish Gaelic (L2) yet had outstanding receptive competence. She referred to these individuals as '*semi-speakers'* (p. 26). Similar issues of productive language abilities have been raised in Irish-immersion studies which have reported that following IE, children's Irish productive skills are less advanced than comprehensive skills (Ó Duibhir, 2009).

Researchers have shown that there is a positive correlation between LP and level of exposure to each language (e.g. Thordardottir, 2011; Cummins, 1979; 2000). If a bilingual child has limited exposure to their L2 during their waking day (e.g. 10 to 20%) this will likely result in incomplete acquisition of that language and Thordardottir (2011) has suggested that bilinguals require at least 40% of waking hour exposure in a language if their vocabulary skills are to be comparable to monolinguals (assessed using standardised assessments if receptive and expressive vocabulary). Furthermore as bilingualism has been shown to foster the development of certain cognitive skills (see Bialystok, 2001; 2009 for review; Cummins 1978), a positive relationship between levels of LP and such skills should be expected (Bialystok, Craik, & Luk, 2012). However, few studies have employed a within-groups design to investigate this question (Duncan & DeAvila, 1979) and issues of LP and cognitive performance are relatively poorly understood (Kroll & Bialystok, 2013; see section 5.3.1). Furthermore, literature in the 1950s and 60s was still debating the best methods to objectively measure LP in quantitative terms. Issues of LP and methods of assessment will be discussed further in Chapter 6 of this thesis.

1.2.3.2. Degree of difference between the two languages

Like LP, information regarding the impact that degree of difference between the L1 and L2 has on children’s linguistic and cognitive abilities is also limited. However, noting the differences and/or similarities between the languages of a bilingual is important to identify the degree of cognitive effort required for the learner to develop language skills, for example morphology, grammar, and phonetics (Diaz, 1983). For instance, it may be easier for a L1 Spanish individual to acquire another Latin-based L2 such as Italian or French than it would be for an L1 Japanese individual (Bialystok, Majumder, & Martin, 2003). Bialystok and Feng (2011) discussed how a child whose languages are structurally similar (and share more cognates) may progress more rapidly in each of their languages. In their study, Chinese-English bilinguals did not display the same advantage as Spanish-English and Hebrew-English bilinguals on a decoding task, indicating that the relationship between languages contributed to task performance (Bialystok et al., 2005). However, in a study comparing Japanese-English and Spanish-English bilingual 6-year-olds, Barac and Bialystok (2012) found no group differences on tasks of executive function (EF) or cognitive control (see also Tao, Marzecová, Taft, Asanowicz, & Wodniecka, 2011).

1.2.3.3. Age of Second Language Acquisition

Research suggests that the cognitive and developmental processes affected by the bilingual experience may depend on the age at which a child is first exposed to their L2 and most studies which report cognitive benefits of bilingualism recruit bilinguals who have learned their L2 relatively early in life (Tao et al., 2011). Consequently, a distinction can be made between *early* and *late* bilinguals (Baetens Beardsmore, 1991). Early bilingualism refers to the acquisition of an L2 in the pre-adolescent phase of development and *simultaneous bilingualism* is used if a child is introduced to an L2 before the age of three (Genesee, Paradis, & Crago, 2004). Other terms used to describe early bilingualism include: *infant bilingualism* (Haugen, 1956), *bilingualism as a first language* (Swain, 1972), *ascribed bilingualism* (Adler, 1977). A second definition based on the age of acquisition is *second language learners* or *sequential bilinguals* who typically are exposed the L2 after establishing a degree of fluency in the L1 and usually acquire the L2 after the age of 3 (De Houwer, 2009; Genesee et al., 2004). Finally, *late bilingualism* or *achieved bilingualism* (Adler, 1977) is defined as the acquisition of an L2 after the age of approximately 11 years old. However, these age-limited cut-offs may be somewhat arbitrary and by themselves do not provide information on how successful individuals have been in acquiring a high standard of L2 proficiency (McLaughlin, 1984).

Some have argued that the lower the age at which the L2 is learned, the greater the chance of long-term proficiency and maintenance of that language, as environmental factors relating to the simultaneous bilingualism may promote high levels of L2 exposure and input (Tao et al., 2011). However, reviews by Marinova-Todd an colleagues (2000; see also Singleton, 2003; Singleton and Ryan, 2004) and Cenoz (2009) indicated that children should not be affected by the age of L2 acquisition and that, contrary to popular belief, younger learners are no more or no less competent than older learners in the L2. Hakuta (2001, p. 11-12) wrote that: "…*the evidence for a critical period for second language is scanty, especially when analysed in terms of its key assumptions. There is no empirically definable end point, there are no qualitative differences between child and adult learners, and there are large environmental effects on the outcome".* In other words, the hypothesis of there being biological constraints or a critical period for L2 development, dependent on age and which is turned off following puberty is not accurate (Mayberry & Lock, 2003; Tao et al., 2011). Instead, it is more reasonable to recognise that although age may correlate with levels of L2 LP; this is more likely to be as a result of social, psychological, and other contextual factors than maturational constraints. Furthermore, bilinguals who are exposed to a L2 between 2-9 years can achieve high levels of L2 skill (morphology and syntax) once extensive and systematic exposure in more than one context is available (e.g. classroom) and will display stage-like L2 development similar to the development of monolinguals acquiring their L1 (Kovelman & Pettitto, 2002; Kovelman & Pettitio, 2003).

1.2.3.4. Second language learning method(s)

There are many ways and methods to develop and learn an L2. L2 acquisition refers to the process of acquiring the L2 in a natural environment and outside of the formal instructional settings, while L2-learning refers to the process of formal education (McLaughlin, 1978). In the early bilingualism scenario, children can acquire their L2 through different methods e.g. one parent-one language, home versus school language, mixed language (see Baker, 2011). However, there is no strong empirical evidence to suggest that any of these differences should produce overt differences in children's cognitive development provided there are opportunities to produce and use the L2 in a variety of contexts (Kovelman & Pettitio, 2003). Instead, researchers should recognise and note how their participants have acquired or learned their L2 to gain a fuller understanding of the linguistic contexts of their bilingual groups.

1.2.3.5. Attitudes towards the second language

Experience with the L1 and L2 can vary greatly depending on factors such as social environments, politics, and religion (Butler & Hakuta, 2004). Early research argued that learning an L2 may threaten a person's self-esteem when the language has a minority status within the community, e.g. L1 Spanish bilinguals in the U.S.A. (Saer, 1923). Factors such as motivation and attitudes towards the L2 need to be understood when assessing the linguistic and cognitive success of the bilingual experience as attitudinal factors can impact the amount of exposure a child receives in their L2 (Ó Duibhir, 2009; Hickey, 1999; Kennedy, 2013; see section 6.1.3. for details of attitudes towards Irish in this research).

1.2.4. Defining the bilingual samples

Baker (1997) argued that although categorisation is necessary for the study of bilingualism, global definitions are useless as it is the purpose of the definition that dictates who is or isn't classified as bilingual. Bearing this and previous sections in mind, this research had to carefully consider how 'bilingual' groups were selected and defined. In the first experimental study (Chapter 5) bilingual children were recruited from full Irish-medium IE programmes in the Republic of Ireland. The aim of such programmes is to develop children's L2 (Irish) to a degree where they may be classified as bilingual, but at no cost to their L1 (English, Cummins, 1978). As the children were first tested at approximately 9 years of age most had been in the IE system for approximately 5 years. As a result, and in line with previous studies of IE (e.g. Nicolay & Poncelet, 2013a), children are described as *sequential bilinguals* and for simplicity will be described hereafter as *bilinguals*. In the second experimental study (Chapter 6) a second bilingual group was selected based on their higher levels of exposure to the L2. These participants grew up in Irish-speaking pockets of Ireland - *Gaeltacht regions* - where Irish was spoken at a higher frequency in the home and in the community than in IE groups. For this reason these participants were considered to be simultaneous bilinguals and for simplicity will be described as '*native bilinguals*'. In Chapter 6, when comparing native bilinguals with bilinguals in Chapter 5, the term *immersion* and *native bilinguals* will be used to avoid confusion.

1.3. BILINGUALISM: A BRIEF HISTORY

To provide a context to the bilingual debate, a brief discussion and overview of the history of bilingualism research is given in the following sections.

1.3.1. Early 20th Century

Although the study of bilingualism dates back to ancient times (Lewis, 1977), the subject only became a focus for psychologists around the turn of the twentieth century (Albert & Obler, 1978). Previously it was predominantly linguists who were

interested in examining differences between bilingual and monolingual individuals (e.g. Ronjat, 1913; Leopold, 1939-1949).
During the first half of the twentieth century researchers were mainly concerned by any negative associations between bilingualism and general intelligence (*g*) and children's scholastic achievement (Hakutu, 1986). Systematic studies of the relationship between bilingualism and intelligence began around 1920 and research during this period was largely carried out with minority, immigrant populations, raising questions regarding the generalizability of findings and negative biases against such populations. Much of the justification for bilingual education (BE) programmes hinged on the resolution of these early intelligence issues (Bialystok, 2001). Many researchers (e.g. Cummins 1976, 1977, Hakuta, 1986, Butler & Hakuta; 2004) have discussed how methodological and design flaws resulted in the majority of the early negative results favouring the monoglot experience. This section highlights some of the key papers from both the twentieth and twenty-first centuries, and explores the history of the field and where it stands at present.

The bilingual writings of the twentieth century can roughly be categorised into two opposing sides and many of the core textbooks on bilingualism structure their background chapter(s) according to these two opposing positions (e.g. Cummins et al., 2001; Romaine, 1995). At one end are the findings from studies, which maintain that bilingualism has overarching negative effects on the development of a child's cognitive abilities and general intellect. The majority of these studies are published around or prior to the 1960s. At the other end are the findings from researchers who believe precisely the opposite, that bilingualism may have positive effects for specific areas of childhod and adult cognitive development. The majority of these studies emerged following the landmark paper by Peal and Lambert (1962), which drew attention to the methodological flaws of earlier work.

Jespersen (1922, p. 148) gave a negative description of the bilingual experience, an opinion common around the time of publication: *"It is, of course, an advantage for a child to be familiar with two languages: but without doubt the advantage may be, and generally is, purchased too dear. First of all the child in question hardly learns either of the two languages as perfectly as he would have done if he had limited himself to one. It may seem on the surface, as if he talked just like a native, but he* *does not really command the fine points of the language…Secondly, the brain effort required to master the two languages instead of one certainly diminishes the child's power of learning other things which might and ought to be learnt".*

In contrast, Lambert (1977, p. 30) summarises the research supporting the positive advantages: *"There is, then, an impressive array of evidence accumulating that argues plainly against the common sense notion that becoming bilingual, that is, having two strings to one's bow or two linguistic systems within one's brain, naturally divides a person's cognitive resources and reduces his efficiency of thought. Instead, one can now put forward a very persuasive argument that there is a definite cognitive advantage for bilingual children in the domain of cognitive flexibility".*

In line with Jespersen (1922), many of the early writings regarded bilingualism as a negative experience and some of the most widely cited studies were from samples of children in Wales. Saer (1922, 1923), Smith (1923) and Lewis (1977) have all described some level of intellectual disadvantage for bilingual individuals. In 1922, Saer cited the significant superiority of monolinguals and concluded that children learning an L2 through play and association with other children suffer fewer cognitive disturbances than those who learn the L2 through schooling. In his large-scale study (1924) of 1,400 Welsh/English bilingual children aged between 7 and 14 in both urban and rural areas of Wales, Saer found a correlation between bilingualism and intelligence. Despite the fact that no differences were present between the urban language groups, he concluded that bilingualism resulted in lower intelligence due to the lower scores obtained by the bilingual children from rural areas. He also interpreted the language handicap as being the result of linguistic confusion, deeply affecting children's intellect and academic performance until the college years. Criticising both the design and conclusions of this study, Bialystok (2001) pointed out that these results were probably more reflective of the lack of opportunity for bilingual children within rural areas to use their L2 as they had fewer hours of contact with English both before and after school than had the urban children. Furthermore, a number of important socioeconomic (SES) factors were not considered in this study.

A further two studies from Wales suggested that bilingualism caused children to lag behind their monolingual peers in both verbal and non-verbal tasks. Smith (1923) found monolingual children made better progress in their power of expression, choice of vocabulary and accuracy of thought. He also found that monolinguals made greater improvements over a two-year period than bilinguals. Smith concluded that bilingualism was a ‘positive disadvantage’ for children’s development (Smith, 1923, p. 81). A number of researchers (e.g. Jones, 1933, 1959) criticised the work of both Smith and Saers, citing a range of methodological shortcomings of their work including issues of SES, language of testing (children tested in their weaker, L2 while monolinguals tested in their L1) and lack of proficiency controls. Following a re-analysis of Saer’s data, Morrison (1958) found that after parental occupation (a strong predictor of SES) was taken into account, no IQ differences existed between rural and urban bilinguals.

Studies around this time appeared to confirm the bilingual disadvantage in tests of verbal IQ (Barke, 1933; Barke & Williams, 1938), non-verbal IQ (Saer, 1931) and mathematical competence (Carrow, 1957; Manuel, 1935) and the case for bilingual education (BE) was not strong (Darcy, 1953; Lewis, 1959; Mead, 1927). Furthermore, monolingual children had been cited as being approximately 3 years ahead of bilingual children in various skills relating to verbal and non-verbal intelligence (Romaine, 1995). Although the academic outlook appeared bleak for bilingual children prior to the 1960s, a number of studies did not implicate bilingualism in the apparent paralysis of cognitive development. In contrast to negative findings, some linguists did find certain cognitive advantages in favour of children who had simultaneous exposure to two languages. For instance, Leopold (1939-1949) and Ranjat (1913) suggested that bilingualism promoted greater metalinguistic understanding, as children learned to distinguish between word sound and meanings sooner than monolinguals. One of the most influential linguistic examinations from the early twentieth century was by Leopold (1939-1949) who closely examined his daughter, Hildegard, and was one of the first to suggest positive cognitive implications of the bilingual experience. He claimed that by exposing his daughter to two languages, her cognitive development had been enhanced, enabling greater flexibility of thought and linguistic competence. He also suggested that the ability of bilinguals to more effectively separate word sound and meaning than monolinguals led to early awareness of the conventionality of words and the arbitrariness of language, or greater metalinguistic skill.

1.3.2. The Changing Tide for Bilingual Research

In Canada around the early 1960s a contrasting socio-political context framed an open and more balanced view of bilingualism. This enabled more controlled experimental groups and designs using English/French bilinguals (Romaine, 1995). In Canada, French was recognised as being equal to English in status and subsequent legislation strengthened its position in society (The Official Languages Act 1968-9; Bastarache, 1987). However, due to global misperceptions, many parents in Canada around this time were concerned about the negative effects that bilingualism might be having on their children. With attitudes towards bilingualism slowly beginning to change, the stage was set for two researchers in the field to reveal their findings. Until this time, psychometric studies had predominantly demonstrated negative associations between bilingualism and intelligence (Diaz, 1983; Hakutu, 1986, see 1.3.1). Following the publication of one of the most prominent bilingualism papers by two Canadian researchers (Peal and Lambert, 1962) it was appropriate then that Canada should go on to become a leading force in the changing of the tide of bilingual opinion and research, internationally and for years to come.

The study was entitled: *The Relation of Bilingualism to Intelligence*, and was conducted by Elizabeth Peal and Wallace E. Lambert (1962). The significance of this paper was that it unearthed advantages to being bilingual as well as highlighting numerous methodological shortcomings of previous studies. Peal and Lambert did not simply compare but controlled for, degree of language proficiency, recruiting bilinguals who had attained a relatively similar degree of skill in both languages, i.e. 'balanced bilinguals'. Their distinction (p. 6) between '*pseudo-bilinguals'* who have a facility in one language much more than the other and '*true bilinguals'* who have mastered both languages from an early age and have communicative facilities in both languages (L1 and L2) is what set their work apart from the rest. Guided by the writings of O'Doherty (1958), they believed that while pseudo-bilingualism may lead to cognitive delays, genuine bilingualism could be an asset to children's intellectual development (Diaz, 1983) and empirical findings were in stark contrast to the majority of previous research. Their group of 10-year old French-English bilinguals displayed superior non-verbal as well as verbal intelligence compared with matched monolinguals. They concluded that bilinguals had a more diversified structure of intelligence and greater mental flexibility. Crucially, their critique of earlier studies which had failed to match bilingual and monolingual participants on confounding variables including SES, L2 proficiency, assessment language, gender, age, and the urban-rural connection led to future studies having to justify and consider more closely, participant selection. They noted that such factors must have played a central role in why studies continued to find negative cognitive effects of the bilingual experience. The emphasis on 'balanced bilinguals' and this more controlled approach to research opened the floodgates for publications implicating bilingualism in positive cognitive outcomes. Their influence is evident, even today, as publishers demand that the highest methodological standards be upheld in any evaluation of LG effects. However, as Albert and Obler noted (1978; see also Kroll & Bialystok, 2013), the employment of such strict inclusion criteria often elicits a minimal understanding of the cognitive abilities in arguably the majority of bilinguals or those dominant in one of their two languages and recruiting such balanced individuals is a challenging task for any researcher interested in studying this area.

Although Peal and Lambert's study continues to be widely cited, the paper has not been free from criticism. Even the authors themselves have raised questions over the comparability of the two groups recruited in their preliminary work (Romaine, 1995). For instance, researchers may anticipate a bilingual disadvantage for receptive L1 vocabulary due to the lower frequency of L1 and L2 use by bilinguals (e.g. Bialystok & Feng, 2011; Oller & Eiler, 2002; see section 3.3.). As Peal and Lambert's bilinguals were matched with monolinguals for vocabulary knowledge, this may have automatically placed bilinguals at an advantage, raising questions of comparability (Ben Zeev, 1977). Anisfeld (1964) decided to test whether or not Peal and Lambert's bilinguals were more intelligent at the outset than monolinguals. Following a re-analysis of their data and the matching of groups for general IQ (Kuhlman-Anderson IQ test) she found positive bilingual effects still emerged for Raven's Progressive Matrices, a test considered high on '*g*'. She did not, however, correct for other questionable selection criteria (e.g. receptive vocabulary).

1.3.3. The 1960s to Date

Following the publication of Peal and Lambert's (1962) paper, studies investigating the positive impact of the bilingual experience began to emerge and research also witnessed a shift from general assessments of cognitive ability to more specific effects of bilingualism on cognitive development (Anisfeld, 1964; Ianco-Worrall, 1972; John, 1970; Peal & Lambert, 1962). Between the 1960s and 80s a wide-range of papers were published evaluating the influence of bilingualism on a number of cognitive and linguistic measures including symbolic development (Ianco-Worrall, 1972), metalinguistic awareness (Cummins, 1978), concept formation (Bain, 1974; Liedtke & Nelson, 1968) and divergent-thinking skills (Carringer, 1974; Torrance, Wu, Gowan, & Allioto, 1970). The positive effects emerging around this time were in both the verbal and non-verbal domains (Hamers & Blanc, 2000), which created a challenge later on for researchers attempting to draw theoretical conclusions and causal relationships from these findings. One area of cognitive skill that continued to display an advantage was "cognitive flexibility". Theory suggested that this skill may be advanced in bilinguals as a result of having to navigate and control two languages in the mind (Balkan, 1970; Cummins, 1976). Although there was little agreement or understanding of how best to conceptualise flexibility of thought, it was believed that it accounted for improved bilingual performance on a wide range of verbal and spatial tasks (Diaz, 1983).

Ben Zeev (1977a) decided to test the concept of cognitive flexibility. His study compared a group of 96 Hebrew-English bilingual and monolingual children, ranging in age from 5-8 years, brought up in middle-class Jewish families with at least one parent having high academic or professional qualifications. Intelligence was not controlled and, unlike Peal and Lambert's study, vocabulary was not used as a criterion variable for bilingualism. Instead, bilingual balance was tested using a translation test. His findings revealed a bilingual advantage in understandings of syntactic structures as well as a more in-depth knowledge of syntax. Ben Zeev postulated that as bilinguals navigate through the verbal complexities of two languages in early life, they develop a heightened facility for seeking out rules and determining how they are required in varying circumstances. For example, in a symbol substitution task, children had to substitute words in a sentence according to the experimenter's instructions. The bilingual advantage found here was only evident in measures directly related to linguistic and verbal ability; no group differences were observed on the Raven's Progressive Matrices. Ben Zeev (1977b) noted that during testing, bilinguals appeared to approach tasks in an analytic fashion, demonstrating more attention to both the structure and details of the task as well as being more sensitive to feedback from the experimenter. By having to develop an awareness of structural similarities and differences between two languages, bilinguals may progress their analytic strategies at a faster rate than monolinguals, and transfer these skills to other domains and cognitive tasks.

Following a culmination of such positive papers, the global outlook for bilingualism was slowly beginning to change as it began to be recognised as a phenomenon of cultural enrichment as well as a valuable intellectual and societal asset. Throughout the 1960s to 80s more interest in bilingualism was generated in cognitive psychology circles. Researchers such as Jim Cummins and Ellen Bialystok began to recruit participants from bilingual countries such as Canada to investigate the theoretical underpinnings of this experience. From the 1990s onwards, researchers began to take a more top-down approach to the studying this field, exploring how specific areas of cognitive and brain development are affected by the introduction of an L2 (Baker, 2011). With a wealth of studies now available, researchers have begun to identify the areas of cognition which can be positively or negatively affected by the bilingual experience. For example, some have identified that bilinguals may be delayed on many forms of verbal tasks compared with monolinguals (e.g. Kaushanskaya & Marian, 2007; Rosselli et al., 2000; Portocarrero et al., 2007; Gollan & Acenas, 2004, see section 3.3.). On the other hand, bilinguals tend to display a cognitive advantage when faced with tasks requiring a high degree of cognitive control or conflict resolution (e.g. Bialystok, 1986; 1988, 2009; Bialystok & Majumder, 1998; Costa et al., 2009; Kroll & Bialystok, 2013; Zelazo, Frye & Rapus, 1996, see section 3.4.). These control advantages are believed to be a result of the bilinguals' strengthened executive function (EF) skills, developed as a result of having to manage two active language within the brain (Bialystok, 2009; Carlson & Meltzoff, 2008; Costa, Hernandez & Sebastian-Galles, 2008, see section 3.4.1.). The next chapter, Chapter 2 will examine the EF in more detail and Chapter 3 will discuss how aspects of cognition are positively and negatively affected by the bilingual experience.

A limitation of previous research is that the majority of studies use cross-sectional and correlational analyses, therefore, causation between bilingualism and cognitive skills cannot be identified in most cases (Diaz, 1983). Subsequently, researchers still struggle to disentangle whether bilingualism fosters greater cognitive development or whether the more naturally capable children develop to become the most proficient or balanced bilinguals (Cummins, 1978). One method of overcoming this issue is by employing a longitudinal design to investigate LG effects in the development of cognitive skills over time. This was the method employed in Chapter 5 of this thesis and the following section outlines a number of previous bilingualism studies which have also employed longitudinal designs.

1.3.4. Longitudinal research on bilingualism

One of the benefits of using longitudinal analysis is that it enables researchers to examine the potential mechanisms which can elicit group changes over time. In terms of bilingualism, although age-group differences have been explored using cross-sectional studies, few have examined age-related cognitive changes in bilingual and IE students. By using a longitudinal design, this thesis hopes to uncover some of the potential mechanisms eliciting EF differences between monolinguals and bilinguals across a three-year period.

Although they are in the minority, a number of studies have employed longitudinal methods to examine the cognitive correlates of bilingualism. Barik & Swain (1976) analysed IQ data collected over a 5-year period from French-immersion programmes in Canada. Using a repeated-measures design they found that bilinguals scored significantly higher than monolinguals at each time point. However, IQ differences existed from the first testing phase, therefore no group X time interactions were found and findings could not be attributed to children's participation in French IE programmes. Diaz (1985) also investigated the bilingual effect over time by dividing his Spanish-English bilinguals according to Cummins' principle of high and low proficiency (1977, 1978). Cummins' threshold hypothesis (TH, 1978) suggested that a 'threshold' level of proficiency exists, separating children who have gained enough proficiency in the L2 to experience cognitive gains, and children who have not. In contrast to the TH, Diaz found that in the low proficiency group, more cognitive gains were experienced over time (6 months). He suggested a new threshold hypothesis where positive effects of bilingualism were related to the initial effort required in understanding and producing the L2 rather than increasingly higher levels of bilingual proficiency. His sample also revealed that degree of bilingualism was confounded by SES factors. This finding is important in light of previous assumptions made by researchers failing to account for SES when comparing groups.

Hakuta (1984) also used longitudinal data over a three-year period to investigate the effects of non-balanced bilingualism on general cognitive skills including verbal (PPVT), metalinguistic (sentence judgment) and nonverbal abilities (Raven's CPM and Thurstone's Primary Mental Ability). Although a significant relationship between Raven's scores and bilingualism was found, the study only used correlational analysis and coefficients differed between measures and across time. Furthermore, adequate controls (e.g. SES, monolingual controls) were not in place in this study therefore findings were not generalizable to common bilingual populations. In a second longitudinal study, Hakuta and Diaz (1985) evaluated the effect of LP on children's cognitive performance. However, their longitudinal study had only two testing phases and assessed minority populations of children within submersion programmes in North America without any monolingual controls. Despite these limitations they concluded that the significant correlation found between bilingualism and nonverbal IQ (Raven's SPM) was the result of a bilingual advantage for nonverbal skills. It is evident that more thorough longitudinal research with adequate controls and sound analysis is needed within the field to begin to establish the relationship between bilingualism and cognitive ability.

More recently, Gersten and Woodward (1995) used a longitudinal design with four testing phases, each one a year apart to compare the academic abilities of children within two forms of BE: transitional and immersion education (IE). Although their findings indicated that transitional and IE are as effective as each other in developing children's L2 skills, the samples tested in their study still fell below population norms in English language skill. As a result, their findings might have been due to the use of English-minority language children within the U.S., an academic language setting which is known not to promote the L1 and L2 in equal measure (Baker, 2011) and raised questions regarding the generalizability of their findings to global IE programmes. Furthermore, a number of their conclusions were based on teacher and student feedback, therefore specific theories regarding the effects of IE on children's EF performance cannot be deduced from this study.

Finally, a study by Engel De Abreau (2011) used longitudinal data to assess children's working memory (WM) skills using span tasks. Findings revealed no advantage for bilinguals over time and a deficit in their verbal skills compared with monolingual peers. She concluded that the underlying mechanisms of fluid intelligence and WM may be different to the control required to choose between competing lexical responses in the bilingual brain and that simultaneous bilingualism in childhood does not impact on the development of WM skills. The study also called for researchers to examine the effects of sequential bilingualism on children's WM skills. These studies demonstrate the need for further longitudinal research exploring the impact of bilingualism on EF skills to help identify the developmental nature of the relationship between bilingualism and cognitive skills.

1.4. IMMERSION EDUCATION

An example of linguistic environments shown to foster bilingualism is early immersion education (IE) programmes. As children in the thesis were recruited from such programmes, the following sections examine issues relating to IE and Irish IE.

1.4.1 Defining Immersion Education

How research defines "bilingual education" (BE) depends largely on social contexts and, more specifically, the country of interest (Brisk, 1998). For instance, in the United States, BE refers to the education of those whose home language is not English. In contrast, Europeans would often refer to BE as a form of schooling where two languages are valued almost equally (Bialystok, 2001). An examination of the literature highlights the wide variety of BE programmes available worldwide (see Mackey's, *90 varieties of bilingual education*, 1970 for details). Baker (2011, p. 222) discussed the distinction between '*weak*' and '*strong*' forms of BE when evaluating the success of these programmes. Despite this variety, Johnson and Swain (1997) identified a core set of features that define the characteristics of *bilingual immersion* *education programmes* in particular. A number of variables can also vary between BE programmes and can impact on their success. Some of these variables will be discussed here.

Early literature identified negative implications associated with the bilingual experience and similarly, the BE experience came under attack during the early half of the twentieth century. Many of the stigmas surrounding BE developed from what Baker (2011) has described as '*weak forms*' of BE. For instance, researchers from the U.S. continuously recruited children from *submersion education programs*. In contrast to IE, submersion programs involve educating language minority children in monolingual schools and are often described as '*mainstreaming*' schools. Here, the language of education is the national or majority language (e.g. English in the U.S.) and not the home language of the pupil (e.g. Spanish). Children are taught alongside majority language pupils and are expected to use the majority language within the classroom. Therefore, the L2 is developed at a cost to students' L1 which, more than likely, is not recognised within school hours (however, certain children may receive extra language support). '*Structured immersion'* programmes (Brisk, 1998) differ from regular submersion in that children are educated alongside other minority language students and may initially be allowed a certain level of contribution from their home languages (L2; August & Hakuta, 1997). Skutnabb-Kangas (1987, 2000) highlighted the stress associated with learning an undeveloped language in mainstreaming schools as learning an L2 has high concentration demands and children must also comprehend information from curriculum content. This method of language submersion also places teachers under pressure, who must develop the LP of both their regular and minority students on top of teaching a packed curriculum to both populations (Carrasquillo & Rodriguez, 2002; Echevarria & Graves, 1998; Echevarria, Vogt, & Short, 2004; Faltis, 1997).

Alternatively, and in contrast to submersion approaches (or '*weak*' forms of BE), IE is an example of a '*strong form*' of BE. IE involves language minority children using their native, ethnic or heritage language as a medium of instruction within school and with the goal of full bilingualism (Baker, 2011). What sets IE programmes apart from submersion programs is that they recognise the importance of developing a child's L1 and aim to foster competence in both the L1 and L2 (Cummins, 1998).

This thesis recruited children in the Republic of Ireland being taught through their heritage language, Irish, but at no apparent cost to the majority language, English. Therefore, some have described the Irish immersion case as an example of *heritage language bilingual education* (e.g. Baker, 1993). There are many examples of heritage language education programmes worldwide e.g. Hawaii, Australia, Wales with the Irish case being similar to the Welsh system where the heritage language is protected and fostered alongside the development of the majority language, in these cases, English (Baker, 1993, 2000; Harris & Murtagh, 1999; Baker & Jones, 2000; Williams, 2003; Ó Mhurchú, 2003; Lewis, 2008). However, due to some of the criteria surrounding heritage language programmes (e.g. that children are predominantly from language minority homes) it may be more appropriate to describe the Irish case simply as a form of *immersion bilingual education* (for further details of Irish IE programmes, see section 1.4.3).

Like BE, IE is an umbrella term used to describe a range of programmes which vary in both their methods and principles. IE has also become a global phenomenon, popular in many countries including: Canada, Finland, Spain and Ireland. A number of distinctions differentiate types of IE. For instance, the age at which teaching in the L2 begins can vary between programmes. Some choose to begin L2 immersion from kindergarten or the infant stages (3-4 years old), known as *early immersion*. Others start the process later, at around 9-10 years old, known as *delayed* or *middle immersion*. Lastly, children may start immersion at secondary level or *late immersion* (Cloud, Genesee, & Hamayan, 2000). The amount of time children spend being immersed in the L2 also varies between programmes. Some schools commence with 100% immersion in the L2, reducing this (after two to three years) to 80% per week for the next three to four years, until they finish primary/junior schooling with approximately 50% immersion in the L2 per week and are classified as *total immersion* programmes. *Partial immersion*, on the other hand provides close to 50% immersion in the L2 throughout the infant and high school years.

In Canada, *early total immersion* has been the most popular programme, followed by late and then partial immersion. Similarly, in the Republic of Ireland, an *early, total or full-immersion* model is implemented in the majority of cases. Again, children are taught 100% of the curriculum through the children's L2 during the infant stages (3-6 years). Here, teachers speak the immersion language from the start and use verbal and non-verbal cues to help children understand what is being said. Learners may go through a 'silent period' in which they develop comprehension skills in the L2, using the L1 for expression (Johnstone, 2007). Research has shown the success of such IE programmes in which speakers become bilingual and biliterate and can achieve academically through two languages with no negative impact on the language and literacy development of their L1 (Cloud, Genesee, & Hamayan, 2000; Genesee, 1994). It has also been suggested that children may develop more positive attitudes towards their L2 by participating in IE programmes (Lindholm-Leary, 2001). Despite being immersed in the L2, children continue to develop their L1, as this environment is conducive to *additive bilingualism* (Cummins et al., 2001) and it is now widely accepted that early total immersion is one of the most successful types of BE for children's academic and cognitive development (Cummins, 1997, Johnstone, 2007).

Johnson and Swain (1997) provided an authoritative list, establishing a number of core and variable features necessary for successful IE programmes. The Irish immersion case has applied each of these criteria set out in the short summary below:

- L2 is the main medium of instruction: students not only learn the L2 to improve proficiency but also learn other academic subjects through the L2

- The immersion L2 curriculum parallels the mainstream L1 curriculum

- Overt psychological and other supports exist for the L1, both from parents and educators

- The IE programme aims for additive bilingualism or to add and strengthen a learner's language repertoire and to help them avoid falling between the two languages into a state of 'semi-lingualism' where they would be unable to use either language

- Exposure in the L2 is largely confined within the school walls as children generally do not live in an area where their immersion L2 is spoken widely in the community, though it may feature in the national media

- L2 learners generally enter with similar (and limited) levels of L2 proficiency

- Teachers are bilingual or native L2 speakers, enabling them to engage with children at all stages and to respond in the immersion language.

1.4.2. Immersion education outcomes

The term '*immersion education*' (IE) was first coined in Canada to describe a new form of BE developed in Quebec during the 1960s. French was being recognised as important for economic viability yet language proficiency (LP) in French among the population was low. This prompted a group of parents and educators in St. Lambert, Montreal to develop IE programmes as a way to improve French (L2) competencies of children without any cost to the English (L1) language development (Johnson & Swain, 1997). These programmes also carried out systematic evaluations of the consequences on children's language learning (Cummins, 1997) and using these French-Canadian samples, Peal and Lambert were able to publish their renowned study of the cognitive consequences of bilingualism (1962, see section 1.3.2.). While it is now widely accepted that IE, if implemented correctly, should elicit no adverse effects on children's cognitive development, Baker (2011) highlighted a number of key issues that have emerged in more recent IE studies. Four outcomes in particular may affect the success of IE programmes and should not be overlooked by researchers.

The first of these four outcomes is the development of the child's first language (L1). Researchers have debated the effects of IE on the development of children's L1 skill and some maintain that through IE, children will develop their L2 at a cost to their L1 (e.g. Macnamara, 1966; see section 1.4.3.). While it is not unreasonable to assume that intensive education in a child's L2 would be at a cost to the L1, most research now shows that children's L1 skill should be comparable to children educated purely in the L1 (Genesee, 1994). However, this has predominantly been a result of educating through the early total immersion model rather than other models of IE (Cummins, 1998). Some studies have shown that children in early total immersion do tend to progress more slowly than their monolingual peers in the L1 but only during the first 3 to 4 years (while the 80-100% L2 exposure is in place; Genesee, 2004). This lag, which has manifested through school tests of reading, spelling and punctuation, does not consider abilities and level of speech production and comprehension in the L2. Furthermore, as children are being intensely exposed to the L2 during the first few years of immersion, these results are somewhat unsurprising (Baker, 2011) and following approximately 6 years of education, students have been shown to catch up with their monolingual peers (Genesee, 2004).

The second of the IE outcomes noted by Baker (2011) is the development of the child's L2 skills. IE should provide children with a high level of receptive skill in the L2 rather than a high level of productive skill (Cummins, 1998; Genesee, 2004; Ó Duibhir, 2009). For instance, in the case of early IE programmes in Ireland, children tend to gain native-like proficiency in the L2 (i.e. Irish) at around 11 years old but only for receptive language skills (e.g. listening and reading, Johnstone, 1999). Levels of productive skill for grammar and vocabulary, in particular are not thought to reach the same degree of proficiency by this age and require further input outside of the school environment (Cummins, 2009; Harris et al., 2006; Kennedy, 2012; Swain & Johnson, 1997; Ó Duibhir, 2009). Subsequently, one of the limitations of IE is that children often fail to use their L2 beyond the school walls, limiting their L2 experience (Murtagh & van der Slik, 2004).

The third outcome is a student's ability to succeed in academic subjects through a dual-language experience. Some have questioned whether the IE experience can cause children to fall behind in other areas of academic achievement e.g. mathematics, science, history and geography, as learning through an L2 may add to the difficulty of such curricula. Evaluations from early total immersion students generally indicate no significant difference between them and mainstream monolingual children in academic ability (e.g. Cummins, 2005; Genesee, 2004). However, children in early IE programmes may display some lag, during the initial stages of learning or 'early total immersion' as L2 skills are not adequately developed to cope with complex curriculum material (Genesee, 2004). If academic assessments are conducted in the L2 and modified to account for the lack of L2 competence, IE students should perform equivalently to monolingual peers. But what are the effects for children who already have lower ability ranges, e.g. children with language impairment? In fact, research has shown that children who display a below average IQ score experience no adverse effects when educated through IE programmes (e.g. Brennan-Wilson, 2013; Bruck, 1982; De Courcy et al., 2002) and gain the additional skill of an L2 without any adverse costs, suggesting a benefit of the IE experience for children with these types of delays (Bruck, 1978, 1982; Brennan-Wilson, 2013).

Finally, the fourth outcome of interest is the influence of IE on children's attitudes and social adjustment. The most positive research outcomes comparing children's motivations, attitudes and study skills come from studies with early total immersion students (Baker, 2011). Parents of children in this type of programme generally express satisfaction with their academic achievements as well as their personal and social behaviour (e.g. in Ireland, Hickey, 1997). A study by Ó Muircheartaigh and Hickey (2008) found that classroom anxiety towards learning Irish was significantly higher in late versus early immersion students in Ireland and suggested a transitional programme for late IE students to address LP and anxiety issues. Overall, studies show that the early IE experience should, at least, have no adverse effects for children's social development and indeed, can be seen as a benefit for children's bicultural and social understanding (Kennedy, 2012).

It is clear that while there are many issues facing BE programmes, if implemented and carried out effectively, IE in particular has the potential to provide children with a range of benefits. Its success can also be seen through the recent expansion of the IE model on a global scale to countries such as: Australia, the Basque Country, Catalonia, Finland, Hungary, Hong Kong, Ireland, New Zealand, Singapore, and South Africa. The following section looks at the IE case in Ireland, where IE has become one of the fastest growing education sectors in the country and where the number of students attending IE schools has increased by almost 40% in the last two decades (Gaelscoileanna, 2013).

1.4.3. Ireland and immersion education

The Irish language is of Celtic origin and has been described an autochthonous or heritage language (Ó Duibhir, 2011). Its features are similar to Scottish Gaelic and more distantly; Welsh (Ó Murchú, 2003; Baker, 1993). As in countries which aim to restore heritage languages through immersion education (IE) programmes (e.g. Wales and Hawaii), the Republic of Ireland has embraced IE to help restore both the native identity of its population as well as to help prevent a complete language shift from Irish to English (Baker & Jones, 1998; Fishman, 1991). Regarding the Irish language, a Government of Ireland statement (2006) cites its aim to ensure that as many citizens as possible are bilingual in Irish and English and that *"…a high standard of all-Irish education will be provided to school students whose parents/guardians so wish".* The all-Irish education they refer to uses an additive IE model (Ó hAinifféin, 2008). As Irish IE schools are aware that most pupils will not speak Irish while at home, strict languages policies are often put in place to encourage Irish as a means of communication, particularly within the schoolyard (Ní Mhalaóin, 2005).

An early study by Stark (1940) concluded that early acquisition of a L2 at school does not necessarily weaken the L1 but may in fact strengthen it. This study recruited participants from an Irish-medium primary school in Dublin. However, the controls were weak and apart from age, Stark did not control for confounding factors such as SES, gender or proficiency levels (Darcy, 1953). Although studies were limited around this time, the work of John Macnamara (1966; Macnamara, Svarc, & Horner, 1976) left a lasting impression on the Irish public's outlook of IE. He claimed that IE had the potential to cause delays in children's academic performance and development of their home language (L1). His findings that IE children were 11 months behind monolingual peers in tests of arithmetic and IQ were widely cited as evidence for the debilitating nature of BE. His theories were also used in theoretical debates surrounding bilingualism and cognition. Based on Piagetian theory, Macnamara (1967) suggested that cognitive schema develop independently of language systems, therefore it does not make sense that bilingualism should have an impact on cognitive structures or processes. In writing this, Macnamara called into question all research in the field as well as leaving a lasting impression on parents in

Ireland who began to question the effectiveness of the Irish IE model (Cummins, 2001). Subsequently, IE schools declined until the 1970s, when opinion again shifted and the number of Irish-medium schools began to increase. During this period, Cummins published a new set of findings, which in contrast to Macnamara found no academic or cognitive deficit for pupils with IE programmes (Cummins, 1977a, 1977b; Cummins & Gulutsan, 1974). Cummins also heavily criticised Macnamara's work, which tested children through their weaker (L2, Irish) language and used a comparison group of British children, where cultural and curricula differences may have confounded effects. Furthermore, when IE children's performance was compared with monolingual children using a national Irish sample, no differences in English skills were obtained.

Since the 1970s, the number of Irish IE or Gaelscoileanna has grown from 17 in 1972 to 140 in 2009 (Ó Duibhir, 2009). The popularity of IE is also increasing in Northern Ireland. However, this research focused on the Republic of Ireland only, due to the unique cultural and political context surrounding the Irish language in Northern Ireland, considerations that are beyond the remit of this thesis. Approximately 5% of students receive their primary education through the medium of Irish and a further 2.5% attend Gaeltacht schools in Irish-speaking communities of Ireland (Máirtín, 2006). The majority of Irish immersion schools employ an early, total immersion policy for the first year of Junior Infants (P1, age 4-5 years), followed by the introduction of English language arts in Senior Infants (P2, age 5-6 years) representing 15% of instructional time. The remaining 85% of instruction is through the medium of Irish and this percentage remains constant. Students in IE come from predominantly English-speaking homes, with a small number (less than 3%) coming from Irish-speaking homes (Ó Duibhir, 2009). Similarly in this thesis, the majority of children in the IE groups were from English speaking homes, while in the Gaeltacht IE groups, a mixture of Irish and English was spoken in the majority of homes. Attrition within IE is not attributed as being a problem and the retention rate from year 1 to year 3 in this thesis was high at 96%.

Unlike the work of Cummins and Macnamara, most recent studies of Irish IE have focused on socio-political, pedagogical and attitudinal issues relating to children's progress in Irish IE rather than specific cognitive implications of the Irish IE experience. For instance, Hickey (1997, 1999, 2001, 2007, Kavanagh & Hickey,

2012) reported that children's level of Irish may improve through attending all-Irish preschools or 'naíonraí'. Hickey also cited the importance of parental involvement in a child's language experience and that communication through Irish within the home should be encouraged to consolidate the progress that children make within IE. Parsons and Lyddy (2009) assessed the most appropriate sequence and best practices for reading instruction in Irish IE primary schools. Their study found that the language of formal instruction for reading is not critical to later L1 reading success (see Cummins, 2001). Their study also found that children from the Gaeltacht had significantly better language skills in the L2 (Irish) than IE and monolingual children and that the delays found in the Gaeltacht group on English linguistic skills did not appear to be maintained following the introduction of formal reading instruction. In a more recent study, Kennedy (2012) assessed the EF skills from a cross-section of children at different ages and within Irish IE. Although his study found no significant cognitive advantages for the IE experience, children's EF skills were not delayed and therefore it was concluded that the IE experience is beneficial for children's academic and linguistic success. Ó Duibhir (2011) pointed out that Irish IE fosters the development of native-like receptive skills in students' L2 (Irish). However, students may not achieve as high a level of competency in their productive language skills. His study examined students' perceptions of their Irish competencies and found there was a mismatch between perceived ability and competency. All of these findings have given researchers and educationalists valuable insights into the effectiveness of Irish IE programmes and ways they can be improved. However, few studies have examined the impact of IE on children's cognitive development over time and how increased exposure to Irish impacts on children's EF skills in particular. The first study in this thesis (Chapter 5) aims to examine these issues further.

Alongside the study examining children's cognitive development within IE, a second facet of this research examined children from the Gaeltacht's cognitive performance at Time 3 of testing. The Gaeltacht is a unique linguistic environment and exists within small pockets of Ireland where Irish has been traditionally used as the means of communication within the community. However, the Gaeltacht is currently experiencing an attrition of spoken Irish. These issues and unique characteristics relating to the Irish Gaeltacht are briefly discussed below.

According to the 2006 Irish census 41% of the Irish population reported an ability to speak Irish. However, the numbers who speak Irish on a daily basis is significantly lower at 1.8%. A further 0.55% (n = 22, 515) reported speaking Irish on a daily basis as result of living within one of the '*Gaeltacht*' or Irish-speaking regions of Ireland. These regions of L1 and L2 (mixed Irish and English) speakers are not an exclusively Irish phenomenon as many other countries have similar regions aiming to develop bilingualism within predominantly monolingual cultures e.g. the Basque region of Spain, Welsh-speaking regions of Wales and French-speaking regions of Canada. Although the Irish Gaeltacht is still active, a number of cultural and societal changes have meant that the level of Irish spoken within these areas is highly variable and decreasing. In 2006, 71% of the Gaeltacht population over 3years of age reported being Irish speakers (n = 91, 862; Romaine, 2008), however only 40% reported speaking Irish on a daily basis (Punch, 2008), a fall from the 2002 Census figure of 54% (Census, 2002; 69, Table 34A).

While in the past it may have been reasonable to conclude that children raised within the Gaeltacht had native-like abilities or balanced bilingualism in Irish and English, children growing up in the modern Gaeltacht have shown varying degrees of proficiency in their L2 as the frequency of Irish spoken within the home reduces (Harris et al., 2006; Ó hIfearnáin, 2008). Furthermore, the number of L1 Irish children has dramatically decreased compared to other bilingual countries (e.g. Wales; Baker, 1993) as the majority of children now come from English-speaking homes (Ó Murchú, 2003). Harris et al. (2006) demonstrated this shift when parents from the Gaeltacht reported that they only occasionally spoke Irish with their children at home. Immigration from English-speaking regions and from other countries has also meant that Irish is spoken less within the Gaeltacht communities (Ó Riagáin, 1997). Despite this decline in spoken Irish, children living within the Gaeltacht are predominantly still taught through the medium of Irish and therefore should have, at a minimum, an equal level of proficiency to peers educated within IE outside of these regions (Ó'hIfearnáin, 2007; MacDonnacha et al., 2005; Parsons & Lyddy, 2009). Parsons and Lyddy (2009) commented that children taught within Gaeltacht schools can vary greatly in their levels of Irish fluency and that the numbers of children commencing primary school with high levels of Irish is decreasing (National Council for Curriculum and Assessment, 2007). Educational researchers have discussed how parents' concerns regarding their children's level of English skills have led to an increase in English being spoken within Gaeltacht homes. However, it is not believed that this technique is as effective as enrolling children in Irish immersion pre-school and primary education (e.g. Hickey, 1999). Furthermore, Cummins' developmental interdependency hypothesis (1979) suggested that L1 competence improves as a function of L2 competence once intensive exposure in the L2 begins. These issues highlight a number of factors that can contribute to the success of IE within the Gaeltacht regions of Ireland. This study examined the EF skills of children living in both the Gaeltacht and those attending IE in comparison with urban IE and monolingual participants to examine whether these contextual differences affect children's cognitive performance.

1.5. SUMMARY

This chapter discussed issues of bilingualism over the last century. One of the key areas of debate relates to how researchers define their bilingual groups. IE is a form of BE, shown to have benefits for children's linguistic, cognitive and cultural development so long as early total immersion programmes are implemented. These programmes aim to add an L2 to children's repertoire of skills at no cost to their L1 and are the IE programmes implemented in the Republic of Ireland. Issues within Irish IE in particular were also discussed. The chapter also noted that the Gaeltacht, or Irish-speaking regions of Ireland face their own challenges when raising bilingual children.

CHAPTER TWO
EXECUTIVE FUNCTION AND BILINGUALISM

2.1. INTRODUCTION

As in the bilingualism literature, researchers continue to debate how best to conceptualise executive function (EF). This chapter discusses some of the definitions and models of EF (section 2.2. and 2.3.) and section 2.4. examines how EF are represented within the brain. As this thesis included a longitudinal study of EF development (Chapter 5), section 2.5. discusses how EF components develop during middle childhood in particular. Sections 2.6. looks at how interventions have been used to improve EF performance as well as how EF skills are crucial in day-to-day life. Finally, limitations of EF models and methods of assessment are examined in section 2.7.

2.2. DEFINING THE EXECUTIVE FUNCTION

Those studying cognitive control and attention continue to debate how to appropriately conceptualise and define the executive function (EF). In spite of recent efforts to clarify the precise nature of EF, the concept continues to evolve (e.g. Burgess, 1997; Diamond, 2013; Smith & Jonides, 1999; Baumeister & Vohs, 2004). Monsell (1996, p. 93) described the nature of EF as one of the '*unsolved mysteries of the mind*'. When synthesising the definitions available it is important to consider the opinions of leading researchers within the field itself. Adele Diamond has pioneered research on the developmental nature of the EF (also known as executive control, attentional control or cognitive control) for many years. She defined the executive functions as (2013, p. 136): "…*a family of top-down mental processes needed when you have to concentrate or pay attention, when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible*". In other words, the principle of EF combines a number of mechanisms or functions that allow an individual to control and modulate their behaviour and cognitive processes. These regulating processes are therefore essential for human cognition (Miyake, Friedman, Emerson, Witzki, Howerter, and Wager, 2000). However, there is little agreement concerning how many functions make up this 'family' of mental processes or how best to group them into a coherent structure, identifiable within specific neural substrates of the brain.

EF has also been described as a multidimensional concept of behaviour. Together the EFs combine to serve a number of higher order cortical functions such as, goal-directed behaviour, attentional control, temporal organisation and planning (Anderson, 1998; Burgess, 1997; Duncan, 1986; Fuster, 1997; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Reitan & Wolfson, 1994, Zelazo & Frye, 1998). Behaviourally these processes enable us to control our impulses, to hold information temporarily in mind while performing another task, to think flexibly and to complete novel or complex day-to-day tasks (Huizinga & Smidts, 2011; Zelazo, Müller, Frye, & Marcovitch, 2003). Consequently, the EFs involve multiple neural networks within the brain, including the thalamus, basal ganglia, and prefrontal cortex or PFC (Bull, Espy & Wiebe, 2008; Fuster, 1997; Middleton & Strick, 2001, 2002; Pennington, 2002; Wilcutt et al., 2007).

Despite disputes regarding the areas of the brain required for executive functioning, the majority of neurological studies cite the use of the PFC during performance of EF tasks. Historically, the PFC had been thought to mature much later in a child's development, resulting in a research vacuum within the area of EF development during childhood (Hughes & Graham, 2002). Furthermore, standardised EF tasks have been purposefully designed to be highly challenging and many have assumed that such tests are inappropriate for use with children. The development of modern neuroanatomical and experimental techniques has brought about a surge in interest exploring the nature of the EF, predominantly from a clinical perspective.

Clinical research has found EF impairments played a central role in a number of developmental disorders including Attention Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD; Ozonoff, 1997). These deficits were also strongly correlated with impairments in the frontal lobes of the brain. Following these discoveries interest in the EF began to filter through to normative research and within the last 30 years, researchers have investigated the developmental trajectories of the EF across childhood. It is now thought that for a young child, utilising their

EFs is challenging due to the under-developed nature of EF during early childhood (e.g. Best, Miller, & Naglieri, 2011). However, EF abilities have still been observed in infants as young as 6 months (e.g. Diamond, 2006; see section 2.5.).

Despite the availability of numerous definitions, modern research continues to struggle with conceptualising and modelling the EF. One of the key differences in opinion within EF literature relates to how unified the individual EFs actually are. While some authors consider the EF as a single entity, conceptually equivalent to the general intelligence factor, *g* (e.g. Duncan, Burgess, & Emslie, 1995; Duncan, Emslie, Williams, Johnson, & Freer, 1996), others prefer to adopt a more fragmented view of the EF (e.g. Diamond, 2013; Miyake et al., 2000). In a fragmented or component-specific model, the EF is seen as a central system, made up of relatively independent sub-functions (e.g. Baddeley, 1996; Burgess, 1997; Lehto et al., 2003; Miyake et al., 2000; Robbins, 1998; Shallice & Burgess, 1996). The next section will discuss these theoretical accounts in more detail.

2.3. MODELLING THE EXECUTIVE FUNCTION

A number of models have been proposed to conceptualise the overarching structure of the EF. Fundamentally, these models differ in the number of functions they include. They also differ in the degree to which each function is separable and whether or not the EF should be viewed as a unitary construct. These debates are often mediated by the importance placed on the Working Memory (WM) mechanism or function. While some have suggested that WM is a function included as part of the overall EF system (Diamond, 2013) others have proposed that the EF is the central mechanism to the overarching WM system (e.g. Baddeley, 1986). This debate tends to divide researchers and most often segregates them according to their areas of interest (i.e. EF or WM).

WM researchers tend to implement Baddeley and Hitch's (1974) *working memory model* as the theoretical basis for their research while EF researchers tend to work from the EF model proposed by Miyake and colleagues (2000; see Diamond, 2013). Beyond this debate of modelling the EF, researchers have also debated the specific functions or EF components that should be included within each model. This review will briefly discuss two prominent EF models with emphasis on those that provide a sound theoretical basis for this research. However, an in-depth discussion of all models and EF debates is beyond the remit of this thesis.

The first model discussed is a classic model of the EF and is arguably one the most historically important models in cognitive psychology. The conceptualisation of the specific EFs also originated from this prominent cognitive framework. The multi-component model proposed by Baddeley and Hitch (1974, see also Baddeley, 1986, 1996, 2000), entitled the working memory (WM) model of cognition comprised of four main components: the phonological loop (dealing with speech-based, phonological information; Gathercole & Baddeley, 1994), the visuospatial sketchpad (concerning visual and spatial information), the episodic buffer (a limited capacity storage system which temporarily integrates information from the visuospatial sketchpad and phonological loop; Baddeley, 2000) and the central executive (CE; or the central control system, concerned with the control and regulation of cognitive processes, in other words, the EFs). The phonological loop, visuo-spatial sketchpad and episodic buffer act as “slave systems” to the CE which controls the allocation of resources across systems, organises multiple cognitive activities, and can revise the content of memory in light of new and relevant information (or updating; Bull, Espy, & Wiebe, 2008). Assessments of the CE involve testing processing and storage skills concurrently using tasks such as listening span (Daneman & Carpenter, 1980), counting span (Case, Kurland, & Goldberg, 1982) or backward digit span. Assessments of the visuospatial sketchpad and phonological loop include short-term memory (STM) tasks, requiring minimal resources from the long-term memory (LTM) store and no additional cognitive demands are present. Examples of STM tasks are the digit recall, word recall, corsi blocks, and visual-patterns task (Baddeley, 1992; Milner 1971; Pickering & Gathercole, 2001).

The work of Norman and Shallice (1980) and their model of attentional control heavily influenced Baddeley’s (1986) WM model. Despite the CE component of the WM model having arguably the most influence on cognitive achievement (i.e. controlling each of the slave systems), for many years it remained the least studied of the four components (Baddeley, 1996). Understanding the mechanisms comprised within the CE is essential as EFs are included within this system. As will be discussed, many researchers maintain that the EFs are independent of each other to some degree (e.g. Lehto et al., 2003, Miyake et al., 2000). Baddeley (1996) proposed that three distinct functions are ascribed to the CE: inhibition, shifting and updating. Inhibition or inhibitory control (IC) refers to an ability to suppress or ignore prepotent or automatic responses as well as ignoring distracting stimuli from the environment. Shifting (or switching) is the ability to shift between mental sets or rules and has also been described as the ability to think flexibly (cognitive flexibility). Finally, updating is the ability to store information while updating, manipulating and alternating information in WM. These EFs enable individuals to focus their attention while dealing with interference and/or conflict present in complex cognitive tasks, e.g. speech production (Miller & Cohen, 2001, Ye & Zhou, 2009).

With the development of modern neuropsychological techniques, researchers have been able to understand the nature of the CE or EF system as well as linking it, anatomically, to specific areas of the brain (i.e. the frontal lobes; Shallice, 1982, 1988). If research can map the neural areas active within the frontal lobes during the performance of complex cognitive tasks, this may lead to an understanding of the neural mechanisms governing the EF processes (e.g. Baddeley, 1996; Duncan, 1986; Shallice & Burgess, 1991, 1993). In spite of new experimental techniques, many still question the specification of EF. For example, should the EF be seen as a single, co-ordinated system serving a number of functions, or should it be defined as a cluster of largely autonomous control processes (Baddeley, 1996)?

In 2000, Miyake and colleagues published what has now become one of the most prominent models of EF. Prior to this model being developed, one of the major challenges for researchers was the lack of agreement regarding the relationships between component EFs (e.g. Best & Miller, 2010; Brocki, & Bohlin, 2004; Jurado & Rosselli, 2007). While Baddeley suggested that component EFs are included within the CE of the WM model (e.g. Baddeley, 1996) others have proposed that a WM component is present as part of the overarching EF model (Diamond, 2013; Lehto et al., 2003; Miyake et al., 2000; Willcutt et al., 2005). Distinguishing between component EFs is extremely challenging for researchers who have attempted to isolate each of the functions during experimentation. Furthermore, due to the nature of EF tasks, researchers often struggle to isolate component EFs during assessment (e.g. Chan, Shum, Toulopoulou, & Chen, 2008; Collette et al., 2005; see section 2.7.). As this thesis is predominantly a study of EF development, the latter conceptualisation (i.e. Miyake et al., 2000) is the model adopted hereafter.

The advantage of Miyake and colleagues model (2000) was that it brought together opposing views, describing the EF as a group of interrelated yet distinct components. They suggested that the EFs have both a unified and diverse nature and attempted to overcome issues of task impurity by employing a range of tasks to test each of the proposed EFs, using latent variable analysis to draw conclusions regarding conceptualisation. Guided by the work of Baddeley (1996), Miyake and colleagues also investigated the validity of inhibition, shifting, and updating (WM) as three separate EFs. Finally, they investigated the unified nature of functions using more complex EF tasks, requiring the collaboration of EF resources. Confirmatory factor analysis (CFA) was used over exploratory factor analysis (EFA), allowing their model of best fit to be guided by theory. Nine EF tasks were employed (3 per function) to test the relationship between the three functions of interest (i.e. inhibition, shifting and updating). Results indicated that a full three-factor model best fitted the data from the nine executive tasks and notably, EFs were not completely independent and shared some commonality between each other. The model (See Fig. 1), which identified a degree of commonality between EFs, was also a better fit for the data than the model assuming complete independence. Figure 1 shows that while individual EFs are separable, inhibition, shifting and WM also share a moderate degree of unity.

In the second part of their analysis, structural equation modelling (SEM) was used to examine how each of the three EFs contributed to more complex tasks, e.g. Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993) and the Tower of Hanoi (ToH; Humes, Welsh, Retzlaff, & Cookson, 1997). SEM results showed that specific EFs contributed differently to performance on these tasks, e.g. shifting contributed heavily to the WCST while inhibition contributed more to the ToH task. As a result of these findings researchers now carefully consider the potential EF skills tapped by their behavioural tasks. Furthermore, it may be the case that another function or strategy beyond those tested contributed to performance on complex EF tasks.

Since publishing their paper, Miyake and colleagues' model of EF has received extensive support with samples of both adults and children (e.g. Anderson, 2002; Asato, Sweeney, & Luna, 2006; Best, Miller, & Naglieri, 2011; Bull et al., 2004; Fisk & Sharp, 2004; Garon et al., 2008; Henry & Bettenay, 2010; Huizinga et al., 2006; Huizinga & van der Molen, 2007; Lehto et al., 2003; Rose, Feldman, & Jankowski, 2011). In line with this conceptual framework, the following sections describe each of the proposed executive sub-functions included as part Miyake et al.'s EF model: inhibition (or inhibitory control; IC), shifting (switching or cognitive flexibility) and working memory (WM or updating).

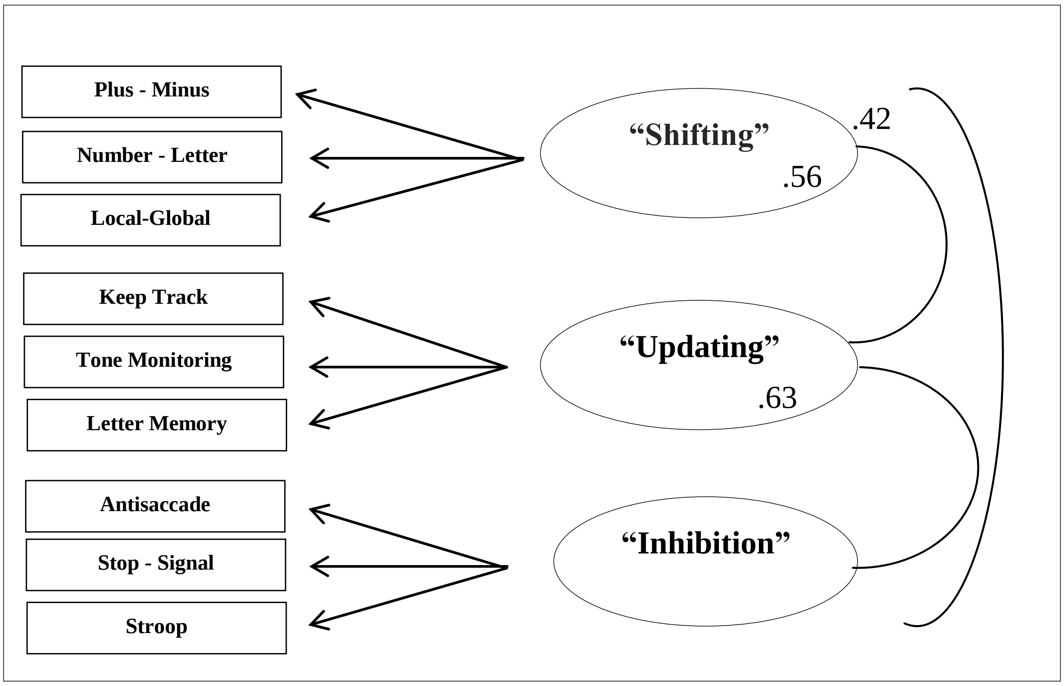


Figure 1 *Miyake and colleages (2000) three-factor model of executive functions*

2.3.1. Inhibition

One function suggested as being dissociable from other EFs is inhibition or inhibitory control (IC; Miyake et al., 2000) and involves being able to deliberately inhibit dominant and autonomic (or prepotent) responses. These autonomic responses may originate from internal predispositions or may be environmental distractors.

Therefore, IC enables individuals to focus their attention, behaviour, thoughts and emotions while avoiding habitual or distracting responses (Diamond, 2013).

There are a number of responses associated with IC as tasks tap a number of inhibitory processes (Nigg, 2000). Perceptually, IC requires ignoring certain stimuli and attending to alternative stimuli of choice. This ability to ignore and focus on task-relevant stimuli is also called interference suppression, selective/focused attention, attentional control or attentional inhibition (Posner & DiGirolama, 1998; Theeuwes, 2010). Response inhibition or cognitive inhibition refers to an ability to suppress prepotent responses or representations. Resisting these unwanted thoughts as well as proactive interference (PI) from responses previously made is what defines response inhibition (Postle et al., 2004). While there are often a number of alternative differentiations made between IC processes, e.g., simple versus complex response inhibition (Best & Miller, 2010; Garon et al., 2008) often these distinctions are vague and lack validity (Diamond, 2013; Garavan et al., 2002; Willcutt et al., 2005). Furthermore, despite conceptual disparities between types of IC being made, mapping IC differentiations onto brain areas has been extremely challenging for researchers (Diamond, 2013). Within the neural networks of the brain there appears to be a lack of difference between regions controlling the IC function, although IC continues to activate the frontal lobes of the brain (Jahanshahi et al., 1998; Kiefer, Marzinzik, Weisbrod, Scherg, & Spitzer, 1998). Friedman and Miyake (2004) used factor analyses to try and dissociate the different aspects of IC (e.g. resisting distractions versus inhibiting a prepotent response) and found no such discernible differences as both loaded strongly onto a single factor. Therefore, it is still reasonable to suggest that IC may be viewed as a single component of the EF.

Common tasks of IC include the Stroop task (Stroop, 1935; MacLeod, 1991), Flanker task (Eriksen & Eriksen, 1974; Mullane et al., 2009), antisaccade tasks (Luna, 2009; Munoz & Everling, 2004), delay of gratification tasks (Kochanska et al., 2001), and go/no go tasks (Cragg & Nation, 2008). The Opposite Worlds task used in this study (based on the more simple day-night task) is proposed to assess response inhibition as children must inhibit a prepotent response (e.g. say “1” when they see the number 2) while activating the correct verbal response (e.g. saying “2” when they see the number 1; Best & Miller, 2010; Gerstadt, Hong, & Diamond, 1994; Manly et al., 2001). The colour-word Stroop is thought to tap the interference suppression dimension of IC as children must choose between alternative cues during the incongruent trials (see 4.5.4. for details). Despite subtle differences in tasks, all involve inhibiting the tendency towards production of a more salient response. However, due to the debate of how best to conceptualise the EF and the difficulties isolating specific EFs during task performance some researchers maintain that it cannot be assumed that any task is a pure measure of IC (e.g. Roberts & Pennington, 1996, see section 2.7.).

2.3.2. Switching

Another function identified within Miyake and colleagues' EF model was cognitive shifting or switching (also known as cognitive flexibility). Switching refers to the ability to switch between multiple tasks, operations and mental sets (Monsell, 1996), and requires an ability to engage with and disengage from different aspects within a task (Lehto et al., 2003; Miyake et al., 2000). Behaviourally, it is important to remain flexible when performing complex tasks in day-to-day life and Diamond (2013) described an overlap between cognitive switching and creativity. Monsell (1996) identified that frontal lobe impairment is linked to difficulties switching between mental sets as patients struggle to switch between tasks (e.g. Luria, 1966; Stuss & Benson, 1986). Like with IC, some researchers have attempted to differentiate between switching skills, e.g. attention (perceptual switching) versus response switching (Rushworth, Passingham, & Nobre, 2005). Although evidence for such distinctions has been obtained from brain-imaging studies, for the purposes of this thesis, switching will be considered as one entity as it has been by other EF researchers (e.g. Brown & Bowman, 2002; Dias et al., 1996, 1997; Fox, Barense, & Baxter, 2003; Rogers, Andrews, Grashby, Brooks, & Robbins, 2000; Rushworth et al., 2005; Sylvester et al., 2003; Wager, Jonides, & Reading, 2004).

Using brain-imaging techniques, researchers have shown that switching incurs measurable temporal costs within the frontal lobes of the brain (e.g. Moulden et al., 1998; Rogers & Monsell, 1995). Although stimuli and operations vary among tasks, all switching tasks aim to have the switching requirement in common. Common switching tasks include dimensional change card sort task (Zelazo, Frye, & Rapus,

1996) number-letter task (Rogers & Monsell, 1995), global-local (Navon, 1977), creature count task (Manly et al., 1999; Milte et al., 2012). However, as with IC, issues surrounding the isolation of switching during task performance often arise within EF research (e.g. Lehto et al., 2003). Some have argued that switching between stimuli involves a degree of WM as well as IC of previous tasks operations (Miyake et al., 2000). Although it seems natural to assume that switching involves the inhibition of the previously activated set, in contrast to IC, switching tasks usually require participants to switch between two or more mental sets (Best & Miller, 2010).

2.3.3. Working Memory

Working memory (WM) or updating which is closely linked to WM refers to the manipulation and control of information while also storing information within subcomponents of the WM system: the phonological loop and the visuo-spatial sketchpad (Baddeley, 1986, 1996; Baddeley & Logie, 1999). It involves concurrently storing and manipulating information without relying on external aids or cues (Alloway, Gathercole, & Pickering, 2006; Baddeley & Hitch, 1994; Best & Miller, 2010; Goldman-Rakic, 1987; Huizinga et al., 2006; Smith & Jonides, 1999). There are two types of WM depending on the type of information being manipulated; non-verbal WM utilises the visuospatial sketchpad while verbal WM relies on the phonological loop.

WM is also distinct from STM. While STM is the process used to store information, WM has the added difficulty of having to manipulate and update the information being stored. Furthermore, STM and WM cluster onto separate factors following research with children and adults (Alloway et al., 2004; Gathercole et al., 2004). The neural substrates activated during WM tasks, as with switching and IC, are located within the frontal lobes of the brain and, more specifically, the PFC (D'Esposito et al., 1999; Smith & Jonides, 1999).
WM assessments often utilise some form of span task. Tasks involve storing information in the STM while manipulating that information in the WM. The backward digit span, for example, requires participants to remember numbers they have heard and to recall these numbers in reverse order (Lezak, 1983). The

manipulation of information (i.e. verbal numbers) requires activation of the WM function and subsequent PFC activity to represent numbers in reversed order (D'Esposito & Postle, 1999). As with all tests of specific EFs, WM tasks have faced criticisms surrounding their ability to isolate only the WM component. Diamond (2013) argued that the span tasks in particular may utilise aspects of IC to avoid uttering a number from a previous set. However, a number of studies support the dissociation between the WM and the IC component (e.g. Anderson & Spellman, 1995; Gernsbacher & Faust, 1991; Hasher, Stoltzfus, Zacks, & Rypma, 1991; Miyake et al., 2000; Zanto et al., 2011). Furthermore, the failure of tasks to isolate specific WM or IC mechanisms does not mean that they do not exist. The issue may be that the nature of any behavioural experiment makes it difficult to isolate specific EFs without input from other cognitive constructs. Of the EFs discussed, WM has also appeared to be the function most related to general intelligence (*g*) (e.g. Alloway & Alloway, 2010; Carpenter et al., 1990; Engle, Kane, & Tuholski, 1999; Friedman et al., 2011).

2.3.4. Unified executive function component

While most EF studies conclude that specific EFs can be dissociated, evidence for the unified nature of EF while performing complex cognitive tasks is becoming commonplace in modern research. The strength of Miyake et al.'s (2000) work was that it recognised both these diverse and unified aspects of EF. Furthermore, researchers have suggested that the high inter-correlation between EF tasks is evidence for this unitary EF component (e.g. Carlson, Mandell, & Williams, 2004; Diamond, Prevor, Callender, & Druin, 1997; Friedman & Miyake, 2004; Friedman & Miyake, 2012; Hughes, 1998; Hughes & Ensor, 2005; Kochanska et al., 1996; Lehto et al., 2003).

Complex cognitive tasks have been utilised to assess this unified EF component. Two examples are the Wisconsin Card Sorting Test (WCST; Heaton et al., 1993) and the Tower of Hanoi task (ToH; Simon, 1975), both thought to require the co-ordination of a number of executive sub-functions. In the WCST participants must monitor information from the examiner in order to decipher how to categorise cards as well as suppressing or inhibiting their previous responses while flexibly switching between category rules (Huizinga & Smidt, 2003; Miyake et al., 2000). Tasks such as the WCST and the ToH also rely heavily on novelty so that participants do not simply rely on LTM stores to perform well.

As day-to-day activities and decisions often require the consideration of numerous possibilities while selecting between environmental stimuli, it is likely that we will use this unitary EF on a regular basis rather than single EF components (Diamond, 2011; 2013). As a result, an understanding of this potentially higher-order, co-ordinating EF is extremely valuable and although identified, it is not conceptually defined within Miyake and colleagues' EF model (2000). While a number of theories have proposed a common mechanism across different EFs no agreement has been made regarding its theoretical underpinnings (e.g. Duncan et al., 1996, Duncan & Owen, 2000; Engle et al., 1999; Kimberg & Farrah, 1993).

Alternative models to Miyake et al.'s three-factor model of the EF may shed light on issues surrounding the conceptualisation of the unified EF component. For instance, a number of studies have not found the structure of the EF to be so definitive and precise. Wiebe et al. (2011) applied a similar three-factor CFA model of EF in 3-year old pre-schoolers. In contrast to Miyake and colleagues, their results favoured a unitary or domain-general EF solution rather than distinct EFs. Unified EF is also proposed as being the most representative model for the EF development of young children (Fuhs & Day, 2011; Wiebe, Espy, & Charak, 2008; Wiebe et al., 2011; Willoughby, Wirth, Blair, & Greenberg, 2010, van der Ven et al., 2012). Other studies have differentiated between some but not all EFs, e.g. distinction between WM and switching but not IC (Huizinga, Dolan, & Van der Molen, 2006; Van der Sluis et al., 2007) and for WM but not IC or switching (Van der Ven et al., 2012). One explanation for such discrepancies is that developmental changes may elicit conflicting results (Van der Ven et al., 2012). Perhaps the EF is unitary during early childhood but becomes more fragmented as children's EF develops over time (Wiebe et al., 2011), although a number of studies have found a three-factor structure of EF to be present in young children (e.g. Hughes, 1998; Senn, Espy, & Kaufmann, 2004). Conflicts may have also arisen through task selection as consensus is yet to be reached on how best to test children's EF, with studies employing a wide and varied range of EF tasks. With such lack of uniformity among studies, it is unsurprising that conflicting models of EF have emerged.

Recently, Miyake and Friedman updated their earlier model of EF (Friedman et al., 2006; Friedman et al., 2008; Miyake et al., 2000; Miyake & Friedman, 2012). Their 2012 model is similar to their original hypotheses of the EF being both uniform and diverse in nature, except that the specificity of IC as an influential EF is now called into question. In their updated model (Fig. 2), entitled the *unity/diversity framework*, a common *unitary* function includes influences from all three *diverse* EFs (Miyake and Friedman; 2012) although no unique variance is left for the IC function as it is strongly correlated with the common or unitary EF function/component. They proposed that the unified EF component relates to the active maintenance of task goals and goal-related information and added weight to their updated model with longitudinal twin data, which highlighted the stability of each of the proposed EF components over time (17 to 23 years).

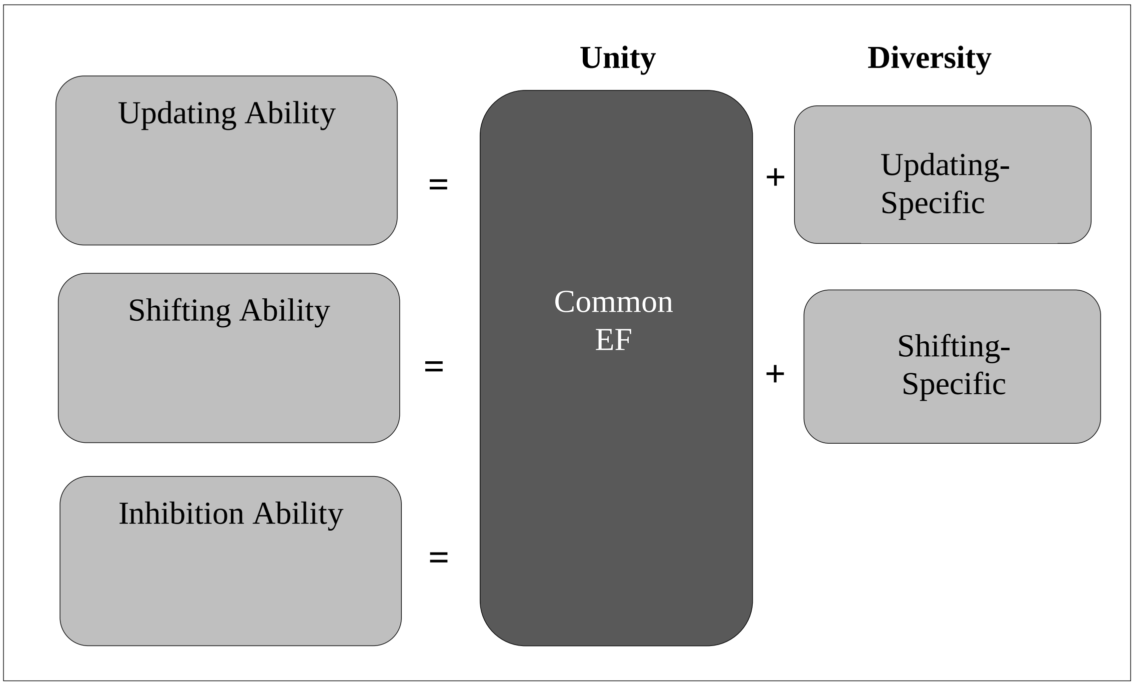


Figure 2 *Schematic representation of Miyake and Friedman's (2012) unity/diversity framework of EF*

2.4. EXECUTIVE FUNCTION AND THE BRAIN

The use of modern neuroimaging techniques alongside behavioural experiments has led to widespread agreement that the EFs are most active in frontal lobes of the brain and more specifically, areas of the prefrontal cortex (PFC) (Hughes & Graham, 2002; Knight & Stuss, 2002). This has been shown through PFC activation during performance on a wide range of cognitive tasks (Duncan & Owen, 2000). Furthermore, patients with damage to the PFC use poor cognitive strategies and display behavioural incoherence (Shallice & Burgess, 1991). An awareness of the neurological circuits within the brain can advance understanding of both behavioural and biological development (Nelson & Bloom, 1997). By applying innovative brain imaging techniques, e.g. Positron Emission Tomography (PET), Functional Magnetic Resonance Imaging (fMRI) and Electroencephalogram (EEG), research has been able examine brain functioning using non-invasive methods.

The PFC lies in the anterior area of the frontal cerebral cortex of the brain (Fuster, 2008) and is instrumental in the regulation of perception, thought, and behaviour (Knight & Stuss, 2002; Shallice, 2002). Specific and general EFs may be employed while performing complex and novel cognitive tasks and it is the PFC brain region that has been shown to be most engaged while performing such tasks. Subsequently, studies have shown that individuals who recruit the PFC while learning novel information perform better than those who do not engage this area of the brain (Duncan & Owen, 2000). On the contrary, when something is no longer new and has become an automatic behaviour (e.g. driving), engaging the PFC while engaging in such a task may not always be helpful. Due to the high mental energy required to engage the frontal lobes, using the PFC during tasks which should be automatic may even hinder performance, e.g. during sport (Chein & Schneider, 2005; Diamond, 2013). The maturation of the frontal lobes of the brain continues much later than other brain areas, beyond adolescence and into early adulthood (Levin et al., 1991; Segalowtiz & Davies, 2004).

As cognitive control or EF is related to the maturation of the PFC (Sowell et al., 2004; Luna et al., 2010), this relationship may explain the developmental trajectory of EF, which continues beyond that of other cognitive functions (Best & Miller, 2010; Lezak, 1995). Despite its protracted development, areas of the PFC, e.g. the dorsolateral prefrontal cortex (DLPFC) have been shown to be sufficiently developed as early as 7 months old and form the basis for higher cognitive processes which continue to develop into adulthood (e.g. Diamond & Goldman-Rakic, 1989; Garon et al., 2008). Although morphological changes within the PFC are usually developed by around puberty, the qualitative nature of the PFC continues to develop throughout adolescence and into early adulthood (Romine & Reynolds, 2010; Stuss, 1992). A number of physical changes have also been identified during the late stage maturation of the PFC including; on-going myelination of nerve fibres (e.g. Giedd et al., 1999), reduction in grey matter (e.g. Sowell et al., 2004), synaptogenesis (Huttenlocher, 2013), and resisting metabolism (Chugani et al., 1987; for a review see Diamond, 2002). The observed changes continuing to adulthood suggests that children and adolescents may lack certain abilities in advanced executive skills (Anderson, 2002).

Within the PFC itself, it appears that maturation is marked by a decrease in grey matter that is first completed within the orbitofrontal cortex (OFC), followed by the ventrolateral PFC (VLPFC) and then by the dorsolateral PFC (DLPFC; Bunge & Crone, 2009; Gogtay et al., 2004). Bunge and Crone (2009) proposed that these differences accounted for the differences in the rate of development of distinct EF sub-processes. To add weight to the argument of EF diversity, neuro-imaging research has also endeavoured to demonstrate how specific areas of the frontal lobes are implicated when engaging each of these EF subcomponents (Miyake et al., 2000; Miyake & Friedman, 2012). For instance, IC deficits have been shown in children diagnosed with ADHD who display impaired performance of IC tasks, e.g., Go-No-Go and Stroop task as well as displaying atypical patterns of PFC activation relative to typically developing children (Durston et al., 2006). FMRI studies have also implicated the PFC while performing IC tasks with normative samples of children and adults (e.g. Go-No-Go task, Casey et al., 1997).

2.5. EXECUTIVE FUNCTION DEVELOPMENT

In line with the neurological evidence implying the prolonged development of the frontal lobes past childhood and extending into early adulthood, this section will review whether the development of the EF follows a similar trajectory. As the frontal lobes are engaged during the regulation and controlling of behaviour it is reasonable to assume that development of the EF should follow a similar trajectory to frontal lobe development (Romine & Reynolds, 2005). Indeed, research indicates that EFs may take longer to develop than other areas of cognition which peak during childhood (e.g. causal reasoning; Fischer 1980 and theory of mind; Piaget, 1936; 1952; Wellman, Cross & Watson, 2001; see also Huizinga & Smidt, 2010), the development of EF extends beyond adolescence and into early adulthood (Best, Miller, & Jones, 2009; Huizinga & Van der Molen, 2007). The protracted development of certain EFs has also been associated with the slow maturation of the PFC (Amso & Casey, 2006; Casey, Tottenham, Liston, & Durston, 2005). Interestingly, it appears that the developmental trajectories of component EFs may deviate slightly from one another (Best et al., 2009; Best, Miller, & Naglieri, 2011; Diamond, 2002; Romine & Reynolds, 2011). Despite the protracted development of EF skills, relatively little is known about the trajectory of children’s EFs as most EF research utilises adult samples (Best et al., 2011; Garon, Bryson, & Smith, 2008).

Research examining EF development across the lifespan has been aided by the innovation of age-appropriate EF tasks developed for younger children (Diamond, 2002; Frye, Zelazo, & Palfai, 1995; Gerstadt, Hong, & Diamond, 1994; Hughes, 1998; Kochanska, Murray, Jacques, Koenig, & Vandeceest, 1996). Such tasks enable developmental researchers to observe EF capabilities as early as infancy (Espy, Jaufmann, McDiarmid & Glisky, 199; Russell, 1999; Welsh & Pennington, 1988). Indeed, some researchers have shown that infants as young as 7 months utilise simple IC EF strategies to perform tasks such as the Piagetian, A-not-B task (e.g. Diamond, 1985; Espy et al., 1999; Piaget, 1954; Russell, 1999; Welsh & Pennington, 1988) compared with older children who must adopt more complex strategies or unified EF skills to perform complex tasks such as the WCST (Miyake et al., 2000).

In line with the debate surrounding conceptualisation of the EF, developmental researchers differ in how they choose to model the EF and subsequently its development across childhood. Those who consider executive functioning to be unitary in nature discuss general changes during childhood. For instance, Munakata (2001) and Morton and Munakata (2002) posit that two types of representations develop as part of a child's EF: latent memory traces and active memory representations. Active representations relate to attention and take the form of sustained activity in the PFC. They develop later than latent representations which correlate with habitual responses and long-term memory stores and are formed in more posterior cortex. Following the work of Luria (e.g. 1961, 1966), Zelazo and colleagues (Zelazo & Frye, 1998; Zelazo & Müller, 2002) consider EF as a unitary construct and describe its development through their theory of Cognitive Complexity and Control (CCC; Zelazo, Carter, Reznick, & Frye, 1997). They proposed that children's EF develops rapidly through the preschool years as rule representations become more advanced and hierarchical. Young children are unable to integrate their representations and therefore perseverate more easily. Towards the end of the preschool period, children have honed their rule system and become able to reflect on rules while integrating elements of knowledge into a more complex rule system (Garon et al., 2008; Zelazo, Qu, Müller, 2005). They also suggested that children's cognitive development can be described in terms of their EF development and that age-related performance improvement in EF mediates improvements on a wide range of cognitive tasks (Zelazo & Frye, 1998; see also Zelazo & Müller 2002, 2010).

In contrast to these more unitary conceptualisations, researchers such as Diamond (2013; Lehto et al., 2003) have adopted dimensional approaches to EF development. Diamond postulated that inhibition, shifting, working memory and the unified EF component may each have unique developmental trajectories (2006, Diamond et al., 1997). In this respect her theoretical perspective compliments the model proposed by Miyake et al. (2000). Like Munakata and Zelazo, Diamond recognised that perseveration occurs in early childhood as a result of an inability to resolve conflicts within the mind. In other words, successful executive functioning requires an ability to overcome automatic, prepotent behaviours despite the draw from previous experience (Diamond, 1985; 2002; 2006).

An understanding of the developmental nature of EF across the lifespan is important for health professionals and educationalists working with children and adolescents (Anderson, 2002). With this knowledge, early detection of cognitive deviations can be identified, age-appropriate interventions developed, and, for researchers, group differences assessed more accurately. A growing number of behavioural and neurological studies have tracked EF development from infancy through to adolescence and into early adulthood. Many of these changes have also correlated with changes in areas of the frontal lobes, which continue to localise and mature with age (Davidson et al., 1997, 2006; Diamond & Taylor, 1996; Hudspeth & Pribram, 1990; Lehto et al., 2003; Thatcher, 1991, 1997; Zelazo, Carter, Reznick, & Frye, 1997).

From preschool through to primary school age (3-7 years) significant gains in EF development have been observed (Best & Miller, 2010; Davidson et al., 2006; Garon et al., 2008; Huizinga, Dolan, & van der Molen, 2006; Huizinga & van der Molen, 2007; Luciana, Conklin, Hooper, & Yarger, 2005; Somsen, 2007, Romine & Reynolds, 2005). There is a particularly marked improvement in EF skill between 3 to 5 years of age, expressed through social cognition (theory of mind; Wimmer & Perner, 1983), moral development (Kohlberg, 1963) and performance on more complex cognitive tasks such as the dimensional change card sort task (DCCS; Zelazo, Reznick, & Pinon, 1995), ambiguous figures (Gopnik & Rosati, 2001), Simon task (Simon, 1969; Davidson et al., 2006), the day-night Stroop-like task (Gerstadt, Hong, & Diamond, 1994) and the Go-No-Go task. For example, using a modified version of the Simon task (Simon, 1969), Davidson et al. (2006) found that children as young as 4 years old were able to perform correctly on 80% of incongruent trials in the Simon task, with accuracy increasing as a function of age.

Subsequent improvement in EF abilities continues from middle to late childhood with rapid improvements seen between 5 and 13 years of age (Brocki & Bohlin, 2004; Diamond, 2006). More complex EF tasks such as the anti-saccade task, the Wisconsin Card Sort Test (WCST), the directional Stroop task, and WM span tasks have been utilised to test EF skill of older children and young adolescents. Children between 6 and 7 years of age have difficulty performing the anti-saccade task (thought to tap the IC) but improve dramatically beyond these ages, yet performance does not peak until around 20 years of age (Luna et al., 2010). Likewise in the WCST, children show improved performance between 5 and 11 years but may not reach adult levels until their early 20s (Rosselli & Ardila, 1993; Welsh et al., 1991). Similarly switching and WM span tasks have shown continued improved performance from childhood to early adulthood although development is more gradual in WM tasks after 11 years (e.g. Meiran, 1996; Cepeda, Kramer, & Gonzalez de Sather, 2001). A number of studies have examined EF development across age groups and found that certain EFs may reach ceiling sooner than others. For instance Welsh et al. (1991) and Luciana and Nelson (2002) found that the unified EF component may develop later than IC, switching and WM (e.g. WCST and ToH performance). These studies are in line with research suggesting that tasks with simple EF demands, e.g. WM span tasks with 2 or 3 items may be developed by middle childhood while performance on more complex tasks and the ability to co-ordinate a number of EF components or multi-tasking continues to develop into early adulthood (e.g. Brocki & Bohlin, 2004; Levin et al., 1991; Luciana & Nelson, 1998; Welsh et al., 1991).

This research aimed to explore the developmental trajectories of children's EF between 8 and 12 years, a brief discussion of the specific and unified EF components' development within these age groups follows. A number of studies have employed similar age samples to explore the development of EF throughout childhood and early-adolescence, with most adopting a component-specific approach to EF, using tasks thought to assess individual EFs. Some have been longitudinal while most use cross-sectional designs.

One study by Lehto et al. (2003) aimed to examine the cognitive development of 108 children aged 8 to 13 years. Their second aim was to assess the organisation of the EF system using latent variable analysis with two testing phases (13 day interval). Their findings suggested agreement with contemporary views that consider the EFs to be simultaneously uniform and diverse (Miyake et al., 2000). Like Miyake and colleagues, they found their EF tasks clustered onto three main component functions of inhibition, switching and WM. However, they found the issue of 'task impurity' to be of major limitation to their findings as the study utilised different EF tasks to those of the Miyake et al. study, decreasing its reliability. Like in Miyake et al., Lehto and colleagues also utilised complex EF tasks to assess the unified EF component and suggested that future studies employ similar complex and simple EF measures when formulating hypotheses regarding the nature of EF. Developmentally, their findings suggested that WM and switching task performance matured with age while IC did not display the same trend and that IQ also did not significantly correlate with IQ (WISC). They suggest that the development of IC skills depends on task demands as more complex IC tasks, e.g. Stroop continue to develop through adolescence (Williams, Ponesse, Schacher, Logan, & Tannock, 1999) while others reach ceiling during childhood, e.g. Go-No Go task performance (Dempster, 1992).

Using a similar experimental design to Lehto and colleagues, Huizinga, Dolan and Van der Molen (2006) also applied Miyake et al.'s theoretical model to assess the developmental trajectories of inhibition, shifting and WM. Adopting a latent variable approach they found two EF components to be dissociable from one another: WM and switching. Their multiple-group confirmatory factor analysis (CFA) found that both components moderately correlated on a range of tasks across four age groups (7-, 11-, 15-, and 21-years old). However, tasks believed to tap IC did not appear to correlate with one another. In other words, the IC component did not load onto a common factor within their EF model. Garon et al. (2008) suggested that the wide range in age groups used by Huizinga and colleagues may have complicated their model, with both development and component EFs adding to the variance. Huizinga and colleagues also found continuation of EF development beyond early childhood and into adolescence, reaching adult levels of performance between 11 and 15 years of age. WM development continued through to early adulthood while switching reached adult levels during adolescence (15 years). Overall there was a stable increase in the dissociable EF components across age groups.

A number of researchers have adopted component specific approaches to examine the development of children's EF skills (see Best & Miller, 2010 for review). For the IC function, development has been shown to improve throughout childhood. In a study with children between 4 – 11 years old, IC skills improved as function of age on the Simon task. However, before age 6, children displayed difficulty in performing incongruent trials (Davidson et al., 2006). Luna and colleagues (2004) also showed performance IC improvements on the antisaccade tasks with performance increasing from childhood (50% errors) to adulthood (10-20% errors). Their findings showed that developmental improvements continued from mid to late adolescence where the percentage or errors begins to reach adult levels (see also, Anderson et al., 2001; Christ et al., 2001; Luna, 2009). However, Best and Miller (2010) described how from 5 years onwards, studies showing IC improvements have been mixed at best and may depend on the type of IC task administered (e.g. Becker et al., 1987; Dempster, 1992; Levin et al., 1991; Nigg, 2000) and that some researchers have maintained that pure IC reaches adult levels at around 12 years old (e.g. Bédard et al., 2002; Huizinga & Ven der Molen, 2007).

In terms of switching development, performance also depends heavily on the task demands (Diamond, 2013). For instance, for tasks requiring participants to switch their way of thinking rather than their attendance to certain stimuli, development occurs later. The ability to switch between multiple dimensions of a task has shown improvement between 7 and 9 years (Anderson et al., 2001). Some researchers have argued that switching abilities may occur later than IC as they require participants to build upon other EF skills such as WM and IC. For example, once a child has made the decision to switch to a certain stimuli they must then inhibit the non-intended response (Clearfield et al., 2006; Garon et al., 2008) and this ability has been shown to emerge around the end of the preschool period and continues to improve across childhood (Cepeda et al., 2001; Diamond, 2013; Diamond et al., 2005; Kloo & Perner, 2005; Kray, 2006).

Luciana and Nelson (1998) used dimensions from the CANTAB to investigate the developmental nature switching abilities across older children and adolescents. They found that most of the improvements occurred between 5 and 6 years of age, although performance increases were observed throughout childhood and early adulthood with the number of stages completed increasing as a function of age. Similarly, Davidson and colleagues (2006) found that switching improved from 4 years through to adolescence. Some researchers have suggested that adult levels of performance for the switching function are attained at around 12 years of age (Cepeda et al., 2001; Crone, Bunge, Van der Molen, & Ridderinkkhof, 2006; Huizinga & Van der Molen, 2007; Kray, Eber, & Lindenberger, 2004) while others such as Huizinga et al. (2006) have found, that abilities did not reach adult level until 15 years of age and that strategies may change between childhood and adolescence with older children using more planning and anticipatory strategies prior to switching (Czernochowski et al., 2009; Munakata et al., 2012). Despite these results there is still relatively little research regarding the developmental nature of cognitive switching, therefore conclusions for how children should perform across certain age groups should be tentative at best (Cepeda et al., 2001; Crone, Bunge, Van der Molen, & Ridderinkhof, 2006; Crone, Ridderinkhof, Worm, Somsen, & van der Molen, 2004; Zelazo, Craik, & Booth, 2004).

The WM function may take longer to develop due to the difficulty of holding and manipulating information in the mind (Diamond, 2013; Crone et al., 2006; Davidson et al., 2006; Luciana et al., 2005) but this ability had been has been shown to be sufficiently developed for use in more complex EF tasks by around 6 years of age (Diamond et al., 1997; Gathercole et al., 2004). Gathercole et al. (2004) suggested that WM skills follow a linear trajectory with performance increasing as a function of age between 4 – 14 years of age. They also suggested that WM skills may reach adult levels at around 14 – 15 years. Other researchers have found that WM performance continues to develop until late adolescence on tasks of verbal and visuospatial WM (Best & Miller, 2010). Hale, Bronik and Fry (1997) found that WM skills showed a linear improvement with age and that 19 year olds outperformed 8 and 10 year olds on both verbal and visuospatial WM tasks.

Using complex EF tasks (e.g. WCST and ToH) to test the unified EF component, researchers have shown that performance improves with age and that developmental improvements may level off during adolescence. In a study of EF development using the WCST, Somsen (2007) found that certain aspects of the task were shown to develop sooner than others in a sample of 259 children between 6 and 18 years of age. For instance, the number of errors made and number of categories completed increased with age and plateaued at around 11 years while perseverative responses continued to improve until late adolescence.

These findings show that both the specific and unified EF skills should improve as a function of age in the sample tested in this thesis (children between 8 and 12 years). However, there are still relatively few longitudinal studies examining the developmental nature of EFs, therefore predictions for their trajectories should be tentative at best.

2.6. IMPROVING EXECUTIVE FUNCTION DEVELOPMENT

There is a growing body of research to suggest that EF training can enhance children’s long-term success, particularly within the academic environment. For example, children who have been taught more effective EF strategies during maths computations and reading comprehensions also improve their overall academic performance (Iseman & Naglieri, in press; Naglieri, Das, & Goldstein, 2012). A number of researchers have picked up on the benefits of EF for long-term academic success and have worked on techniques and training methods to help foster its development. Diamond & Lee (2011) discussed the importance of targeting children’s EF from an early age for school readiness and academic success, as well as general mental and physical health (Kusche, Cook, & Greenberg, 1993). It has also been argued that targeting the EF is more important than IQ for school readiness (Blair & Razza, 2007). Diamond goes on to list a number of approaches that have been shown to improve the performance of children’s EF during early school years. These include: computerized training (e.g., Best et al., 2011; Holmes, Gathercole, & Dunning, 2009; Morrison & Chein, 2011; Muraven, 2010; Shipstead et al., 2012), aerobic exercise and sport (Hillman, Erickson, & Kramer, 2008; Tuckman & Hinkle, 1986), martial arts and mindfulness practices (Lakes et al., 2013; Lakes & Hoyt, 2004) and Montessori (Montessori, 1949; cited in Tong & Rodriguez, 2013). Diamond also discussed how any interventions that have elicited EF improvements, do so in the more complex EF tasks requiring higher EF demands. Group differences begin to emerge during these more complex EF measures, as the EFs must continuously be challenged in order to see improvements. In other words, task demands must become increasingly difficult for interventions to be effective (Diamond, 2011). Best et al. (2011) suggested that although EF interventions have elicited improvements with young children, it is unclear what influence such interventions may have for older children, whose EFs are further along in their development.

Apart from potential EF intervention, certain early childhood experiences may also enhance - as well as hinder - the development of the EF. For instance, bilingualism is one such experience believed to affect the development of the EF (Bialystok, Craik, Green, & Gollan, 2009; Diamond, 2010; see section 3.4.). During infancy, exposure to two languages has been examined as a method of improving EF development (Kovacs & Mehler, 2009; Wass et al., 2011). Bilingualism has been shown to accelerate EF development in children and to delay the onset of EF decline during cognitive aging (e.g. Bialystok & Viswanathan, 2009). EF has also been shown to develop earlier in bilinguals compared with monolingual peers, and bilingual individuals continue to outperform monolinguals on a range of EF tasks (Adi-Japha, Berberich-Artzi, & Libnawi, 2010; Bialystok, 2010; see Bialystok, Craik, Green, & Gollan, 2009 for review; Prior & MacWhinney, 2010). However, what is not known is whether bilingualism affects component EFs or the unified EF component (Bialystok, 2011).

Crucially, there appears to be a strong link between the EF and children's ability to succeed within day-to-day life, particularly within the academic environment (see Diamond, 2013 for a review). This finding has important implications for educational policies, which may underestimate the influence of EF skills. Some have argued that more emphasis should be placed on the development of EF skills within the classroom in order to foster children's long-term success (Meltzer, 2011; St. Clair-Thompson & Gathercole, 2006). Furthermore, environmental experiences and EF training can affect how children perform, particularly in the promotion of maths and reading skills (Best et al., 2011). Early experiences such as bilingualism may modify the developmental trajectory of EF by accelerating or decelerating its progress and this is discussed further in Chapter 3.

2.7. LIMITATION OF EXECUTIVE FUNCTION RESEARCH

Despite the wide range of studies examining the EF, a number of methodological and conceptual issues continue to challenge researchers. Many of these issues relate to task impurity, conceptualisation and the inconsistency in tasks employed to assess EF skills (e.g. Lehto et al., 2003; Miyake et al., 2000). Although a variety of tasks have been employed in studies of EF, many agree that a degree of novelty, complexity and integration of information must be present during assessment, as well as a strong rationale for how the task is tapping into each or a number of EF processes (Anderson, 1998; Henry & Bettenay, 2010; Shallice, 1982; Walsh, 1978). However, defining what is complex or novel is not straightforward as these concepts may vary between individuals (Stuss & Alexander, 2000; Anderson, 2002). Furthermore, task novelty is decreased through repeated exposure to EF tasks (e.g. for longitudinal studies) and high reliability coefficients have been difficult to obtain by researchers. Henry and Bettenay (2010) suggested that due to the nature of EF tasks, researchers must accept these lower reliability levels and that once a person becomes familiar with a task its ability to measure novelty is diminished. However, they also noted that from the 6 EF batteries they assessed, the Test of Everyday Attention for Children (TEACh; Manly et al., 1999; see section 4.5.3.) battery was one of the most consistently reliable, although no single battery offered a complete assessment of EF skill.

To overcome issues of task impurity, researchers have been encouraged to utilise a range of EF measures rather than a single-task design to tap selected EFs. However, despite this method, it is difficult to be certain that the selected tasks do not test a function beyond the one of interest (Anderson, 2002; Miyake et al., 2000). In other words, a researcher can give no guarantee that a WM task will only tap the processes underlying WM and that no other function will be utilised by participants during task performance (Best & Miller, 2010; Bialystok, 2011). This issue of *task impurity* is widely cited as being one of the major limitations to EF research (see Lehto et al., 20003; and Hughes & Graham, 2002 for discussion). Issues of task impurity have also arisen from the lack of consensus regarding the conceptualisation of specific EFs as some suggest that subcomponents of EF are overly broad and weakly defined (e.g. Willcutt et al., 2007).

One example of this conceptual debate is the distinction made between WM and IC (e.g. Diamond, 2013). Researchers have suggested that these components depend on the same limited-capacity store (e.g. Engle & Kane, 2004; Wais & Gazzaley, 2011). A number of studies have also argued that the function of IC is indistinguishable from other components of EF (Egner & Hirsch, 2005; Hanania & Smith, 2010; Huizinga, Dolan, & Van der Molen, 2006; Nieuwenhuis & Yeung, 2005; Miyake & Friedman, 2012; Miyake & Friedman, 2012; Munakata et al., 2011). Even within different IC tasks, validity has been questioned as researchers failed to isolate inhibition among common measures believed to tap into this component. Shilling, Chetwynd, and Rabbitt (2002) found weak and varied correlations (-.13 and .22) between a set of Stroop-like tasks widely cited as requiring IC. Davidson et al. (2006) conducted a longitudinal study of children from 4 to 13 years, administering a number of experimental tasks designed to tap IC and WM separately as well as together (e.g. Simon task tapping inhibitory control with low WM load, Abstract-Shapes task requiring WM and little or no inhibition). Contrary to their initial prediction of WM and inhibition being separable, they found that RTs for tasks tapping WM and inhibition correlated highly with one another (0.82).

In line with Baddeley and colleagues' earlier work (1986, 1992), some researchers define IC as part of the WM system rather than WM being a component of the EF system. For instance, Kane and Engle (2000, 2002; Diamond, 2013) define WM as: "*an ability to maintain selected information in an active, retrievable state while inhibiting distractors and interference (i.e., short-term memory and interference control at the attentional and cognitive levels"*. Neurologically, there is also evidence that IC and WM utilise similar mechanisms. The prefrontal-parietal cortex has been implicated when performing tasks of WM requiring participants to focus and maintain the task goal in mind. Similarly, tuning out irrelevant information and thoughts in order to selectively attend to task-relevant stimuli requires the use of the prefrontal, parietal system (e.g. Awh et al., 2000; Awh & Jonides, 2001; Gazzaley & Nobre, 2012; Ikkai & Curtis, 2011; Nobre & Stokes, 2011).

Despite numerous accounts stipulating a lack of distinguishability between WM and IC, some researchers maintain that IC and WM are, in fact, dissociable from one another (e.g. Anderson & Spellman, 1995; Diamond, 2002; Gernsbacher & Faust, 1991; Hasher et al. 1991; Zanto et al., 2011). Evidence for this distinction is evident from studies outlining the unique correlations between component EFs, and other cognitive skills, e.g. WM and IQ (*g*; Alloway & Alloway, 2010; Friedman et al., 2011; Salthouse et al., 2003). Beyond this, a failure to isolate a specific IC component within studies of EF does not mean that IC is an academic fallacy. While the construct of IC may exist, the issue for researchers may be that isolating it through behavioural experimentation is too difficult as it may be highly correlated with other constructs as well as showing a lack of reliable individual differences after early childhood (Van der Sluis et al., 2007).

To help overcome task impurity and conceptualisation issues Hughes and Graham (2002) suggested that children be recruited for studies of EF. With children there is less need for a 'sequence' of acts when attempting to tap into EFs. Furthermore, as experimental manipulations have a greater impact on children than adults (due to limited processing), interactive effects and causation may be explored more easily within experiments. Using assessments with children a number of researchers have been able to isolate specific EFs from one another, e.g. WM from IC (Best & Miller, 2010; Diamond, 2002; Hughes, 1998; Welsh et al., 1991). Despite these benefits, EF research with children is not without its own difficulties. The late development of literacy skills creates problems for tasks requiring the use of written language skills e.g. Stroop (Hughes & Graham, 2002). Zelazo (1999, 2000) argued that language plays a crucial role in EF development as linguistic formulations may be directly responsible for developmental improvements on EF tasks, creating issues for researchers trying to untangle the relationship between EF and linguistic development. Furthermore, Henry, Messer and Nash (2012) found that EFs were delayed in children with specific language impairments, particularly WM and IC skills. Such findings highlight the importance of adequately controlling for verbal skills when assessing children's EF development.

2.8. SUMMARY

This chapter examined the debate surrounding conceptualisations of the EF. The adopted EF model was developed by Miyake and colleagues (2000) and was influenced by the earlier work of Baddeley and Hitch (1974). Here, the EF is both uniform and diverse with subfunctions IC, switching and WM acting independently as well as together during complex tasks. EF development occurs later than other cognitive skills, extending into early adulthood, offering researchers the opportunity to explore its trajectory across wide age-ranges. Finally, a number of experiences have been shown to affect the development of EF although limitations such as task impurity and low test-retest reliabilities may impact on these results. The next chapter applies the EF model propsed by Miyake (2000) to an experience believed to positively impact on children and adult's EF skills, bilingual or multilingualism (Diamond, 2011). The impact of bilingualism on the specific as well as unified EF components will also be discussed in the following chapter (Chapter 3).

CHAPTER THREE
BILINGUALISM AND COGNITIVE DEVELOPMENT

3.1. INTRODUCTION

The aim of this chapter is to discuss the impact that the bilingual experience has on children's cognitive development. Before discussing the potential cognitive effects of bilingualism, section 3.2. considers how language, and in particular two languages, develops within the mind as well as examining how bilingual individuals process language differently to monolinguals. While sections 3.3. and 3.4. discuss general cognitive disadvantages and advantages associated with the bilingual experience, section 3.4.1 focuses on the area of cognition widely cited as being most positively affected by bilingualism, the executive function (EF). Finally section 3.5. attempts to make sense of different cognitive outcomes linked to bilingualism as well as providing a summary of previous sections.

3.2. LANGUAGE AND BRAIN DEVELOPMENT

3.2.1. Language development

Before discussing how children might develop two languages, it is important to recognise how language is acquired in general. Although the developmental stages appear to be reasonably robust for certain language characteristics, e.g. lexicon, syntax, phonology and pragmatics (starting from birth to 3 months and continuing through to adolescence and early adulthood), there is still considerable debate surrounding the mechanisms responsible for their development (Bialystok, 2001). A summary of the main theories of language is beyond the remit of this thesis but it can be said that most theories fit into one of two categories - the formal (e.g. Chomsky, 1965, 1995; Pinker, 1984; 1994) and the functional (e.g. Langacker, 1986; 1991; Van Valin, 1991) approaches to language development. Briefly, the formal approach regards language as an innate process, biologically predisposed and ready to develop as soon as is it activated by environmental inputs. It describes language as a modality distinct from other cognitive functions, the development of which derives from a universal grammar, controlled by a language acquisition device, present within all of us (nativist account). In contrast, the functional approach postulates that language emerges from a child's interactions and experiences with their environments (behaviourist account). Input variations will therefore change a child's representations and construction of language. In the functional account, languages emerge from a need to fulfil cognitive, social and communicative functions. Therefore, it assumes that language is fully integrated with all other cognitive domains and depends entirely on them for its structure, meaning and use.

With these arguments in mind, how can such opposing approaches be reconciled and, perhaps more crucially, how can they enhance our understanding of bilingual language acquisition? Bialystok (2001) made the point that both accounts offer valuable insights into how children may acquire two languages. She also suggested that they should be used in unison rather than in competition, describing formal accounts as the "attempt to explain language" and functional accounts as the "attempt to explain communication" (p. 50). Under formalist approaches, the age at which a child acquires their L2 (e.g. simultaneous or sequential) is irrelevant as their development is guided by an in-built mechanism of acquisition and by constraints of universal grammar. However, in the functionalist approach, age of acquisition is of importance as social interactions and previous knowledge drive children's language acquisition, separating the experiences of simultaneous and sequential L2 learning.

As in the general language development literature, describing the mechanisms of bilingual language acquisition is not a simple task. It has been argued that acquisition of two languages may be both similar and distinct from monolingual language development and depends on a range of factors including age and levels of proficiency (e.g. Butler & Hakuta, 2004). Furthermore, synthesising the formal and functional accounts may also be more valuable for understanding bilingual language acquisition than simply choosing one view or another (Bialystok, 2001). Grosjean (1998) suggested that parents should not expect the patterns of development in each of their bilingual child's two languages to mimic patterns displayed by monolinguals, as the representational systems for the two languages are different. For instance, it is common for bilingual children to mix and switch between their L1 and L2 during acquisition as a natural stage in their development. However, as a result of the numerous historical accounts outlining bilingual disadvantages (e.g. Macnamara, 1966, see 1.3.), many parents and educators still assume that young children acquiring two languages simultaneously run the risk of having language delays. In fact, children learning two languages develop at much the same rate as a child learning one (American Speech-Language-Hearing Association, 2009). Although a bilingual child's vocabulary in either language may be less equivalent to a monolingual peer, this finding is not unsurprising as, when taken together, total vocabulary from both languages is comparable and often exceeds monolingual peers (e.g. Marchman & Fernald, 2010; see 3.3.). Furthermore, as a bilingual child's linguistic experiences are fundamentally different in each language it is perhaps unfair to treat and compare a bilingual child as if they were two monolinguals (Grosjean, 1998, 2001; Romaine, 1999).

A method used to further understand language development focuses on how we represent language in the mind. With the help of modern neuroimaging techniques, e.g. ERP, fMRI, PET, some questions regarding language representation can now be addressed in more detail. While the areas of the brain used for language processing are similar for monolinguals and bilinguals (regardless of age or language; Klein et al., 1995), researchers continue to debate how languages are represented and processed. As with theories of language acquisition, two competing approaches have emerged to explain how bilinguals represent language in the mind. The first is the unitary language system, which assumes that bilingual children use the same system to represent and process both of their languages. This can lead to difficulties for young bilinguals (under 3 years of age) who may struggle to differentiate between the L1 and L2 (e.g. Pettito & Dunbar, 2004; Redlinger & Park, 1980; Vihman, 1985; Volterra & Taeschner, 1978). The second approach is the dual or differentiated language system where each language has a unique representation, similar to that of a monolingual (David & Wei, 2008; Genesee, 1989; Pettito & Dunbar, 2004). Most evidence suggests a differentiated language system although bilingual language processing may combine both approaches. For instance, representations of the L1 and L2 may be distinct within the mind yet both may be active during language use (Bialystok, Craik, & Luk, 2012; Kroll & Sunderman, 2003). Furthermore, a number of alternative factors may affect individuals' linguistic experiences such as proficiency levels and learning environments and such factors may result in varying degrees of language representation and organisation within the mind (Bialystok, 2001). In conclusion, bilingual children's representations of two languages are unique yet are not simply the sum of two monolingual minds. While each language and its representation may be distinct, they may also share a number of elements including space within the brain and levels of activation. As the bilingual brain has been shown to activate both languages during speech production, this may result in a number of functional differences compared with monolinguals as the bilingual becomes practiced in controlling two language representations (Abutalebi & Green, 2007). Some of these proposed changes are discussed in the following section.

3.2.2. The bilingual brain

In line with questions surrounding language acquisition, many people wonder whether the bilingual brain functions equivalently to monolinguals. One way in which they do differ is that bilinguals appear to have two language representations active at all times which interact with one another, regardless of the linguistic context (Grosjean, 1998; Bialystok et al., 2012). Evidence for this dual-language activation has come from a number of disciplines, some of which are discussed here (e.g., Abutalebi & Green, 2007; Bialystok et al., 2012; Costa & Caramazza, 1999; Gollan & Acenas, 2004, Rodriguez-Fornells et al., 2002). Furthermore, if it is shown that both the L1 and L2 are active then bilinguals must adapt to managing these representations in order to effectively communicate in either language. Individuals must control and inhibit the non-intended language, which constantly competes for recognition and selection and, using principles of neuroplasticity, it is reasonable to suggest that the constant competition bilinguals face may impact on other areas of cognitive function, beyond linguistics alone (e.g. Abutalebi & Green, 2007; Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Craik, & Ryan, 2006; Diamond, 2010).

Behavioural experimentation has attempted to demonstrate dual-activation of the L1 and L2 within the bilingual brain but without any definitive conclusions (Bialystok et al., 2012, Kroll & Bialystok, 2013). For instance, the picture-naming task has shown that words semantically related to a picture delay children's responses in naming that picture even when they are not in the intended language of response (e.g. Costa &

Caramazza, 1999; Hermans et al., 1998). This is evidence that both languages are competing for selection, with bilinguals having to resolve these competitions while picture naming in one language or another (Abutalebi & Green, 2007). Another behavioural finding is that during language switching tasks, bilingual children take longer to switch back into their dominant language (L1; Meuter & Allport, 1999). Here it is suggested that as individuals must inhibit the L1 so strongly when responding in the L2, this suppression continues when they switch back into the L1, resulting in delayed responses (Green, 1998). However, these asymmetrical switch costs have not been found in all cases as delays may be reduced with proficient bilinguals, following practice or as a result of stimulus type (Costa & Santesteban, 2004; Finkbeiner, Almeida, Janssen, & Caramazza, 2006; Meuter & Allport, 1999; Verhoef, Roelofs, & Chwilla, 2009). In summary, while the behavioural tasks used as evidence for dual-activation in the bilingual brain have been mixed, most agree that both the L1 and L2 representations are active to some degree. Furthermore, patients with frontal lesions to the brain have also provided researchers with behavioural evidence of dual-language activation as these patients have displayed intrusions from the non-intended language or inappropriate language switching (Fabbro et al., 2000).

Modern neuroimaging techniques enable researchers to uncover in more detail how two languages are represented within the bilingual mind and are used to map areas of the brain active during language processing. While ERP can tell us *when* things are happening, PET and fMRI can tell us *where* in the brain these processes occur (Craik & Bialystok, 2006). These techniques allow researchers to detail the mechanisms of control and language representation, complementing behavioural investigations. Abutalebi and Green (2007) used neuroimaging data in combination with bilingual aphasia data and found that the neural representation of participants' L2 overlapped with the neural representation of the L1 (see also Perani et al., 1998; Rodriguez-Fornells et al., 2002). Furthermore, their study showed that the majority of L1 and L2 activation occurs in the prefrontal areas of the brain (e.g. Rodriguez-Fornells et al., 2005). Kovelman, Baker and Pettito (2008) used fMRI to investigate whether bilinguals had differing neural representations for language processing than monolinguals. Behavioural and fMRI data showed that bilinguals processed their L1 (English) and L2 (Spanish) differently. In other words, bilinguals had differentiated neural representations for their two languages. Monolinguals and bilinguals showed activation areas typically associated with language processing such as the left inferior frontal cortex (LIFC), an area comprising the classic Brocas and Brodmanns areas (Price, 2000, Hagoort, 2005). However bilinguals showed increased activation and recruitment of the LIFC during L1 production compared with monolinguals. Kovelman and colleages concluded that highly proficient, early bilinguals control each of their languages differently within the brain although there is a high level of overlap and similarity between the areas of neural activation of both languages. The difference found between L1 and L2 processing in their bilinguals was the level of LIFC activation used to process each language. Furthermore, bilinguals recruited areas of the LIFC more than their monolingual peers, which may indicate that through practice, bilinguals have adapted to utilise their LIFC with greater intensity than monolinguals.

Kim et al. (1997) used fMRI to show that the areas of the brain utilised during language production may differ between early (L2 acquired < 3 years) and late bilinguals. Their findings showed that languages acquired during adulthood (late bilinguals) were spatially separated from native languages yet activated adjacent areas (mostly around Broca’s area). Alternatively, when the L1 and L2 were acquired in the early stages of language development (early bilinguals), the L1 and L2 tended to be represented in common frontal cortical areas. However, their sample size was small (n = 12), different language pairs made up the bilingual group and researchers have struggled to replicate their findings (e.g. Bialystok, 2001, Perani et al., 1998). Other studies have shown no such neural difference between early and late bilinguals and that levels of proficiency in the L1 and L2 have a more significant impact on brain representations than age of acquisition (Perani et al., 1998). A note of caution must be made when interpreting neuroimaging results. Although they can map brain activity, they do not offer explanations for mechanisms of thought (Rodriguez-Fornell et al., 2005). De Bot (2008) also criticised neuroimaging research, concluding that it has failed to provide any breakthroughs in accounts of the multilingual brain. He commented that researchers from purely neuroscience backgrounds often have a limited understanding of psycholinguistic issues and therefore often fail to adequately control their language groups.

Neuroplasticity refers to the human brain's ability to alter and change itself in response to environmental experiences and demands (Draganski et al., 2004). The brain's ability to change is now a broadly recognised phenomenon and has been shown in a wide variety of studies, in clinical and non-clinical samples, from juggling to stroke rehabilitation (Dayan & Cohen, 2011; Dimyan & Cohen, 2011). The principles of neuroplasticity have also been used as an explanation for the changes or differences observed in the brains of bilinguals compared with monolingual peers (Bialystok et al., 2011). As a result of bilinguals having to represent and control two languages in the brain, might this lead to changes within the brain affecting skills beyond language competence? One study found that learning an L2 may increase the density of grey matter in the left inferior parietal area of the bilingual brain (Mechelli et al., 2004). After comparing 25 monolinguals, 25 early and 33 late bilinguals, Mechelli and colleagues found that early bilinguals displayed an increase in grey matter density over the other groups. Furthermore, levels of grey matter increased as a function of L2 proficiency and decreased as the age of acquisition increased. Their study implicated an effect of bilingualism and L2 acquisition on the structure of the brain. The left inferior parietal area has also been implicated in a variety of other cognitive functions, such as attention (Nebel, Wiese, Stude, de Greiff, Diener, & Keidel, 2005) and auditory comprehension (Yeatman, Ben-Shachar, Glover, & Feldman, 2010). Bialystok, Craik, Klein and Viswanathan (2004) also suggested that the process of constantly controlling two languages in the mind enhances bilinguals' overall cognitive control abilities as a result of the brain areas activated during both activities (see section 3.4.).

3.3. DISADVANTAGES OF BILINGUALISM

While controlling two languages in the mind, bilinguals may experience a number of cognitive changes, not all of which result in positive implications for task performance. For instance, a number of cognitive skills appear to be disadvantaged by the bilingual experience. Although it seems counterintuitive that some of these disadvantages relate to children's linguistic skills, particularly as children often acquire their L2 with ease, this is what many studies have shown (Sorace, 2011). Common findings include bilingual children controlling smaller vocabularies in each of their languages compared with monolinguals, having slower lexical retrieval and being delayed during speeded tasks requiring language processing e.g. picture naming task (e.g. Bialystok, Craik, & Luk, 2008; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Sandoval, Gollan, Ferreira, & Salmon, 2010). Bilingual disadvantages and some of the theories detailing how they may arise will be discussed briefly in this section.

As mentioned above, a number of studies have shown that bilingual children control smaller vocabularies in each of their two languages compared with monolingual children (e.g. Barac & Bialystok, 2012; Bialystok, 2009, Bialystok & Feng, 2011; Bialystok, Luk, Peets, & Yang, 2010; Mahon & Crutchley, 2006; Marchman & Fernald, 2010; Oller & Eiler, 2002). These findings are reasonable as children processing two languages will hear and speak each of their languages (L1 and L2) at a lower frequency than a monolingual will speak and hear their one language (L1). Two of the most prominent measures of vocabulary size have been the MacArthur Communicative Development Inventory (CDI; Fenson et al., 1993) for productive vocabulary and the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) for receptive vocabulary. In 2002, Oller and Eilers conducted a large-scale study of almost 1,000 primary school children from Miami, chosen to represent differing language backgrounds, SES backgrounds, educational experiences and proficiency levels, alongside monolingual controls. They assessed vocabulary using a battery of four standardised tests of oral language ability, word comprehension and word production. Monolinguals outperformed bilinguals across all age groups from 5 to 10 years old, indicating a persistent deficit in oral proficiency for bilingual children. Similarly, when Bialystok and Feng (2011) pooled PPVT-III data from 16 studies comparing 963 monolingual and bilingual participants (ranging in age and languages spoken) they found a significant main effect of language group (LG) with monolinguals scoring significantly higher than bilinguals at all age groups. This signified a persistent delay in vocabulary acquisition for bilingual children (Fig. 3.). Adult bilinguals have also displayed deficits in vocabulary acquisition compared with monolingual comparisons (Craik & Bialystok, 2006; Fernandes, Craik, Bialystok, & Kreuger, 2007; Perani et al., 2003; Portocarrero, Burright, & Donovick, 2007) although some studies have found no such difference (Bialystok, Craik, Klein, & Viswanathan, 2004) and studies using smaller sample sizes have also shown no significant difference between monolingual and bilingual receptive vocabulary sizes (e.g. Allman, 2002; Cromdal, 1999; Yan & Nicoladis, 2009).

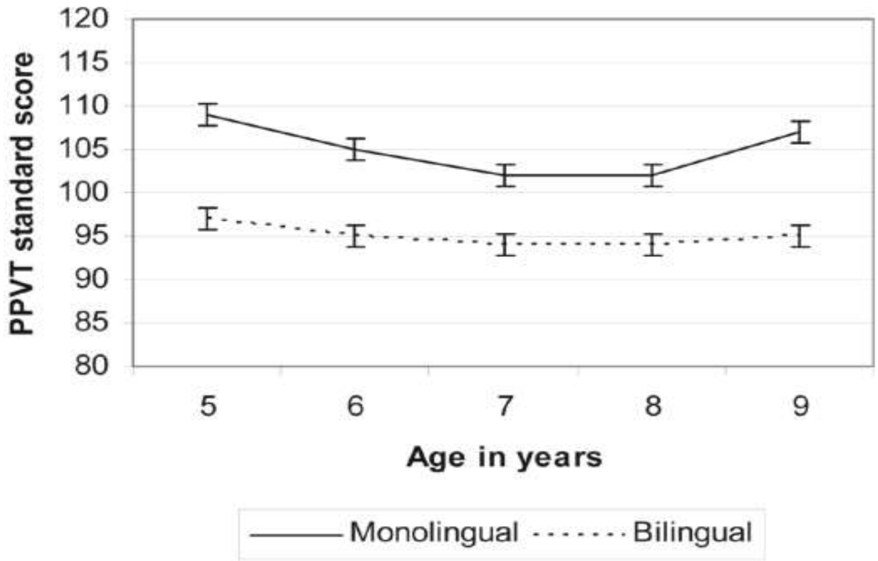


Figure 3 *Mean PPVT-III standard scores and standard errors of 963 children combined from 16 studies (taken from Bialystok & Feng, 2011).*

Frequency of use has already been mentioned as a mechanism for this bilingual disadvantage (Abutalebi & Green, 2007; Ben Zeev, 1977; Michael & Gollan, 2005; Oller, Pearson, & Lewis, 2007). Evidence in favour of the frequency theory has shown that vocabulary size is influenced by the length of exposure a child has had to his/her languages with higher exposure associated with larger vocabularies (David & Wei, 2008; Gathercole, 2002; Pearson, Fernández, Lewedag, & Oller, 1997; Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2013; Turian & Altenberg, 1991). Although this disadvantage appears commonplace when comparing bilingual and monolingual vocabularies in the L1, often when the sum of children's L1 and L2 vocabularies are compared with monolinguals' total vocabulary, no group LG differences are found (Junker & Stockman, 2002; Marchman et al., 2010; Pearson et al., 1993; Oller et al., 2007; Poulin-Dubois et al., 2013). Furthermore, Bialystok and Feng (2011) have questioned the generalizability of vocabulary acquisition findings due to wide variations within individual vocabulary scores and argued that studies of vocabulary should, but cannot, account for how bilingual children's vocabularies would fare had they been learning one language only. Furthermore, the specific languages being considered may be a significant predictor of outcomes as vocabulary delays may decrease if children's two languages are similar in structure and share similar cognates, e.g. Spanish and English (Barac & Bialystok, 2012; Bialystok, Majumder, & Martin, 2003).

A second hypothesis for why vocabulary deficits have been found in bilinguals is the 'weaker link' hypothesis (Gollan, Montoya, Cera, & Sandoval, 2008). This suggests that due to bilinguals using either of their two languages at a lower frequency than monolinguals using their one language, weaker links have been formed between the neural networks connecting words and concepts (Michael & Gollan, 2005).
The weaker link hypothesis has also been put forward as a possible explanation for why bilinguals often display deficits in tasks of lexical retrieval (Gollan & Acenas, 2004). For instance, in tasks of verbal fluency, lexical decision-making and lexical retrieval (e.g. picture naming task), bilinguals have shown delayed responses and lower accuracy rates compared with monolingual peers (e.g. Bialystok, Craik, & Luk, 2008; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ransdell & Fischler, 1987; Rosselli et al., 2000; Roberts, Garcia, Desrochers, & Hernandez, 2002; Portocarrero et al., 2007). Behaviourally, researchers have also observed that bilingual children and adults often encounter 'tip-of-the-tongue' (TOT) states more often than monolingual peers (Gollan & Acenas, 2004; James & Burke, 2000; Michael & Gollan, 2005).

An alternative explanation for this delay is the "hard problem" hypothesis put forward by Finkbeiner, Gollan and Caramazza (2006). As discussed (section 3.2.2.), researchers now believe that bilingual children maintain separate representations in the mind for their L1 and L2. As a result, the alternative lexical representations must compete for selection, making it difficult to quickly select the correct target lexical node. The closer these representations are in meaning the more difficult selection becomes, for example, it would be more difficult to select between the synonymy of couch and sofa than between couch and chair. As almost every concept is associated with a synonymous lexical node in the bilingual mind, the hard problem may explain why bilinguals display slower RTs during lexical retrieval tasks compared with monolinguals. Highly proficient bilinguals should also display even slower RTs as synonyms become even more salient and indeed some studies have found this (e.g.

Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007). However, for the most part, proficient bilinguals do not appear to display any increased delays when speaking one language over another compared with less proficient bilinguals (Finkbeiner et al. 2006; Ivanova & Costa, 2008). Subsequently, a number of theories have been proposed to make sense of how bilinguals deal with the hard problem (e.g. Costa & Caramazza, 1999; Finkbeiner et al., 2006; Green, 1998; La Heij, 2005). One hypothesis developed by Green (1986, 1993, 1998) and known as the inhibitory control model (ICM) suggested that the hard problem in bilinguals is dealt with by using control mechanisms managing both language and action processes. Task schemas are used to control the bilingual lexicon by activating and inhibiting lexical nodes. Therefore if the task demands that participants name an object in their L1, a relevant task schema activates the lexical representations of the L1 and suppresses the L2 task schema. These suppressions or inhibitory processes will naturally incur timing costs, resulting in bilingual performance delays for RT tasks. Bilingual children and adults must also inhibit or suppress the non-target language during language production (David & Wei, 2008; Dikstra, 2005; Grosjean, 2008; Poulin-Dubois et al., 2013). Proficient bilinguals become experienced and expert at dealing with these internal conflicts (Costa & Santesteban, 2004; Meuter & Allport, 1999) and indeed it is not unusual to observe bilinguals' lack of errors while speaking as well as the ease at which they switch between language codes (e.g. Kroll et al., 2008; Muysken, 2000; Myers-Scottom, 2002).

In summary, there are a number of disadvantages resulting from the bilingual experience but most relate to areas of linguistic processing. Highly proficient bilinguals may be able to overcome some of these disadvantages through mechanisms of control while less proficient bilinguals may be less experienced in controlling two languages, resulting in performance delays on a range of lexical retrieval tasks compared with monolingual peers.

3.4. ADVANTAGES OF BILINGUALISM

In contrast to the disadvantages associated with bilingualism, a number of cognitive advantages are also anticipated for those who develop high proficiencies in two languages. Apart from some more apparent advantages such the enhanced ability to learn new languages beyond the L1 and L2, and a deeper understanding of language structures, many bilingual advantages extend beyond these linguistic domains (Sorace, 2011). Some of these advantages and theories for how they develop will be discussed here. A number of researchers have commented that too few studies of bilingualism discuss the mechanisms associated with any apparent bilingual advantages, choosing instead to simply describe LG effects (e.g. Cummins, 1976; Cummins et al., 2001; Hakuta & Diaz, 1985). Bearing this in mind, this section describes some of the theories and mechanisms postulated to result in bilingual advantages.

Although most advantages are non-linguistic, certain linguistic areas do appear to be advanced through bilingualism, e.g. metalinguistic and phonological awareness (e.g. Adesope et al., 2010; Bruck & Genesee, 1995; Cambell, Ruth, & Sais, 1995; Cummins, 1978, 1979; Yelland, Pollard, & Mercuri, 1993). Metalinguistic awareness refers to the ability to the think about language(s). It relates to an awareness of linguistic form and structure as well as an understanding of how changes to form and structure can influence meaning. For example, in the Wug task (Berko, 1958) children must identify the nonsense word within a sentence and manipulate it to conform to the morphological rule of English (Bialystok & Barac, 2012; Cazden, 1974; Adesope et al., 2010). By learning two languages (and subsequently two meanings) it is postulated that bilinguals develop a greater understanding and insight into the arbitrary and symbolic nature of language (e.g. Ben-Zeev, 1977; Bialystok, 1986, 1988; Bialystok, 2009; Cromdal, 1999; Cummins, 1978; Duncan, 2005; Galambos & Goldin-Meadow, 1990; Ianco-Worrall, 1972). Although bilinguals have shown modest advantages on a range of metalinguistic tasks (Bialystok & Majumder, 1998; Yelland, Pollard, & Mercuri, 1993), recently Bialystok and Craik (2010) have suggested that these advantages may be mediated by a number of factors such as the relationship between the L1 and L2, the language of education, and bilinguals' executive function (EF) skills (Bialystok, Majumder, & Martin, 2003). Furthermore, metalinguistic skills have been closely associated with English language abilities in bilinguals but not with increased levels of bilingualism or L2 proficiency (Bialystok & Barac, 2012).

Problem solving, enhanced learning strategies, creative and divergent thinking have all been put forward as being advantaged through bilingualism (Landry, 1974; Lee & Kim, 2011; Peal & Lambert, 1962; Ricciardelli, 1992). Using tests of non-verbal problem solving skills (e.g. The Colour-Progressive Matrices) a number of researchers have proposed that bilinguals may show advantages in problem-solving abilities (e.g. Bamford & Mizokawa, 1989). One theory for this advantage is that bilingual children may be more practised in selectively attending to relevant - and disregarding irrelevant - information as a result of speaking and hearing two languages (Adesope et al., 2010).

The acquisition of a L2 may also enhance children's creative and divergent thinking skills (e.g. Galinsky, 2010; Ho, 1987; Konaka, 1997; Leung et al., 2008; Lee & Kim, 2011; Maddux et al., 2010; Ricciardelli, 1993; Srivastava, 1991). Csikszentmihalyi (1999) argued that creativity is more a cultural and social phenomenon rather than cognitive, as different cultures express creativity in different ways (Torrance & Sisk, 1997). As bilinguals may be exposed to different cultures as well as different languages, this may lead to creativity benefits. Lee and Kim (2011) tested this hypothesis with Korean-English bilinguals using the Torrance Tests of Creative Thinking (Torrance, 1998). They found a positive relationship between bilinguals' levels of proficiency and their creativity scores. As language balance increased, so too did creativity levels (regardless of age or gender). They concluded that multiculturalism and BE should be encouraged by educators to promote creativity. Cushen and Wiley (2011) argued that it is the combination of these enhanced problem-solving and creative skills that enable bilinguals to perform better on a number of thinking or insightful problem-solving tasks. Although there were few LG differences, their findings revealed that in tasks requiring increased cognitive flexibility, bilinguals performed better relative to their own performance on conditions requiring minimal flexibility of thought. On the other hand, monolinguals performed worse under conditions of cognitive flexibility relative to other task demands and it appeared that when tasks required an increased use of the EF or attentional control systems, bilinguals thrived where monolinguals struggled.

In summary, although bilinguals show advantages in general aspects of cognition, most advantages relate to, or may be mediated by, EF or cognitive control skills. As a result, the following sections focus on the possible benefits of bilingualism for executive functioning skills and the potential mechanisms for how these skills have been developed through L2 acquisition.

3.4.1. Bilingualism and the executive functions

In recent years, the study of bilingualism and cognitive development has focused heavily on EF skills as bilinguals have repeatedly shown advantages in tasks of EF compared with monolingual peers (e.g. Bialystok, 2006, 2009; Blair, Zelazo, & Greenberg, 2005; Colzato et al., 2008; Costa, Hernández, & Sebastián-Gallés, 2008; Kroll & Bialystok, 2013; Zelazo & Muller, 2002). As a result, researchers have been drawn to this cognitive domain, hoping to develop theories of both second language acquisition (SLA) and executive control. EF skills involve the ability to control and focus attention while ignoring distracting or irrelevant information and its influence on scholastic achievement and academic success is well established (Diamond, 2013; Diamond & Lee, 2011; Miyake et al., 2000; see sections 2.2. and 2.6.). These findings make it an even more appealing topic for developmental researchers attempting to develop programmes to help children improve scholastic and life-long success (Diamond & Lee, 2011). Not only have EF skills appeared earlier within bilingual individuals (Bialystok, 2001; Carlson & Meltzoff, 2008; Kovacs, 2009) but bilinguals have also shown slower effects of cognitive aging as a result of having enhanced EF skills (Bialystok, Craik, & Freedman, 2007; Bialystok et al., 2004).

The role of the EF is to regulate and maintain information within our brains during complex cognitive tasks. As both languages appear to be active in the bilingual brain during language production (e.g. van Heaven, Schriefers, Dijkstra, & Hagoort, 2008; see 3.2.2.) highly proficient bilinguals will have conflicts and interferences during task performance. This presents the additional challenge of determining which language to use while at the same time, preventing word production in the unselected language (Green, 1986, 1998). It is thought that as a result of having to manage and monitor cross-linguistic interferences from an early age, bilingual children may subsequently develop enhanced EF abilities (Adesope et al., 2010; Bialystok, 2009; Bialystok & Martin, 2004; Bialystok & Viswanathan, 2009; Blair, Zelazo, & Greenberg, 2005; Carlson & Meltzoff, 2008; Gathercole et al., 2010; Mezzacappa, 2004; Prior & MacWhinney, 2010; Yang, Yang, & Lust, 2011; Zelazo & Muller, 2002).

The enhanced ability to attend to different language representations may be key to unlocking bilinguals' improved performance in EF tasks (Adesope et al., 2010; Bialystok, 2001; Bialystok, Martin, & Viswanathan, 2005; Yoshida, 2008). However, identifying the mechanisms causing LG effects can be difficult, as explanations cannot simply be derived from group differences alone. A number of studies look at specific aspects of the EF in an attempt to unearth the causal mechanisms behind LG differences. Despite a number of issues surrounding the separation and classification of EF (see 2.7.), the following sections will describe how researchers have linked the bilingual experience to each of the specific EFs - *Inhibition* (3.4.1.1.), *Switching* (3.4.1.2.) *and Working Memory* (3.4.1.3.) - as well as the *Unified Executive Function Component* (3.4.1.4.).

3.4.1.1. Inhibition

Inhibition or inhibitory control (IC) involves being able to suppress irrelevant or distracting cues and attend to relevant information in order to perform a specific task (see section 2.3.1.). Some have suggested that the key to the bilingual advantage in executive control lies in the bilinguals' enhanced IC function (e.g. Bialystok et al., 2004; Green, 1998; Levy, McVeigh, Marful, & Anderson, 2007). As a result of dual-language activation, bilinguals must constantly control and suppress the unintended language during speech production, resulting in an enhanced ability to suppress irrelevant information and distracting stimuli (Colzato et al., 2008; Costa, 2005; Prior & Mac Whinney, 2010).

A number of experimental paradigms have been used to demonstrate the enhanced IC skills of bilingual children and adults, such as the Stroop (Gathercole et al., 2010), Simon (Bialystok, Martin, & Viswanathan, 2005; Martin-Rhee & Bialystok, 2008), flanker (Emmorey, Luk, Pyers, & Bialystok, 2008; Mezzacappa, 2004; Yang, Shih, & Lust, 2005), anti-saccade (Bialystok, Craik, & Ryan, 2006) and ANT (Carlson & Meltzoff, 2008; Tao et al., 2011) tasks. In each case, bilinguals outperformed monolinguals in some facet of the task (e.g. accuracy and/or timing). Bialystok, Craik and Luk (2008) found that bilinguals displayed reduced inhibitory cost on both the Stroop and Simon tasks, even after controlling for reading speed (automaticity).

Marian and Blumenfeld (2011) found that bilinguals also demonstrated enhanced IC skills with higher accuracy and smaller Stroop effects compared with monolinguals on a non-linguistic version of the Stroop task.
Green (1986, 1993, 1998) was one of the first to provide a strong theoretical argument for how IC was facilitated by the bilingual experience. His IC model was based on the earlier work of Norman and Shallice (1988) who proposed an action model which regulated language through external sensitivity (exogenous) and internal direction (endogenous) as methods of control. Their model also drew a distinction between systems controlling routine (direct control over behaviours) and non-routine behaviour (employment of supervisory attention system, SAS; Shallice & Burgess, 1996). Green also argued that inhibition is not a unitary process and distinguished between active and reactive inhibition. Active inhibition of competing codes relies on the use of a central inhibitory system rather than local inhibitory connections. With experience, this inhibitory system may extend from language codes to non-language codes, which may explain the bilingual advantage on other non-verbal executive tasks. With reactive inhibition (Desimone & Duncan, 1995), the target code can be directly strengthened through the support of a facilitative link between the hypothetical goal system and the target code. By supporting the target code, its activation may not only be increased, but it could also mean that the non-target code may be inhibited via the inhibitory link.

Colzato et al. (2008) have also adopted this view of differing aspects of IC, and argued that bilinguals only show an advantage in the latter, reactive inhibition. They investigated various versions of the IC hypothesis by comparing groups on tasks believed to tap different aspects of IC and suggested that bilinguals may not differ from monolinguals in terms of active inhibition but that they have acquired a stronger ability to maintain their action goals and to use them to bias goal-related information. Indirectly, they proposed that this may lead to reactive inhibitory capabilities for irrelevant information. In other words, bilinguals may not be particularly good inhibitors but, instead, are good at maintaining the task goal, providing stronger support for target representations (Colzato et al., 2008; Prior & MacWhinney, 2010).

The work of both Green and Colzato et al. (2008) along with previous work by Bialystok and colleagues (e.g. 2001, 2004) has sparked a lively debate regarding the level and nature of IC used by bilinguals to suppress the non-target language as well as the influence of bilingualism on tasks of IC which have not shown signs of enhancement through linguistic experience (e.g. Bialystok, 2011; Carlson & Meltzoff, 2008; Paap & Greenberg, 2013). Furthermore, researchers are making distinctions between aspects of IC with little consensus and minimal discussion of the mechanisms behind these classifications, with bilingual advantages shown for some aspects but not others (Martin-Rhee & Bialystok, 2008). For instance, Bunge et al. (2002) differentiate between two types of IC: interference suppression and response inhibition. Recent evidence suggests that, as with active and reactive inhibition, children may show different developmental trajectories for each type. As such, bilinguals may display advantages for certain aspects of IC but not others (e.g. Bialystok et al., 2008). Bialystok and colleagues (2004) proposed different LG effects depending on the type of IC task used. Interference suppression involves attending to the relevant features of a stimulus and avoiding or ignoring irrelevant features when choosing between a two or more stimulus to respond correctly while response inhibition requires avoiding a cued or salient response and may be more of a motoric level of executive control (Nicolay & Poncelet, 2013a; Kroll & Bialystok, 2013). As bilingualism is not proposed to require the withholding of habitual responses in the manner required during response inhibition, no advantages have been found for such tasks, e.g. on the Day-Night task (Bunge et al., 2002; Carlson & Meltzoff, 2008; Colzato et al., 2008; Luk et al., 2010; Martin-Rhee & Bialystok, 2008; Tao et al., 2011). On the other hand, the ability to resolve a perceptual conflict by inhibiting a misleading cue, as in interference suppression, is required during language production and subsequently has been found to be advantaged in bilinguals e.g. on the Stroop task (Bialystok, 2001; Bialystok et al., 2008; Costa et al., 2008, 2009; Green, 1998; Martin-Rhee & Bialystok, 2008; Paradis, 1984). However, recent research has suggested that IC alone cannot account for the bilingual advantage in EF tasks (Bialystok, 2011; Bialystok et al., 2004, Bialystok et al., 2012; Costa et al., 2009; Hilchey & Klein, 2011; Kroll & Bialystok, 2013; Tao et al., 2011; Nicolay & Poncelet, 2013a). Some have proposed that rather than IC advantages on mixed-block tasks (congruent and incongruent stimuli) such as the colour-word Stroop, bilinguals have an enhanced ability to make quick judgments regarding the type of stimuli presented or enhanced conflict monitoring skills (e.g. Bialystok et al., 2011; Costa et al., 2009; Hilchey & Klein, 2011; Hernandez et al., 2010; Hernandez et al., 2012).

In summary, as a result of bilinguals suppressing and inhibiting the unintended language during speech production, certain aspects of IC may be advantaged through the bilingual experience. However, a number of studies have not found significant results in favour of bilinguals on tasks of IC (e.g. Costa et al., 2009; Colzato et al., 2008; Bialystok et al., 2006; Kroll & Bialystok, 2013). Consequently, care must be taken before generalising findings across bilingual groups. Furthermore, there is yet to be clarity regarding the mechanisms causing any LG effects on tasks of IC.

3.4.1.2. Switching

The function of switching has also shown signs of improvement as a result of the bilingual experience (e.g. Bialystok, 2009). Behaviourally, bilinguals have demonstrated a strong ability to switch mental sets as they can be seen shifting with ease and speed from one language to another, making only minimal errors (Muysken, 2000; Myers-Scotton, 2002). The switching function enables bilinguals to control which language they use and assists them in code-switching or readily changing between languages (Hernandez et al., 2001; Price et al., 1999; Rodriguez-Fornells et al., 2005, see section 2.3.2. for switching definitions). Through their practice of day-to-day language switching, it is conceivable that this will benefit bilinguals' switching function (Bialystok, Craik, Klein, & Viswanathan, 2004) and some studies have found that this switching practice transfers to benefits on non-linguistic switching tasks (e.g. Prior & MacWhinney, 2010).

Penfield and Roberts (1959) were some of the first to suggest that subconscious mental switches must be made by bilinguals when attempting to speak in one language or another. Macnamara and Kushnir (1971) took this work further and distinguished between input switching and output switching. In both cases, they proposed that a language system (or subsystem) is either on or off. For example, if a bilingual was translating words from L1 to L2, they can only comprehend input from the L1. Paradis (1981) challenged this assumption both theoretically and empirically by demonstrating that individuals were influenced by the non-selected language or language being switched off (Paradis, 1984; Grosjean, 1988, 1997, 1998). This led to the development of models examining different levels of language activation and control within the bilingual brain.

Most researchers now agree that while speaking, both languages are active within the bilingual brain and neither is completely switched off (see 3.2.2., e.g. Colomé, 2001; Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Poulisse, 1997, 1999). Language-switch costs are cited as evidence for this dual-activation of both languages in bilingual speakers (Meuter and Allport, 1999). If this is the case then bilinguals' ability to switch between languages with ease is somewhat counterintuitive as a large degree of conflict must be present within the bilingual brain. One explanation for the bilingual advantage may be that as a bilingual becomes increasingly proficient in their L2, selecting and switching between languages becomes more practised (Green, 1998). As a result of such practice, Meuter and Allport (1999) proposed that the bilingual advantage in tasks of EF was due to their enhanced cognitive flexibility. Due to persistent switching between languages, bilinguals unconsciously enhance their cognitive flexibility skills, which in turn may lead to higher performance in tasks with high cognitive switching demands. Indeed, the switching delays shown during language switching, mirror the delays found during non-linguistic task switching or 'switch costs' (Meuter, 2005; Monsell, 2003; Prior & Gollan, 2011).

Switching advantages have been shown for bilinguals in a number of studies. For example, Bialystok and Martin (2004) found that bilingual pre-schoolers were successful on the dimensional change card sort task (DCCS; Zelazo, Resnick & Pinon, 1995) earlier than their monolingual counterparts. Prior and MacWhinney (2010) used a dual-task paradigm modelled on the task used by Bialystok, Craik and Ruocco (2006). They assessed monolingual and bilingual adolescents' switching abilities using 'bivalent' stimuli (two competing responses rather than one in the original paradigm) in an attempt to draw a direct comparison between local switch costs and general mixing or monitoring costs. They claimed that their task schema paralleled language selection in bilinguals as competition needed to be resolved before the task was performed. They found a pronounced reduction in switching costs for the bilingual adolescents and concluded that (p. 259):

"...enhanced bilingual executive function has been ascribed to the constant need to select the appropriate language; a process which involves achieving coordinated...activation of the interrelated features of the chosen language".
Another explanation for why bilinguals show advantages in cognitive switching tasks relates to the processes used to switch between the L1 and L2. When the language of response is altered from trial to trial, RT responses are slowed compared with switching between trials with no language switch (Meuter & Allport, 1999). Interpretations of such findings suggest that participants must establish a new language set and overcome the language set inertia of the language used in the previous trial. Such processes are very similar to the ones described in the general task-switching literature therefore bilinguals have become more practiced in overcoming these types of challenges (Meiran, Chorev, & Sapir, 2000; Prior & MacWhinney, 2010). Furthermore, neuroimaging research has shown an overlap between areas of the brain used during task and language switching (Abutalebi & Green, 2007; Garbin et al., 2010; Prior & Gollan, 2011).

Using a meta-analysis Luk and colleagues (2012) identified the neural correlates of bilingual cognitive control in language switching. Most of the areas active during language switching were located in the frontal regions of the brain. Their findings were also in line with research by Abutalebi and Green (2008) who noted that a number of areas responsible for controlling switching abilities were active in the bilingual brain during language switching tasks e.g. the left dorsolateral prefrontal cortex (DLPFC), the anterior cingulate cortex (ACC), the caudate nucleus, and bilateral supramarginal gyri (SMG). These brain areas are also utilised during EF tasks, particularly the DLPFC and the bilateral SMG (Toro, Fox, & Paus, 2008). The ACC and caudate nucleas have also shown activation during tasks of cognitive control (see Luk et al., 2012). Following their meta-analysis, and in contrast to Abutalebi and Green's previous research, Luk et al. (2012) found enhanced activation in bilingual brain areas which have also been activated during the performance of demanding tasks of response control, performance monitoring, error detection, feedback, and related processes (e.g. left side frontal lobe regions). Such neuroimaging studies provide an insight into the mechanisms behind enhanced executive control in bilinguals, who must continually utilise their switching functions to shift between language representations.

Many researchers have argued that switching skills play a key role in bilinguals' ability to efficiently communicate in either language. Bialystok et al. (2011) argued that the tasks most likely to show advantages for bilinguals are those resembling the types of switches that bilinguals must make on a day-to-day basis (e.g. changing modalities on the DCCS). Consequently, findings have revealed a switching advantage for bilinguals on a range of tasks (e.g. Bialystok et al., 2009; Kroll et al., 2008). However, some researchers maintain that processes beyond switching are also necessary for bilinguals to overcome the competition and conflict they face when selecting a target language, for example, IC or unified EF skills (Kroll & Bialystok, 2013; Bialystok et al., 2012). Such difficulties of isolating specific EF components using complex cognitive tasks makes it difficult for researchers to conclude the benefits of bilingualism for specific aspects of EF such as switching (e.g. Paap & Greenberg, 2013), particularly as research often uses a combination of languages to make up bilingual groups (Bialystok et al., 2008; Prior & MacWhinney, 2010) and examine specific EF components rather than a combination of functions. Furthermore, neuroimaging studies have shown that language switching utilises a number of brain regions rather than one region in particular, preventing conclusions of a specific function resulting in a bilingual advantage (Luk et al., 2012; Wang et al., 2007). Prior and Gollan (2011) found that the reduced task-switching costs were not present in their bilingual groups, and that proficiency and SES factors played an important role in their findings with Spanish-English bilinguals outperforming Mandarin-English bilinguals with lower proficiency in the L2 highlighting the importance of factors beyond LG categorisations.

3.4.1.3. Working Memory

In terms of the working memory (WM) function, there continues to be debate surrounding the effects of bilingualism on this component. At least two competing and contrasting hypotheses exist regarding the effect of bilingualism on WM.

The first hypothesis suggests that due to bilinguals needing to constantly manage their two languages, added demands are placed on the WM system. Bilinguals' dual-language activation may therefore result in less efficient information processing during WM tasks due to the high cognitive load (Van Merrienboer & Sweller, 2005).

The alternative hypothesis suggests that as a result of bilinguals' well-developed ability to inhibit one language while using another, this may increase the efficiency of their WM function. In other words, bilinguals' WM resources are, effectively, being mediated by their IC processes (Bialystok et al., 2004; Bialystok, Craik, & Luk, 2008; Fernandes, Craik, Bialystok & Kreuger, 2007; Just & Carpenter, 1992; Michael & Gollan, 2005; Rosen & Engle, 1997).
These competing hypotheses have led to inconclusive findings with regard to the effect of bilingualism on WM performance, and results seem to depend on the nature of the task (Adesope et al., 2010; Bialystok et al., 2008). A number of studies have found no LG differences in WM performance between monolinguals and bilinguals (e.g. Bialystok & Feng, 2010; Bonifacci, Giombini, Bellochhi, & Contento, 2011; Engel de Abreau, 2011), while in tasks requiring a high degree of attentional control, bilinguals do appear to show enhanced WM capacity compared with monolinguals (e.g. Bialystok et al., 2004; Engle, 2002; Kane, Bleckley, Conway, & Engle, 2001). Conversely, for tasks requiring less attentional control or attention-aided tasks, this bilingual advantage often disappears (Yang, Yang, Ceci, & Wang, 2005). A recent longitudinal study by Engel de Abreu (2011) found no WM advantage emerged in 6 - 8 year old bilinguals tested over a three-year period. This study suggested that the processes required for WM and IQ are different than those required during EF control tasks and that the complex nature of EFs means that bilingualism may affect aspects of but not all EF processes. De Abreu also proposed that WM skills in sequential bilinguals (who have acquired their L2 at a later age) should be examined, as language switching may have become an automatic process in her sample of simultaneous bilinguals. Similar studies have also found no bilingual advantage on WM skills, for example: Bajo, Padilla, and Padilla (2000), Bialystok and colleagues (2008) and Namazi and Thortardottir (2010).

In conclusion, studies examining the WM performance of bilinguals have been inconclusive regarding whether or not WM, like inhibition and switching, is enhanced through bilingualism but recent studies seem to suggest that no bilingual advantage should be expected for this function. Furthermore, the WM question may relate directly back to attentional control and how much attention or unified EF skills are required for each task.

3.4.1.4. Unified executive function component

While the previous sections attempted to identify specific functions of the EF influenced by bilingualism, recent research has begun to shift towards a more unified or co-ordinated explanation for the EF advantages found within certain bilingual groups. Rather than bilinguals showing advantages in specific EF domains, researchers are beginning to consider that the attentional control advantage may lie in bilinguals' ability to co-ordinate and combine their EF skills in order to perform complex cognitive tasks more effectively (e.g. Costa et al., 2009; Kroll & Bialystok, 2013; Bialystok, 2011). Others have suggested that bilinguals have advantages in a broader concept, which have been called conflict monitoring (e.g. Costa et al., 2009; Hilchey & Klein, 2011). Morales, Calvo and Bialystok (2012) commented that many of the EF tasks used in research to demonstrate the benefits of bilingualism should more accurately be described as tasks of unified EF components.

In a now widely cited study, Carlson and Meltzoff (2008) discussed the benefits of bilingualism relating to specific EF advantages. They claimed that their study had uncovered an effect of bilingualism on the EF component of IC. The study compared native and immersion bilinguals with monolingual controls on a battery of EF tasks. However, LG effects were only displayed on two tasks: the Advanced Dimensional Change Card Sort (Zelazo et al., 1996) and the Comprehensive Test of Non-Verbal Intelligence (C-TONI; Hammill, Pearson, & Wiederholt, 1997). Both of these tasks, rather than testing for IC, are complex cognitive tasks requiring the use of a number of functions. As non-verbal IQ is often used as a control measure in studies of bilingualism it seems unusual to cite the C-TONI task as a measure of IC only. Furthermore, their factor analysis confirmed this with LG differences in favour of the bilingual group found for the conflict tasks rather than the delay or IC-specific tasks (e.g. ANT, delay of gratification, Simon). Although they concluded that their study demonstrated an EF advantage for bilinguals, they failed to provide an explanation for the potential mechanisms causing such effects and did not acknowledge that the tasks used as evidence for LG effects required the use of more than one specific EF component. In another study of bilingual effects on EF components, Colzato and colleagues (2008) concluded that the bilinguals were not more efficient at inhibiting unwanted responses but instead were more capable of selectively focusing their attention and had superior skills in general cognitive control (see also Adesope et al., 2010; Bialystok, 1999; Bialystok & Majumder, 1998; Kroll et al., 2008).

Another argument for why a more unified EF component may be the area most affected through bilingualism is that on a day-to-day basis bilinguals will rarely use their EFs in isolation. Therefore, it does not make sense that they should be advantaged in isolated EF components. For instance, during speech production bilinguals must decide on which language to use (L1/L2) while controlling two active language representations in mind (WM), attend to the target language while ignoring competition from the non-intended language (IC) and change between languages depending on context (switching). These skills may result in overall EF advantages for bilinguals rather than an advantage in any single EF component (Bialystok, 2011; Kroll & Bialystok, 2013). Recent trends in EF research have also recognised the importance of an underlying EF mechanism used to combine skills from each component (e.g. Best & Miller, 2010; Garon et al., 2008; Lehto et al., 2003). Bialystok (2011) used a dual-task paradigm (adapted from Bialystok et al., 2006) to assess bilinguals' and monolinguals' ability to manage complex EF demands. Bilinguals showed an advantage in overall performance on the visual but not the verbal version of the task. Morales et al. (2012) also found an advantage in bilingual participants' ability to co-ordinate their EFs during the most complex task stimuli in a task of complex working memory and executive control.

In summary, recent literature in the field of bilingualism and cognitive control (EF) is beginning to move in line with the current general EF research (e.g. Miyake & Friedman, 2012, see section 2.3.4.). These studies suggest that rather than specific components of the EF being positively affected by the bilingual experience, bilinguals may have a general advantage in their overall ability to monitor attention and co-ordinate the specific functions of executive control as a result of having to monitor and control two active languages on a daily basis (e.g. Bialystok, 2011; Costa et al., 2009; Hilchey & Klein, 2011; Morale et al., 2012).

3.5. SUMMARY

Sections 3.3. and 3.4. discussed the advantages and disadvantages associated with the bilingual experience. But how can these mixed findings be reconciled with one another? The term “bilingual paradox” has been used to describe the advantages and disadvantages produced by the bilingual experience (Petitto & Kovelman, 2003; Sorace, 2011). In one respect, the ease with which children can acquire and develop their L2 is compelling, yet there are still questions concerning the language delays and confusion bilingual children display during language production and processing in particular (Sorace, 2011). This summary tries to interpret these contradictory results.

While performing a linguistic task, participants must often make speedy decisions on dimensions such as register, collocation and synonymy and the bilingual speaker is faced with the added challenge of choosing the correct target language (Barac & Bialystok, 2012; Bialystok et al., 2012). If dual-activation is an accurate hypothesis, these languages will have to compete for selection, resulting in the timing delays often shown by bilinguals during speeded or reaction time tasks (e.g. Abutalebi & Green, 2007; Bialystok, 2009; Bialystok et al., 2009; Costa et al., 2009; Meuter & Allport, 2009). While anecdotally, language selection is not an obvious problem for proficient bilinguals, it may make ordinary linguistic processing more effortful, leading to disadvantages in certain tasks of linguistic skill (Bialystok et al., 2012). On the other hand, bilinguals have shown advantages in tasks of non-verbal processing (e.g. Bialystok, 2001; 2011; Carlson & Meltzoff, 2008; Engel et al., 2012). These findings may also be the by-product of having both languages active within the brain (Abutalebi et al., 2012; Green, 1998; Luk et al., 2010; Luk et al., 2011). As proficient bilinguals must constantly monitor and control both of their languages, they should become practiced in dealing with these internal cognitive conflicts (Luk & Bialystok, 2013). The cognitive mechanism most attuned to dealing with these internal conflicts is the EF. An example of how the dual-language conflict may serve to improve EF skill is taken from the function of IC. Some have suggested that the selection of the target language for the bilingual requires global inhibition of the non-target language or local inhibition of an irrelevant, language specific distractor, e.g. translation equivalent. As IC is also a function of the overarching EF system, bilinguals' practice with inhibiting the non-intended target language may result in advantages for tasks of non-verbal IC (Bialystok et al., 2012; see 3.4.1.1.).

Bialystok and Feng (2009) were interested in examining issues relating to the bilingual paradox. Using a proactive interference (PI) paradigm, they discovered that the bridge between verbal and non-verbal discrepancies was language proficiency (LP) and once vocabulary level was controlled, bilinguals outperformed monolinguals. They concluded that performance on PI depends on both verbal ability (to access semantic memory and hold words in mind for later recall) and on executive control (to monitor the words and update the lists presented and avoid repetition).

In conclusion, there appears to be little doubt amongst researchers that there are positive effects of the bilingual experience on children's cognitive development. A key finding within this research is that in order for bilinguals to lexicalize or produce words in the intended language (ignoring the unintended language) a degree of language control is needed (Costa et al., 2006, 2009; Finkbeiner et al., 2006; Gollan et al., 2005; Green, 1998). However, agreement has not been reached regarding the mechanisms used to do this. In fact, it may be that different language control mechanisms are employed, depending on individuals' level of proficiency in the L1 and L2 (Costa et al., 2006). Furthermore, there is still much discrepancy regarding the classification and interpretation of the EFs and researchers have yet to identify which of the EFs is most affected by the bilingual experience (Bialystok, 2009).

Developmental research has found that EF components may emerge at different ages; therefore, bilingualism may have differing influences at different stages of development (Carlson, 2003; Diamond, 2002; cited in Bialystok & Viswanathan, 2009). This thesis aims to examine some of these issues further by employing a battery of EF tasks and by using a longitudinal design to explore issues of cognitive development and LP to help identify some of the control mechanisms employed by language groups.

CHAPTER FOUR
METHODOLOGY FOR THESIS STUDIES

4.1. INTRODUCTION

The following chapter outlines the design and procedure used in the subsequent experimental chapters of this thesis (see Chapters 5 and 6). As the majority of tasks used classic paradigms and standardised assessments, procedures and methods of administration remained constant for each testing phase (i.e. at Time 1, Time 2 and Time 3) and between studies. Therefore, this methodology chapter can be used as a reference point for how each task was administered within each of the experimental chapters to follow. Methodologies specific to each study will be briefly outlined here but will be presented more fully in the relevant chapters (e.g. socioeconomic status and proficiency categorisations). This chapter focuses on the general administrative and procedural protocols that remained constant across time and studies.

4.2. DESIGN

Chapter 5 employed a longitudinal design to compare children from monolingual and bilingual backgrounds and from middle and low socioeconomic backgrounds. This mixed design had two between subjects variables, each with two levels (Language Group: *Monolingual and Bilingual* and Socioeconomic Status: *Low and Middle*) as well as one within-subjects variable, Time, which had three levels (*Time 1*, *Time 2* and *Time 3*). The longitudinal studies in this paper consisted of three testing phases, each one approximately 1 year apart. At Time 3, a further cross-sectional study was employed, comparing children's Time 3 results with age-matched children from a Gaeltacht region of Ireland (see section 1.4.3.) or '*early/native'* bilinguals. This cross-section of participants was recruited to compare differing levels of language proficiencies and will be discussed in detail in Chapter 6.

4.3. PARTICIPANTS

In total, 147 children were recruited at the initial phase of testing. A number of children (*n* = 25) were excluded from the analysis as a result of having general (e.g. dyslexia, autism spectrum disorder) and/or specific (e.g. reading or maths delays) learning impairments. Children who had high proficiency in a language other than English or Irish (*n* = 5) were also excluded. Due to the nature of longitudinal work, a small percentage of attrition occurred between Time 1 and Time 3, although overall the retention rate of participants was high (96%) compared with other longitudinal studies (e.g. Capaldi & Patterson, 1987).

4.3.1. Language Groups

Monolingual children were taught all lessons (with the exception of their 3.5 hour, weekly, Irish language class) through the medium of English and spoke only English at home[1].

Bilingual children were taught through an early, total immersion model in Irish (see section 1.4.3. for more details). Some children in the immersion education (IE) and Gaeltacht bilingual groups spoke Irish within the home (e.g. with one parent/sibling). Most children in the immersion bilingual group spoke only English in the home but a higher frequency of Gaeltacht bilingual children spoke Irish within the home (see 6.5.1.). The degree of Irish spoken within the home was assessed through language proficiency (LP) questionnaires administered to parents and children at Time 3 of testing (see Appendix I and III respectively and 6.3.1. for further discussion). Degree of English and Irish proficiency within the classroom was assessed through LP questionnaires administered to teachers at Time 3 of testing (Appendix II).

The participants recruited from the Gaeltacht, or *'native bilinguals'* were also educated through an early total immersion Irish model. These children were recruited from a small area of Ireland, where Irish is used as a mode of communication within

[1] Certain children in both the bilingual and monolingual groups spoke a language other than Irish or English at home. These children were excluded.

the community. Gaeltacht group demographics are described in more detail in sections 1.4.3. and 6.5.1.. Compared with the bilinguals in the longitudinal studies, a significantly higher proportion of these children spoke either Irish only or both Irish and English with at least one parent within the home. Chapter 6 explores the similarities and differences between Gaeltacht pupils compared with bilinguals and monolinguals of the same age in more urban, English dominant areas of Ireland.

Table 1 *Participant group information*

	Time 1	Time 2	Time 3
Class Group	3rd Class (P5)	4th Class (P6)	5th Class (P7)
Age	8-9 years	9-10years	10-11years
Total Children	126	123	140
Monolingual	56	55	54
Bilingual	70	68	67
Low-SES	40	40	40
Mid-SES	86	83	81
Gaeltacht Comparison	-	-	19

4.3.2. Socioeconomic Status

Rather than control for socioeconomic status (SES) as many researchers in the past have done, this thesis aimed to compare any differences in children from differing SES backgrounds. Initially this was done through proxy or aggregated measures of SES. Here, schools were approached based on their geographical locations: middle income catchment areas and low-income catchment areas of Dublin city. Schools were then chosen according to the Irish Department of Education and Skills (An Roinn Oideachais agus Scileanna) standardised grading system for primary schools. Schools classified as disadvantaged or *Deis* and which required extra financial and social supports were selected for low-SES groups. Mid-SES schools received no additional governmental support.

In the Republic of Ireland, primary and secondary schools are ranked by the Department of Education and Learning on a banded system called '*DEIS*', which stands for 'Delivering Equality of Opportunity in Schools'. DEIS provides a standardised system for identifying levels of disadvantage and utilises an integrated School Support Programme (SSP; Department of Education and Skills). In Ireland, almost 670 primary schools are currently included under the programme (342 urban/town schools and 327 rural primary schools). The DEIS Action Plan (2005) outlines the support disadvantaged schools received and defines educational disadvantage as:
"…the impediments to education arising from social or economic disadvantage which prevents students from deriving appropriate benefit from education in schools."

Schools in the DEIS scheme are divided on an urban-rural, band-rated system, with *Band 1* schools requiring the highest level of social and educational support from the scheme. Two '*Band-1*' urban schools were selected for assignment to the low-SES groups. Therefore, both schools received a similar range of extra governmental support including: a smaller than average pupil-teacher ratio (PTR) of 20:1, financial support in the form of a DEIS grant (based on levels of disadvantage and enrolment figures), access to home-schools community liaison services, access to schools meals programme, access to a range of support under the school completion programme, literacy/numeracy support such as Reading Recovery, Maths Recovery, First Steps, Ready Steady Go Maths, planning support, a range of professional development support as well as additional funding under the School Books Grant Scheme (Department of Education and Skills). On the other hand, middle-socioeconomic (Mid-SES) schools received no extra support from the Irish government and were situated in areas known to have a more balanced proportion of low-, middle- and high-income families (see section 5.2.1. for more details).

At Time 3 of testing it was decided that more detail regarding the children's individual background demographics would be beneficial in order to more accurately determine children's SES. Therefore, background questionnaires were distributed to the children's parents/guardians in order to obtain information about parental occupations and highest level of academic attainment (Appendix I).

4.3.3. Language Proficiency

The linguistic backgrounds of the bilingual children were assessed through proficiency questionnaires distributed to the children's main caregivers at Time 3 of the longitudinal study and to the Gaeltacht participants. The questionnaires required parents/guardians to rate the degree of English/Irish used by children and the children's level of skill in each language including reading, writing, listening, speaking (Appendix I). Parents/guardians were instructed to mark the child's level of proficiency using a 10cm horizontal line ranging from no ability at all (e.g. *He/she cannot read at all* = 0cm) to exceptionally high ability (e.g. *Exceptionally high reading ability* = 10cm; Fig. 4.).

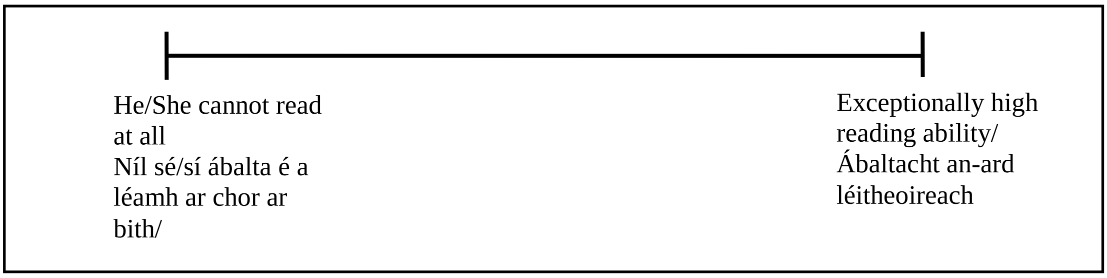


Figure 4 *Example of the LP questionnaire 'reading' item for parents/guardians*

Questionnaires were also administered to each of the participants' teachers at Time 3 (Appendix II) with questions relating to children's English and Irish language ability within the academic setting (i.e. reading, writing, comprehension and speaking). Finally, self-rated questionnaires were administered to immersion and native bilingual children themselves at Time 3. Ratings for these questionnaires used a similar rating scale with happy and sad faces, representing children's beliefs regarding their abilities in each language (Fig. 5.). Assistance was provided for each child, individually, when filling out these questionnaires (Appendix III). Levels of proficiency for each group are examined and discussed in Chapter 6.

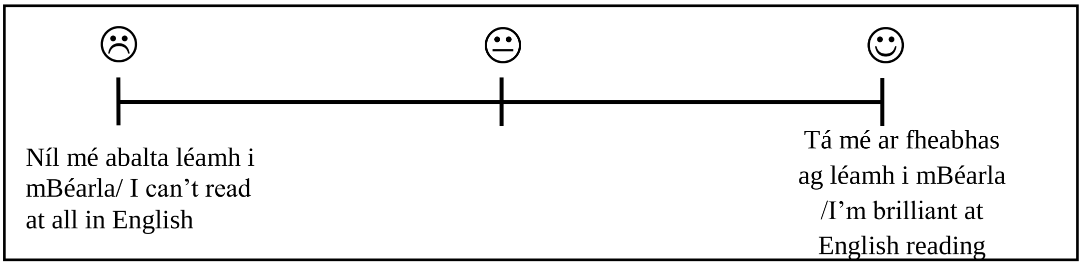


Figure 5 *Example of the LP questionnaire ‘reading’ item for children*

4.4. PROCEDURE

The following procedure outlines how each task was administered, at each time point and across groups. However, it must be noted that not all tasks were administered at each time point (see Table 2.). For instance, control measures {English receptive vocabulary (PPVT-4) and non-verbal IQ (Raven’s SPM)} were administered at Time 1 only, while EF assessments, e.g. the colour-word Stroop, were administered at all three time points (i.e. Time 1, Time 2 and Time 3).

For all tasks (apart from the Raven’s SPM), participants were tested on a one-to-one basis and in a quiet space within the school and outside of the classroom. Tasks were presented in a fixed order across all participants. For tasks that were administered in both English and Irish (e.g. colour-word Stroop) the order of administration and language of instruction was counterbalanced. Each testing session lasted approximately 1 hour, excluding administration of the Raven’s SPM (RSPM). Due to the level of difficulty associated with the Wisconsin Card Sorting Test (WCST) for children, this task was administered on a separate day from other tasks.

Given the length and nature of the task, the RSPM was administered to each child in a larger group setting. For the mid-SES groups the RSPM was administered within the classroom, as one of the daily classroom activities. For the low-SES groups, the RSPM was administered in smaller groups of approximately 6-8 children, in a quiet area of the school and outside of the classroom. Smaller groups were used here to help maintain the children’s focus in what is a relatively long task, requiring a high level of concentration (approximately 40 minutes).

Table 2 *Tasks administered at Time 1, Time 2 and Time 3*

Task	Time 1	Time 2	Time 3
Peabody Picture Vocabulary Scale	Yes	No	No
Raven's SPM	No	Yes	No
Creature Count	Yes	Yes	Yes
Opposite Worlds	Yes	Yes	Yes
Colour-Word Stroop	Yes	Yes	Yes
Trail Making Test	Yes	Yes	Yes
Wisconsin Card Sorting Task	No	No	Yes
Working Memory Test Battery-for Children	No	No	Yes

4.5. TASK ADMINISTRATION

4.5.1. Peabody Picture Vocabulary Test-4th Edition

Children's English receptive (hearing) vocabulary was assessed using the Peabody Picture Vocabulary Test-4 (PPVT; Form B) (Dunn & Dunn, 2007). The PPVT scale is an untimed test of English receptive vocabulary rather than a speed test. Because sets that are too easy or too difficult are not given, the form can usually be administered in about 10-15 minutes.

Two training items were used to teach the children the nature of the task and how to respond before testing began. The administrator then said a word aloud to the child and asked that they point to, or say the number of the picture that best represented the meaning of that word from a simple, four-picture page layout.

Each participant was tested only over his or her *critical range of items*, that is, the items that are of appropriate difficulty for the participant. This range included the Basel Set, the Ceiling Set, and all item sets in between. Following the two training items the administrator began with a Start Item recommended for the participant's age. These starting points were established so that about 85% of the examinees in the designated age group would meet the Basal Set criterion (11 or 12 items correct) in the first administration set. The Basal Set for a participant is the lowest item set administered that contains one or zero errors. The administrator must establish the Basal Set before testing can continue with more difficult sets. If one or more errors are made in the first set administered, the examiner must complete the administration of that set and then drop back to the previous set (if there is one) and administer all 12 items in that set. Continuation of testing in this manner continues until the Basal Set Rule is met.

After establishing a participant's Basal Set, testing continues forward until establishing the Ceiling Set. The Ceiling Set for a participant is the highest set of items administered containing eight or more errors. Once the Ceiling Set was established, testing discontinued as the individual had been tested over his or her critical range.

Scoring the PPVT-4

The raw score for the test was obtained by subtracting from the number of Ceiling Item, the participant's total number of errors over his or her critical range (i.e. from the Basal Item to the Ceiling Item). Standard scores and percentile ranks were then calculated from the participant's raw score and age equivalent. The PPVT-4 age norms are based on a representative sample of 3, 540 people aged from 2 years 6 months through to 90 years and older from across the United States (see Dunn & Dunn, 2007). Furthermore, the PPVT-4 standard score scale is the same as the scale used in many other tests, allowing for direct comparison of PPVT-4 with scores obtained on tests of language, achievement, and ability.

Reliability and validity of the PPVT-4

Dunn and Dunn (1997) reported that test-retest reliabilities as high with a median of .93 with a range of .92 to .96 (this highlights its resistance to factors such as fatigue and illness) although this task was only assessed at Time 1 of the longitudinal study. They also give details regarding the validity of the PPVT-4 which has been shown to correlate highly with tests of oral and written language skills.

4.5.2. The Raven's Standard Progressive Matrices

The Raven's Standardised Progressive Matrices test (RSPM; Raven, Raven, & Court, 2004) was constructed to measure the eductive component of intelligence (*g*) as defined in Spearman's theory of cognitive ability. As Raven and colleagues (2004, pp. SPM I) explain, eductive ability is the ability to… *"...forge new insights, the ability to discern meaning in confusion, the ability to perceive, and the ability to identify relationships. Since perception is primarily a conceptual process, the essential feature of eductive ability is the ability to generate new, largely non-verbal, concepts which make it possible to think clearly".*

The test is made up of five Sets, or series, of diagrammatic puzzles exhibiting serial change in two dimensions simultaneously. Each puzzle has a part missing, which the participant has to find among the options provided. The standard test consists of 60 problems divided into five Sets (A, B, C, D, and E), each made up of 12 problems with six to eight possible choices for answers. In each Set the first problem is intended to be self-evident. The problems which follow build on the argument of those that have gone before and become progressively more difficult. Children indicated their responses by circling or marking the appropriate answer on individual answer sheets. To ensure sustained interest and freedom from fatigue, each problem is boldly presented and as far as possible, is pleasing to look at.

Scoring the RSPM

Participants' answer sheets are corrected by the examiner and given a total score out of 60 trials. Raw scores from the RSPM are converted to age-scaled equivalents rather than means and standardisations. In other words, children's raw scores (marked out of 60) correspond to the 5^{th}, 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles, according to their age in the 1979 British standardisation group (Raven, Court, & Raven, 2000, p. 78). The raw scores at the 50^{th} percentile are an approximate measure of the means of each age group and allow comparison between different standardisations.

Reliability and validity of the RSPM

Reliability of the RSPM has been assessed using a number of age-ranges, clinical and non-clinical samples (Court and Raven, 1995). The test-retest reliability of the RSPM has been estimated at .96 (Burke, 1972). The general trend is of good reliability in terms of internal consistency and retest reliability. The internal consistencies and internal reliability of the RSPM has been reported as high by a number of studies (e.g. U.S. standardisation, 1993).

Concurrent validity between the Wechsler Adult Intelligence Scale Full Scale IQ and the SPM has been estimated at .57 (McLaurin & Farrar, 1973) and correlations have been high with tests of non-verbal IQ rather than verbal and the RSPM (e.g. Vincent & Cox, 1974). Predictive validity is usually assessed using scholastic achievement assessed sometime after administration of the SPM. Validity coefficients reported in studies with English and non-English-speaking children and adolescents generally range up to about .70 (Elley & MacArthur, 1962).

The RSPM has been described as one of the purest and best measures of *g* or general intellectual functioning available. Evidence for this claim comes from several factor-analytic studies involving both children and adults. Studies investigating these effects with British children (Emmett, 1949; Gittins, 1952; Nisbet, 1953) reveal high loadings of up to .83 on *g*. In the U.S. a loading of .81 was reported (Zagar, Arbit, & Friedland, 1980). Studies have also shown no loading on verbal-educational and numerical ability factors.

4.5.3. Test of Everyday Attention for Children

The Test of Everyday Attention for Children (TEA-Ch; Manly, Robertson, Anderson, & Nimmo-Smith, 1999) is a standardised and normed clinical battery for children and adolescents aged 6-16 years. The task includes nine subtests of children's attentional control (EF) abilities, two of which were administered for this thesis, subtest 3: '*Creature Counting'* task and subtest 8: '*Opposite Worlds'* task. Both subtests are believed to assess participants' attentional control and cognitive switching abilities. In other words, these subsets are believed to measure how well children can control their attention in order to achieve goals that will be useful for them to complete the task. Each subtest provides a standard age-scaled score (M = 10, SD = 3) based on the normative sample. Age-scaled scores are converted to percentiles (like in the RSPM) and are presented in Table 3. TEA-Ch was adapted for use with 6 to 16-year-olds from the original Test of Everyday Attention (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1995) and normative data have also been reported from 293 healthy UK children (Manly et al., 2001). These data allow for relative assessment across different attentional capacities.

There is often considerable overlap between terms such as '*executive function*' (EF) and '*attentional control*'. In assessment, a general difference has been that tasks used to tap EF have employed more complex tasks, which may have more than one solution, emphasising the co-ordination of EF skills, or requires planning to resolve them. For these reasons, Manly et al. (1999) have chosen the term 'attentional control/switching' to describe the Opposite Worlds (OW) and Creature Count (CC) tasks from the TEA-Ch battery. Of the nine TEA-ch subtests, several are believed to implicate the EFs (Walk, Don't Walk, Creature Count, Opposite Worlds and Sky Search Dual Task). Opposite Worlds is said to assess the EF of IC while Creature Count is believed to assess the more general construct of cognitive flexibility or switching.

Table 3 *Age-scaled scores and percentile band conversions*

Age-Scaled Score	Percentile Band
19	>99.8
18	99.4-99.8
17	98.5-99.4
16	96.7-98.5
15	93.3-96.7
14	87.8-93.3
13	79.8-87.8
12	69.2-79.8
11	56.6-69.2
10	43.4-56.6
9	30.9-43.4
8	20.2-30.9
7	12.2-20.2
6	6.7-12.2
5	3.3-6.7
4	1.5-3.3
3	0.6-1.5
2	0.2-0.6
1	<.02

4.5.3.1. Creature Count

The CC subtest aimed to test children's attentional control and switching abilities. Breaking from doing one task to begin another or changing the way a task is performed is generally associated with a delay before optimal levels of performance are achieved. By using two simple skills, counting up and counting down, and making explicit the occasions when a switch is needed, this test provides a relatively simple measure of this capacity.

Children had to repeatedly switch between counting upwards and counting downwards. Each trial consists of a page with between 9 and 21 'creatures' placed along a path. Between 3 and 6 arrows pointing either up or down are interspersed among these creatures on each trail. Participants were asked to count these creatures along the path. If they come across an arrow the direction in which they were counting either upwards or downwards was changed, as directed by the arrow (Fig. 6.). At the end of the path, they must recite the number at the final creature. This number is recorded on scoring sheets. Participants received an accuracy score (number of trials correct) and, if at least 3 out of 7 trials are correct receive a total timing score. Therefore, both time taken and accuracy were scored in this subtest. Prior to testing, it was established that children could count from one to twelve and from twelve down to one. All participants in this sample had no difficulty with counting and so subtests could be administered to all participants. As it was a difficult task to explain, the examiner took the children through the trial task carefully to ensure their understanding. Participants were instructed that this was a game about counting and will require them to count up, *"like one…two…three, and to count down, like three…two…one. We will be counting these creatures in their burrow and we always follow the burrow around from the top to the bottom like this"* (administrator outlined one of the burrows, traced with their finger).

Instructions continue: *"these arrows tell you the direction in which you have to count. So we start off counting up from one. Follow my finger…one, two, three, four, five, six…then the arrow tells us to start counting down from six…so it would be five, four, three, two…then the arrow tells us to start counting up again from two…three, four, five. So the answer at the end is five. Watch me do that again and notice that when I come to the arrows I say 'up' or 'down' to remind myself of what they mean. OK. Now you try".* Following these instructions, children completed two practice trials with reminders from the examiner if errors were made. Once there was satisfaction that children understood what was required of the task, testing could begin. No further help was provided.

Scoring the Creature Count

Stopwatch timing began when the participant counted '*one*' and stopped when the final number was said. Self-correction of a verbal response was permitted. The task involves 7 trials (or trails), each with a unique number of switches (ranging from 3 to 5 switches). The time taken to complete each trial is taken only if the participant completes the trial without error. If errors are made within a trial then the trial is not counted in the calculation of participants' timing score. The total time taken for all correct trials is calculated by dividing the total time spent on correct trials by the number of arrows in those trials. This generated a 'Timing Score', which was then standardised according to the TEA-Ch (Manly et al., 1999) manual.

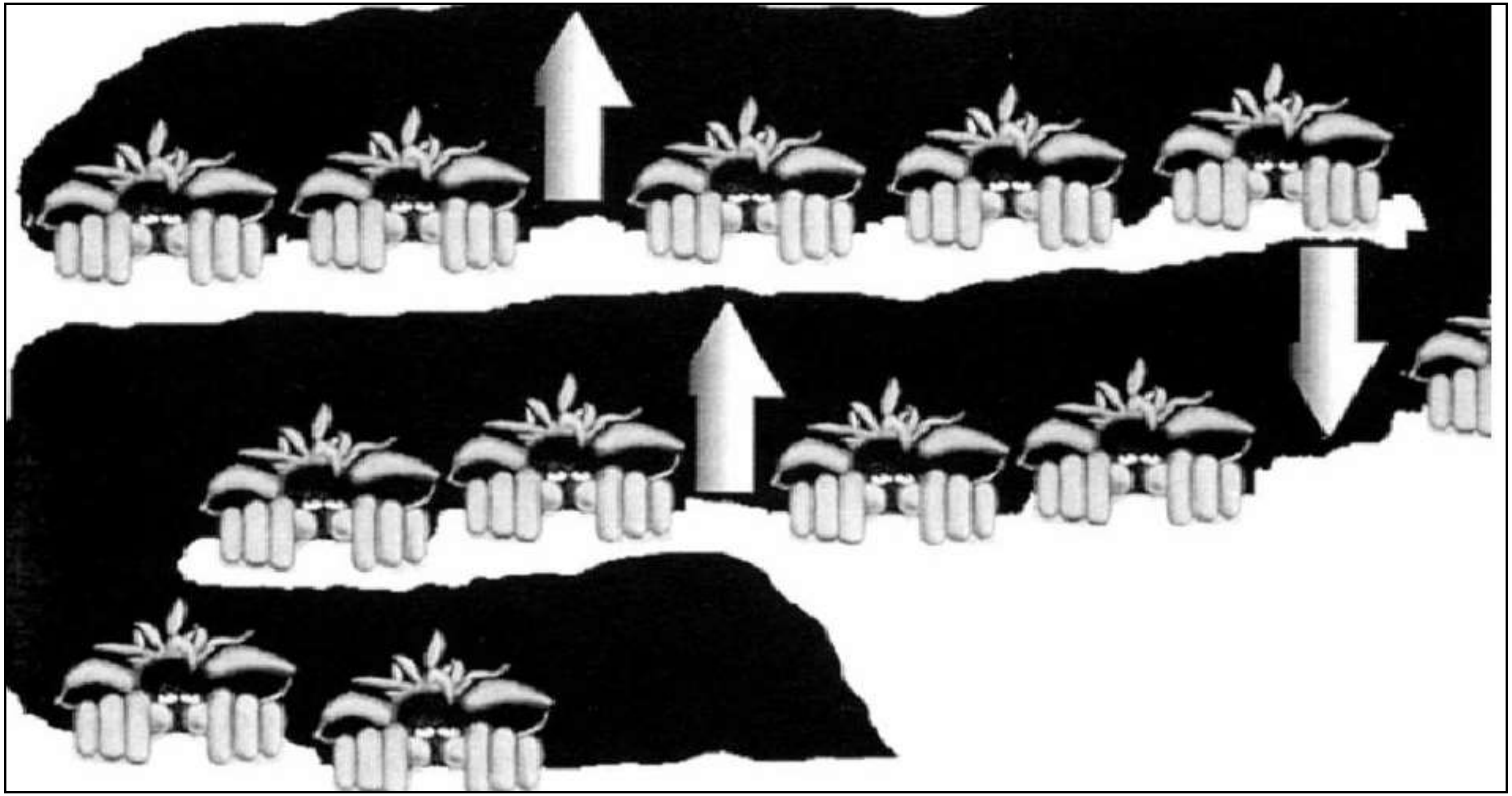


Figure 6 *Creature Count task example*

4.5.3.2. Opposite Worlds

The OW subtest aimed to test the children's attentional control (EF) and specifically their IC function. Performing a task in a novel way, particularly when there is a much more routine way of performing it, is an ability that has been associated with a higher-level of EF skill. This skill is emphasised in the OW task. It should take participants significantly longer to complete the Opposite Worlds (OW; incongruent) condition of this task than the Same World (SW; congruent) condition. As both task conditions (congruent and incongruent) require switching between numbers 1 and 2, the extra time taken to perform the OW trials (incongruent) reflects the timing 'cost' of producing a non-obvious verbal response ('two' for 1 and 'one' for 2) and does not simply reflect demands of task switching. This cognitive cost was examined further by calculating the difference in time taken to complete the OW over SW trials, calculated as the 'Worlds' Difference' condition. This was not a condition present in the Manly and colleagues (1991) version of the task but is considered similar to the Stroop effect. In this respect the OW and Worlds Difference conditions examined the ability to suppress an automatic or 'prepotent' verbal response or participants' IC abilities. The OW task has similarities with common measures of IC including the 'day and night task' (Gerstadt, Hong, & Diamond, 1994; Passler, Isaac, & Hynd, 1985). In the 'day and night task' the verbal response '*day*' has to be given to a dark card or picture of the moon, while the response '*night*' had to be given to a light card or picture of the sun (Passler, Isaac, & Hynd, 1985; Gerstadt, Hong, & Diamond, 1994).

The task consisted of four pages each with a long sequence of squares, numbered '1' or '2'. Participants had to read along the line of numbers, saying the appropriate number for the condition as quickly as possible. There were three task conditions: a SW or congruent condition, and an OW or incongruent condition and a Worlds Difference condition, assessing the difference in reaction times (RTs) between the congruent and incongruent conditions. The SW condition where children followed the sequence naming the digits '1' and '2' in the manner they appear. In the OW trials, participants performed the sequence saying the opposite of each number: 'one' when they saw 2 and 'two' when they saw 1 (Fig. 7.). The speed at which the children performed this cognitive reversal is the crucial measure from this subtest.

The sum of time taken to complete each of the two SW and OW trials was converted to an age-scaled score for each condition. For the Worlds Difference scores, each participant's total raw timing scores from the two Same World trials was subtracted from their total OW raw timing scores, giving a total reaction time (RT):

Worlds Difference = Sum of Opposite Worlds – Sum of Same World

Two practice trials were provided for participants before the four test trials began (two for the same world and two for the opposite world conditions). The order of the presentation of practice trials was counter-balanced. Experimental trials were performed in the same order for each child: SW, OW, OW, and SW. Children were reminded which trial was incongruent and which trial was congruent before timing began.

Scoring the Opposite Worlds

Stopwatch timing began once the child said '*Start*' and finished when the child said '*Stop*'. Separate timing scores for each condition, congruent and incongruent, were calculated by adding the timing scores for each of the two trial types. Number of errors was not scored. However, if an error was made participants had to return to that square to say the correct answer, delaying their total timing score.

The OW task was translated into Irish so that bilingual children's performance could be compared in both English and Irish. Stickers with Irish translations were placed over the English version of each trial e.g. 'Tús' over 'Start' and 'Críoch' over 'Finish'. Order of administration in English and Irish was counterbalanced for the sample. Therefore, in total, monolinguals performed 4 trials (2 congruent and 2 incongruent) and bilinguals performed 8 trials (2 English congruent, 2 English incongruent, 2 Irish congruent, 2 Irish incongruent).

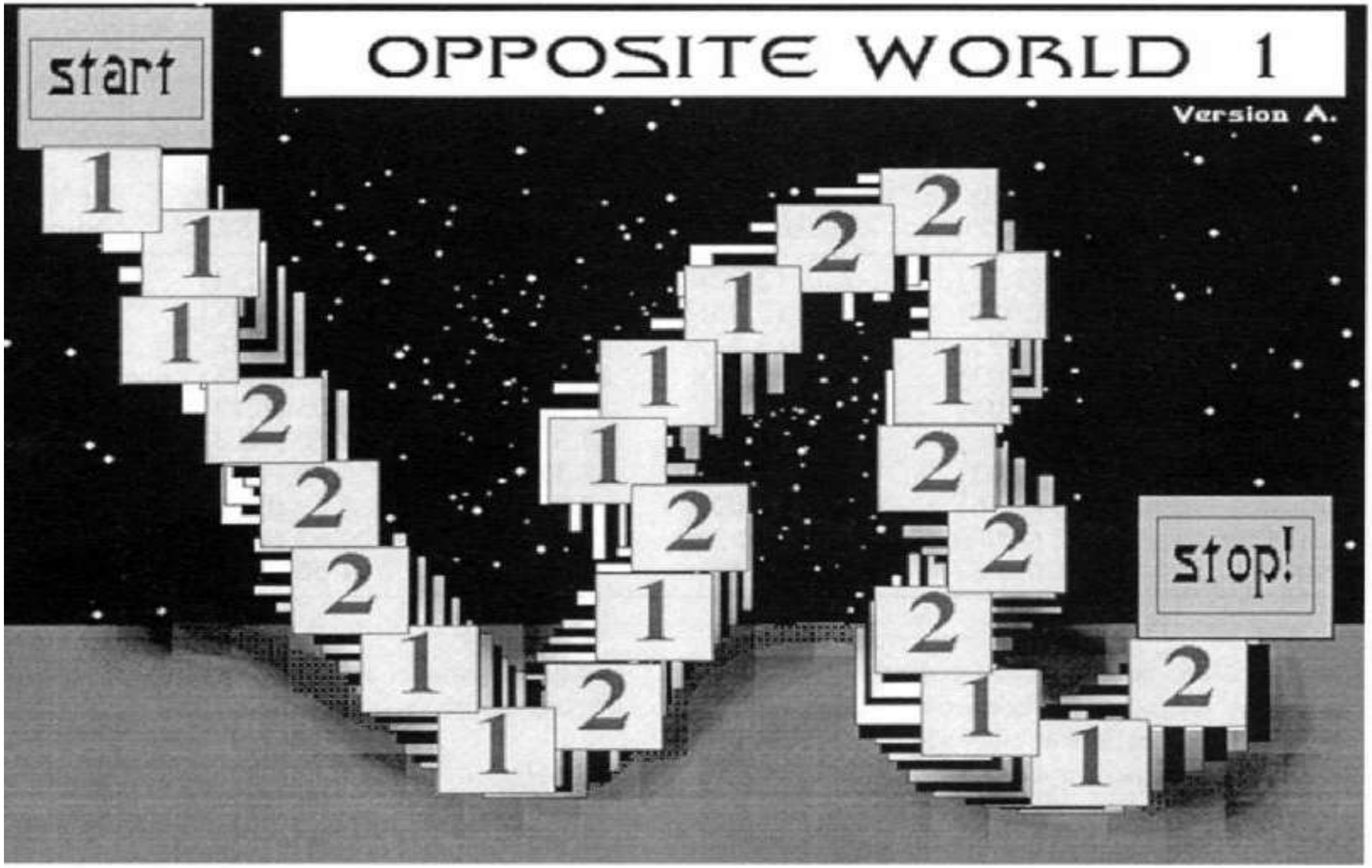


Figure 7 *Opposite Worlds task example*

Reliability and validity of the TEA-Ch

Test-retest reliability for the TEA-Ch task has been reported as moderate to high (.65 to .85) in a random subgroup of 55 children (Manly et al., 2001). Henry and Bettaney (2010) reported correlations for the Opposite Worlds to be high at .92. Test-retest correlations were also high for the Creature Count at .69 for accuracy and .73 for time. Manly et al. (1999) reported these retest correlations as slightly lower with a Creature Count accuracy coefficient of .71 and timing score .57. For the Same Worlds condition they reported a correlation of .87 and for Opposite Worlds .85 (once age was partialled out).

The TEA-Ch manual (Manly et al., 1999) offers evidence of construct validity by reporting significant and high correlations between subtests and widely used neuropsychological measures ranging from .71 to .85. Convergent validity is also reported as being strong as subsets of TEA-Ch have correlated with a range of measures of executive function (EF). Evidence for discriminant validity was evident as correlations were not present between TEA-Ch and IQ (Manly et al., 2001).

In the Opposite Worlds task, the child must inhibit their prepotent response, the correct digit name and time taken to do this is the dependent measure of interest. Manly and colleagues (2001) reported modest correlations between it, the Stroop test (.24) and the Matching Familiar Figures Test (.25). Test-retest reliability was also high (.92). Evidence for the validity of the CC task has been modest but positive correlations have been found with the Stroop task (.31), Trails B of the TMT (.21) and with the Matching Familiar Figures Test (.35; Manly et al., 2001). By including verbal counting and visuospatial arrow symbols, this task requires both verbal and visuospatial processing. Test-retest reliability have been reported as reasonably good (.69 accuracy, .73 timing).

4.5.4. The Colour-Word Stroop

A common measure of EF is the Stroop Interference Test, originally developed by Stroop in 1935 to measure selective attention and cognitive flexibility. However, the test is now believed to best measure a participant's IC skills (IC; Archibald & Kerns, 1999; Boone, Miller, Lesser, Hill, & D'Elia, 1990). IC refers to an ability to inhibit an automatic, prepotent response (i.e. dominant response; for details see Section 2.3.1., Spreen & Strauss, 1998).

In the colour-word version of the Stroop task participants must name the colour of an item with varying levels of difficulty. In the neutral condition, participants name the colour of simple colour patches (e.g. XXXX in red, blue, green or yellow ink; Fig. 8.). In the congruent condition participants name the colour of words which are in agreement with one another (e.g. YELLOW written in the colour yellow, participant must say "yellow"; Fig. 9.). This congruent condition of the task is believed to reflect participants' basic reading rate and is affected by speech production or language delays. The final incongruent condition requires participants to name the colour of the ink in which colour words are written. However, unlike the congruent condition, the coloured ink and colour words are not in agreement with one another (e.g. YELLOW written in colour RED, participant must say "red"; Fig. 10.). The ink colour and colour word are always in conflict for incongruent trials and the incongruent condition is considered the condition to best assess mental flexibility and the ability to inhibit the dominant response (Wecker, Kramer, Wisniewski, Delis, & Kaplan, 2000; Homack & Riccio, 2004). This study used a computerised version of the task where colour words were displayed to children on a computer screen. Children responded by speaking into a microphone. The microphone was connected to a computerised, automated-response box, which was programmed to record RTs as soon as responses were made.

Since its original publication, numerous versions of the Stroop task have been developed. Subsequently, there is no available standardised version of the test. Variants of the task include; languages used, the use of dots or XXXs for neutral trials, the number of items per task, methods of administration (e.g. motor or verbal responses) and the elimination or inclusion of differing trial types (Spreen & Strauss, 1998).

This study used E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) and an automated-response time box to design a computerised version of the Colour-word Stroop task and in order to obtain precise measurement of participants' RTs for each individual trial. Participants were presented with 60 trials, 10 neutral (i.e. XXX in colours red, blue, green or yellow), 20 congruent (e.g. the word 'red' in red ink) and 30 incongruent (e.g. the word 'red' written in green ink). Participants completed 20 practice trials prior to the experimental trials to allow for practice with the microphone and to test for colour-blindness. Trials were presented to participants on a computer screen and were generated in a random order by the computer programme.

Bilingual groups performed both English and Irish versions of the test. For the Irish version of the test, colour-words were translated (blue = *gorm*, red = *dearg*, yellow = *buí*, green = *glas*) and presented in the same manner as the English version and with the same number of trials. The order of administration for language versions of the test was counter-balanced for bilingual participants.

Scoring Colour-Word Stroop

Following the protocol used in the Californian Stroop Test, a subtest of the Delis-Kaplan Executive Function Scale (D-KEFS: Delis, Kaplan, & Kramer, 2001), participants' median time taken to complete each trial type (neutral, congruent, incongruent) was recorded as their score. This is in contrast to the original Stroop test where the number of words read in a given time limit is used as the participants' score. An automated-response box recorded participants' RTs as soon as they responded verbally through a microphone. Once a response was recorded the next trial was automatically generated on screen. Accuracy was assessed by hand recording participant responses and verifying these against the colours produced by the computer. If a participant made a sound that was not in response to a trial but was recorded by the microphone, e.g. coughing or sneezing, these trials were marked as void by the examiner to prevent a RT bias.

Calculating the Stroop effect

Five RT scores were analysed from the Colour-word Stroop task: neutral RT, congruent RT, incongruent RT, facilitation RT and inhibitory control (IC) RT. Rather than using the mean RTs, median RTs were used to reduce the weight given to extreme values and is the procedure recommended for Stroop analysis (e.g. Heathcote, Popiel, & Mewhort, 1991). Although there are a number of methods used to calculate the Stroop effect or IC RTs, the method used here was in line with Golden and Golden (2002) and MacLeod (1991). Here, IC RTs were calculated by subtracting each participants' median neutral RT from their median incongruent RT, giving the difference in RTs between neutral (ink colour only) and incongruent trials or the amount of extra attentional skill required to respond to incongruent over congruent trials. Facilitation RTs were also calculated by subtracting participants' median congruent RT from their median neutral RTs (see MacLeod, 1991 for details). Percentage accuracies for each trial type (neutral, congruent, incongruent) were calculated for each individual. Although there is a large amount of literature and normative data for the colour-word Stroop in relation to adults, no such normative data exist for children (Homack & Riccio, 2004).

Colour-word Stroop Reliability and Validity

As the Colour-word Stroop task does not have unified standardisation, reliability and validity scores are difficult to generalise across tests. For the three trial types measured - neutral, congruent and incongruent - research suggests that temporal reliability is good ($r > 0.80$; Connor, Franzen, & Sharp, 1988; Graf et al., 1995; Sacks, Clar, Pols, & Geffen, 1991) with some practice effects evident across all three trials (Feinstein, Brown, & Ron, 1994). Reliability scores have not been demonstrated with any of the derived scores, i.e. facilitation and IC RTs. In the Golden version of the test (Golden, 1978), *t*-scores for congruent, incongruent and IC or interference trials were generated and show a high test-retest reliability reported at .89, .84, .73 for a group version of the test and .86, .82 and .73 for the individual test.

Factor analysis has been used as a method for establishing construct validity for the Stroop test. Performance on Stroop has been found to load onto the same factor as the Block Design, Digit Symbol, Similarities, and Digit Span subtests of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981).

Assessing the construct validity of neuropsychological measures can also be identified through their sensitivity to dysfunctions of the central nervous system (CNS) or EF areas of the brain (i.e. the frontal lobes). In 1976, Golden used the Stroop with 141 brain injured and normal participants. He found that when all three trial types were performed poorly, participants usually had left hemisphere or diffuse injuries. Those who scored poorly on the colour naming (neutral) trials but had average scores on the other two trial types displayed a pattern of right-hemisphere injuries that may have caused an inability to verbally name colours.

Figure 8 *Neutral condition example from colour-word Stroop*

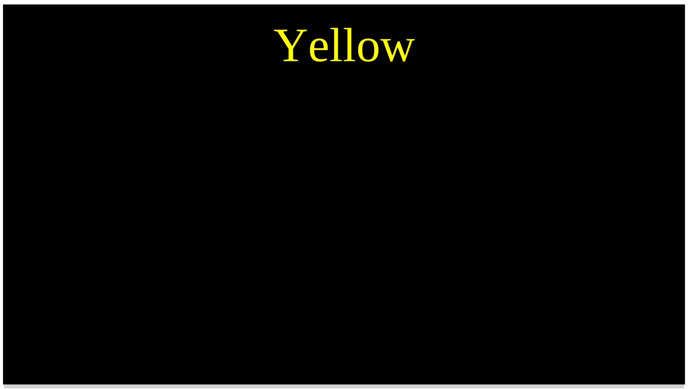


Figure 9 *Congruent condition example from colour-word Stroop*

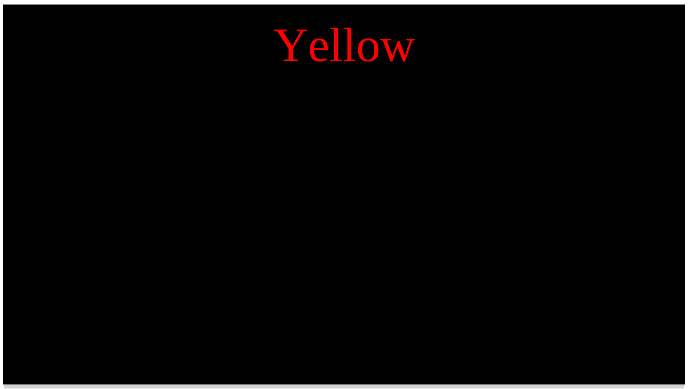


Figure 10 *Incongruent condition example from colour-word Stroop*

4.5.5. Working Memory Test Battery for Children

The Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001) provides a broad-ranging assessment of working memory (WM) capacities that is suitable for use with children aged between 5 and 15 years. This test battery comprises of nine subtests designed to tap the three main components of WM using methods that are simple and quick to administer and that have been extensively employed by psychologists to investigate memory function in both children and adults (Pickering & Gathercole, 2001). The test is also based upon extensive literature concerning the triarchic structure of WM, consisting of a central executive and devoted visual and verbal 'slave' systems (Baddeley, 1986; Gathercole & Baddeley, 1993).

For this study, three subtests of the WMTB-C were used: forward digit recall, backward digit recall and forward blocks recall. A fourth subtest was added in this thesis as a measure of children's visuo-spatial WM - backward block recall. This was not an official subtest of the WMTB-C but used the same apparatus as the forward block recall and similar instructions to the backward digit recall. A brief description of each of the subtests follows.

4.5.5.1. Verbal short-term memory task: forward digit recall

The digit recall task required participants to recall a series of digits. The participant is presented with spoken sequences of digits and has to immediately repeat them in the order heard. Digit sequences were read by the examiner at a rate of about one digit per second. List lengths increased across blocks and the number of correctly recalled lists served as the dependent variable with a maximum score of 54. All digits were presented in an even monotone and at a rate of 1 per second.

Three practice task blocks were presented prior the first block of experimental trials. The maximum span of numbers required in the practice blocks was 3 (range 1 – 3). The experimental task began at the highest span (1-3) at which the participant was successful in the practice blocks. Each experimental block had a maximum of six trials per span. If the child responded correctly to four trials within a block, the examiner proceeded to the next block after giving credit to any omitted trials. The test is finished if the participant made 3 or more errors within any block or if all trials in the task were completed. The maximum span tested was 8 digits in length and the mean span for participants was 5.

Scoring the forward digit recall

Responses to each trial were scored 0 or 1 point. Scores were summated to give an overall Trials Correct score. If trials were omitted as a consequence of moving to the next block of tests, the child is given a credit of 1 for each unadministered trial. Within the final block that was administered, a score of 1 is given for any trials that were correctly recalled up to the point at which the discontinuation rule came into force. No credit is given for any correct recalled trials after three errors have been made in a particular block.

Span score is also given as the number of digits making up the trial sequence in the last block before the discontinuation rule came into force, i.e., the span corresponding to the penultimate block of trials administered. The range in span score is 1-9.

4.5.5.2. Verbal working memory: backward digit recall

As in the forward digit recall task, the backward digit recall involved the spoken presentation of sequences of digits for immediate recall. The child is required to recall the list in reverse order, i.e., the recalled list should begin with the last item heard and end with the first item heard (e.g., ‘1, 2, 3’ would become ‘3, 2, 1’). Two practice trials are provided for the task before test trials begin.

Practice trials were presented in the same manner as the forward digit recall. The move on rule was applied whenever the child achieved four out of six correct trials in block. The test ended if 3 or more errors were made within any block.

Scoring backward digit recall

The backward digit recall was scored in the same way as the forward digit recall except that the maximum score possible was 42, rather than 54. The span score for the backward digit recalled ranged from 2 to 7 and participants had a mean span of 4.

4.5.5.3. Visuospatial short-term memory task: forward block recall

The block recall task involved the presentation of sequences tapped out on the block recall board (see Fig. 11.). Initially children are asked to recall the location of only one block. Following this, sequences of two and more blocks were presented. It is important that each sequence is recalled in exactly the same order as it was seen. Therefore, in the event of the examiner making an error in the administration order of a block sequence, the child’s response was to be scored relative to the actual sequence rather than the intended one. Three practice trials were provided prior to the first block of test trials. Testing began at the block of trials corresponding to the greatest span (2 or 3) that the participant successfully completed in the practice trials.

Before starting the test, the block recall board was placed between the examiner and the participant so that both can touch the block easily. The numbers (1-9) on the blocks faced the examiner so as not to be visible to participants. Each trial was made up of a series of numbers from 1 to 9, corresponding to the order in which the blocks on the board must be tapped. It was important to make sure that the examiner’s hand did not obscure the board as the sequence was being presented. The blocks were pointed to in a smooth and steady fashion and at a rate of 1 block per second.

Scoring forward block recall

If the child responded correctly to 4 trials within a block, the tester moved on to the next block after giving credit for the omitted trials. If the child made 3 or more errors within any block the test was stopped. The block recall task was scored in the same way as the digit recall tasks where each trial was given a score of 1 or 0. The maximum possible score is 54 for the block recall task. The span range for participants was 3 to 6 with an average of 5.

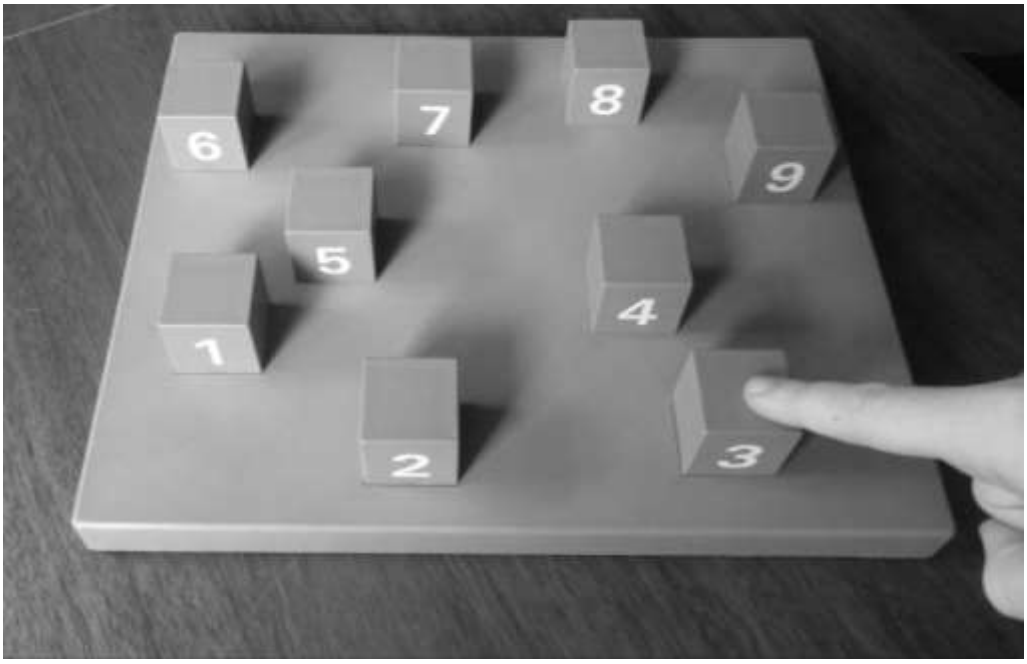


Figure 11 *Corsi-blocks testing tool (examiner's perspective)*

4.5.5.4. Visuospatial working memory task: backward block recall

As in the forward block recall task, the backward block recall task involved the presentation of sequences tapped out on the block recall board. However, instead of repeating the sequence as it was presented, participants had to recall the sequence in reverse order i.e. the recall should begin with the last block presented and end in the first block presented (e.g. blocks 1, 2, 3 would become 3, 2, 1,). Sequences were derived using variants of the forward blocks sequences. Three practice trials were provided before beginning at the block of trials corresponding to the greatest span (2 or 3) that was successfully completed in the practice trials.

Scoring backward block recall

As in the forward block recall, if the child responded correctly to 4 trials within a block, the tester moved on to the next block after giving credit for the omitted trials. If the child made 3 or more errors within any block the test is stopped. As the backward block recall is an unstandardized task, derived from the block recall task, it is scored in the same way i.e. each trial is given a score of 1 or 0 and there is a maximum score of 54 for the task. The span range was 3 to 6 blocks with a mean of 4 for participants in this study.

Reliability and validity of the WMTB-C

Engel de Abreu, Conway and Gathercole (2010) carried out a three-year longitudinal study with 122 kindergarten children from Luxembourg to explore the effects of WM in multilingual children. Internal reliability estimates for the scores on different measures were calculated using Cronbach's alpha. Reliability coefficients were high for both the digit recall (.84, .91, .89) and the backwards digit recall (.85, .84, .80) tasks across all three time points. Packiam Alloway (2007) reported test-retest reliability for the forward digit recall test as .84 (the strongest correlation from the three span tasks utilised by her). For children ages 4.5-11.5 years test-retest reliability for the backward digit recall test was .64. In the forward block recall task, test-retest reliability was .83.

4.5.6. Trail Making Test (TMT)

Originally included as part of the *Halstead-Reitan Battery* which was developed for the *Army Individual Test Battery* (1944), the Trail Making Test (TMT; Armitage, 1946; Reitan & Wolfson, 1985; Spreen, 1998) is now one of the most popular neuropsychological tests and is included in most cognitive test batteries (Lezak, Howieson, & Loring, 2004). The TMT provides information on visual search, scanning, speed of processing, mental flexibility and EFs (Tombaugh, 2004). This task is mainly thought to tap selective attention/visual search and the capacity to switch attention between two different sorts of targets or cognitive flexibility. Bialystok (2010) argues that even Trails A requires executive demands as children must hold in mind their current place in the sequence (e.g. 4) while searching for the next element (i.e. 5) through distracting space filled with other digits. This process is not as effortful for adults as the counting sequence is deeply engrained and automatic. Others report that the Trails A is a measure of processing speed and relies on visual perceptual abilities and motor speeds rather than EF skill (e.g. Crowe, 1998; Waldman, 2005).

The task is made up of two parts. Participants are given brief instructions on how to complete the task and short practice versions for each of the Trails. The specific administration procedure for the TMT is provided in Reitan's (1979) *Manual for Administration of Neuropsychological Test Batteries for Adults and Children.* In Trails A, participants are presented with 25 encircled numbers, randomly distributed on a sheet of paper (see Fig. 12.). As each circle contains a number, the aim is to connect the circles following a rule of ascending number sequence using a pencil. Participants are instructed to draw lines connecting the numbers in order, beginning on 1, without lifting the pencil from the page. In Trails B of the test, the task requirements are similar. However the task is made more difficult with the page presented containing both numbers, 1 to 12 and letters, A to L (see Fig. 13.). Participants are asked to join the 25 circles based on an alternating rule of ascending numbers and letters in alphabetical sequence (e.g., 1, A, 2, B, 3, C…12-L).

Scoring the TMT

The scoring procedure introduced by Reitan (Reitan & Wolfson, 1985) involves recording the total time taken for participants to complete each part of the task. Two scores are obtained (Trails A and Trails B timing scores) reflecting the total time (in seconds) to complete each task. In Reitan's (1979) administration format, errors are not scored, but when they occur, the participant is alerted to their mistake and instructed to correct it, thus slowing their overall performance time. Depending on the child's ability, the task will lasts approximately 5 minutes for a completed administration.

There are a wide number of normative samples available which vary depending on age, demographics and nationality of participants. This has led to certain inadequacies and inconsistencies when scoring the TMT (Lezak et al., 2004; Mitrushina, Boone, & D'Elia, 1999; Spreen & Strauss, 1998). Mitrushina et al. (2005) have provided a detailed description and discussion of normative data available for the TMT. However, due to the level of difficulty associated with the TMT no normative data or standardisations are available for children and subsequently, the ages of participants tested as part of this thesis. Wecker et al. (2000) reported that the majority of normative data samples generally range from 15 to 90 years. Although Delis et al. (2001) discussed normative data for ages 8-89 years in the development of the Delis-Kaplan Executive Function System (D-KEFS), their version used a variant of the TMT and therefore is not directly comparable to data from this study which used the classic TMT paradigm.

Reliability and validity of the TMT

According to surveys of test usage in neuropsychological practice, the TMT is one of the most frequently used neurological assessments and is sensitive to a variety of cognitive impairmenst and processes (Lezak, 1995; Spreen & Strauss, 1998; Tombaugh, 2004). Dikmen et al. (1999) reported test-retest reliability of 0.79 for Trails A and 0.89 for Trails B over a 9-month interval. McCaffrey et al. (2000) also presented data following repeated administrations and Abe et al. (2004) proposed that the TMT has a strong enough test-retest reliability to be used within longitudinal studies.

Sánchez-Cubillo et al. (2009) carried out a comprehensive literature review, correlation and regression analysis in order to clarify which cognitive mechanisms underlie performance on the TMT. Their study examined the performance of clinical and non-clinical samples on the TMT and concluded that Trails A mainly required visuoperceptual abilities while Trails B primarily tapped into working memory and secondary task switching. The Trails B-Trails A difference minimises visuoperceptual and working memory demands, providing a relatively pure indicator of executive control or EF abilities. Following extensive reviews, Trails B is now believed to differ significantly from Trails A in cognitive demands, length of time taken to complete trail and perceptual complexity. Trails B has been shown to place additional demands on the ability to alternate (Crowe, 1998; Gaudino et al., 1995;

Salthouse et al., 2000), to flexibly modify a course of action (Arbuthnott & Frank, 2000; Kortte et al., 2002) as well as being able to inhibit the prepotent response to continue with either number or letter (Arbuthnott & Fran, 2000). Others attribute the added cognitive difficulty of Trails B to having to maintain two response sets simultaneously in mind (Eson et al., 1978; Lezak et al, 2004; Reitan, 1971). In other words, it appears Trails B places additional demands on the EF system. This is supported by the fact that Trails B also loads on an attention factor (O'Donnell et al., 1994) as well as evidence from clinical populations with frontal lobe damage or traumatic brain injury. These participants show impaired performance on Trails B compared with non-clinical samples (Cicerone & Azulay, 2002; Corrigan & Hinkeldey, 1987; Reitan, 1971).

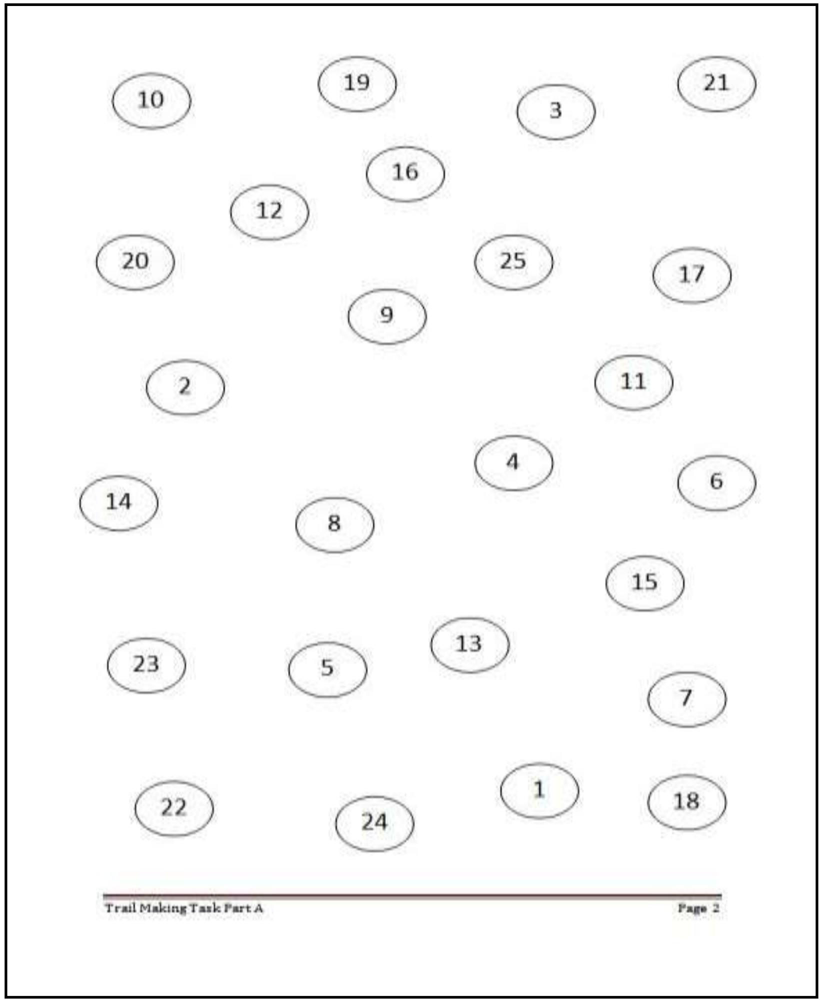


Figure 12 *Trails A example*

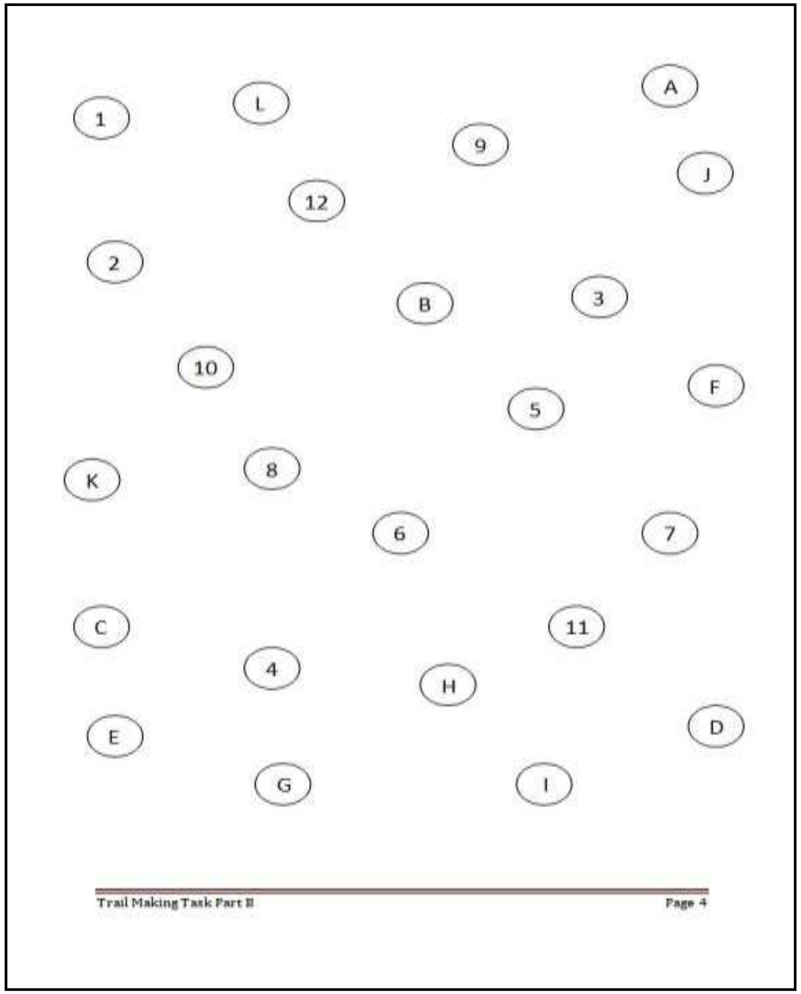


Figure 13 *Trails B example*

4.5.7. Wisconsin Card Sorting Test

One of the most frequently used tasks to assess executive control in clinical and nonclinical populations is the standardised, Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay & Curtis, 1993). Originally developed by Grant and Berg (1948), the test aims to assess abstract reasoning, concept formation, and response strategies to changing contextual contingencies (Nyhus & Barceló, 2009). In 1981, Heaton developed a standardised and shortened version of the WCST (Heaton et al., 1993). This version is now widely cited as a measure of general "*executive function*" (EF; Grant & Berg, 1948; Heaton et al., 1993). The task requires participants to achieve their goal through the development and maintenance of an appropriate problem-solving strategy across changing stimulus conditions (Luria, 1973; Shallice, 1982). As it is a measure of complex cognitive functioning, the WCST utilises a range of EFs including planning, organisation, shifting of cognitive sets, goal-directed behaviour and inhibition of impulsive responses (Chelune & Baer, 1986; Gnys & Willis, 1991; Welsh & Pennington, 1988).

The WCST consists of four stimulus cards and 128 response cards that depict figures of varying forms (crosses, circles, triangles, or stars), colours (red, blue, yellow, or green) and numbers of figures (one, two, three, or four). The four stimulus cards with the following characteristics are placed before the participants in left-to-right order: one red triangle, two green stars, three yellow crosses and four blue circles (see Fig. 14.). The participant is then handed a deck of 128 cards and instructed to match each consecutive card from the deck with one of the four stimulus cards, whichever one he or she thinks it matches. The participant is told only whether each response is right or wrong and is never told the correct sorting principle (or category). Therefore the task required participants to find the correct classification principle by trial and error and examiner feedback. Once the participant has made a specified number of consecutive "correct" matches to the initial sorting principle (colour) the sorting principle is changed (to form or number) without warning, requiring the participant to use only the examiner's feedback to develop a new sorting strategy. Participants must maintain the new sorting principle (or set) across these changing stimulus conditions while ignoring the other, now irrelevant stimulus dimension. The task proceeded in this manner through a number of shifts in set (i.e. sorting principle) and among three possible sorting categories (Colour, Form, and Number; Heaton et al., 1993).

Figure 14 illustrates the layout of the stimulus cards as seen from the perspective of both the participant and examiner. While the cards are being placed on the table the following instructions are given to each participant (Heaton et al., 1993, p. 5):

"This test is a little unusual because I'm not allowed to tell you very much about how to do it. You will be asked to match each of the cards in these decks (point to the response card deck) *to one of these four key cards* (point to each of the stimulus cards in succession, beginning with the red triangle). *You must always take the top card from the deck and place it below the key card you think it matches. I cannot tell you how to match the cards, but I will tell you whether you are right or wrong, simply leave the card where you have placed it and try to get the next card correct. There is no time limit on this test. Are you ready? Let's begin".*

The first sorting principle is **Colour**. As the participants began to sort the response cards, the examiner responded with "correct" or "right" each time the participant matched according to **Colour**, and with "incorrect" or "wrong" each time they matched to a stimulus dimension other than **Colour**. The process continued until the participant had produced 10 *consecutive* **Colour** responses. Without comment or any other indication, the examiner then changed the correct sorting category to **Form**. The change between sorting categories must be a smooth and undetectable transition, both verbally and nonverbally. **Form** remained as the correct sorting category (principle) until the participant had again attained 10 consecutive correct responses. Without warning or cues, the examiner then changed the correct sorting category to **Number**. After 10 consecutive correct responses to **Number**, the examiner would then switch back to **Colour** as the correct sorting category, then to **Form**, and then to **Number** in the manner previously described. The test continued until the participant has successfully completed six categories or until both decks of response cards had been used, whichever occurred first.

The WCST is not timed and this is explained to the participant. While participants will vary in their time taken to complete the task, most complete the test within 20-30 minutes.

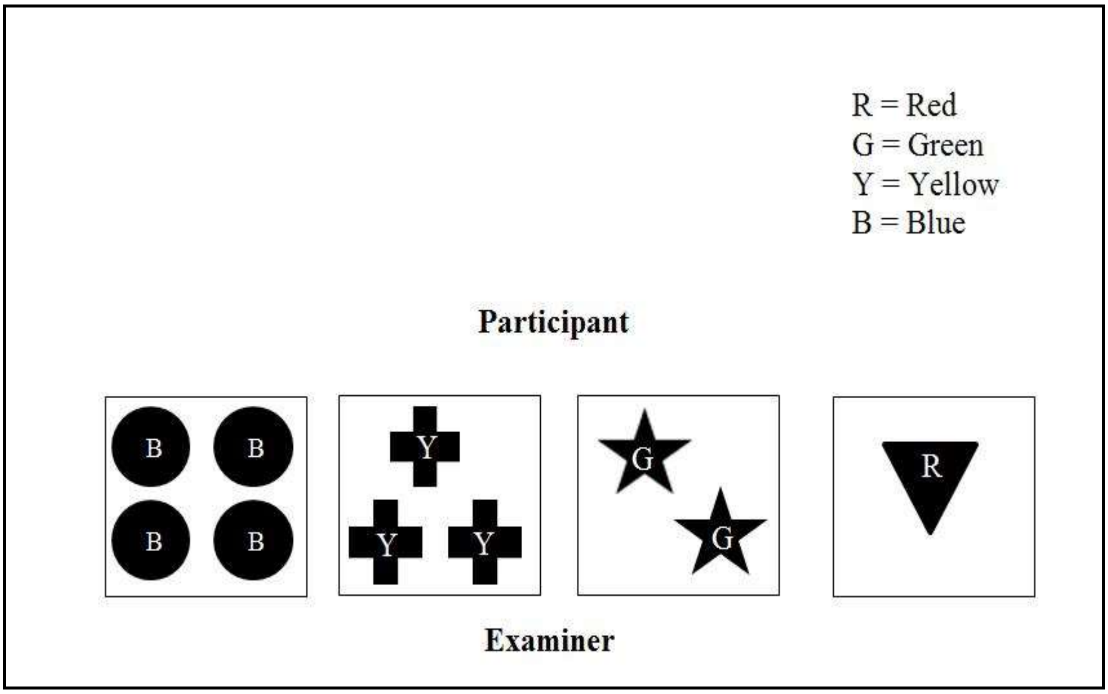


Figure 14 *Orientation of WCST stimulus (key) cards*

Scoring WCST

Accurate recording of the participants' responses is a critical element to the subsequent scoring of the WCST. As each response is unique and stimulus and response cards may match on more than one dimension (e.g. both Colour and Form), it is crucial that the examiner recorded responses accurately. For example, if the response and stimulus card were identical then they would match on Colour, Form and Number and must be recorded as so. The response dimensions are recorded in the same manner for each item, irrespective of whether the response is correct or incorrect with respect to the current correct sorting category.

Scoring of the WCST has been a source of difficulty for many researchers due to variable or incorrect application of scoring rules set out by Heaton (1981). Therefore the task was administered, recorded and scored according to the procedure outlined in the Heaton manual (Heaton et al., 1993). Their manual provides a detailed account of how to accurately score the test.

Despite the fact that Heaton's correction norms offer many different scores (16 in total), due to the internal structure of the tests, many researchers tend to only evaluate a selection of these scores as an indicator of participants' performance. The majority of researchers consider two or three scores, including: number of categories completed, number of perseverative errors, and number of non-perseverative errors (Barceló & Knight, 2002; Bowden et al., 1998; Nyhus & Barceló, 2009). Perseverative error responses are defined as errors resulting from persistence in responding to a stimulus characteristic that is no longer correct (Barceló & Knight, 2002; Heaton et al., 1993). The following scores were examined in this thesis:

1. **Total number of errors**
2. **Perseverative responses**: responses which occur when a participant continues to sort cards according to the same rule despite negative feedback, reflecting EF difficulties on skills such as switching, IC and WM.
3. **Perseverative errors**: failure to change mental rule after receiving negative feedback so that participant continues sorting cards according to the previous-category dimension despite feedback indicating their response was wrong.
4. **Non-perseverative errors**: the normal errors needed to learn a new rule, reflecting an attitude to change the response after receiving disconfirming feedback (Barcélo, 1999).
5. **Number of Categories Complete**: the number of sequences of 10 consecutive correct responses matched to the criterion sorting strategy. The number of categories ranges from 0 to 6.
6. **Trials to Maintain First Category**: this was the number of trials taken to find the first correct sorting rule.
7. **Failure to Maintain Set**: this refers to participants' susceptibility to distraction and interference or with problems integrating temporally separating events.

Heaton and colleagues (1993) derived normative data for the WCST from 899 non-clinical participants from the United States of America. This sample was divided into six distinct groups, the largest being a group of 453 children and adolescents between the ages of 6 years, 6 months to 17 years, 11 months (52% males, 48% females).

WCST Reliability and Validity

According to Heuttener et al. (1989), when the WCST is administered and scored according to the procedures given by Heaton (1981; Heaton et al., 1993), interscorer reliability coefficients were high ranging from .90 to 1.000 from the 11 scorers evaluated. Intrascorer reliability coefficients ranged from .83 to 1.000. These findings of high interscorer and intrascorer reliability of WCST data obtained from children and adolescents are consistent with those of Axelrod et al. (1992), who examined WCST data collected from an adult sample.

Due to the nature of the WCST, practice effects have led to low test-retest reliability scores. Paulo, Axelrod and Tröster (1995) evaluated test-retest stability of the WCST in a sample of 87 older adults. Test-retest coefficients were low ranging from .12 on Learning-to-Learn to .66 on Total Number or Errors. All coefficients fell below the minimum standard of .80 for tests used in clinical decision-making. Average retest gains were 5-7 standard score points for the 1-year period. In a longitudinal twin-study (n = 747) examining performance on the WCST in adolescents tested at 12 and 14 years, test-retest correlations obtained were suggested to be in the lower bound estimate of test-retest reliability (r = .037; Anokhin, Golosheykin, Grant, & Heath, 2010). Due to low test-retest reliability coefficients, this study examined WCST differences at Time 3 only.

The WCST is currently used as a measure of general executive functioning, combining a number of different skills, e.g. selective attention, self-monitoring, and tendency to perseverate. It is also frequently used as a measure of a participant's ability to shift response or set shifting (Goldstein & Green, 1995). Its influence is evident with over 115 studies employing the WCST in both clinical and nonclinical populations between 1948 and 2004 (Axelrod, Greve, & Goldman, 1994; Romine, Lee, Wolfe, Homack, George, & Riccio, 2004). The WCST has also been considered a useful tool in the evaluation of developmental changes in cognition (Dempster, 1992; Heaton et al., 1993; Shute & Huertas, 1990). Although the WCST has been used to assess frontal lobe impairments in clinical samples (e.g. Drewe, 1974; Robinson, Heaton, Lehman, & Stilson, 1980) some have argued that rather than engaging one specific brain area, the WCST requires a distributed neural network of both cortical and subcortical brain structures which dynamically integrate with one another (Barceló, 2003; Fernandez-Duque & Posner, 2001; Posner & Petersen, 1990). As well as general executive functioning, set shifting is also thought to be a central cognitive mechanism for performance on the WCST that appears to be specific to prefrontal functioning (Barceló, 2001; Braver, Reynolds, & Donaldson, 2003; Nyhus & Barceló, 2009; Rubinstein et al., 2001; Shallice et al., 2008).

CHAPTER FIVE
THE IMPACT OF BILINGUALISM AND SOCIOECONOMIC STATUS ON CHILDREN'S EXECUTIVE FUNCTION DEVELOPMENT

5.1. INTRODUCTION

This study aims to address issues from previous chapters (Chapters 1, 2 and 3) by examining the effects of language and socioeconomic status on children's cognitive performance over time. The study focused one area of cognitive development in particular, the executive function(s) (EFs; see section 2.2.).

Chapter 1 discussed bilingualism from a historical perspective as well as issues within current research. One of the conclusions was that it is unfair to compare monolingual children with bilingual children whose L2 is not recognised as being equal to that of their L1, for example children within submersion education programmes. On the other hand, the linguistic environment provided by early, total immersion education (IE) programmes has been shown to improve children's L2 skills (e.g. reading, writing, comprehension, speaking) at no cost to their L1 skills (e.g. Baker, 2011, Cummins, 1979, 2000; Genesee, 1987, Swain & Johnson, 1997, Swain & Lapkin, 1991; see section 1.4.2.). As a result, children in IE who are developing their L2 proficiency with age and experience have participated in a number of bilingual studies, with positive cognitive effects reported for these types of emerging bilinguals (e.g. Bialystok, 1986; Bialystok et al., 2008; Moreno et al., 2010; Nicolay & Poncelet, 2013a). At the same time, immersion students' linguistic development may be delayed relative to their monolingual peers, although, with time, students have been shown to catch up (e.g. Bialystok et al., 2010, Genesee, 2004; Nicolay & Poncelet, 2013b).

It is assumed that children within early, total or full-IE programmes are not as balanced in their L1 and L2 as children who have been brought up with an L1 and L2 from birth or simultaneous bilinguals, particularly in productive language skills (e.g. Genesee, 2004; Hermanto, Moreno, & Bialystok, 2012; Swain, 1984; Swain & Lapkin, 1998). However, labelling children who have completed a 'full-IE programme' as *"bilingual"* is not an unfair supposition as many will develop their

L2 skills to a level, high enough to be classified as bilingual (Baker, 2003; Bialystok, Peets, & Moreno, 2012). A number of studies utilising IE samples have found cognitive benefits in favour of bilinguals (e.g. Bialystok et al., 2012; Hermanto et al., 2012, Nicolay & Poncelet, 2013a), although in the context of general bilingual research, such studies remain scarce (Jared et al., 2011). The following study aims to examine the EF development of children in full Irish-IE programmes (for details see sections 1.4.3 and 4.3.) and will use the term "*bilingual"* to describe participants within these groups. While the children in this study are labelled as bilingual, we recognise opinion which has argued that IE groups be classified as *'second language learners'* or '*developing bilinguals"* (e.g. Cummins, 1991, p. 72; Gort, 2006, p. 325; McVeigh, 2012; Swain & Lapkin, 1998). However, it is argued that *'bilingual'* is the most appropriate terminology in this circumstance as testing commenced in the latter half of primary school, at approximately 8-9 years of age, with children having had approximately 5 years of full-immersion in the L2 (and in some instances were educated through Irish during the pre-school stages).

Due to extensive practice using EFs to control two languages, bilinguals have displayed advantages relative to their monolingual peers on a number of EF tasks. Researchers have linked these advantages to specific as well as unified EFs including IC, switching and more complex EF skills (e.g. Bialystok, 2009; Bialystok & Martin, 2004; Bialystok et al., 2004; Bialystok, 2009; Diamond, 2011; see 3.4.). Such findings have inspired the following study which aims to investigate the impact of the IE experience on children's developing EFs by comparing monolingual and bilingual children's cognitive performance over a three-year period using a battery of common EF tasks. The cognitive tasks were selected to target specific as well as unified EFs proposed within Miyake et al.'s model of EF (2000; updated by Miyake and Friedman, 2012; see 2.3.).

Although there are numerous studies implicating bilingualism in advancing children's and adult's EF skills, researchers have yet to understand how these effects change and develop over time (Kroll & Bialystok, 2013). Cummins (1979) argued that although researchers have found positive effects associated with bilingualism, few have uncovered the mechanisms causing such effects or have examined the cognitive-bilingual relationship from a developmental perspective. Longitudinal designs are one method of unpicking these issues of causation. However, the bilingualism literature distinctly lacks this type of experimental design (see section 1.3.4.). EF researchers have made efforts to investigate the developmental nature of the EF, with a number of longitudinal studies outlining its developmental trajectory (e.g. Anderson, 2002; Berlin, Bohlin & Rydell, 2004; Bull et al., 2008; De Luca et al., 2003; Erickson et al., 2008; Hoff et al., 2005; Mazzocco & Kover, 2007; Riggs, Blair, & Greenberg, 2004). Unlike other cognitive functions (e.g. causal reasoning and theory of mind; Fischer, 1980; Piaget, 1936, 1954; Wellman, Cross, & Watson, 2001l Sobel & Kirkham, 2006), the EF displays a protracted trajectory with development continuing throughout late childhood and early adulthood (e.g. Anderson et al., 2001; Best & Miller, 2010; see section 2.5.). As a result of this late-staged development and to allow adequate time for children's L2 to develop (Cummins, 2000; Genesee, 2004), children in this longitudinal study were recruited at approximately 9 years of age, with follow-up testing sessions at 10 and 11 years.

It is evident that more research is needed to evaluate the developmental effects of bilingualism on children's cognitive functions. By implementing a longitudinal design, this study hopes to understand more clearly how aspects of the EF are affected by the bilingual experience across a three-year period. Using this design the study also aims to be able to make certain causal links between aspects of bilingual and cognitive development within the immersion sample.

5.1.1. The impact of socioeconomic status on language group effects

The influence of socioeconomic status (SES) and environmental background on children's psychosocial and cognitive development has been widely researched within the last century. Subsequently, the link between SES and child development is now well established and findings suggest that low-SES backgrounds and social deprivation may have negative implications for a number of lifestyle and developmental outcomes (see Bradley & Corwyn, 2002 for review; Diamond, 2013; Farah et al., 2006; Hackman & Farah, 2009; McCall, 1981; McLoyd, 1998). Consequently, SES must be a key consideration when designing any study of child development, particularly within the educational setting where individual SES backgrounds are often overlooked. Indeed, more and more studies now control for demographic backgrounds compared with traditional experimental research which often failed to recognise its importance; so much so, that SES may have confounded significant results (e.g. Cummins, 2000; Hackman et al., 2010). However, measuring SES is not straightforward and is persistently difficult due to a lack of standardised measurement tools and uniformity within definitions. Although there is no standard definition, Müeller and Parcel (1981) described SES as:

"…an individual's or a family's ranking on a hierarchy according to access to or control over some combination of valued commodities such as wealth, power, and social status".

Within the developmental literature, studies have found that children from lower SES backgrounds often display performance delays in tasks of language ability and EF compared with children from higher SES backgrounds although a number of studies have produced mixed effect sizes for the link between EF and SES (e.g. Farah et al., 2006; Noble et al., 2005, 2007). In terms of language abilities a number of studies have observed lower vocabulary levels, phonological awareness and syntax in children from lower-SES groups (e.g. Hart & Risley, 1995; Hackman & Farah, 2008). Although a number of studies have found SES disparities in EF tasks (e.g. Diamond, 1990; Lipina et al., 2005), the association between SES and EF may reflect the complex nature of SES as a construct, with multidimensional influences. For instance, Noble et al. (2007) investigated the effects of SES (measured along a scale from high to low) on children's language and EF abilities. Their results were in line with previous findings (e.g. Noble et al., 2005) that language (measured using the PPVT) is significantly associated with SES. Although there was a general association between SES and EF ability, variance in both the go-no-go task and delayed non-match tasks was not explained by SES, two common tasks of EF. Furthermore, the association between SES and working memory tasks (e.g. memory picture pairs) was mediated by factors such as home and school variables. These findings are in line with previous work (Noble et al., 2005) that has found the association between SES, language and EF to be disproportionately weighted towards language rather than EF measures. Noble and colleagues (2007) concluded that there might be a complex interplay between SES and neurocognitive development requiring further investigation. Using ERP techniques, Stevens, Lauinger and Neville (2009) showed that children whose mothers had lower levels of academic attainment (no college experience) showed reduced effects of selective attention on neural processing while listening and responding to competing tones, relative to children whose mothers had higher levels of academic attainment (at least some college experience). Their results also indicated a reduced ability to suppress or inhibit responses to sounds in the unattended channel in children with lower SES and were not related to differences in receptive language skills. The benefit of using ERP in this study was that they could localize the mechanism (distractor suppression) of the attentional deficits in the children from lower SES groups. They concluded that children from lower SES backgrounds experience difficulty with tasks that require IC or suppression of an unintended response. Some researchers have speculated that the differences in EF performance between low and typical SES groups may be the result of protracted development of the PFC (Farah et al., 2006; Mezzacappa, 2004; Noble et al., 2007; Noble et al., 2005). High levels of the stress hormone cortisol have been shown in children from lower SES backgrounds and may be affecting this delayed development of the PFC (Lupien, King, Meaney, & McEwen, 2000; Lupien et al., 2001).

Historically, the bilingual experience was often linked with negative cognitive outcomes (e.g. Macnamara, 1966). More recent research has questioned the methods of early findings and found that debilitating results often arose from poor SES controls (Baker, 1988; Bialystok, 2001; Peal & Lambert, 1962; Romaine, 1995; for discussion see 1.3.). Despite improved experimental techniques and control, questions surrounding the role that SES may play in findings linking bilingualism to certain cognitive advantages are still being asked (e.g. Gathercole et al., 2010; Kroll & Bialystok, 2013; Mezzacappa, 2004; Morton & Harper, 2007; Noble, McCandliess, & Farah, 2007; Paap & Greenberg, 2013). For instance, Morton and Harper (2007) argued that advantageous results may arise from a lack of SES controls. As many of the modern studies connecting positive cognitive outcomes with bilingualism rely on immigrant populations in North America (i.e. Canada), they assert that these populations have SES advantages over control groups. Their study alludes to the increase in stricter immigration policies, which now select candidates based on their already high academic achievements, language(s) and occupations (Morton & Harper, 2007; Statistics Canada, 2003). Furthermore, following their controls for SES, Morton and Harper reported no bilingual advantage on the Simon task of inhibitory control (IC) and conclude that studies of bilingualism should recognise the impact of high SES on their samples. Gathercole et al. (2010) also highlighted the need for well-controlled experimental studies of language effects and suggested that the role of confounding factors such as SES on children's performance is still relatively unknown. Their findings revealed that SES contributed to the overall group (monolingual and bilingual) performance on a number of EF task outcomes (tapping and Stroop task) rather than for the bilingual groups in particular. In reply to Morton and Harper (2007), Bialystok (2009) argued that bilingualism and SES are independent of one another and that bilingualism may affect EF performance regardless of SES backgrounds, citing studies with low-SES bilingual groups as evidence for this (e.g. Carlson & Meltzoff, 2008, Mezzacappa, 2004). She also argued that the lack of studies carried out with SES as an independent variable is the result of SES being adequately controlled.

De Abreu et al. (2012) examined the cognitive impact of bilingualism in children from disadvantaged or impoverished backgrounds. They argued that although previous research may have controlled for SES, bilingual advantages may still only emerge in higher SES groups as few have examined the cognitive effects of bilingualism by comparing low-SES bilinguals with low-SES monolingual peers. Their study compared low-SES Portuguese monolingual children at 8 years to Luxembourgish-Portuguese bilinguals who were either first or second generation immigrants to Luxembourg from Portugal. Executive functioning was measured using tasks from the TEA-Ch and the Automated Working Memory Assessment (Alloway, 2007) as well as a modified version of the Flanker task (Rueda et al., 2004). Although the groups' abstract reasoning (Raven's Colored-Progressive Matrices; Raven, Court, & Raven, 1986) and WM results did not differ, bilingual children performed significantly more quickly on the flanker task and sky search attentional control tasks than their monolingual peers. They concluded that even for children from impoverished or low-SES backgrounds, bilingualism may still give them an advantage over monolingual peers for tasks of executive control rather than general cognitive abilities.

Whether researchers decide to control for or investigate separately the role of SES, there is no question that ignoring its influence on children's development may be detrimental to the generalizability of any subsequent findings. Furthermore, there seems to be a complex and yet undetermined interplay between SES, language abilities and EF skills within the child development literature. It is for these reasons that this study decided to explicitly examine the role of SES in children's EF development as well as its influence on language group effects. Monolingual and bilingual children from both low and typical SES backgrounds were recruited for this study and their EF development was assessed over a three-year period to explore any EF developmental differences within each group.

5.1.2. Classifying socioeconomic groups

Despite the widespread use of SES as an independent variable and as a control in a variety of research disciplines, there is no standard empirical measurement of SES across studies for children and adolescents (Bornstein & Bradley, 2003). Many choose a tripartite approach to classify SES, incorporating parental income, parental education, and parental occupation as indicators of SES in children and adolescents (Duncan, Featherman, Duncan, 1972; Sirin, 2005). Parental education has been the most stable aspect of SES as it tends not change over time and is a predictor of parental income (Hauser & Warren, 1997). Occupation is also a strong indicator of SES as it tends to be ranked according to both education and income but also carries information regarding prestige and culture (Sirin, 2005). Aggregated measures of SES are also common within educational research, e.g., schools attended (Caldas & Bankston, 1997) and neighbourhood locations (Brooks-Gunn, Duncan, & Aber, 1997). School-based SES usually considers the number of students eligible for free or reduced-price school meals while neighbourhood SES is usually measured as the proportion of residents aged at least 20 years old who have not completed secondary education (according to census data, Brooks-Gunn, Denner, & Klebanov, 1995). Both these proxy measures of SES are thought to share the underlying definition of SES as a contextual indicator of social and economic well being, beyond the SES resources available within the students' homes (Sirin, 2005).

Taking these issues into account, this study decided to classify children using both individual and aggregated SES measures. At the individual-level, background questionnaires were given to each child's parents/caregiver assessing parental occupation and highest level of academic attainment (see Appendix I). Occupations were coded based on the International Standard Classifications of Occupations, 2008's (ISCO-08) major grouping units. The ISCO describes the occupational classification system as: *"a tool for organising paid employment jobs into detailed categories for comparison to be further aggregated into broader groups…"* and provides a coding system for the variable occupation.

Average years in education were also considered as a predictor of SES and for occupation and educational levels. Data from the parent/caregiver with the highest level of attainment only was retained for analysis. Due to confidentiality agreements with the parents/guardians and in line with school regulations, parental income was not assessed in this study. At the aggregated-SES level, school rankings were considered and grouped according to the level of extra governmental support received. Neighbourhoods were also carefully considered to compare groups of low- and mid-SES (for details see 4.3.2.).

5.1.3. Research Questions/Hypotheses

This study is a longitudinal study of monolingual and bilingual children's cognitive development. As few studies have used this type of design to investigate language group effects in children from low and mid SES backgrounds, it was difficult to develop specific hypotheses. However, drawing on multidisciplinary research from areas such as cognition, child development, education and bilingualism, a number of general predictions can be made regarding age, SES and language status.

- Developmental research supposes that the EFs have a protracted development relative to other aspects of cognition. This study aims to assess children's cognitive development using a range of EF tasks across a three year period from 9 to 11 years old. It is predicted that children's performance on each of the EF tasks will improve across each testing phase from Time 1 to Time 3.

- SES research has shown that low-SES backgrounds may result in a number of developmental delays, including tasks of cognitive skills. However, there is a lack of clarity regarding the effects of SES on children's EF performance in particular. Therefore, although a general trend in favour of children within the mid-SES group is expected across all measures of EF, no prediction is made regarding specific SES outcomes on each EF task. Language on the other hand has been shown to be strongly associated with SES therefore it is expected that children within the low-SES groups will perform more poorly on the PPVT test of English receptive vocabulary.

- Bilingualism research suggests that children display disadvantages relative to their monolingual peers on a range of linguistic tasks. As bilingual children use each of their languages at a lower frequency than a monolingual uses their one language, it is predicted that bilinguals will obtain lower scores on the PPVT-4.

- Inhibitory control: using previous bilingual findings it was predicted that bilinguals would outperform monolinguals on the IC tasks of interference suppression (Stroop) rather than response inhibition (Opposite Worlds). Bilingual effects are predicted to be particularly evident in the incongruent conditions requiring higher levels of IC.

- Switching: as bilinguals must constantly switch between language sets it is predicted that they will outperform monolinguals on the Creature Count switching task.

- Working memory: As research remains uncertain of the bilingual influence on WM skills and as recent studies have found no effect of bilingualism on WM tasks, this study does not predict a bilingual advantage for STM or WM measures.

- Unified executive function: Although this function is a relatively new area of research, studies do suggest that bilinguals may show superior skills on more complex cognitive tasks, therefore it is predicted that for the unified EF tasks, bilinguals will outperform monolinguals.

5.2. Method

5.2.1. Participants

A total of 147 children were assessed at Time 1 of this study. Exclusion criteria included: recipients of extra learning supports (ELS) in reading or mathematics, bilingualism (beyond English or Irish), participants with scores below the 10^{th} and above the 95^{th} percentile band on the RSPM and participants with developmental disorders such as autism spectrum disorder, dyslexia, dyspraxia and social and emotional problems. Following these exclusions, 117 children were included for analysis at Time 1.

Participants were matched for gender (50% females in mid-SES monolingual group, 56% females in mid-SES bilingual group, 44% females in low-SES monolingual group, 57% females in low-SES bilingual group), SES, and languages spoken (Irish and English for bilinguals and English only for monolinguals). There was a 96% retention rate of participants from Time 1 to Time 3 of testing. Main participant characteristics across time are presented in Table 4. The mean age of participants at Time 1 was 9 years 6 months (*SD* = 5.0 months; range 8 years 9 months – 10 years 6 months).

At Time 2 a total of 114 participants were assessed (97% retention rate from Time 1 from Time 2). The mean age of participants at Time 2 was 10 years 4 months (*SD* = 4.2 months; range 9 years 8 months – 11 years 4 months) with a mean gap of 10 months between Time 1 and Time 2. At Time 3, 112 participants were assessed (96% retention rate from Time 1 to Time 3). The mean age of participants at Time 3 was 11 years and 4 months (*SD* = 4.5 months; range 10 years 6 months – 12 years 1 month) with a mean gap of 10 months between Time 2 and Time 3.

In order to categorise participants into appropriate SES groups, individual and aggregate measures of SES were considered. To ensure that no further SES differences existed within each SES group, individual data from parent/guardian background questionnaires (see Appendix I) were also analysed. At the individual level, occupations were coded based on the International Standard Classifications of Occupations 2008 (ISCO-08) major grouping unit and were coded as: managers (1), professionals (2), technicians and associate professionals (3), clerical support workers (4), service and sales workers (5), skilled agricultural, forestry and fishery workers (6), craft and related trades workers (7), plant machine operators, and assemblers (8), elementary occupations (9). Another code was added to represent the sample as those who were unemployed were coded as 10. The parent/guardian with the highest occupational code was retained for analysis. Mean years in education were also calculated for each group (ISCED-97, UNESCO, 1997). Individuals who completed schooling to Junior Certificate level (GCSE) received approximately 11 years of formal education, Leaving Certificate completers (A-Levels) received approximately 14 years of formal education, undergraduate level 17 years, postgraduates/professional qualifications 18 years and PhD level or professional and postgraduate training 19 years.

Table 5 displays the means (and standard deviations) for each group's highest achieving parent/caregiver's average years in education and percentage rates within groups for occupational codes. Independent t-tests showed that there was a significant difference between mid-and low-SES groups' average years in education, t (36) = -4.87, p < .01, d = .40. The low-SES groups' highest academic attainment was significantly lower than that of the mid-SES group. These results showed that on average, participants in the low-SES group did not attain Leaving Certificate level (A-Level) of education, while the mid-SES group completed on average three years of further education beyond Leaving Certificate level. Within-group analysis (ANOVAs) showed that there was no significant difference between language groups' average years in education within the mid-SES or the low-SES group. Chi-squared results indicated a significant association between SES and occupation, χ^2 (9, N = 42) = 26.01, p < .01. The data file was split to examine language group differences within each SES group and no association between occupation code and language group was found within the mid-SES, χ^2 (4, N = 23) = 1.51, p = .83 or low-SES, χ^2 (7, N = 19) = 7.51, p = .38 groups.

Table 4 *Participant numbers, class year and ages across time points for longitudinal study*

Group Demographic	Socioeconomic Group	Language Group	Time 1	Time 2	Time 3
Class Group			3rd Class (P5)	4th Class (P6)	5th Class (P7)
Age Range			8-10years	9-11years	10-12years
n	Mid-SES	Monolingual	32	32	32
		Bilingual	48	46	45
		Total	80	78	77
	Low-SES	Monolingual	18	18	17
		Bilingual	19	18	18
		Total	37	36	35
	Total		117	114	112

Table 5 *Means for highest occupation and academic achievement demographics for each socioeconomic group*

		Low-SES Group			Mid-SES Group		
		Monolingual	Bilingual	Total	Monolingual	Bilingual	Total
Response Rate		44%	53%	49%	50%	28%	39%
Mean Years in Education		13 years	14 years	13.5 years	17 years	16 years	16.9 years
(SD)		(2.67)	(2.27)	(2.45)	(1.82)	(2.15)	(1.91)
Occupation Code	Occupation Category						
1	Managers	-	-	-	19%	29%	22%
2	Professionals	-	-	-	44%	43%	44%
3	Technicians and associate professionals	-	18%	11%	19%	14%	17%
4	Clerical support workers	25%	9%	16%	6%	14%	9%
5	Service and sales workers	13%	27%	21%	13%	-	9%
6	Skilled agricultural, forestry and fishery workers	13%	9%	11%	-	-	-
7	Craft and related trades workers	13%	-	5%	-	-	-
8	Plant machine operators, and assemblers	25%	9%	16%	-	-	-
9	Elementary occupations	13%	18%	16%	-	-	-
10	Unemployed	-	9%	5%	-	-	-

5.2.2. Apparatus and procedure

For details of the apparatus and procedure for this longitudinal study see Chapter 5 (section 5.4. and 5.5).

5.2.3. Analysis

For tasks generating standardised scores, mixed Analyses of Variance (ANOVA) were used to analyse data over time (5.3.2.1. and 5.3.3.). For tasks generating raw scores, individual growth curve models (IGCs) were used to investigate the change in scores across time (see section 5.3.2.2. and 5.3.5.1.).

Growth Curve Analysis

This study used hierarchical linear modelling (HLM) growth curve analyses to assess the raw data from the following outcome variables: Opposite Worlds (SW, OW and Worlds Difference conditions), Trail Making Test (Trails A, Trails B and Trails Difference conditions) the Colour-Word Stroop (Neutral, Congruent, Incongruent, Facilitation and Inhibition conditions). Growth curve modelling was used as a method to overcome some of the limitations enforced by traditional statistical measures of change (e.g. listwise deletion, assumptions of independence of observations in ANOVA). Unlike ANOVA, individual growth curve (IGC) models allow researchers to model within-person systematic change as well as between-person differences over time (Shek et al., 2011). IGC is also helpful when using unbalanced data over time (e.g. inconsistent time intervals and missing data). In contrast to other statistical models, IGC allows for varied numbers and spacing between measurement intervals. It also enables researchers to: assess intra- (within-person) and inter-individual (between groups) differences in growth parameters, examine the influence of predictors on individual growth, and explore the causal links between these predictors and explore changes in outcome variables across time. Raw scores rather than standardised scores were used to assess participants' true change over time (Stoolmiller & Bank, 1995). Using HLM

IGC this study examined the changes over time in each of the outcome variables individually using a two-level IGC model.

Longitudinal Data Set

The multilevel model approach to IGC modelling was used in this study to decipher language group and SES group trajectories (Singer, 1999; Singer & Willet, 2003). The data were restructured and analysed using SPSS in line with the protocol laid out by Shek and colleagues (2011). Most of the missing cases within the data set were as a result of participant absence on the day of collection, although a number of participants were lost through attrition between Time 1 and Time 3 (96% retention rate). Person-level data were converted to person-period data or a univariate format where each individual has multiple records, one for each time point (i.e. Time 1, Time 2, Time 3), with four types of variable: subject identifier, time indicator, outcome variable and predictor variables (Language Group and SES Group). As mentioned, the strength of IGC is that it allows for irregularity in the number of and spacing between, time points. To prevent against age effects at Time 1, an ‘age centred’ variable was calculated to control for any age differences. This age centred variable was then added as a predictor to each of the growth curve models but was not significant for any of the measures tested.

Data Analytic Strategy

The data were analysed using a mixed effects model with restricted maximum likelihood (REML) estimation. REML was selected as it typically provides less biased estimates of the variance components of the model than full information maximum likelihood (ML) and is particularly useful when using a small number of groups (Raudenbush & Bryk, 2002). As the analysis examined individual changes over time as well as the effects of predictors (Language and SES Groups) a two-level hierarchical model was created. To investigate individual changes over time (intra-individual change), individual trajectories were summarised by regressing observed values on to the average trajectory or by fitting them to a specific form of parametric model using ordinary least squares (OLS) regression (Singer & Willet, 2003). Group differences (interindividual changes) are explored through the heterogeneity of changes or how much the proposed trajectory relates to the predictor variables. Therefore there are two levels in the IGC model. The Level-1 model relates to the within-person or intra-individual change model (repeated-measurement over time). It examined the individual variation or developmental changes for individuals over time. The Level-1 model, below, estimated the average within-person initial status and rate of change over time without the inclusion of predictors:

$$Y_{ij} = \beta_{0i} + \beta_{1i}\,(\text{Time}) + r_{ij}$$

β_0 refers to the initial status or intercept (e.g. Time 1) representing individual participants' ($_i$) outcome variable score. β_1 is the linear rate of change or time slope for individual $_i$ and r_{ij} is the residual in the outcome variable for individual $_i$ at Time $_t$. If the effect of linear growth (Time, β_1) is statistically significant then nonlinear individual growth trajectories or higher-order polynomial trends can be explored (i.e. quadratic and cubic slopes).

The Level-2 model examined whether the rate of change varies or remains stable across individuals. The growth parameters (within-subjects intercept and slope) are used to predict the between-subjects variables at Level 2. The equation below demonstrates the Level-2 model where an explanatory variable (W_j) is included to analyse the predictor's effect on interindividual variation on outcome variables:

$$Y_{ij} = \gamma_{0i} + \gamma_{1i}\,(\text{Time}) + \gamma_{2i}W_j + r_{ij}$$

In the above example, Y*ij* represents the grand mean for the outcome variable for the entire sample at Time *t*. γ_{0i} is the initial status of the outcome variable for the whole sample at Time *t*. γ_{1i} is the linear slope of change relating to the outcome variable for the whole sample at Time *t*. γ_{2i} is used to test whether each of the predictors (i.e. Language and SES) are associated with the growth parameters (i.e. initial status, linear growth). r_{ij} refers to the random effects (i.e. amount of variance) that are explained by the predictor (s).

This study tested whether Language Group (monolingual and bilingual) and SES Group (mid-SES and low-SES) were predictive of the developmental trajectories of children's EF scores. EF was examined using a number of outcome measures: Opposite Worlds, Trail Making Test and Colour-Word Stroop. Using the method set out by Singer and Willett (2003) a number of models were tested to explore children's trajectories: unconditional mean model (Model 1) to test the amount of intra- and interindividual variance explained by the outcome variable, unconditional growth model (Model 2) to explore whether growth curves were linear or curvilinear, higher-order polynomial model (Model 3) to test if the rate of change accelerated or decelerated across time, conditional model (Model 4) to investigate the relationship of predictors to growth parameters (initial status, linear growth, quadratic growth). The intercept and linear slope were allowed to vary across individuals and missing cases were handled through pairwise deletion.
For growth modelling two main parameters are important for interpretation: the "intercept parameter" representing the initial status of the outcome variable and the "slope parameter" which describes the rate of growth. Both can be interpreted as regular regression coefficients (Mai Luu et al., 2009). The intercept estimate referred to the initial outcome variable at Time 1 of testing and the slope estimate expressed the rate of change in the outcome variable score from Time 1 to Time 3 of testing. Growth-curve analysis also produces a random effect component (or variance component) describing the residual variability within and between individuals. This enabled a computation of the explainable variation accounted for by predictors (Singer & Willet, 2003).

To select model of best fit, -2 log likelihood (i.e., likelihood ratio test/deviance test), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) were used (model fit statistics for outcome variables presented in Appendix V).

IGC also examines the proportion of total outcome variation that is related to interindividual differences (i.e., intraclass correlation coefficient, ICC). ICC describes the amoung of variance in the outcome variable that is attributed to differences between individuals and the higher the value the higher the estimated average stability of the dependent variable over time. If the estimated average stability (ICC) of the outcome variable is below 0.25 or 25% than IGC might not be better than traditional methods of analysis (i.e. ANOVA) in estimating fixed effects (De Leeuw & Kreft, 1995; Singer & Willett, 2003).

5.2.4. Grouping of tasks for results

To examine aspects of the EF in more detail, tasks were grouped according to the function they were proposed to test. This grouping system aims to assist the reader in light of the amount of data being presented. Organisation of the results section is as follows:

Section 5.3. Control Measures

Section 5.3.1. Results Table

Section 5.3.2. Inhibitory control tasks
5.3.2.1. Opposite Worlds
5.3.2.2. Colour-word Stroop

Section 5.3.3. Switching task
5.3.3.1. Creature Count

Section 5.3.4. Working memory tasks
5.3.4.1. Working Memory Test Battery-for Children

Section 5.3.5. Unified executive function tasks
5.3.5.1. Trail Making Test
5.3.5.2. Wisconsin Card Sort Test

Section 5.3.6. Relationships between tasks

5.3. RESULTS

Control measures were used to assess children's English receptive vocabulary (PPVT) and non-verbal intelligence (RSPM). Descriptive statistics for these measures are shown in Table 6.

T-tests indicated that language groups (monolingual and bilingual) did not significantly differ on the PPVT, t (115) = -.52, p = .61, d = .002 or RSPM, t (103) = 1.62, p = .11, d = .02. However, there was a significant difference between SES groups on both the PPVT, t (115) = -3.55, p < .01, d = .10 and RSPM, t (103) = -3.0, p < .01, d = .08. Mid-SES participants' performance on both the PPVT and RSPM was better than Low-SES groups.

Within the mid-SES group, monolinguals and bilinguals did not differ significantly on the PPVT; although monolinguals had a higher mean score overall than bilinguals. There was also no significant difference between monolinguals and bilinguals on the RSPM although means for bilinguals were higher than for monolinguals. In the low-SES groups, no difference was observed between language groups on the PPVT or on the RSPM although mean scores indicated that bilinguals scored higher on both measures than monolinguals.

Table 6 *SES and language group means (and standard deviations) for control measures*

Task	SES Group	Monolinguals	Bilinguals	Total
PPVT	Mid-SES	104.91 (10.01) *n* = 32	101.06 (12.67) *n* = 48	102.60 (11.77) *n* = 80
	Low-SES	92.56 (13.94) *n* = 18	94.26 (16.89) *n* = 19	93.43 (15.33) *n* = 37
	Total	100.46 (12.91) *n* = 50	99.13 (14.20) *n* = 67	99.70 (13.62) *n* = 117
R-SPM	Mid-SES	48.28 (25.89) *n* = 29	54.02 (26.49) *n* = 41	56.74 (26.93) *n* = 70
	Low-SES	32.50 (17.93) *n* = 18	40.88 (21.52) *n* = 17	36.57 (19.92) *n* = 35
	Total	42.23 (24.22) *n* = 47	50.17 (25.67) *n* = 58	46.62 (25.23) *n* = 105

Note: PPVT standard score 100 = average per population. RSPM standard score 50 = average per population.

5.3.1. Results Table

Table 7 *Summary of tasks, sections and effects for all outcome measures*

Task	Section	Time Effect	Language Group Effect	SES Effect	Within Group Effects
PPVT	5.3. Control	Not applicable	No	**Yes** Low-SES scores lower than Mid-SES	None
Raven's SPM	5.3. Control	Not applicable	No	**Yes** Low-SES scores lower than Mid-SES	None
Opposite Worlds	5.3.2.1. IC Task				
Same World (congruent)	5.4.2.1.1.	**Yes** Scores increased across time	No	No	**Yes** MSBs outperformed MSMs at Time 3
Opposite Worlds (incongruent)	5.3.2.1.2.	**Yes** Scores increased across time	No	No	**Yes** LSBs outperformed LSMs at Time 1
Worlds Difference	5.3.2.1.3.	No	**Yes** Bilinguals faster timing scores than monolinguals, particularly at Time 3	No	**Yes** LSBs had significantly faster timing scores at Time 3 than LSMs. LSMs had slower timing scores than MSMs at Time 1 and Time 3
Colour-Word Stroop	5.3.2.2. IC Task				
Neutral RT	5.3.2.2.1.	**Yes** RTs decreased across time	No	**Yes** SES predictor of linear decrease in Neutral RTs	**Yes** MS group improved RTs at a faster rate than Low-SES group
Congruent RT	5.3.2.2.2.	**Yes** RTs decreased across time	No	**Yes** SES predictor of linear decrease in Neutral RTs	None MS group improved RTs at a faster rate than Low-SES group
Incongruent RT	5.3.2.2.3.	**Yes** RTs decreased across time	No	No	None
Facilitation RT	5.3.2.2.4.	**Yes** Timing scores (ms) decreased across time	No	No	None
Inhibition RT	5.3.2.2.5.	**Yes** Timing scores (ms) decreased across time	No	No	**Yes** MSBs had significantly slower Inhibition times at Time 1 but caught up to MSMs by Time 2 and Time 3.
Creature Count	5.3.3.1 Switching Task				
Timing		**Yes** Scores increased across time	No	No	None
Creature Count Accuracy	5.3.3.1.1.	**Yes** Standardised scores increased across time as raw timing scores decreased	No	**Yes** Mid-SES groups performed significantly better than Low-SES group	**Yes** MSBs outperformed MSMs at Time 3.

Task	Section	Time Effect	Language Group Effect	SES Effect	Within Group Effect
WMTB-C	5.3.4.1. WM Task				
Forward Digit	5.3.4.1.	Not applicable	No	No	None
Backward Digit	5.3.4.1	Not applicable	No	**Yes** MS group better	None
Forward Blocks	5.3.4.1.	Not applicable	No	No	None
Backward Blocks	5.3.4.1.	Not applicable	No	**Yes** MS groups better	None
Trail Making Task	5.3.5.1. Unified EF tasks				
Trails A	5.3.5.1.1.	**Yes** Raw timing scores showed linear decrease across time	No	**Yes** SES predictor of change (Mid-SES groups improved their timing scores at a faster rate than Low-SES groups)	None
Trails B	5.3.5.1.2.	**Yes** Raw timing scores showed quadratic growth (rate of linear decrease in timing scores slowed from Time 2 to Time3)	**Yes** Language group was a significant predictor of the linear but not quadratic change in timing scores. Bilinguals showed linear decrease while monolinguals reduction in timing scores slowed from Time 2 to Time 3	**Yes** SES predictor of initial status but not change. Low-SES timing scores began and remained slower than Mid-SES participants	None
Trails Difference	5.3.5.1.3.	**Yes** Raw timing scores significantly decreased across time.	**Yes** Bilinguals had significantly reduced Trails Difference times compared with monolinguals, particularly at Time 2 and Time 3	**Yes** Low-SES participants had significantly higher timing scores than Mid-SES participants across time	**Yes** LSBs outperformed the LSMs at Time 2 and Time 3.
Wisconsin Card Sort Task	5.3.5.2. Unified EF tasks				
Total Errors		Not applicable	Yes Bilinguals outperformed monolinguals	No	Yes MSBs outperformed MSMs
Perseverative Responses		Not applicable	Yes Bilinguals outperformed monolinguals	No	Yes MSBs outperformed MSMs
Perseverative Errors		Not applicable	Yes Bilinguals outperformed monolinguals	No	Yes: MSBs outperformed MSMs
Nonperseverative Errors		Not applicable	Yes Bilinguals outperformed monolinguals	No	Yes MSBs outperformed MSMs
No. of Categories Complete		Not applicable	Yes: Bilinguals outperformed monolinguals	No	Yes: MSBs outperformed MSMs
No. of Trials to Complete 1st Category		Not applicable	Yes: Bilinguals outperformed monolinguals	No	Yes: MSBs outperformed MSMs
Failure to Maintain Set		Not applicable	Yes: Bilinguals outperformed monolinguals	No	Yes: MSBs outperformed MSMs

5.3.2. Inhibitory control tasks

5.3.2.1. Opposite Worlds Task

This standardised task had two conditions: Same World or congruent and Opposite Worlds or incongruent. Three-way mixed ANOVAs were used to compare changes over time and any differences between language and SES groups. The analyses performed were 3 x 2 x 2 mixed ANOVAs with two between-groups variables, each with two levels: Language Group (Monolingual and Bilingual) and Socioeconomic Status (Mid-and Low-SES) and one within-subjects variable of Time which had three levels (Time 1, Time 2 and Time 3). Descriptive statistics for the Same Worlds and Opposite Worlds standardised timing scores are displayed in Table 8 and Table 10 respectively (raw timings scores in seconds) for the Same and Opposite Worlds conditions are displayed in Appendix IV).

Raw scores from the task were converted to age-scaled scores according to the procedure laid out by Manly et al. (1999) and in line with other cognitive tasks e.g. the WISC-III (Weschler, 1991). Age-scaled scores have a mean of 10 with a standard deviation of 3. However, percentile band scores may be more meaningful for generalisation therefore age-scaled scores can be converted using Table 3. A child who scores in the 43.4th-56.6th percentile, for example, performs better than approximately 50% of children of his/her age and gender. A third condition, ‘Worlds Difference’ scores was also assessed (see section 4.5.3.2.). This variable was proposing to show the total amount of extra EF skill necessary to perform the incongruent conditions of the task compared to the congruent, and is similar to the inhibition condition generated in the Colour-word Stroop task (see section 5.5.4).

Opposite Worlds IGC Results

Raw reaction times (RTs) for the Same World, Opposite World and Worlds Difference conditions were analysed using IGC. As Language and SES groups did not predict any of the changes in the Opposite World task these results are not reported here.

Standardised Same World

Standardised Same World scores were calculated using the converted raw timing scores derived by adding the raw times from the two Same World trials (Manly et al., 1999). A total of 111 participants were included in the analysis; 76 participants in the Mid-SES group (32 monolingual and 44 bilingual) and 35 participants in the Low-SES group (17 monolingual and 18 bilingual). Mauchly's test indicated that the assumption of sphericity had not been violated so the Sphericity Assumed was employed (see Table 8 for descriptive statistics).

Table 8 *Means (and standard deviations) for Same World age-scaled scores across time*

Socioeconomic Status	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (*n* = 32)	8.56 (1.74)	9.66 (1.56)	10.03 (1.81)
	Bilingual (*n* = 44)	8.93 (2.39)	10.18 (2.33)	11.18 (2.47)
	Total (*n* = 75)	8.78 (2.13)	9.96 (2.04)	10.70 (2.28)
Low-SES	Monolingual (*n* = 17)	8.71 (1.40)	9.35 (1.69)	10.59 (1.91)
	Bilingual (*n* = 18)	9.44 (2.71)	9.39 (1.82)	9.89 (1.91)
	Total (*n* = 35)	9.09 (2.17)	9.37 (1.73)	10.23 (1.91)
Total	Monolingual (*n* = 49)	8.61 (1.62)	9.55 (1.60)	10.22 (1.84)
	Bilingual (*n* = 62)	9.08 (2.47)	9.95 (2.21)	10.81 (2.38)
	Total (*n* = 111)	8.87 (2.14)	9.77 (1.96)	10.55 (2.17)

There was a significant three-way interaction between time, SES and language group, F (2, 214) = 4.01, p = .02, η_p^2 = .04. There was no significant main effect of language group, F (1, 107) = .96, p = .33, η_p^2 = .01, or SES, F (1, 107) = .30, p = .59, η_p^2 = .003, and no significant interaction between SES and language group {F (1,

107) = .83, p = .36, η_p^2 = .01} for Same World scores. Figure 15 shows the three-way interaction observed between time, language group and SES. Age-scaled timing scores improved from Time 1 to Time 3 across all groups. Overall, the Mid-SES bilingual (MSB) group had the largest increase in standard scores from Time 1 to Time 3. However, the group with the smallest increase in scores was the Low-SES bilingual (LSB) group whose age-scaled scores and percentile band scores did not change from Time 1 to Time 3. No other interactions or main effects were found.

There was a significant main effect of Time, F (2, 214) = 29.26, $p < .01$, η_p^2 = .22. Pairwise comparisons indicated that age-scaled scores increased significantly ($p <$ .01) from Time 1 to Time 2 and from Time 2 to Time 3 ($p < .01$). Overall, participants' percentile band groupings increased from within the 30.9th-43.4th percentile band at Time 1 to within the 56.6th -69.2nd percentile band at Time 3. Participants in the mid-SES group performed below average at Time 1 (within the 30.9th -43.4th percentile band) but above average at Time 3 (within the 56.6th -69.2nd percentile band). Participants in the Low-SES group performed below average at Time 1 (within the 30.9th -43.4th percentile band) but at average by Time 3 (within the 43.4th-56.6th percentile band).

SES Groups across Time

In the mid-SES group, there were no significant differences between language groups' Same World scores at Time 1, F (1, 74) = .55, p = .46, η^2 = .01 or Time 2, F (1, 74) = 1.23, p = .27, η^2 = .02. At Time 3 there was a significant difference between monolinguals' and bilinguals' scores, F (1, 74) = 4.99, p = .03, η^2 = .06; bilinguals had a higher age-scaled mean at Time 3 than monolinguals (Fig. 16.). In the low-SES group, there was no significant difference between language groups' standardised Same World scores at Time 1, F (1, 33) = 1.01, p = .32, η^2 = .03, Time 2, F (1, 33) = .004, p = .95, η^2 = 1.08 or Time 3, F (1, 33) = 1.18, p = .29, η^2 = .03.

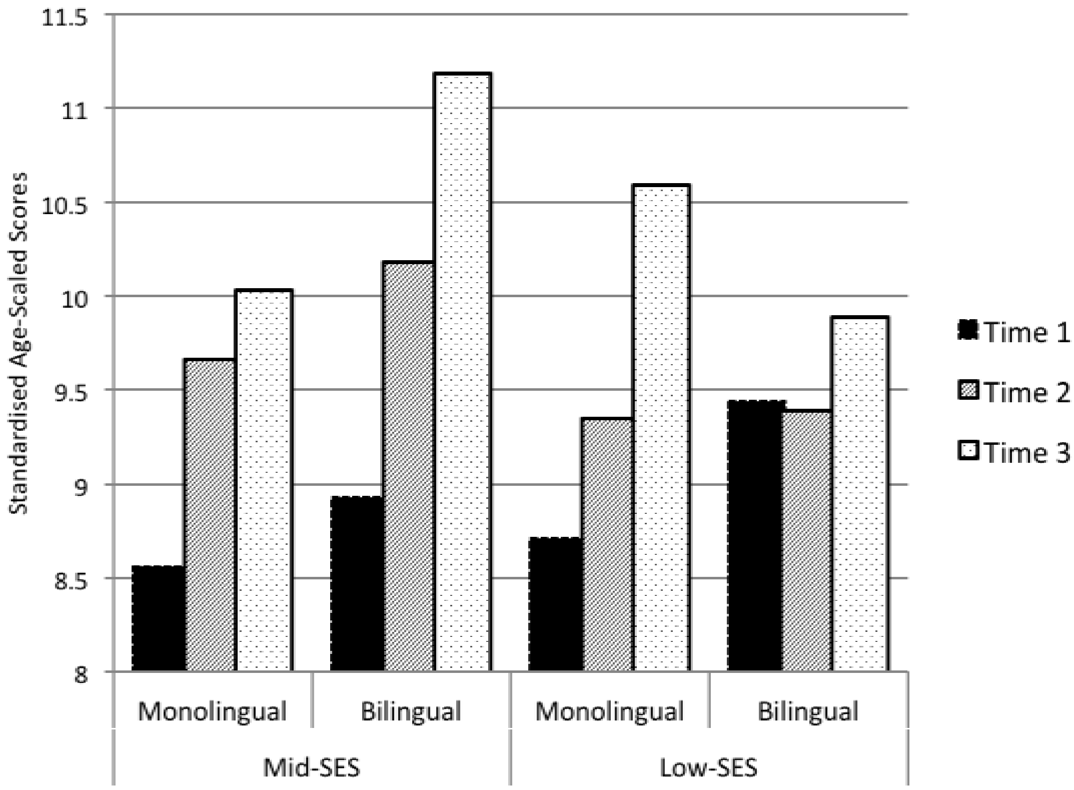


Figure 15 *Time x SES x Language group interaction for Same Worlds timing scores*

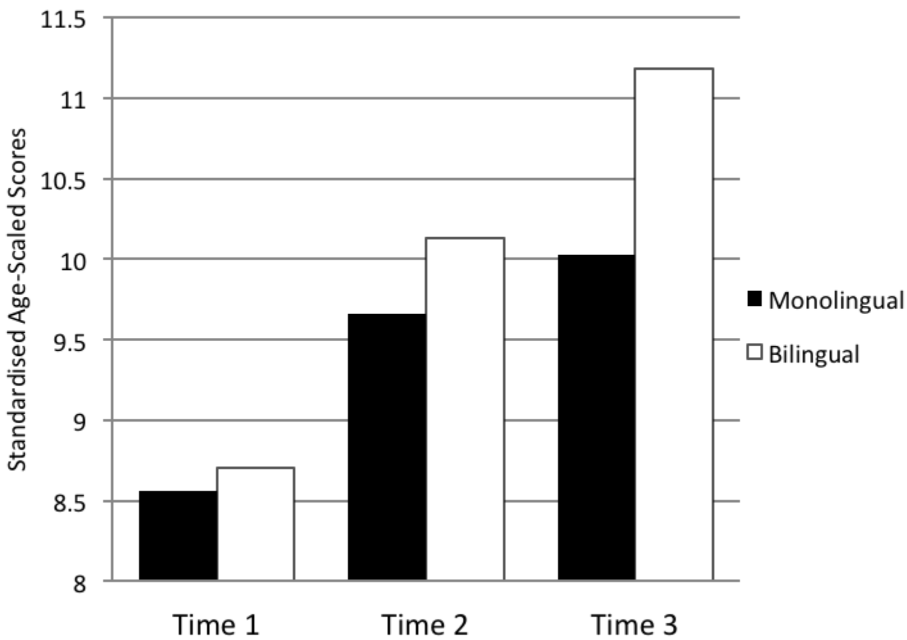


Figure 16 *Mid-SES language group differences in Same World scores across time*

Between groups comparison

One-way ANOVAs were used to explore how each set of Same World scores differed between groups at each time point (Time 1, Time 2, and Time 3). The between-groups variable had four levels: Low-SES Monolingual (LSM), Low-SES Bilingual (LSB), Mid-SES Monolingual (MSM), and Mid-SES Bilingual (MSB). No significant difference was found between groups at Time 1, F (3, 107) = .69, p = .56, Time 2, F (3, 107) = 1.17, p = .33 or Time 3, F (3, 107) = 2.51, p = .06 (see Table 8).

Individual Group Changes across Time

To examine how each group's Same World scores changed over time, the data file was split according to group and paired-sample t-tests compared scores between all three time points: Time 1 – Time 2 (pair one), Time 2 – Time 3 (pair two) and Time 1 – Time 3 (pair three). See Table 9 for breakdown of paired-samples t-tests results within each group.

The ***MSM*** group scores increased significantly between Time 1 and Time 2 and between Time 1 and Time 3. Times 2 and Time 3 did not differ. The ***MSB*** group showed a significant change in scores over time. There was a significant increase in scores between Time 1 and Time 2, Time 2 and Time 3 and Time 1 and Time 3. In the ***LSM*** group, no significant differences were found Time 1 and Time 2. However a significant difference was observed between Time 2 and Time 3 and between Time 1 and Time 3. The ***LSB*** group showed no differences between Time 1 and Time 2, between Time 2 and Time 3 or between Time 1 and Time 3.

Table 9 *Paired-sample t-tests comparing Same World age-scaled scores*

Group	Pair One Time 1 – Time 2	Pair Two Time 2 – Time 3	Pair Three Time 1 – Time 3
Mid-SES Monolingual			
t (31)	3.37**	-1.40	-4.33**
Effect size	.27	.38	-
Mid-SES Bilingual			
t (43)	-4.24**	-3.73**	-8.09**
Effect size	.29	.24	.60
Low-SES Monolingual			
t (16)	-1.58	-2.65*	-4.80**
Effect size	-	.30	.59
Low-SES Bilingual			
t (17)	.09	-1.23	-.65
Effect size	-	-	-

*Note: *p < .05, **p < .01. Effect sizes are Cohen's d.*

Standardised Opposite Worlds

A mixed ANOVA was used to examine participants' standardised Opposite Worlds timing scores over time. Standardised scores were calculated using the converted raw timing scores derived by adding the raw times from the two Opposite Worlds trials (Manly et al., 1999, see section 4.5.3.2.). A total of 110 participants were included in the analysis; 75 participants in the Mid-SES group (31 monolingual and 44 bilingual) and 35 participants in the Low-SES group (17 monolingual and 18 bilingual). Descriptive statistics are presented in Table 10. Mauchly's test indicated that the assumption of sphericity had not been violated so the Sphericity Assumed was employed.

Table 10
Means (and standard deviations) for the Opposite World age-scaled scores across time

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (*n* = 44)	8.84 (2.15)	9.97 (1.98)	10.23 (2.08)
	Bilingual (*n* = 31)	9.11 (2.37)	10.52 (2.61)	11.23 (1.70)
	Total (*n* = 75)	9.00 (2.27)	10.29 (2.37)	10.81 (2.28)
Low-SES	Monolingual (*n* = 17)	7.41 (1.91)	9.29 (1.11)	9.88 (2.23)
	Bilingual (*n* = 18)	9.39 (3.48)	9.67 (2.09)	10.22 (1.70)
	Total (*n* = 35)	8.43 (2.96)	9.49 (1.67)	10.06 (1.96)
Total	Monolingual (*n* = 48)	8.33 (2.16)	9.73 (1.74)	10.10 (2.12)

Bilingual (*n* = 62)	9.19 (2.72)	10.27 (2.48)	10.94 (2.21)
Total (*n* = 110)	8.82 (2.51)	10.04 (2.20)	10.57 (2.20)

As in the Same World condition, there was a significant three-way interaction between Time, SES and Language Group, F (2, 212) = 4.33, p = .01, η_p^2 = .04. There was no significant main effect of language group, F (1, 106) = 3.58, p = .06, although this figure was approaching significance, no significant main effect of SES, F (1, 106) = 2.85, p = .10, and no significant interaction between SES and language group, F (1, 106) = .13, p = .72. No other main effects or interactions were found.

Figure 17 shows the three-way interaction between Time, Language Group and SES. Although the scores appeared to increase across time, the three-way interaction may have arisen from the LSB group's scores. While this group had the highest mean overall at Time 1, their scores did not increase at the same rate as the other three groups. At Time 2 and Time 3 the MSM and MSB groups had higher mean scores than the LSB group. The effect size of the three-way interaction was also small at .04.

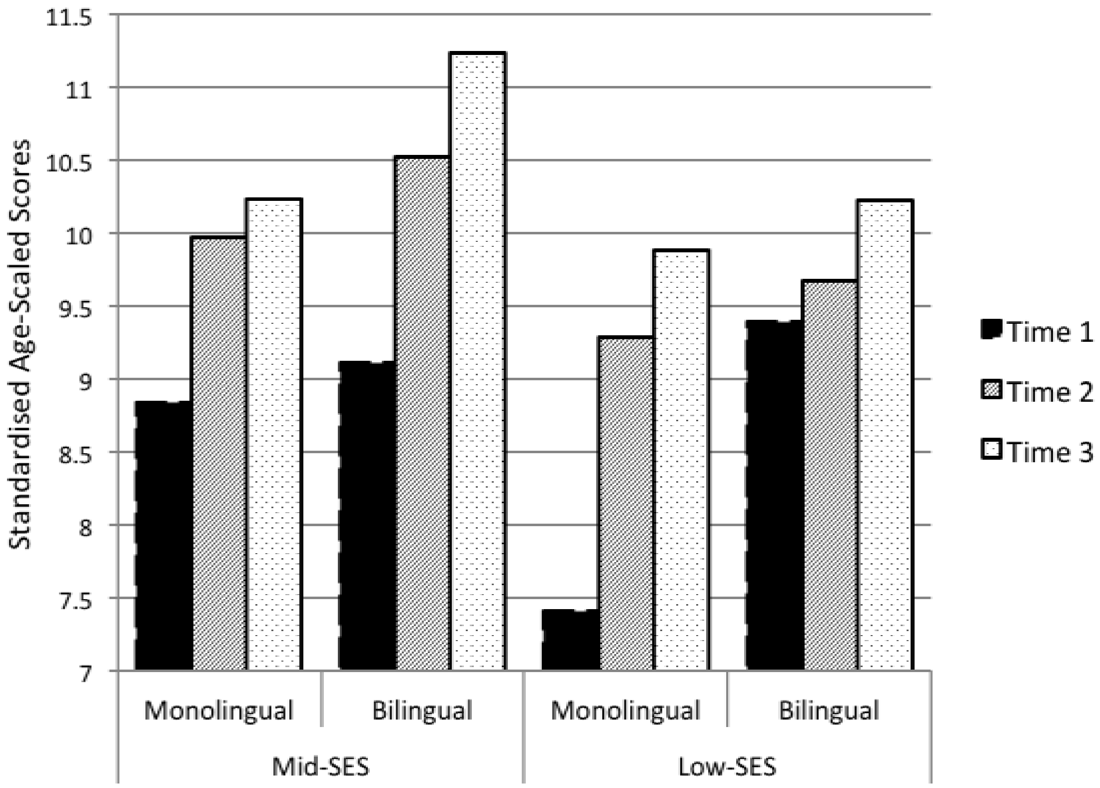


Figure 17 *Time x SES x Language group interaction for Opposite Worlds timing scores*

There was a significant main effect of Time, *F* (2, 212) = 33.62, $p < .01$, $\eta_p^2 = .24$. Pairwise comparisons revealed that age-scaled scores increased significantly ($p < .01$) from Time 1 to Time 2 and from Time 2 to Time 3 ($p < .01$). However, the main effect of Time was modified by the significant three-way interaction. Participants scored below average at Time 1 (30.9th - 43.4rd percentile band), at around average at Time 2 (43.4th-56.6th percentile band) and above average at Time 3 (56.6th-69.2nd percentile band). The below average result at Time 1 may be confounded by the LSM group who scored within the 12.2nd to 20.2nd percentile band at Time 1, well below population average, although they did improve their scores and by Time 3 were performing at average (43.4th-56.6th percentile band). Overall, the MSB group had the highest percentile band scores and by Time 3 and were 1 age-scaled score above the MSM group performing above average on the Opposite Worlds condition (56.6th-69.2nd percentile band).

SES Groups across Time

In the mid-SES group there were no significant differences between language groups' age-scaled timing scores at Time 1, F (1, 73) = .26, p = .61, Time 2, F (1, 73) = .10, p = .32 or Time 3, F (1, 73) = 3.65, p = .06 although scores at Time 3 were approaching significance. In the low-SES group there was no significant difference language groups' age-scaled scores at Time 2, F (1, 33) = 43, p = .52 or Time 3, F (1, 33) = .26, p = .61. However at Time 1 there was a significant difference between language groups, F (1, 33) = 4.27, p = .05, η^2 = .11, with bilinguals scoring significantly higher than monolinguals (Fig. 18.).

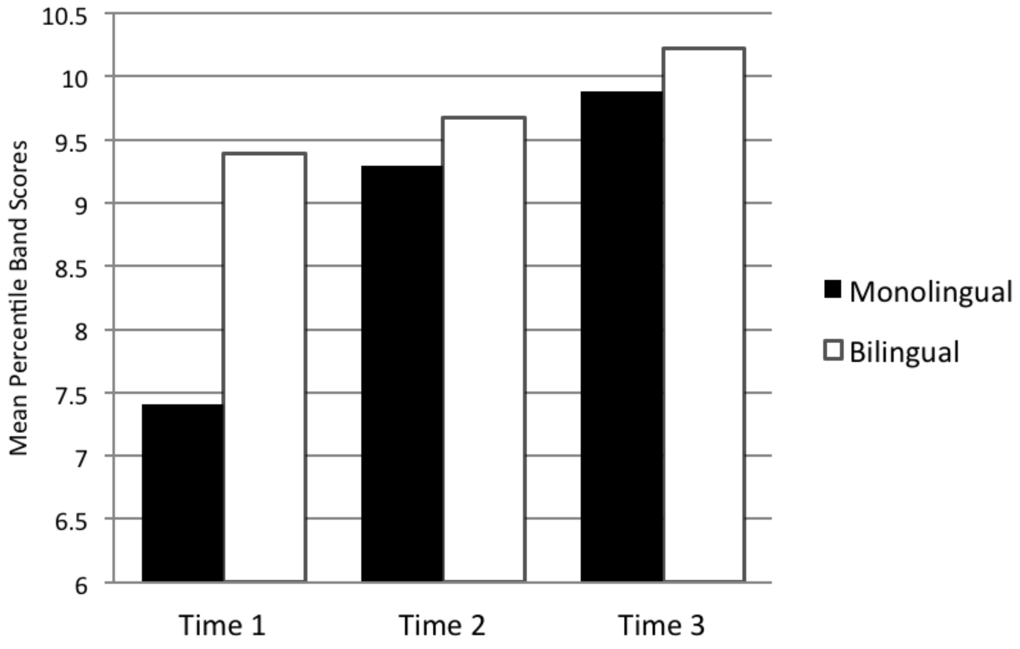


Figure 18 *Low-SES Opposite Worlds scores for Language Groups across Time*

Between groups comparison

One-way ANOVAs explored how each set of age-scaled Opposite World scores differed between all four groups (MSM, MSB, LSM, LSB) at each time point (Time 1, Time 2, and Time 3). No significant difference was found between groups at Time 1, F (3, 106) = 2.37, p = .07, Time 2, F (3, 106) = 1.57, p = .10, or Time 3, F (3, 106) = 2.35, p = .08 (see Table 10 for means).

Individual Group Changes across Time

To examine how each group's scores changed across time, paired-sample t-tests were used to compare scores between all three time points: Time 1 – Time 2 (pair one), Time 2 – Time 3 (pair two) and Time 1 – Time 3 (pair three). See Table 11 for paired-samples t-tests results within each group. The ***MSM*** group's scores increased significantly from Time 1 to Time 2 and between Time 1 and Time 3. Time 2 and Time 3 scores did not differ. The ***MSB*** group showed a significant change in scores between all three pairs. There was a significant increase in scores between Time 1 and Time 2, Time 2 and Time 3 and Time 1 and Time 3. In the ***LSM*** group, significant differences were found between Time 1 and Time 2 and Time 1 and Time 3. Scores did not differ between Time 2 and Time 3. The ***LSB*** scores did not differ across time.

Table 11 *Time differences between for the Opposite Worlds standardised scores*

Group	Pair One Time 1 – Time 2	Pair Two Time 2 – Time 3	Pair Three Time 1 – Time 3
Mid-SES Monolingual			
t (30)	-4.54**	-.74	-4.58**
Effect size	.41	-	.41
Mid-SES Bilingual			
t (43)	-4.05**	-2.23*	-7.99*
Effect size	.28	.10	.60
Low-SES Monolingual			
t (16)	-4.92**	-1.19	-4.34**
Effect size	.60	-	.54
Low-SES Bilingual			
t (17)	-.37	-1.66	-1.19
Effect size	-	-	-

*Note: *p < .05, **p < .01. Effect sizes are Cohen's d.*

Worlds Difference raw timing scores

There was no significant main effect of time, F (2, 212) = 3.88, p = .69, η_p^2 = .004, although mean scores did show a general decrease from Time 1 to Time 3. There was no significant main effect of SES, F (1, 106) = 3.66, p = .06; η_p^2 = .03, and no interaction between SES and language group, F (1, 106) = 3.23, p = .08; η_p^2 = .03. However, there was a significant main effect of language group, F (1, 106) = 4.26, p

= .04; η_p^2 = .04. Mean scores (Table 12) and Figure 19 show that bilinguals had significantly lower Worlds Difference timing scores than monolinguals across time and particularly at Time 3 of testing. No other interactions or main effects were found.

Table 12 *Means (and standard deviations) for Worlds Difference response times (sec) across time*

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (n = 32)	6.72 (3.78)	6.08 (3.35)	5.26 (2.60)
	Bilingual (n = 44)	7.14 (3.96)	5.80 (3.73)	4.80 (2.20)
	Total (n = 76)	6.96 (3.87)	5.92 (3.55)	5.00 (2.37)
Low-SES	Monolingual (n = 17)	10.12 (8.86)	7.0 (2.76)	7.16 (2.73)
	Bilingual (n = 18)	7.67 (6.82)	6.20 (3.35)	4.26 (2.13)
	Total (n = 35)	8.86 (5.81)	6.59 (3.06)	5.67 (2.82)
Total	Monolingual (n = 49)	7.90 (4.27)	6.40 (3.16)	5.92 (2.77)
	Bilingual (n = 62)	7.29 (4.91)	5.92 (3.60)	4.65 (2.18)
	Total (n = 111)	7.56 (4.63)	6.13 (3.41)	5.21 (2.53)

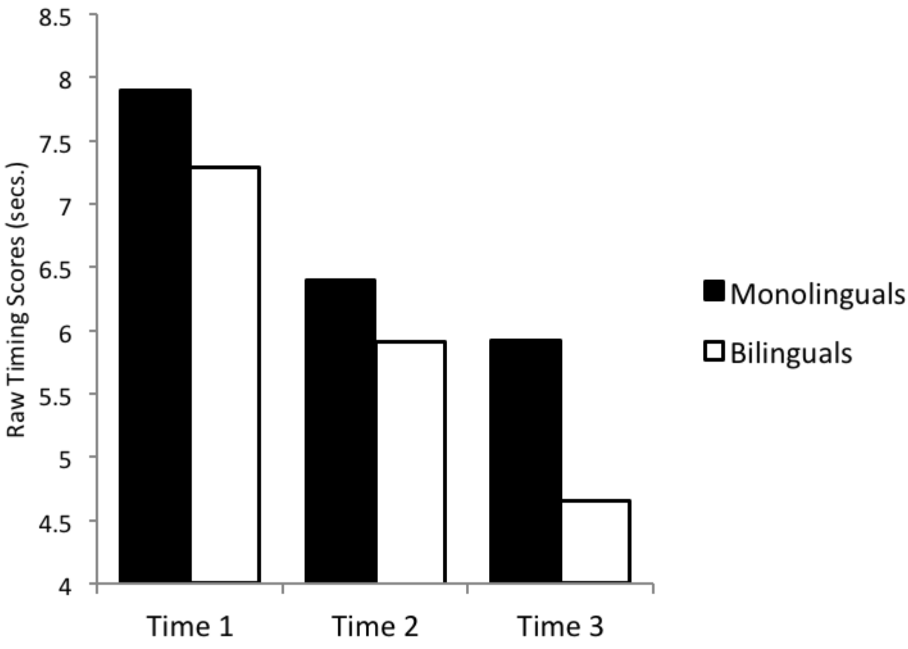


Figure 19 *Worlds Difference raw timing scores for language groups and across time*

To assess the Language difference further, Worlds Difference RTs were investigated at each time point (Time 1, Time 2 and Time 3) using one-way ANOVAs. Although bilinguals mean RTs were faster than monolinguals at all three time points, only results at Time 3 were significant, F (1, 110) = 7.40, p = .01; η^2 = .06. It is possible that the LG effect in the previous mixed ANOVA was heavily influenced by the Time 3 results.

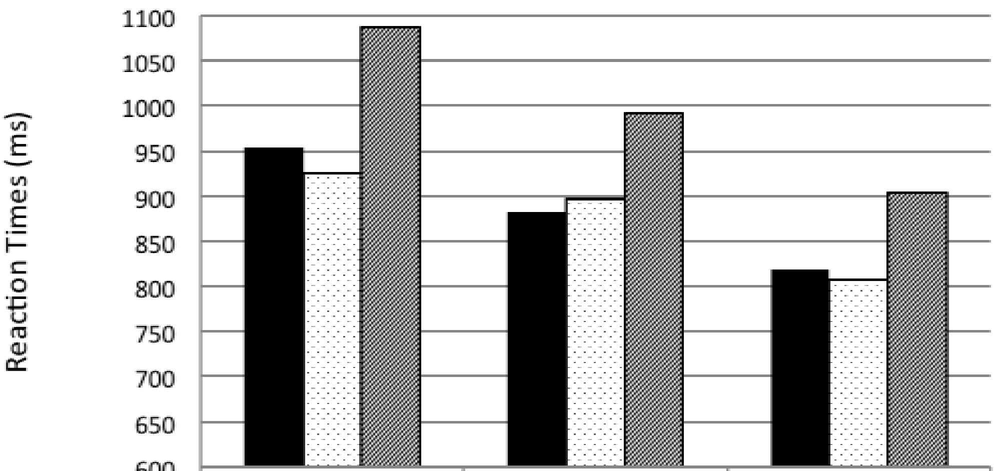
The next set of analyses examined language group difference within each of the SES groups. The only significant effect was found in the low-SES group where there was a LG difference Time 3, F (1, 34) = 12.40, p = .001; η^2 = .27. Here, bilinguals had significantly faster timing scores than monolinguals. The LG effect found at Time 3 was influenced heavily by the LSB group who showed a significantly reduced difference between their Same and Opposite Worlds raw timing scores.

The data file was then split by LG to investigate SES differences within language groups. No main effect of SES emerged within the bilingual group. However, a significant main effect of SES was present in the monolingual group at Time 1, F (1, 48) = 8.08, p < .01; η^2 = .15 and Time 3, F (1, 48) = 5.72, p = .02; η^2 = .12. The MSM group had significantly faster timing scores at Time 1 and Time 3 than the LSM group.

5.3.2.2. Colour-Word Stroop Task

The colour-word Stroop task has five task conditions or outcomes variables (all automated verbal reaction time responses, measured in milliseconds): neutral, congruent, incongruent, facilitation and inhibition (see Section 4.5.4.). Group differences were examined for each of the five conditions using growth curve analysis and t-tests compared overall condition times.

Figure 20 presents the total mean reaction times (milliseconds) for each condition across time. The graph shows that reaction times in the incongruent condition were slower than in the congruent and neutral conditions. Congruent trials were performed more quickly than the neutral trials at Time 1 and Time 3 but not Time 2.



	Time 1	Time 2	Time 3
Neutral	953.91	882.3	818.18
Congruent	925.55	897.26	806.75
Incongruent	1087.81	992.55	903.76

Figure 20 *Mean reaction times (ms) for Stroop task conditions across time*

Paired samples t-tests were used to compare trial conditions at each time point. No significant difference was found between congruent and neutral trials apart from at Time 1, t (107) = 2.24, p = .03, d = .04, where congruent trials were performed more quickly than neutral trials. T-tests for congruent and incongruent RTs showed a significant difference between these conditions at Time 1, t (107) = -11.68, p < .01, d = .56, Time 2, t (102) = -8.63, p < .01, d = .42, and Time 3, t (98) = -11.18, p < .01, d

= .56. At all three Time points, RTs were significantly slower in the incongruent condition.

Table 13 presents the total mean accuracy scores, represented as the percentage correct for each condition across time. It seems that all conditions had reached ceiling as mean percentage scores correct did not fall below 95% on any condition. However, paired t-tests revealed a significant improvement in incongruent accuracy scores from Time 1 (*M* = 96%, *SD* = 4) to Time 2 (*M* = 98%, *SD* = 2), t (101) = -5.06, $p < .001$ and from Time 1 to Time 3 (M = 98%, SD = 3), t (98) = -5.51, $p < .001$.

Table 13 *Total mean percentage trials correct for Colour-word Stroop conditions across time*

Assessment Period	Time 1	Time 2	Time 3
Neutral	99%	99%	99%
Congruent	99%	99%	99%
Incongruent	96%	98%	98%

5.3.2.2.1. Neutral condition

The significant values of the initial status (β = 952.62, *SE* = 14.48, $p < 0.01$) and linear slope (β = -77.58, *SE* = 8.1, $p < 0.01$) indicated a significant change in the initial status and Level-1 linear trajectories. The negative value in the linear slope indicated that Neutral RTs decreased across Time by an average of 78ms (see Table 14 for means).

Table 14 *Means (and standard deviations) for Colour-Word Stroop Neutral condition*

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual *n* = 32	984.84 (174.01)	913.82 (177.03)	826.19 (119.44)
	Bilingual (*n* = 42)	950.13 (168.29)	867.86 (127.46)	785.30 (111.89)
	Total (*n* = 74)	965.14 (170.49)	887.93 (151.69)	803.99 (116.38)
Low-SES	Monolingual (*n* = 17)	941.97 (131.19)	882.56 (108.93)	852.17 (94.63)
	Bilingual (*n* = 18)	958.33 (199.56)	851.78 (127.01)	851.84 (121.79)
	Total (*n* = 35)	950.39 (167.57)	867.64 (117.21)	852.00
Total	Monolingual (*n* = 49)	969.97 (160.37)	902.75 (155.79)	834.48 (111.75)
	Bilingual (*n* = 60)	952.59 (176.55)	863.27 (126.39)	805.02 (117.81)
	Total (*n* = 109)	960.40 (168.92)	881.49 (141.41)	818.73 (115.41)

The level-2 model found that Language Group was not a significant predictor of Neutral RTs initial status (β = - .75, *SE* = 17.17, *p* = .97) or of linear change (β = - 3.53, *SE* = 9.37, *p* = .71; see Table 15 for coefficients). Fig. 21 indicates that monolinguals and bilinguals improved their Neutral RTs at a similar rate across time.

Table 15 *Stroop neutral RT Level-1 and Level-2 model coefficients*

		Coe	SE	t-ratio
Level – 1 Linear	Intercept	952.62	14.48	63.16
	Slope	-77.58	8.10	-9.08
Level – 2	Intercept	958	16.36	58.55
Predictors	Slope	-72.28	9.37	-7.72
	Language Group	- .75	17.17	0.23
	Socioeconomic Status Group	16.01	16.10	0.22
	Time X Language Group	-3.53	9.37	-0.55

Linear X SES Group	-23.86	8.97	-1,84

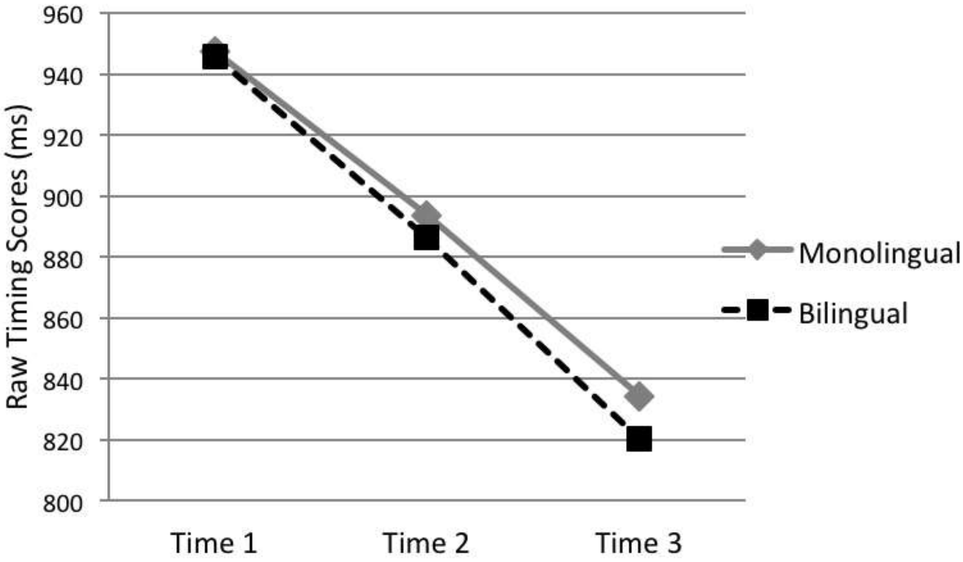


Figure 21 *Neutral RT fitted linear trajectory with predictor language group*

SES was not a significant predictor of Neutral RTs initial status (β = 16.01, *SE* = 16.1, *p* = .32). However, SES was a significant predictor of the linear decrease in Neutral RTs (β = -23.86, *SE* = 8.97, *p* = .01). Fig. 22 and Table 14 show that the RTs of the mid-SES participants improved (decreased) at a faster rate than the low-SES participants despite the fact that at Time 1 the low-SES group were faster than the mid-SES group.

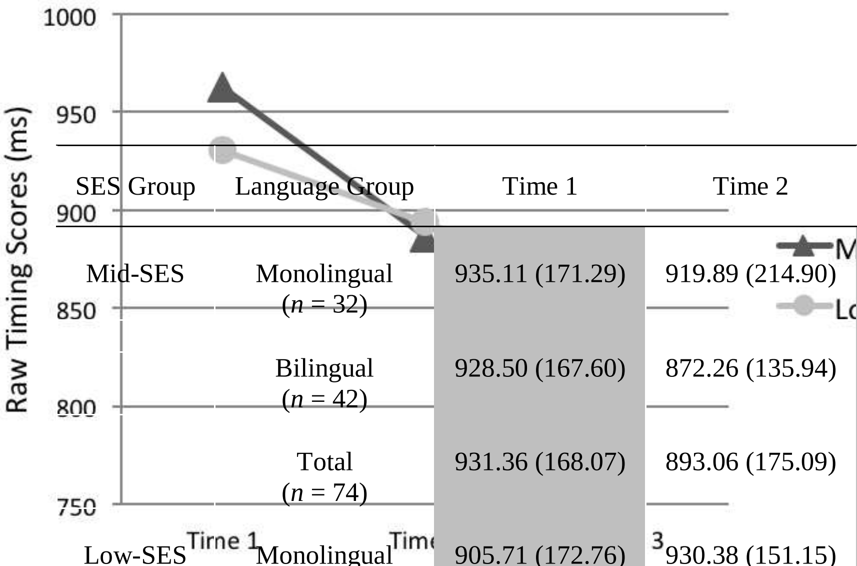


Figure 22 *Neutral RT linear growth model with predictor SES*

5.3.2.2.2. Congruent condition

Table 16 *Means (and standard deviations) for Stroop Congruent RTs (ms)*

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (*n* = 32)	935.11 (171.29)	919.89 (214.90)	824.14 (125.92)
	Bilingual (*n* = 42)	928.50 (167.60)	872.26 (135.94)	764.21 (111.14)
	Total (*n* = 74)	931.36 (168.07)	893.06 (175.09)	791.61 (121.05)
Low-SES	Monolingual (*n* = 17)	905.71 (172.76)	930.38 (151.15)	857.63 (128.47)
	Bilingual (*n* = 18)	926.67 (198.96)	879.31 (163.36)	825.38 (130.96)

	Total (*n* = 35)	916.49 (184.25)	905.62 (156.85)	840.98 (128.63)
Total	Monolingual (*n* = 49)	924.91 (170.58)	923.60 (193.09)	834.83 (126.33)
	Bilingual (*n* = 60)	927.95 (175.86)	874.28 (142.80)	782.33 (119.47)
	Total (*n* = 109)	926.58 (172.71)	897.04 (168.85)	806.76 (124.89)

The significant values of the initial status (β = 935.24, *SE* = 15.83, *p* < 0.01) and linear slope (β = -65.43, *SE* = 9.08, *p* < 0.01) indicated a significant rate of change in between-subjects variation and linear trajectories of individuals (see Table 17 for coefficients). The negative value in the linear slope indicated that Congruent RTs decreased across time by an average of 65ms.

Language Group was not a significant predictor of the initial status (β = 8.9, *SE* = 18.43, *p* = .63) or of linear change (β = -12.82, *SE* = 10.71, *p* = .23). Figure 23 and Table 16 indicate that Congruent RTs changed at a similar rate for monolinguals and bilinguals, although bilinguals slower Congruent RTs at Time 1 improved by Time 3.

	Coe	SE	t-ratio

Table 17 *Stroop congruent RT Level-1 and Level-2 model coefficients*

Level – 1 Linear	Intercept	935.24	15.83	59.28
	Slope	-65.43	9.08	-7.42
Level - 2	Intercept	930.14	17.26	53.88
Predictors	Slope	11.70	31.41	.37
	Language Group	8.90	18.43	0.23
	Socioeconomic Status Group	3.54	17.28	0.22
	Time X Language Group	-12.82	10.71	-0.55
	Linear X SES Group	-20.09	10.22	-1.84

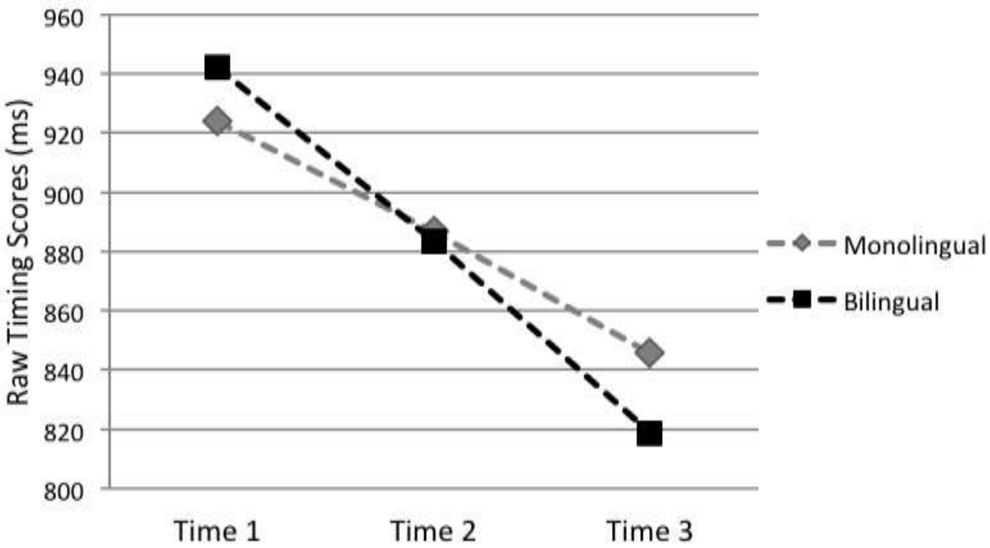


Figure 23 *Fitted trajectory for congruent RTs with predictor language group*

Although SES did not predict the initial status (β = 3.54, *SE* = 17.28, *p* = .84), it was a significant predictor of the linear decrease in Congruent RTs (β = -20.09, *SE* = 10.22, *p* = .05). Figure 24 and Table 17 indicated that RTs decreased at a faster rate in the mid-SES group than in the low-SES group.

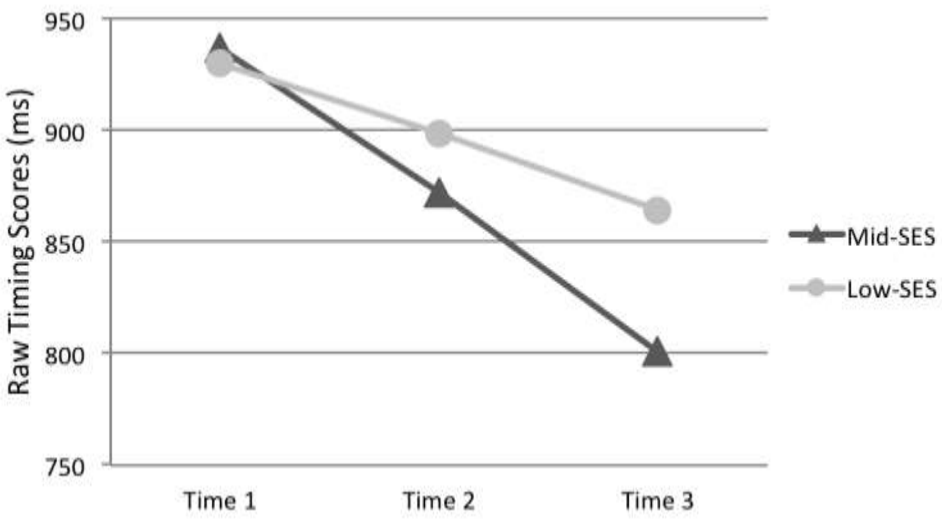


Figure 24 *Fitted trajectory for Stroop neutral RTs with predictor SES*

5.4.2.2.3. Incongruent condition

Table 18 *Means (and standard deviations) for Stroop Incongruent RTs (ms)*

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (*n* = 32)	1061.20 (163.78)	1007.19 (185.35)	901.08 (148.54)
	Bilingual (*n* = 42)	1104.29 (209.22)	980.68 (159.22)	876.42 (128.53)
	Total (*n* = 74)	1085.66 (190.89)	992.25 (170.49)	887.69 (137.57)
Low-SES	Monolingual (*n* = 17)	1070.74 (173.48)	1017.24 (184.35)	934.47 (145.27)
	Bilingual (*n* = 18)	1129.11(195.49)	972.03 (124.17)	942.09 (105.49)
	Total (*n* = 35)	1100.76 (184.79)	995.32 (157.31)	938.40 (124.21)
Total	Monolingual (*n* = 49)	1064.51 (165.46)	1010.75 (183.27)	911.73 (146.76)
	Bilingual (*n* = 60)	1111.73 (203.87)	978.21 (148.98)	895.88 (124.90)
	Total (*n* = 109)	1090.50 (188.23)	993. 23 (165.67)	903.26 (135.05)

In the incongruent condition of the Stroop task, the values of the initial status (β = 1086.91, *SE* = 17.55, $p < 0.01$) and linear slope (β = -103.35, *SE* = 8.67, $p < 0.01$) indicated a significant change in between-subjects variation (see Table 19 for coefficients). The negative value in the linear slope indicated that participants Incongruent RTs decreased across time by an average of 103ms.

The level-2 model found that Language Group was not a significant predictor of initial status (β = 22.93, *SE* = 21.32, $p = .29$) or of linear change (β = -12.18, *SE* = 10.29, $p = .24$). Monolinguals and bilinguals improved their RTs at a similar rate (Figure 25 and Table 18) although bilinguals were slower at Time 1 but by Time 3 had caught up with monolinguals.

Table 19 *Stroop incongruent RT Level-1 and Level-2 model coefficients*

		Coe	SE	t-ratio
Level – 1 Linear	Intercept	1086.91	17.55	62.10
	Slope	-103.35	8.67	-12.26
Level - 2	Intercept	1093.68	19.09	-10.87
	Slope	-100.37	9.23	-10.87
	Language Group	22.93	21.32	1.18
	Socioeconomic Status Group	-3.89	19.99	-.58
	Time X Language Group	-12.18	10.29	-1.34
	Linear X SES Group	-14.19	9.84	-1.14

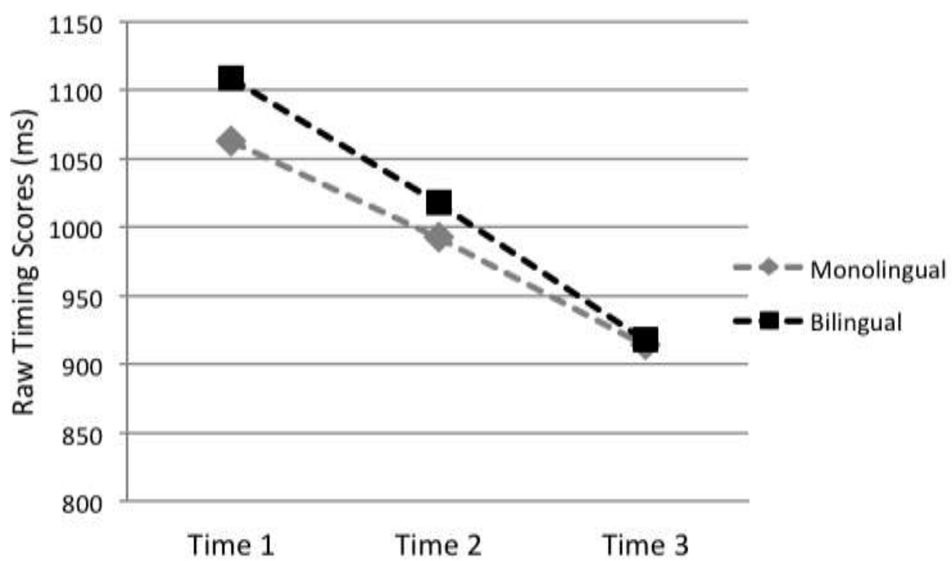


Figure 25 *Fitted trajectory for incongruent RT with predictor language group*

SES was not a significant predictor of the initial status (β = -3.89, *SE* = 19.99, *p* = .85) or of the linear decrease in Incongruent RTs (β = -14.19, *SE* = 9.84, *p* = .15). Figure 26 and Table 18 show that the mid-SES group decreased their RTs at a similar rate to low-SES participants and that both groups began at much the same point at Time 1.

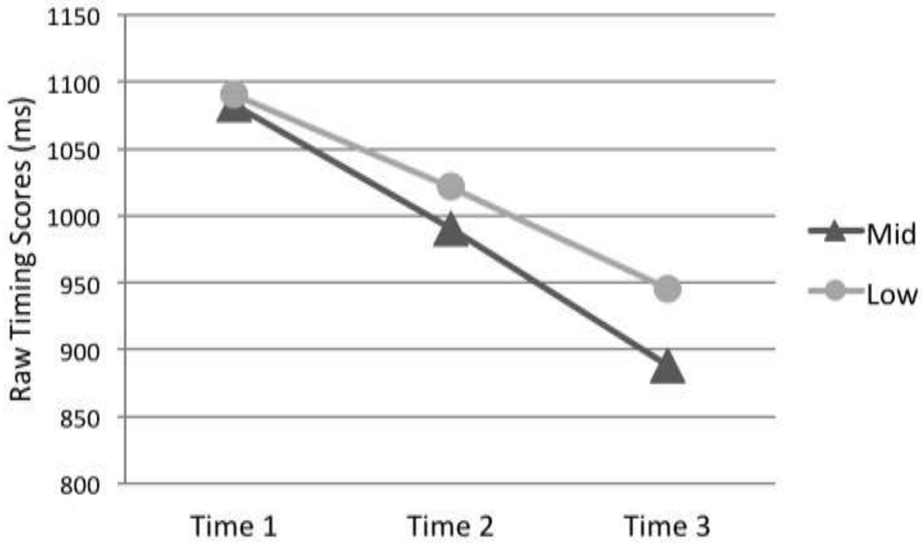


Figure 26 *Fitted trajectory for incongruent RT with predictor SES*

5.3.2.2.4. Facilitation

Results from the unconditional (Level-1) mean model (Model 1) indicated an ICC score of .10, suggesting that only 10% of the total variation in the Facilitation timing scores was due to inter-individual differences (see section 5.2.3.). As a result, ANOVA was considered a more appropriate analysis for this variable. See Table 20 for mean Facilitation RTs.

A mixed ANOVA showed no significant interactions or main effects between Language and SES groups apart from Time, F (2, 184) = 8.10, $p < .01$, $\eta_p^2 = .08$. Facilitation timing scores decreased from Time 1 to Time 2 but increased from Time 2 to Time 3. At Time 2 there was no facilitation effect as the congruent trials were performed more slowly than the neutral trials, leading to a negative mean Facilitation timing score.

Table 20 *Means (and standard deviations) for Stroop Facilitation timing scores (ms)*

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (n = 31)	43.37 (106.76)	-6.06 (96.03)	2.56 (57.04)
	Bilingual (n = 34)	33.18 (136.90)	1.37 (82.93)	16.21 (64.15)
	Total (n = 65)	38.22 (122.62)	-2.18 (88.78)	9.70 (60.78)
Low-SES	Monolingual (n = 15)	37.20 (166.92)	-51.83 (104.72)	-5.47 (57.62)
	Bilingual (n = 16)	53.03 (198.96)	-27.53 (77.19)	26.47 (61.50)
	Total (n = 31)	43.37 (181.27)	-39.29 (90.83)	11.02 (60.85)
Total	Monolingual (n = 46)	41.61 (127.58)	-20.99 (100.15)	-.05 (56.71)
	Bilingual (n = 50)	39.53 (157.57)	-7.88 (81.49)	19.49 (62.87)
	Total (n = 96)	40.53 (143.24)	-14.16 (90.66)	10.13 (60.48)

5.3.2.2.5. Inhibition

As in the Facilitation timing scores, results from the unconditional (Level-1) mean model (Model 1) indicated an ICC score of .21, suggesting that only 21% of the total variation in the Inhibition timing scores was due to inter-individual differences therefore ANOVA was used to analyse Inhibition RT scores (see section 5.2.3.). See Table 21 for inhibition RT means.

Table 21 *Means (and standard deviations) for Stroop Inhibition RT scores (ms)*

SES Group	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (*n* = 31)	82.3 (113.15)	93.37 (93.83)	76.02 (83.04)
	Bilingual (*n* = 34)	145.93 (140.40)	114.23 (116.84)	89.62 (83.57)
	Total (*n* = 65)	115.60 (131.11)	104.32 (91.61)	82.98 (84.74)
Low-SES	Monolingual (*n* = 15)	140.70 (141.45)	147.23 (116.84)	82.30 (89.31)
	Bilingual (*n* = 16)	146.84 (179.27)	120.25 (107.59)	90.25 (78.46)
	Total (*n* = 31)	143.87 (159.42)	133.31 (111.12)	86.40 (82.56)
Total	Monolingual (*n* = 46)	101.37 (124.60)	110.93 (103.77)	78.07 (84.19)
	Bilingual (*n* = 50)	146.22 (152.03)	116.20 (94.76)	89.62 (83.57)
	Total (*n* = 96)	124.73 (140.65)	113.68 (98.69)	84.08 (83.62)

A mixed ANOVA revealed no significant interactions or main effects apart from a significant main effect of Time, F (2, 184) = 4.59, p = .01, η_p^2 = .05. Inhibition timing scores decreased from Time 1 to Time 2 and from Time 2 to Time 3.
To check for any within groups differences, the file was split according to SES. No LG differences were found in the low-SES group. There was a significant main effect of language group in the mid-SES group at Time 1, F (1, 64) = 3.99, p = .05,

η^2 = .06, with monolinguals having a significantly lower Inhibition timing score compared with bilinguals. In other words, at Time 1, monolinguals switched more easily between congruent and incongruent trials than bilinguals. However, by Time 3 this advantage had gone and monolinguals and bilinguals performed almost equivalently, although monolinguals still had lower mean scores (see Table 21). When the file was split according to language group, no significant main effects of SES were observed in the monolingual or bilingual groups. Finally, one-way ANOVAs examining Inhibition timing scores between groups (MSM, MSB, LSM, LSB) revealed no significant main effect of group across time.

5.3.2.3. Inhibitory control task summary

In order to try and evaluate group effects in inhibitory control (IC) two common EF tasks were used: Opposite Worlds and Colour-Word Stroop. Each of these tasks consisted of a number of conditions, measuring performance on congruent and incongruent trials as well as examining the extra attentional capacity (time) needed to perform the incongruent over congruent trials.

All of the task variables were measured using children's timed responses and results indicated that performance on each task condition improved over time, therefore children's IC skills improved between ages 9 to 11 years.
Results from the standardised Opposite Worlds revealed that at Time 1, participants performed slightly below average, between the 40^{th} and 43^{rd} percentile bands. However, performance significantly improved across time and by Time 3 participants performed above average, between the 57^{th} and 69^{th} percentile. Although no standardised scores are available, participants also improved on the Colour-Word Stroop as RTs on all five task conditions decreased from Time 1 to Time 3.

Across these tasks of IC, no consistent differences emerged between LGs although results indicated that bilingual children's performance may improve relative to monolinguals performance. For instance, on the congruent condition of the Opposite Worlds task, the MSB group did not perform better than the MSM group until Time 3, where they showed significantly higher standardised scores. Likewise, in the

Opposite Worlds Difference condition, bilinguals did not outperform monolinguals at Time 1 or Time 2 but by Time 3 they had significantly lower attentional delays compared with monolinguals, resulting in an overall LG difference in favour of bilinguals. Although no LG effects were present in any condition of the Colour-Word Stroop task, the Inhibition RT scores revealed that MSBs had higher attentional delays at Time 1 compared with MSMs but by Time 3 had caught up with monolinguals and performed almost equivalently.

In terms of SES, it appeared that no SES group differences were evident for the IC tasks. SES was a significant predictor of general response time (RTs) performance on the Colour-Word Stroop task (neutral and congruent conditions) with Low-SES groups performing more slowly than Mid-SES groups. However, when additional attentional demands were included in this task (incongruent, IC condition) no effect of SES emerged. The LG effect found at Time 3 for the Worlds Difference condition was more prevalent in the Low-SES group where bilinguals outperformed monolinguals and at Time 1 of the Opposite Worlds condition, LSBs outperformed LSMs. These results indicate that LG effects may be more salient in the Low-SES group.

5.3.3. Switching Task

5.3.3.1. The Creature Count task

The standardised Creature Count produced two scores for analysis: accuracy and timing. Two 3 x 2 x 2 mixed ANOVAs were performed to examine any changes in scores over time. There were two between-groups variables, each with two levels: Language Group (LG; Monolingual and Bilingual) and Socioeconomic Status (Mid- and Low-SES). The within-subjects variable was Time which had three levels (Time 1, Time 2 and Time 3). Descriptive statistics for the standardised Creature Count timing and accuracy scores are presented in Table 22.

5.3.3.1.1. Standardised Creature Count Timing

Standardised Timing scores were only calculated if participants had performed a minimum of three trials correctly. As in the Opposite Worlds, raw scores were converted to age-scaled scores which convert to percentile band scores (see Table 3). A total of 76 participants (31 monolingual and 45 bilingual) made up the mid-SES group and 35 participants (17 monolingual and 18 bilingual) made up the low-SES group. Mauchly's test indicated that the assumption of sphericity had not been violated so Sphericity Assumed was employed.

There was a significant main effect of time, F (2, 214) = 59.02, $p < .01$, $\eta_p^2 = .36$, with both mid- and low-SES groups displaying an improvement in scores from Time 1 to Time 3 (see Table 22). Pairwise comparisons revealed a significant difference between Time 1 and Time 2 scores ($p < .01$), Time 1 and Time 3 scores ($p = .01$) and Time 2 and Time 3 scores ($p <.01$). No other significant interactions or main effects were found. Age-scaled timing scores increased from Time 1 to Time 2 and from Time 2 to Time 3.

Table 22 *Means (and standard deviations) for Creature Count scores across time*

Standardised Creature Count	Socioeconomic Status	Language Group	Time 1	Time 2	Time 3
Timing	Mid-SES	Monolingual (*n* = 31)	8.77 (2.23)	10.45 (2.06)	10.97 (1.91)
		Bilingual (*n* = 45)	8.93 (2.79)	10.64 (2.18)	11.20 (1.82)
		Total (*n* =76)	8.87 (2.56)	10.57 (2.12)	11.11 (1.84)
	Low-SES	Monolingual (*n* = 17)	8.00 (2.15)	9.53 (1.55)	11.18 (2.16)
		Bilingual (*n*= 18)	9.28 (2.78)	10.11 (1.94)	10.56 (2.04)
		Total (*n*= 35)	8.66 (2.54)	9.83 (1.76)	10.86 (2.09)
	Total	Monolingual (*n* = 48)	8.50 (2.21)	10.12 (1.93)	11.04 (1.98)
		Bilingual (*n* = 63)	9.03 (2.77)	10.49 (2.11)	11.02 (1.89)
		Total (*n* = 111)	8.80 (2.55)	10.33 (2.03)	11.03 (1.92)
Accuracy	Mid-SES	Monolingual (*n* = 45)	9.94 (2.54)	11.23 (2.53)	9.90 (2.18)
		Bilingual (*n* = 31)	10.16 (2.53)	10.84 (2.40)	10.87 (2.09)
		Total (*n* = 76)	10.07 (2.52)	11.00 (2.44)	10.47 (2.16
	Low-SES	Monolingual (*n* = 17)	9.88 (3.18)	9.35 (2.94)	10.06 (2.46)
		Bilingual (*n* = 18)	8.44 (2.20)	10.22 (2.80)	9.44 (2.41)
		Total (*n* = 35)	9.14 (2.78)	9.80 (2.86)	9.74 (2.42)
	Total	Monolingual (*n* = 48)	9.92 (2.75)	10.56 (2.80)	9.96 (2.26)
		Bilingual (*n* = 63)	9.67 (2.55)	10.67 (2.51)	10.46 (2.26)
		Total (*n* = 111)	9.77 (2.63)	10.62 (2.63)	10.24 (2.26)

Individual Group Changes across Time

To examine how each group's scores changed across time, paired-sample t-tests compared scores between all three time points: Time 1 – Time 2 (pair one), Time 2 – Time 3 (pair two) and Time 1 – Time 3 (pair three; see Table 23).

For the ***MSM*** group score significantly increased between Time 1 to Time 2. A similar difference was observed between Time 1 to Time 3. No significant difference was present between scores at Time 2 and Time 3. The mean increase in scores between Time 1 and Time 3 was 2 age-scaled scores or two percentile bands (see Table 3). For the ***MSB*** group a significant difference was present between Time 1 and Time 2, and between Time 1 and Time 3. There was no significant difference between scores at Time 2 and Time 3. Like the MSM group, the mean increase from Time 1 to Time 3 was two percentile bands. There was a significant difference in the ***LSM*** group's scores between all three time points and the increase between Time 1 and Time 3 was had a mean increase of three percentile bands. The ***LSB*** group showed no significant difference in timing scores between Time 1 - Time 2, Time 2 - Time 3, and Time 1 - Time 3.

Table 23 *Paired-sample T-Tests results for the Creature Count Timing Scores across groups*

Group	Pair One Time 1 – Time 2	Pair Two Time 2 – Time 3	Pair Three Time 1 – Time 3
Mid-SES Monolingual			
t (30)	-6.17**	-1.97	-6.51**
Effect size	.56		.55
Mid-SES Bilingual			
t (45)	-5.23**	-1.96	-8.00**
Effect size	.38		.27
Low-SES Monolingual			
t (16)	-3.11*	-3.00*	-5.91**
Effect size	.36	.36	.69
Low-SES Bilingual			
t (17)	-1.45	-.94	-1.89
Effect size			

Note: *$p < .05$, **$p < .01$. Effect sizes are Cohen's d.

5.3.3.1.2. Creature Count Standardised Accuracy

A 3 x 2 x 2 mixed ANOVA was used to examine standardised age-scaled scores, calculated from participants' raw accuracy scores (see section 4.5.3.1.). A total of 111 participants were included for analysis. Of this sample, 76 participants were in the mid-SES group (31 monolingual and 45 bilingual). Thirty-five participants made up the low-SES group (17 monolingual and 18 bilingual). Descriptive statistics are presented in Table 22. Mauchly's test indicated that the assumption of sphericity had not been violated so Sphericity Assumed was employed.

There was a significant three-way interaction between time, SES and language group, F (2, 214) = 3.12, p = .05, η_p^2 = .03, see Figure 27. The MSB and LSM monolingual groups improved their scores from Time 1 to Time 3. However, the scores of the MSM group decreased from Time 2 to Time 3. Likewise the LSB group decreased their accuracy scores from Time 2 to Time 3. There was a significant main effect of SES group, F (1, 107) = 7.59, p = .01, η_p^2 = .07. Low-SES groups had significantly lower accuracy scores than mid-SES groups across time. There was also a significant main effect of time, F (2, 214) = 2.98, p = .05, η_p^2 = .03. Pairwise comparison revealed a significant difference between Time 1 and Time 2 accuracy scores (p = .02) but no significant difference between Time 1 and Time 3 (p = .19) or Time 2 and Time 3 scores (p = .28). Accuracy scores increased from Time 1 to Time 2 and decreased slightly from Time 2 to Time 3. No other interactions or main effects were obtained.

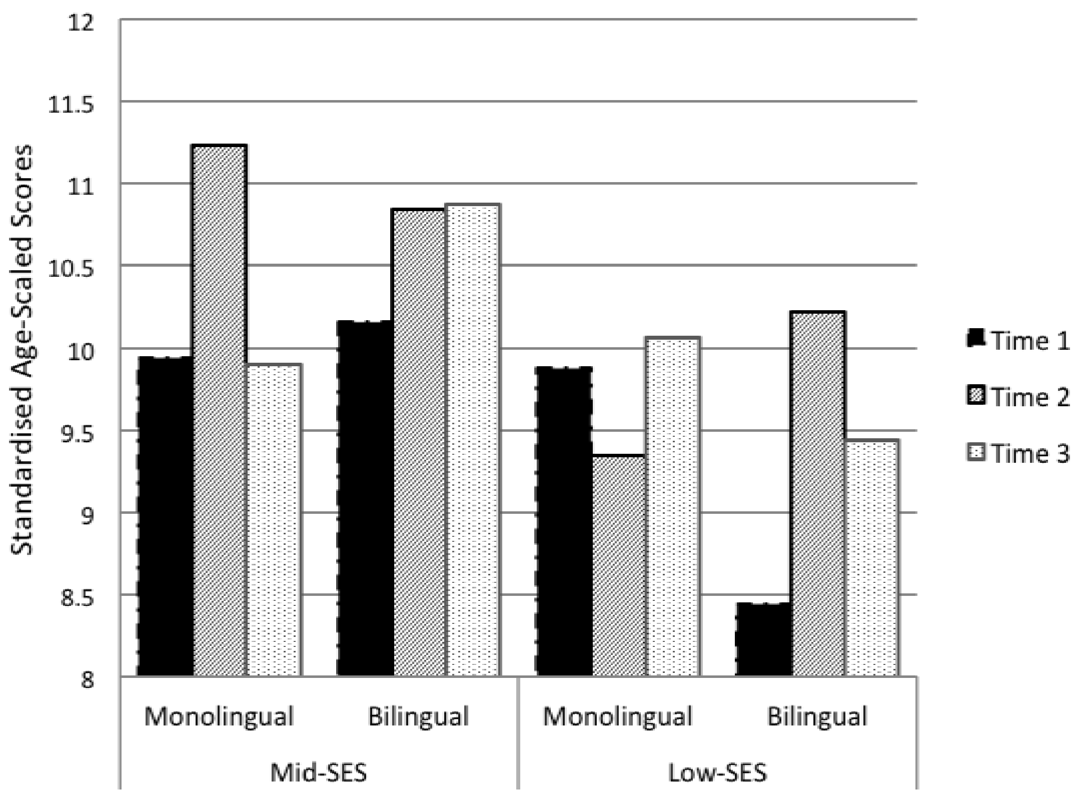


Figure 27 *Time x SES x Language group interaction for Creature Count accuracy scores*

SES Accuracy Scores across Time

One-way ANOVAs were conducted separately for each SES group and at each time point to investigate language group effects more closely.

Within the mid-SES group, there was no significant main effect of LG at Time 1, Time 2 or Time 3. While bilinguals had higher means at each time point than monolinguals, these differences were very small (see Fig. 27). When a second set of ANOVAS was conducted to include those cases excluded in the mixed-ANOVA analysis due to listwise deletion (Monolinguals N = 1; Bilinguals N = 3), the language group effect at Time 3 was significant, F (1, 75) = 3.87, p = .05, with bilinguals scoring higher than monolinguals.

There was no significant difference between low-SES language groups' accuracy scores at Time 1, Time 2 or Time 3. Monolinguals at Time 1 and Time 3 had higher mean scores compared to bilinguals at Time 1 and Time 3. However at Time 2, bilinguals scored higher than monolinguals (see Table 22).

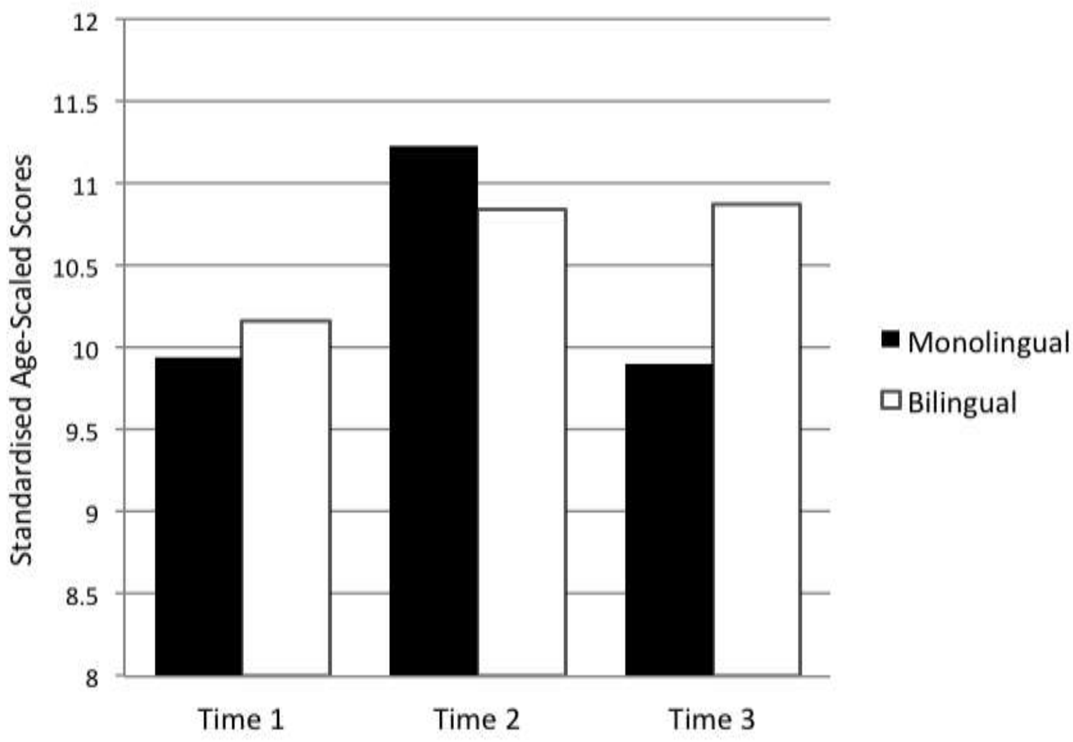


Figure 28 *Mid-SES Creature Count Accuracy scores for Language Groups across time*

Differences between SES and Language Groups

ANOVAs explored how each set of accuracy scores differed between groups at each time point. A significant main effect of group was found at Time 1, F (3, 116) = 2.87, p = .04, η^2 = .07. Post-hoc analysis using Bonferroni test indicated that the MSB group had significantly (p = .02) higher mean accuracy scores than the LSB group at Time 1. All other groups did not differ significantly from one another and no significant differences between groups were present at Time 2 or Time 3 (see Table 22 and Figure 27).

Individual Group Changes across Time

To examine how each group’s accuracy scores changed across time, paired-sample t-tests compared scores between all three time points: Time 1 – Time 2 (pair one), Time 2 – Time 3 (pair two) and Time 1 – Time 3 (pair three). See Table 24 for t-tests results and effect sizes within each group.

For the ***MSM*** group, accuracy scores significantly increased from Time 1 to Time 2. A similar difference was observed between Time 2 and Time 3. However, this difference indicated a decrease in accuracy scores by one age-scaled score or percentile band (see Table 3 for conversion). No significant difference was present between Time 1 and Time 3. For the ***MSB*** group no significant differences were observed for accuracy scores between Time 1 - Time 2, Time 2 - Time 3 or Time 1 - Time 3. Likewise in the ***LSM*** group, no significant differences were found for the accuracy scores between Time 1- Time2, Time 2 - Time 3 or Time 1 - Time 3. The ***LSB*** group showed a significant increase in their accuracy scores between Time 1 and Time 2. The effect size was large with a mean increase in accuracy scores of two percentile bands. There was no significant difference between scores at Time 2 and Time 3 or Time 1 and Time 3.

Table 24 *Paired-sample T-Tests results for the Creature Count Accuracy Scores across groups*

Group	Pair One Time 1 – Time 2	Pair Two Time 2 – Time 3	Pair Three Time 1 – Time 3
Mid-SES Monolingual			
t (30)	-2.33*	2.41*	.00
Effect size	.15	.16	
Mid-SES Bilingual			
t (45)	-1.41	-.06	-1.43
Effect size			
Low-SES Monolingual			
t (17)	.83	-.77*	-.16
Effect size			
Low-SES Bilingual			
t (17)	-2.76*	1.0	-1.19
Effect size	.31		

Note: **p* < .05, ***p* < .01. Effect sizes are Cohen’s d.

5.3.3.2. Switching task summary

The Creature Count (CC) task is said to measure aspects of EF and, in particular, requires participants to switch between their directions of count while maintaining in mind the number they had reached before the switch (Manly et al., 1999). The task assessed both accuracy and timing using age-scaled standardised scores, which can be converted to percentile bands (see Table 3).

As in tasks of IC, participants' performance significantly improved across time for CC timing but not accuracy. Timing improved from 40^{th}-43^{rd} percentile bands at Time 1 to the 57^{th}-69^{th} percentiles at Time 3. Accuracy scores increased from the 43^{rd}-57^{th} at Time 1 to the 57^{th}-69^{th} percentile by Time 2 but decreased back to the 43^{rd}-57^{th} percentile bands at Time 3. The MSM and LSB groups improved their accuracy scores from Time 1 to Time 2 but not from Time 2 to Time 3. These unusual accuracy scores may be as a result of the relatively few CC trials (7). Furthermore, the differences in accuracy scores were minimal across time as participants remained at approximately population average (43^{rd}-57^{th} percentile bands) across time.

No clear LG effects were present, although the MSB group did outperform the MSM group at Time 3 on the accuracy condition of the switching task. There was a significant main effect of SES for the accuracy condition only with mid-SES participants outperforming low-SES participants.

5.3.4. Working Memory Task

5.3.4.1 The Working Memory Test Battery for Children

The results from the Working Memory Test Battery for Children (WMTB-C) recorded children's raw and standardised scores on a number of phonological and visuo-spatial short-term memory (STM) and working memory (WM) tasks: forward digit recall, backward digit recall, forward block recall and backward block recall. However, as the backward block recall task is an unstandardized measure of visuo-spatial WM, only raw scores on this task were recorded (for details see section 4.5.5.). As the WMTB-C was only administered at Time 3 of the study, one-way ANOVAs were used to compare Language Groups, SES Groups and Groups overall (MSM, MSB, LSM, LSB). Descriptive statistics for WM measures are displayed in Table 25.

Language Group Comparison

No differences were observed between monolingual and bilingual language groups for any of the WM measures, both standardized and unstandardized.

Socioeconomic Group Comparison

No significant differences were observed between SES groups' standardised forward digit and forward blocks recall scores. However, significant differences were observed between low- and mid-SES groups for standardised backward digit recall scores, F (1, 89) = 6.31, p = .01, η_p^2 = .07, with the mid-SES groups scoring significantly higher than low-SES groups on this test of verbal WM (see Fig. 29).

Table 25 *Means (and standard deviations) for working memory measures*

Task	SES Group	Monolinguals	Bilinguals	Total
Forward Digit Recall	Mid-SES	100.28 (14.70) (*n* = 25)	98.44 (12.33) (*n* = 34)	99.22 (13.30) (*n* = 59)
	Low-SES	98.33 (12.36) (*n* = 15)	93.44 (7.03) (*n* = 18)	95.67 (9.96) (*n* = 33)
	Total	99.55 (13.74) (*n* = 40)	96.71 (10.98) (*n* = 52)	97.95 (12.27) (*n* = 92)
Backward Digit Recall	Mid-SES	101.56 (20.65) (*n* = 25)	99.97 (16.81) (*n* = 34)	100.64 (18.38) (*n* = 59)
	Low-SES	94.40 (15.66) (*n* = 15)	88.31 (10.62) (*n* = 16)	91.26 (13.38) (*n* = 31)
	Total	98.88 (19.05) (*n* = 40)	96.24 (15.94) (*n* = 50)	97.41 (17.34) (*n* = 90)
Forward Corsi Blocks	Mid-SES	100.20 (11.83) (*n* = 25)	98.85 (13.08) (*n* = 34)	99.42 (12.48) (*n* = 59)
	Low-SES	98.67 (10.58) (*n* = 15)	100.06 (12.04) (*n* = 17)	99.41 (11.22) (*n* = 32)
	Total	99.63 (11.27) (*n* = 40)	99.25 (12.64) (*n* = 51)	99.42 (11.99) (*n* = 91)
Backward Corsi Blocks Correct (raw scores)	Mid-SES	26.92 (3.79) (*n* = 25)	27.29 (3.69) (*n* = 34)	27.14 (3.70) (*n* = 59)
	Low-SES	25.00 (4.44) (*n* = 15)	25.59 (3.43) (*n* = 17)	25.31 (3.88) (*n* = 32)
	Total	26.20 (4.10) (*n*= 40)	26.73 (3.66) (*n* = 51)	26.49 (3.85) (*n* = 91)
Backward Corsi Blocks Span (raw scores)	Mid-SES	4.48 (.82) (*n* = 25)	4.53 (.83) (*n* = 34)	4.51 (.82) (*n* = 59)
	Low-SES	4.13 (.74) (*n* = 15)	4.29 (.69) (*n* = 17)	4.22 (.72) (*n* = 91)
	Total	4.35 (.80) (*n* = 40)	4.45 (.78) (*n* = 51)	4.41 (.79) (*n* = 91)

A significant difference was also observed between the unstandardized backward blocks task. Although there was no significant difference in the highest span reached by participants, there was a difference in the number of items recalled correctly, F (1, 89) = 4.86, p = .03, η_p^2 = .05. On average, the mid-SES groups performed more trials correctly than the Low-SES groups (see Table 25).

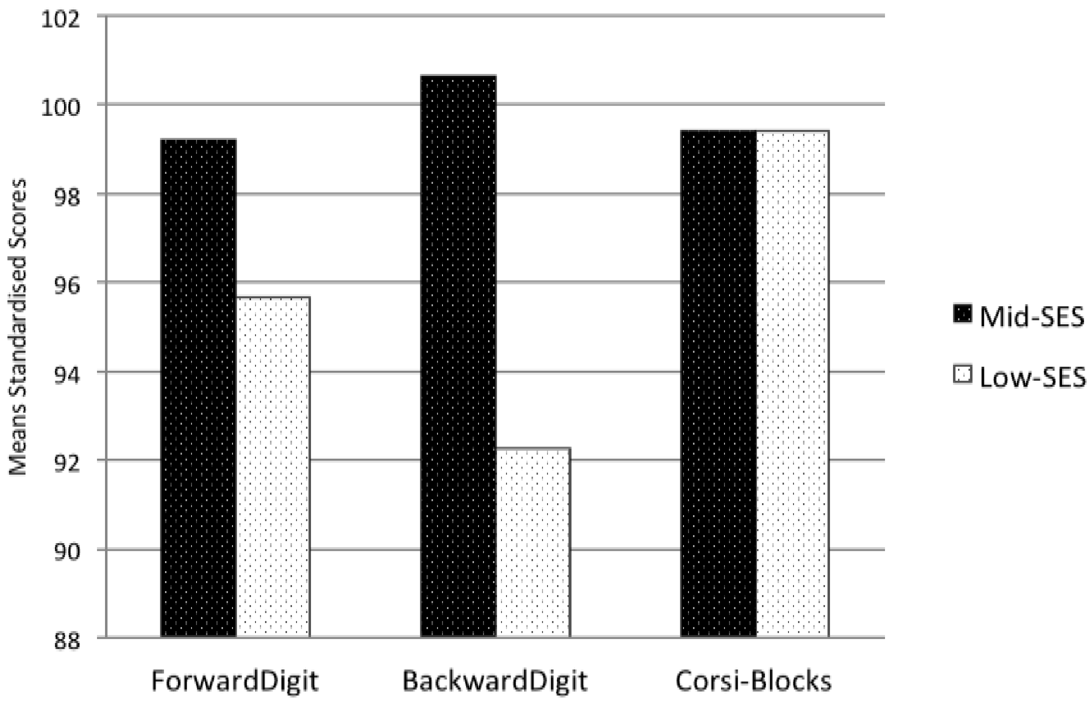


Figure 29 *Mean standardised working memory scores for SES groups*

Between-groups analysis

A between-groups ANOVA showed no significant difference between groups (MSM, MSB, LSM, LSB) on any of the WMTB-C variables. To examine whether each group performed equivalently in the visuo-spatial and verbal tasks, paired t-tests were run to compare each groups' (MSM, MSB, LSM, LSB) performance on the verbal versus visuospatial versions of the STM and WM tasks. Each group performed equivalently on the verbal and visuospatial STM tasks apart from the LSB group who had significantly higher scores on the visuospatial than the verbal STM task, t (16) = -2.17, p = .05 (see Fig. 30). All groups performed better on the visuospatial than the verbal WM tasks, t (24) = -6.97, p < .001 (see Fig. 31).

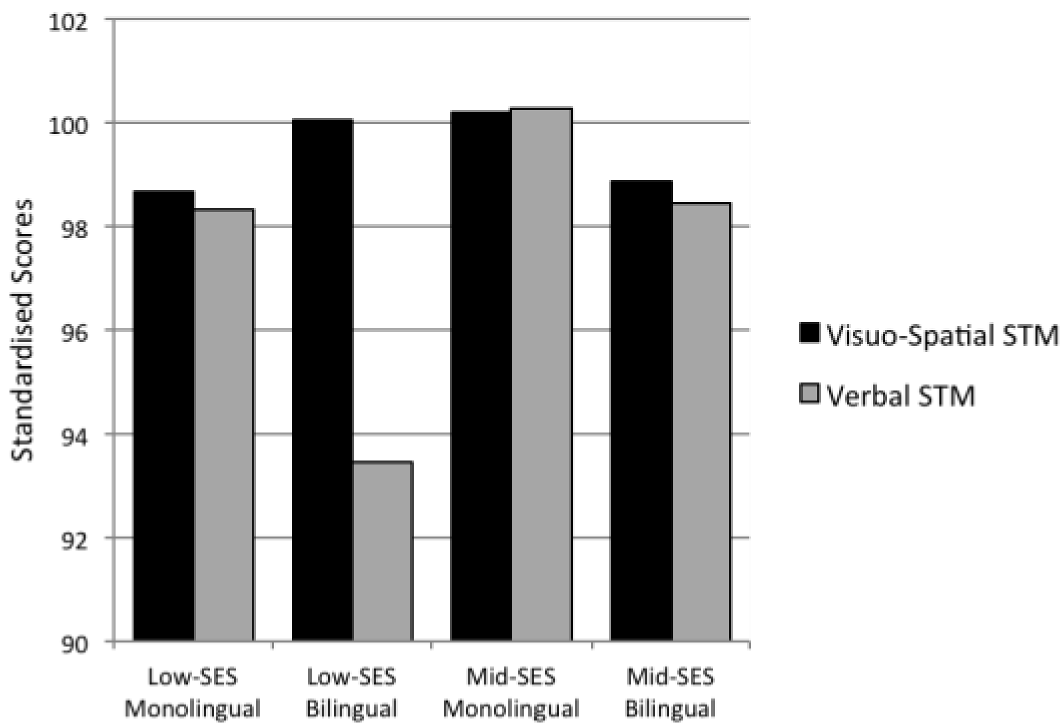


Figure 30 *Comparison of group performance on VS and verbal STM tasks*

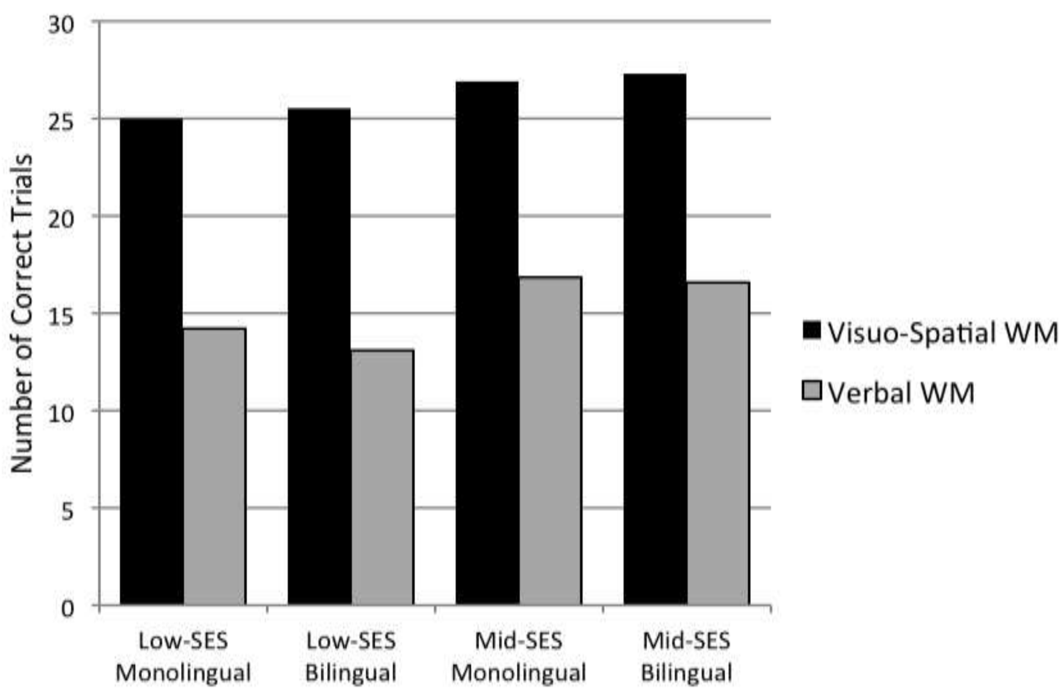


Figure 31 *Comparison of group performance on VS and verbal WM tasks*

5.3.4.2. Working Memory task summary

The four tasks assessed from the WMTB-C were: the forward digit (verbal STM), backward digit (verbal WM), forward blocks (visuo-spatial STM) and backward blocks recall (visuo-spatial WM). Performance on these tasks was assessed at Time 2 only; therefore conclusions regarding the developmental nature of WM could not be made although standardised scores showed that both LG groups were performing at around population average (M = 100; SD = 3).

There was no significant main effect of LG on any of the STM and WM tasks administered. However, there was a significant main effect of SES on standardised backward digit and unstandardized backward blocks recall. Results indicated that the mid-SES groups outperformed the low-SES groups on these tasks of verbal and visuo-spatial WM. Visuo-spatial WM tasks were performed better across all groups and results showed that the LSB group found the verbal STM task significantly more difficult than the visuo-spatial STM task.

5.3.5. Unified executive function tasks

5.3.5.1. Trail Making Task

The Trail Making Test (TMT) is a timed test (seconds) and has two conditions: Trails A or congruent condition, and Trails B or incongruent condition (for details see section 4.5.6.). As standardised scores were not available for either condition, growth curve modelling of response times was used to explore group effects. A third condition, *'Trails Difference'* was also computed by subtracting each participant's Trails A response time from their Trails B response time. This Trails Difference condition was used as a measure of the amount of additional attentional resources needed to complete the more difficult Trails B condition. Growth curve was not used to assess this variable as it did not fulfill the criteria for GC analysis (see section 6.3.3.). Each trial condition will be examined separately.

Figure 32 displays graphically, the times taken to complete each task condition across time. At each time point, the Trails B took significantly longer, $F(1, 107) = 327.78$, $p < .01$, $\eta_p^2 = .75$ than the Trails A trials.

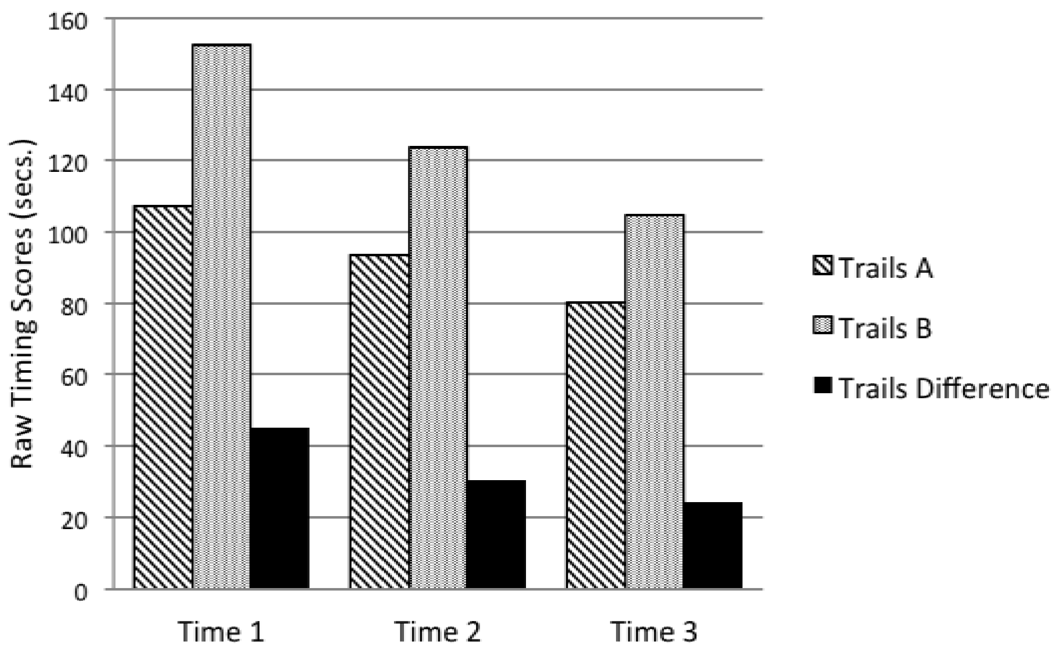


Figure 32 *Mean Trail Making Test response times across time*

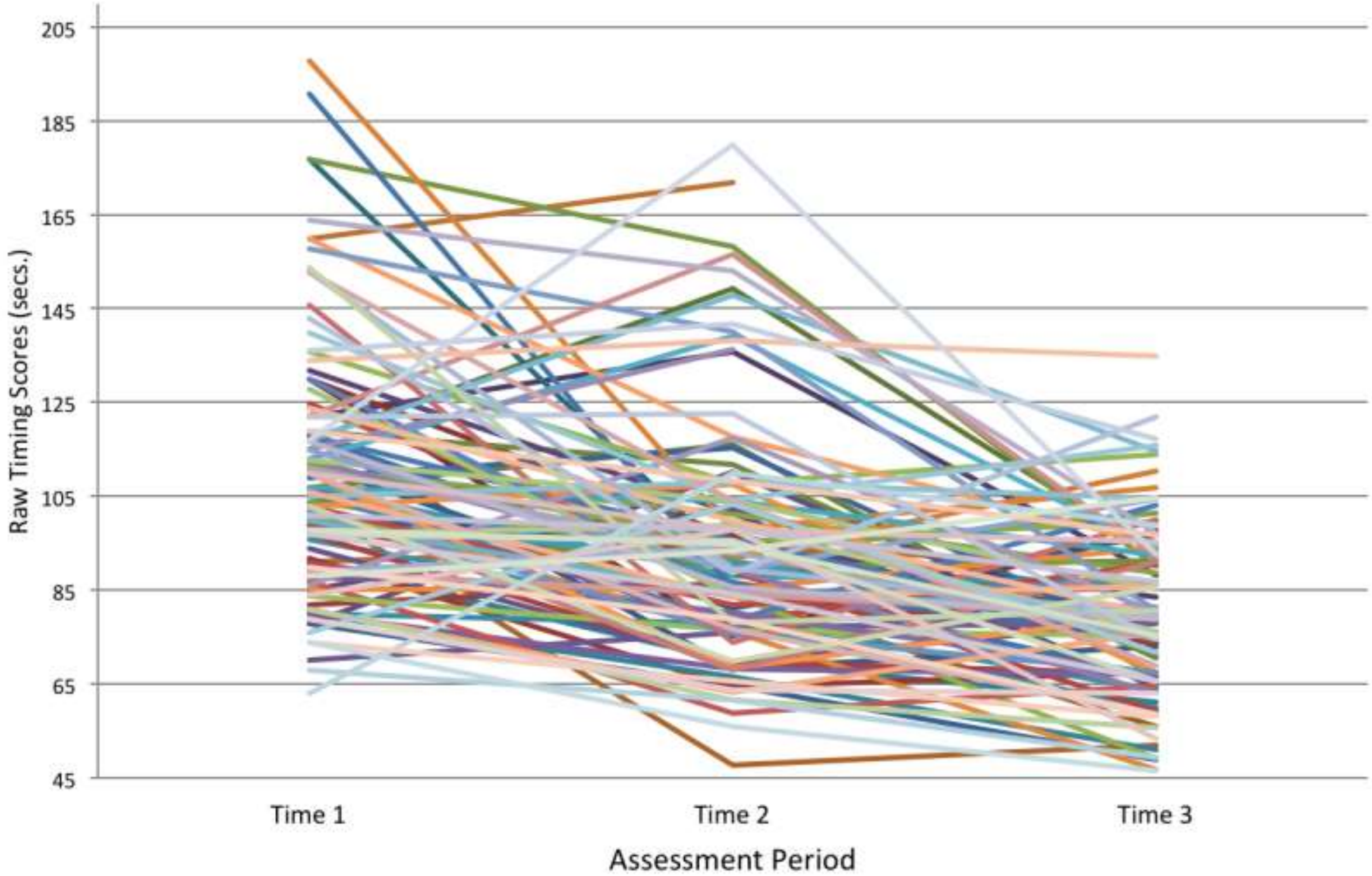


Figure 33 *Empirical growth plot for Trails A response times*

5.3.5.1.1. Trails A

Figure 33 displays the empirical growth plot for the Trails A condition of the TMT. On examining this plot it would be predicted that a linear change will emerge as the model most likely to fit the data. Descriptive statistics for the Trails A response times are presented in Table 26.

Table 26 *Means (and standard deviations) for Trails A scores (sec.) over time*

Socioeconomic Status	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual	102.78 (24.08) (*n* = 32)	93.15 (23.11) (*n* = 32)	80.79 (22.15) (*n* = 32)
	Bilingual	105.13 (23.18) (*n* = 47)	90.13 (22.33) (*n* = 45)	78.86 (16.29) (*n* = 42)
	Total	104.18 (23.42) (*n* = 79)	91.39 (22.55) (*n* = 77)	79.69 (18.93) (*n* = 74)
Low-SES	Monolingual	114.11 (22.49) (*n* = 18)	103.19 (27.71) (*n* = 18)	88.23 (13.14) (*n* = 16)
	Bilingual	113.21 (21.31) (*n* = 19)	92.41 (25.03) (*n* = 18)	75.35 (15.01) (*n* = 18)
	Total	113.65 (21.59) (*n* = 37)	97.80 (26.59) (*n* = 36)	81.41 (15.40) (*n* = 34)
Total	Monolingual	106.86 (23.93) (*n* = 50)	96.77 (25.06) (*n* = 50)	83.27 (19.78) (*n* = 48)
	Bilingual	107.45 (22.79) (*n* = 66)	90.79 (22.95) (*n* = 63)	77.80 (15.87) (*n* = 60)
	Total	107.20 (23.19) (*n* = 116)	93.43 (23.98) (*n* = 113)	80.23 (17.84) (*n* = 108)

The unconditional linear growth model (Level – 1) showed that there was a significant decrease in response times (β = -15.00, *SE* = 1.16, $p < 0.01$) and that the initial status was also significant (β = 106.91, *SE* = 2.03, $p < 0.01$). This indicated that the average participant had a linear fitted trajectory with an initial Trails A timing score of 106.91 seconds which decreased by an average of 15 seconds per time tested.

Table 27 *Trails A RT Level-1 and Level-2 model coefficients*

		Coe	SE	t-ratio
Level – 1 Linear	Intercept	106.91	2.03	62.10
	Slope	-15.0	1.16	-12.26
Level - 2	Intercept	100.58	2.39	46.31
	Slope	-17.23	1.44	-11.94
	Language Group	-1.20	2.46	1.18
	Socioeconomic Status Group	-4.22	2.25	-.58
	Time X Language Group	-1.12	1.51	-0.69
	Linear X SES Group	-14.19	9.84	-1.14

The Level - 2 growth model included predictor variables: Language (monolingual or bilingual) and SES (mid-SES, low-SES) groups (see Table 27 for coefficients). Language group was not a significant predictor of the initial status (β = -1.20, *SE* = 2.46, *p* =.63) or the linear change in Trails A response times (β = -1.12, *SE* = 1.63, *p* = .49). Figure 34 displays the fitted trajectory of monolingual and bilinguals raw times (secs.). The rate of change in Trails A response times appeared to decrease at a similar rate for LGs.

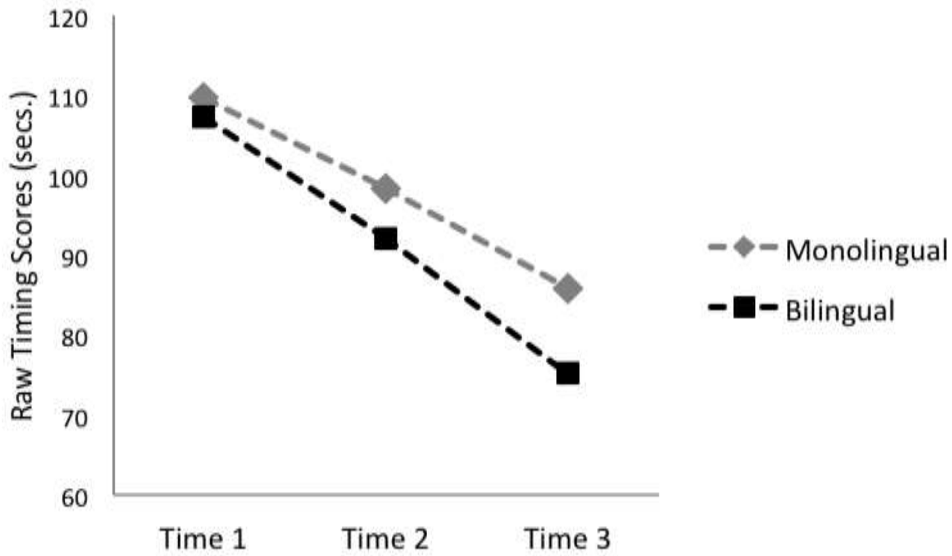


Figure 34 *Fitted trajectory for Trails A with added predictor language group*

Although SES was not a significant predictor of the initial status in Trails A response times (β = -4.22, *SE* = 2.25, *p* =.06) this figure was approaching significance. SES was, however a significant predictor of the linear decrease in timing scores (β = -2.38, *SE* = 1.41, *p* = .02). Low-SES participants were slower at Time 1 and mid-SES participants reduced their Trails A response times at a faster rate than low-SES participants (Table 27 and Figure 35).

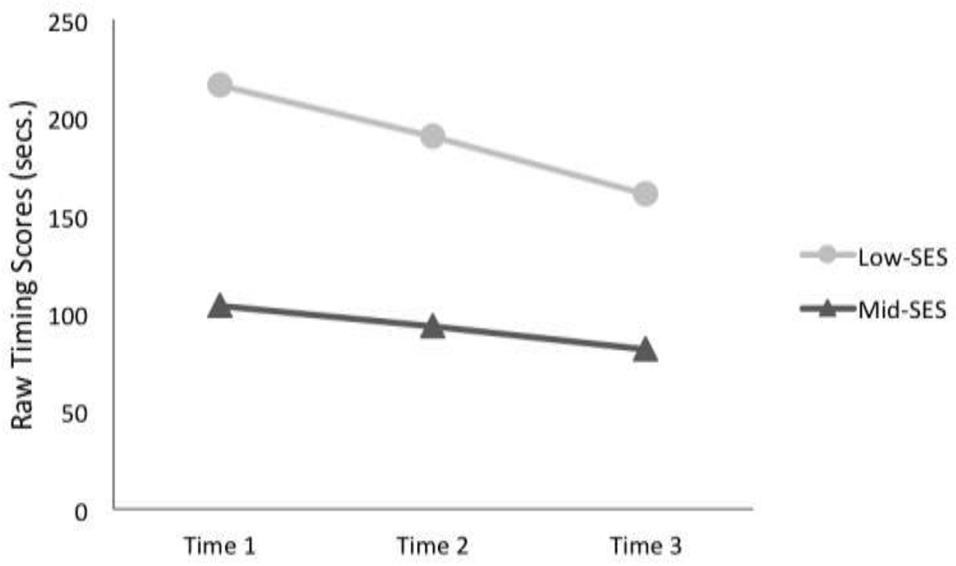


Figure 35 *Fitted trajectory for Trails A with added predictor SES*

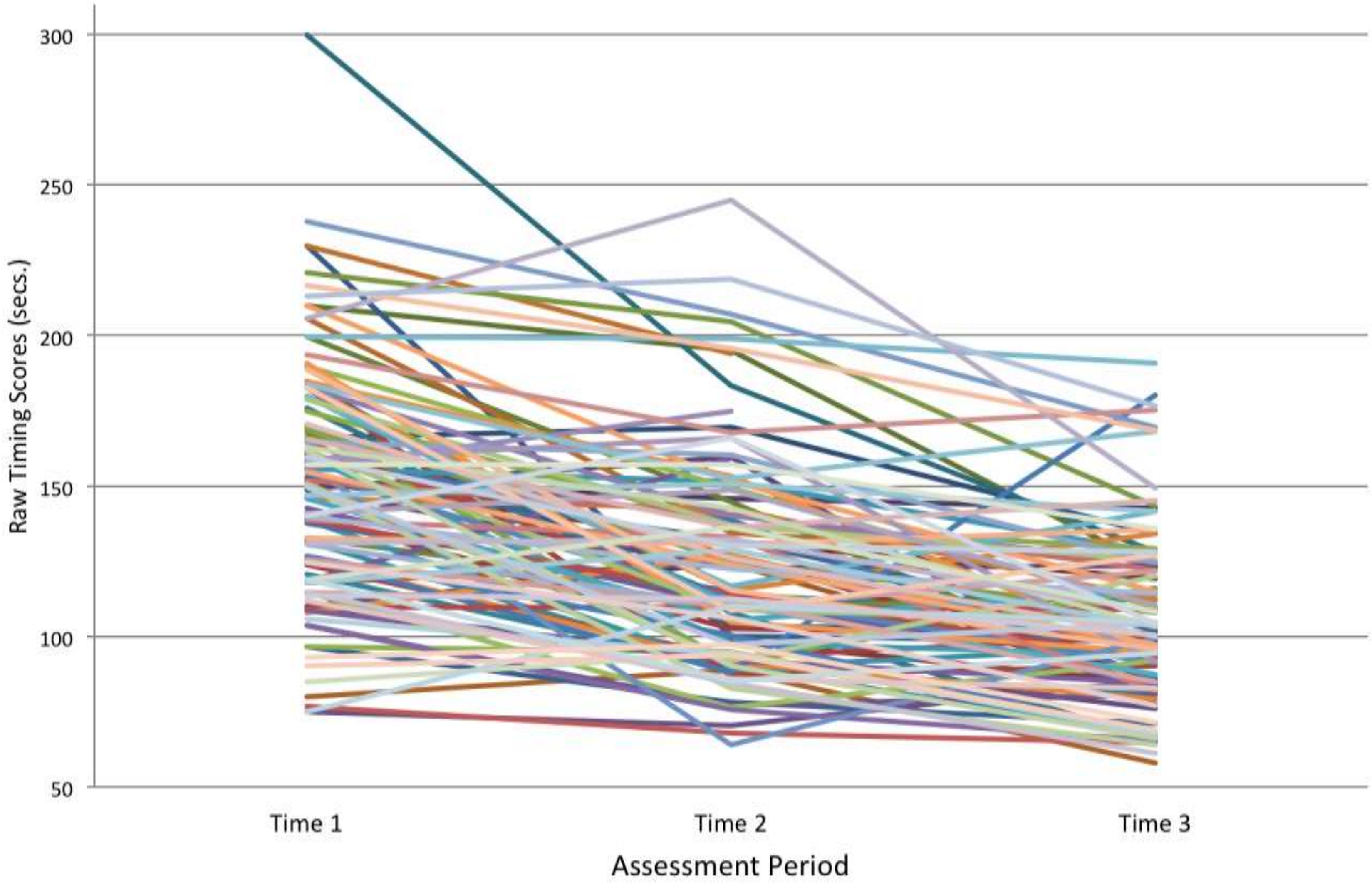


Figure 36 *Empirical growth plot for Trails B response times*

5.3.5.1.2. Trails B

Figure 36 displays the empirical growth plot for the Trails B condition of the TMT. Descriptive statists for the Trails B over time are displayed in Table 28.

Table 28 *Means (and standard deviations) for Trails B scores (seconds) across*

Socioeconomic Status	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual	140.56 (36.34) (*n* = 32)	125.69 (38.74) (*n* = 32)	105.09 (30.43) (*n* = 32)
	Bilingual	147.04 (34.85) (*n* = 47)	111.53 (30.45) (*n* = 45)	96.16 (18.95) (*n* = 42)
	Total	144.42 (35.41) (*n* = 79)	117.41 (34.62) (*n* = 77)	100.03 (24.79) (*n* = 74)
Low-SES	Monolingual	170.83 (28.03) (*n* = 18)	145.22 (30.68) (*n* = 18)	129.94 (31.47) (*n* = 16)
	Bilingual	167.74 (52.32) (*n* = 19)	128.76 (32.55) (*n* = 18)	100.51 (24.82) (*n* = 18)
	Total	169.24 (41.74) (*n* = 37)	136.99 (32.27) (*n* = 36)	114.36 (31.46) (*n* = 34)
Total	Monolingual	151.46 (36.43) (*n* = 50)	132.72 (36.95) (*n* = 50)	113.38 (32.66) (*n* = 48)
	Bilingual	153.00 (41.31) (*n* = 66)	116.46 (31.78) (*n* = 63)	97.47 (20.76) (*n* = 60)
	Total	152.34 (39.12) (*n* = 116)	123.65 (34.96) (*n* = 113)	104.54 (27.74) (*n* = 108)

Trails B displayed a significant quadratic rate of change as results indicated that all growth parameters were significant (intercept = $p < .01$, linear change = $p < .01$, quadratic change $p = .01$; see Table 29 for coefficients). The initial status was 152.34 ($\beta = 152.34$, $SE = 3.61$, $p < .01$) and the linear effect was negative ($\beta = -40.35$, $SE = 5.59$, $p < .01$) revealing that the rate of linear change decreased over time. The significant quadratic effect was positive ($\beta = 8.05$, $SE = 3.02$, $p = .01$) indicating the rate of quadratic growth increased over time. Compared to the linear change trajectory (-40 seconds), the rate of quadratic growth was small (8 seconds) which indicated that the rate of change decreased after Time 1 but slowed from Time 2 to

Time 3. Although the linear decrease slowed between Time 2 and Time 3, it does not necessarily indicate that times increased. The mean difference in timing scores was 30 seconds between Time 1 and Time 2 and 19 seconds between Time 2 and Time 3.

Table 29 *Trails B neutral RT Level-1 and Level-2 model coefficients*

		Coe	SE	t-ratio
Quadratic	Intercept	152.30	3.61	42.21
	Linear slope	-40.35	5.59	-7.22
	Quadratic slope	8.05	3.02	2.67
Level - 2	Intercept	156.34	3.76	41.56
	Linear Slope	-40.06	6.02	-6.66
	Quadratic Slope	7.43	3.27	2.28
	Language Group	-0.06	4.29	-2.98
	Socioeconomic Status Group	-11.67	3.91	-2.98
	Linear X Language Group	-15.43	6.88	-2.24
	Linear X SES Group	4.12	6.31	0.65
	Quadratic X Language Group	6.51	3.72	1.75
	Quadratic X SES Group	-0.33	3.43	-0.1

LG was not a significant predictor of the initial status (β = -0.06, *SE* = 4.29, *p* =.99). LG did significantly predicted the linear change in response times (β = -15.43, *SE* = 6.88, *p* = .03) but not the quadratic rate of change (β = 6.51, *SE* = 3.72, *p* = .08) although this figure was approaching significance. Mean scores (Table 28) and Figure 37 show that the monolinguals and bilinguals recorded similar response times at Time 1 but that bilinguals' reduced their times at a faster rate than monolinguals. Bilinguals demonstrated a linear decrease in Trails B response times. Although monolinguals RTs appeared to reduce, their growth was more quadratic as the rate of change slowed from Time 2 to Time 3. This may explain why LG was a significant predictor of the linear but not quadratic growth model.

SES was a significant predictor of initial status (β = -11.67, *SE* = 3.91, *p* < .001) but not of the linear (β = 4.12, *SE* = 6.31, *p* = .52) or quadratic (β = -0.33, *SE* = -0.1, *p* = .92) change in times. Figure 38 show that the low-SES groups took significantly longer than the mid-SES groups to complete the Trails B although the rate of change in both groups did not differ.

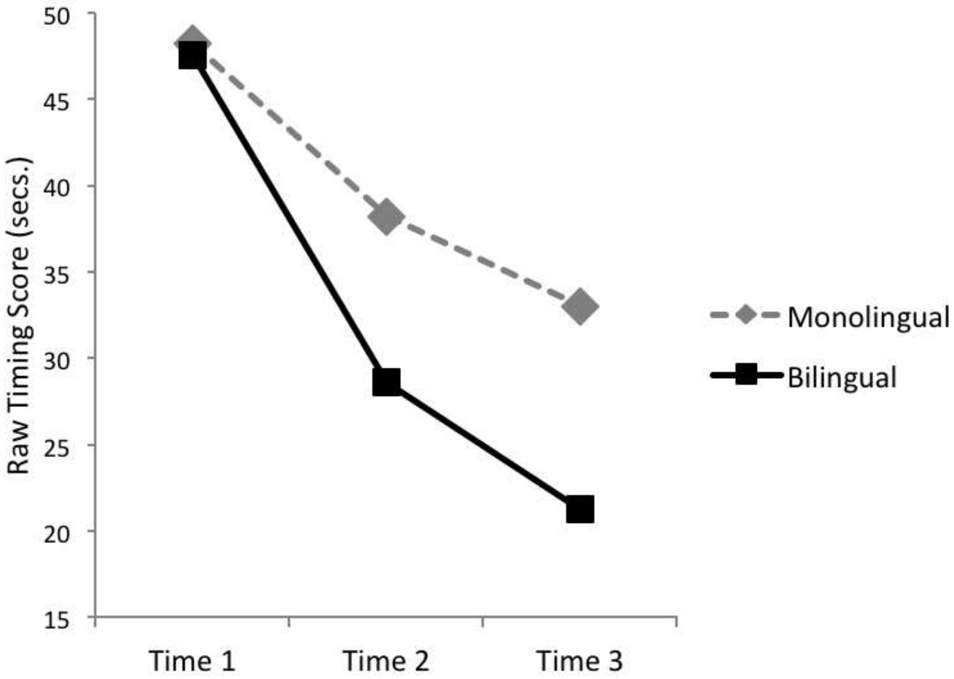


Figure 37 *Fitted trajectory for the quadratic growth model of Trails B with added predictor language group*

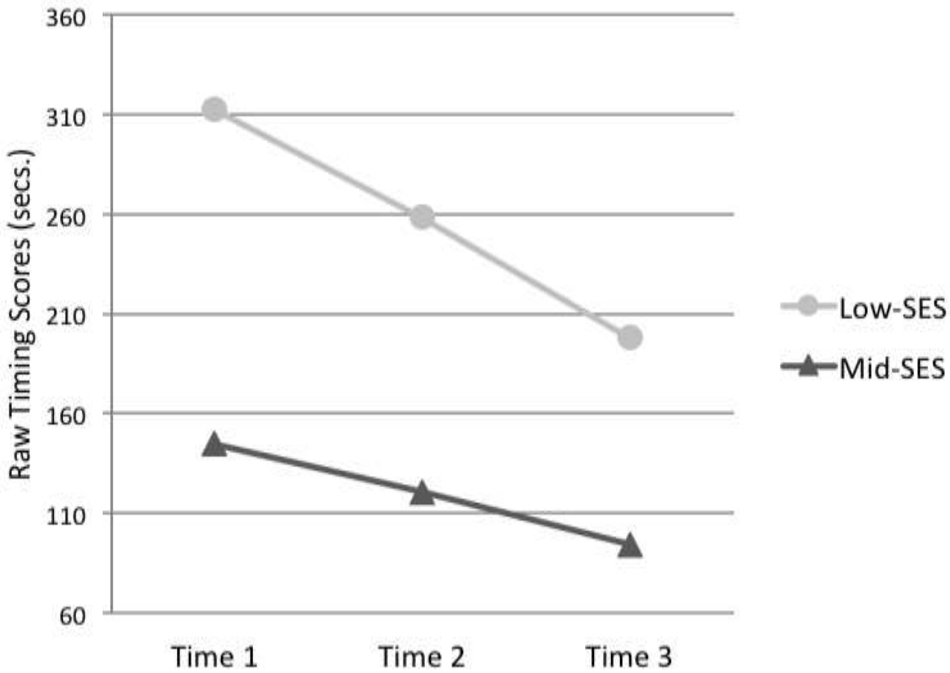


Figure 38 *Fitted trajectory for the quadratic growth model of Trails B response times with added predictor socioeconomic status*

5.3.5.1.3. Trails Difference

To calculate the Trails Difference scores each participant's response times on the Trails A were subtracted from their Trails B response times:

Trails Difference = Trails B – Trails A

Descriptive statistics for the Trails Difference response times are presented in Table 30. Results from the unconditional mean model indicated an ICC score of .17, suggesting that 17% of the total variation in the Trails Difference timing scores was due to inter-individual differences. As this percentage is below the minimum of 0.25 (or 25%) recommended when using IGC, ANOVA were used to estimate fixed effects.

Table 30 *Mean (and standard deviations) for Trails Difference response times (sec) across time*

Socioeconomic Status	Language Group	Time 1	Time 2	Time 3
Mid-SES	Monolingual (*n* = 32)	37.78 (23.76)	32.53 (33.81)	24.31 (20.98)
	Bilingual (*n* = 42)	42.71 (24.86)	20.82 (20.90)	17.31 (16.23)
	Total (*n* = 74)	40.58 (24.35)	25.89 (27.66)	20.33 (18.63)
Low-SES	Monolingual (*n* = 16)	58.63 (22.52)	43.87 (21.62)	41.71 (28.16)
	Bilingual (*n* = 18)	52.44 (44.04)	36.35 (23.66)	25.16 (19.55)
	Total (*n* = 34)	55.35 (35.20)	39.89 (23.66)	32.95 (25.05)
Total	Monolingual (*n* = 48)	44.73 (25.16)	36.31 (30.53)	30.11 (24.74)
	Bilingual (*n* = 60)	45.63 (31.76)	25.48 (22.73)	19.66 (17.50)
	Total (*n* = 108)	45.23 (28.89)	30.29 (26.90)	24.30 (21.56)

A 3 x 2 x 2 mixed ANOVA revealed a significant main effect of time, *F* (2, 208) = 22.71, $p < .01$, $\eta_p^2 = .18$, SES, *F* (1, 104) = 13.69, $p < .01$; $\eta_p^2 = .12$, and language group, *F* (1, 104) = 3.88, $p = .05$; $\eta_p^2 = .04$. Pairwise comparisons and mean scores (Table 30) revealed that Trails Difference timing scores significantly decreased between Time 1 and Time 2 ($p < .01$), Time 2 and Time 3 ($p = .04$) and between Time 1 and Time 3 ($p < .01$). No other interactions or main effects were found.

To assess Language and SES group differences, one-way ANOVAs were conducted across time. There was a main effect of language group at Time 2, *F* (1, 107) = 4.46, $p = .04$; $\eta^2 = .04$, and Time 3, *F* (1, 107) = 6.58, $p = .01$; $\eta^2 = .06$, but not at Time 1. Mean scores indicated that at both Time 2 and Time 3, bilinguals had significantly lower Trails Difference response times than monolinguals. At Time 1 however, monolinguals had slightly lower response times than bilinguals (see Fig. 39). This pattern of results showed that bilinguals experienced a reduction in the additional EF demands necessary to complete the more complex Trails B across time and overcame their disadvantage at Time 1.

There was a significant main effect of SES at Time 1, *F* (1, 107) = 6.40, $p = .01$; $\eta^2 = .06$, Time 2, *F* (1, 107) = 6.65, $p = .01$; $\eta^2 = .05$, and Time 3, *F* (1, 107) = 8.54, $p < .02$; $\eta^2 = .07$. At all three time points, the low-SES group had significantly higher response times than the mid-SES group (Fig. 40). This indicated that the difference between Trails A and Trails B required significantly higher levels of EF demand for the low-SES group compared with the mid-SES.

To file was split according to SES to examine LG differences within each SES group. The only main effect found was in the low-SES group at Time 3, *F* (1, 33) = 4.04, $p = .05$; $\eta^2 = .11$. Here bilinguals had significantly lower timing scores than monolinguals.

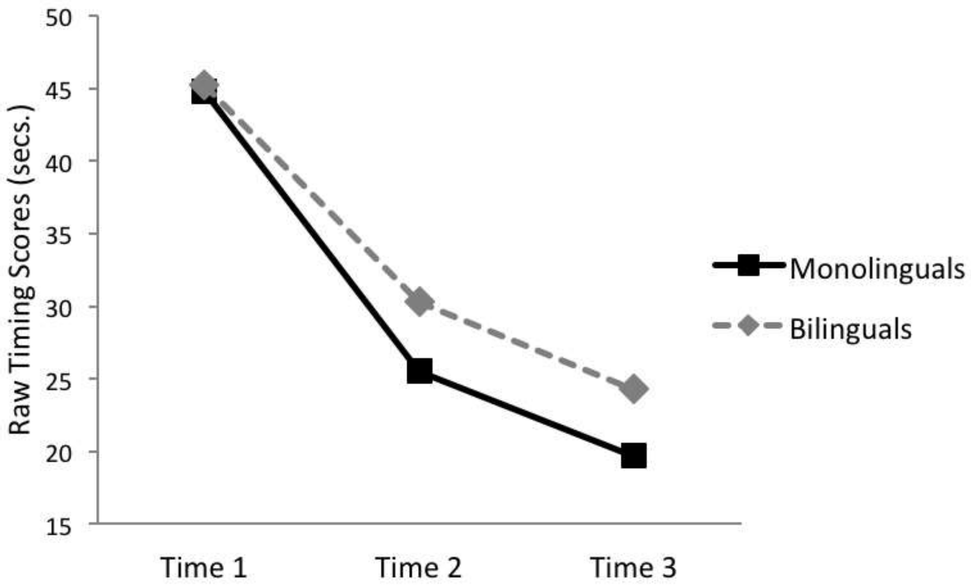


Figure 39 *Trails Difference response times (sec) between language*

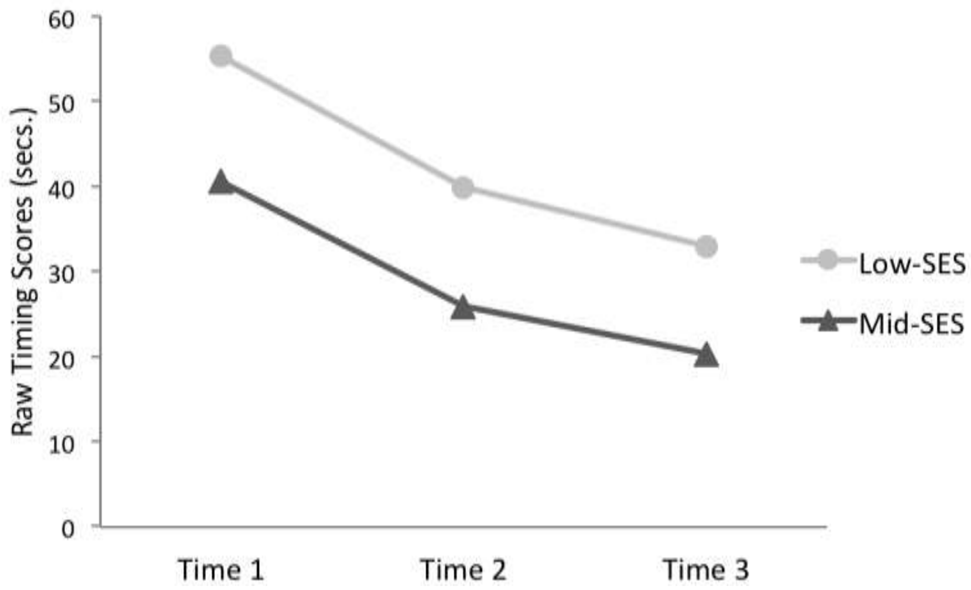


Figure 40 *Trails Difference response times (sec) between socioeconomic groups*

5.3.5.2. Wisconsin Card Sorting Test

As the Wisconsin Card Sorting Test (WCST) was administered at Time 3 only, one-way ANOVAs were used to compare groups. A number of outcomes variables were provided by the WCST (for details see section 4.5.7.): *Total Errors, Perseverative Responses, Perseverative Errors, Nonperseverative Errors, Trials to Complete First Category, Categories Complete and Failure to Maintain Set*. For each of these outcome variables standardised scores are calculated using raw scores and age at time tested and descriptive statistics are presented in Table 31.

Table 31 *Means (and standard deviations) for Wisconsin Card Sort Test*

Task Outcome	SES Group	Monolinguals (n = 40)	Bilinguals (n = 59)	Total (n = 99)
Total Errors	Mid-SES (n = 70)	96.61 (14.73)	102.98 (10.76)	100.43 (12.79)
	Low-SES (n = 29)	95.17 (13.60)	103.71 (11.96)	100.17 (13.15)
	Total (n = 99)	96.18 (14.24)	103.19 (11.02)	100.35 (12.83)
Perseverative Responses	Mid-SES	97.07 (13.19)	104.67 (10.34)	101.63 (12.07)
	Low-SES	97.17 (15.06)	102.82 (12.70)	100.48 (13.76)
	Total	97.10 (13.58)	104.14 (10.99)	101.29 (12.53)
Perseverative Errors	Mid-SES	97.96 (13.80)	104.69 (10.77)	102.00 (12.52)
	Low-SES	97.25 (15.86)	102.82 (12.81)	100.52 (14.16)
	Total	97.75 (14.24)	104.15 (11.32)	101.57 (12.90)
Nonperseverative Errors	Mid-SES	94.75 (12.42)	100.36 (9.55)	98.11 (11.05)
	Low-SES	94.58 (12.75)	102.53 (9.87)	99.24 (11.64)
	Total	94.70 (12.36)	100.98 (9.61)	
Categories Complete	Mid-SES	4.72 (1.61)	5.74 (.59)	5.33 (1.21)
	Low-SES	5.25 (1.22)	5.65 (.61)	5.48 (.91)
	Total	4.88 (1.51)	5.71 (.59)	5.37 (1.13)
Trials to Complete 1st Category	Mid-SES	16.57 (11.89)	12.31 (3.84)	14.01 (8.28)
	Low-SES	16.08 (11.73)	12.94 (3.68)	14.24 (8.02)
	Total	16.43 (11.69)	12.49 (3.78)	14.08 (8.16)
Failure to Maintain Set	Mid-SES	1.43 (1.67)	.71 (86)	1.0 (1.29)
	Low-SES	.92 (1.08)	1.0 (.87)	.97 (.94)
	Total	1.28 (1.52)	.80 (.87)	.99 (1.19

There was a significant main effect of LG on each outcome variable (see Table 31): total errors, *F* (1, 98) = 7.36, *p* = .01; η^2 = .07, perseverative responses, *F* (1, 98) = 7.66, *p* = .01; η^2 = .07, perseverative errors, *F* (1, 98) = 5.80, *p* = .02; η^2 = .06, nonperseverative errors, *F* (1, 98) = 4.95, *p* = .03; η^2 = .04, categories complete, *F* (1, 89) = 12.55, *p* < .001; η^2 = .11, trials to complete first category, *F* (1, 98) = 5.76, *p* = .02; η^2 = .06 and failure to maintain set, *F* (1, 98) = 5.76, *p* = .04; η^2 = .04. In all outcomes, bilingual groups outperformed monolinguals. However, there was no main effect of SES and no significant interactions for any of the WCST outcome variables.

To examine LG effects in more detail the data were split according to SES group and independent t-tests were used to analyse effects. Within the low-SES group there was no significant difference between LGs for any of the outcome variables (see Table 32). Although not statistically significant, the Nonperseverative Errors variable was approaching significance between language groups, *t* (27) = 1.98, *p* = .06, with bilinguals scoring higher than monolinguals. Within the low-SES group, bilinguals had higher mean scores than monolinguals on all WCST outcome variables.

Table 32 *T-test results comparing low-SES language groups*

	t	*df*	*p (alpha)*	*Eta Squared*
Total Errors	1.79	28	.09	-
Perseverative Responses	1.10	28	.28	-
Perseverative Errors	1.05	28	.31	-
Non-Perseverative Errors	1.98	28	.06	-
Categories Complete	.66	28	.51	-
Trials to Complete First Category	-1.13	28	.27	-
Failure to Maintain Set	.08	28	.94	-

Within the Mid-SES group there was a significant difference between language groups on outcome variables: Total Errors, Perseverative Responses, Perseverative Errors, Categories Completed, Trials to Complete First Category and Failure to Maintain Set (see Table 33). There was no significant main effect for the Nonperseverative Errors variable.

Table 33 *Independent t-test results for language group differences within mid-SES group*

	t	*df*	*p (alpha)*	*Eta Squared*
Total Errors	2.09	69	.04	.06
Perseverative Responses	2.60	69	.01	.09
Perseverative Errors	2.18	69	.03	.04
Non-Perseverative Errors	1.43	69	.16	-
Categories Complete	2.78	69	.01	.10
Trials to Complete First Category	-1.81	69	.08	-
Failure to Maintain Set	-2.45	69	.04	.08

5.3.5.3. Unified executive function task summary

The TMT and WCST tasks aimed to assess the unified EF component as each were thought to require the co-ordination of a number of EF skills, adding to task complexity. The TMT, believed to tap the switching, updating and inhibitory control components of the EF is an unstandardized, timed task with three trial conditions. The WCST is a standardised assessment also requiring number of EF components.

As the TMT was administered at all three time points, age-related changes could be assessed using this task and response times improved across time for all task conditions. LG differences were found in the Trails B and Trails Difference conditions with bilinguals outperforming monolinguals as they got older. LG also predicted the change in participants Trails B response times with bilinguals improving at a faster rate compared with monolinguals. SES group differences were found in the Trails B and Trails Difference conditions with low-SES groups having more difficulty with this task as shown by their longer response times (sec) compared with mid-SES groups, a result that was sustained over time.

The WCST results indicated that participants were performing at around population average (M = 100) on all task outcome variables. There was no SES effect on this task however the LG effect was significant. Results showed that mid-SES bilinguals outperformed mid-SES monolinguals on all outcome variables of the WCST.

5.3.6. Relationships between tasks

To examine the relationship between control and EF tasks (Table 34) as well as EF tasks at each time point (Table 35), Pearson's correlation analyses were carried out for each EF task and across groups and time.

Table 34 shows that the control measures PPVT and RSPM positively correlated with one another. The PPVT also correlated with verbal but not visuospatial short-term and WM tasks. The RSPM on the other hand correlated with all WM measures as well as with the perseverative responses from the WCST and the switching task accuracy scores.

The Stroop IC condition did not significantly correlate with any of the other EF tasks, questioning its potential validity in assessing children's EF skills. On the other hand, the TMT was shown to correlate with a number of cognitive tasks including PPVT, RSPM, Verbal STM, VS STM, VS WM, switching and IC task (Opposite Worlds). This showed that it may be a useful assessment of children's general or unified EF function. It did not however, correlate with the perseverative responses from the WCST which may have been the result of the nature of these tasks (one timed, one untimed). Furthermore, other outcomes from WCST may have been correlated with the TMT as only perseverative responses and errors were assessed here.
The Opposite Worlds task also correlated with a number of outcomes including the PPVT, VS WM, switching task timing but not accuracy condition, Trails B and Trails Difference. This indicates that a level of IC may have been required during performance on the TMT task.

Table 35 shows that there were high test-retest reliabilities between all EF tasks tested at Time 1, Time 2 and Time 3 with positive high correlations between time points for each measure. The Stroop however, did not display such high correlations and only Time 1 and Time 2 RTs correlated with one another.

Table 34 *Relationships between control measures and executive function tasks at Time 1*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. PPVT	1	.27**	.28*	.28**	.20	.16	.20	.20	.09	.17	.15	.23*	-.22*	-.18	.10
2. RSPM	.27**	1	.28**	.36**	.30**	.36**	.21*	.20	.03	.30**	.03	.03	-.24*	-.21*	.08
3. Verbal STM	.28**	.28**	1	.65**	.18	.16	.14	.17	.16	.14	.09	.15	-.26	-.18	-.10
4. Verbal WM	.28**	.36**	.65**	1	.15	.07	.16	.18	.14	.29**	.00	.10	-.13	-.09	-.02
5. VS STM	.20	.30**	.18	.15	1	.46**	.10	.11	.04	.20	.06	.21	-.26*	-.18	-.05
6. VS WM	.16	.36**	.16	.07	.46**	1	.10	.09	.07	.07	.05	.22*	-.22*	-.16	.01
7. WCST Perseverative Responses	.20	.21*	.14	.16	.10	.10	1	.98**	-.09	.04	-.09	-.05	.07	.16	.11
8. WCST Perseverative Errors	.20	.20	-17	.18	.11	.09	.98**	1	-.11	.06	-.09	-.05	.06	.14	.05
9. Switching (Time)	.09	.03	.16	.14	.04	.07	-.10	-.11	1	.26**	.44**	.59**	-.38**	-.28**	.17
10. Switching (Accuracy)	.17	.30**	.14	.29**	.20	.07	.04	.06	.26**	1	.05	.16	-.27**	-.19*	.08
11. Same World	.15	.03	.09	.00	.06	.05	-.09	-.09	.44**	.05	1	.65**	-.28**	-.19*	.02
12. Opposite World	.23*	.03	.15	.10	.21	.22*	-.05	-.05	.59**	.16	.66**	1	-.40**	-.26**	.01
13. Trails B	-.22*	-.24*	-.26*	-.13	.26*	-.22*	.07	.06	-.38**	-.27**	-.28**	-.40**	1	.81**	-.03
14. Trails Difference	-.18	-.21*	-.18	-.09	-.18	-.16	.16	.14	-.28**	-.19*	-.19*	-.26**	.81**	1	-.08
15. Stroop IC	.10	.08	-.09	-.02	-.05	.01	.12	.05	.17	.08	.02	.01	-.03	-.08	1

*Note: * $p < .05$, ** $p < .01$ (2-tailed)*

Table 35 *Relationships between executive function tasks across time*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	.27
1. Verbal STM	1	.65**	.18	.16	.14	.17	.16	.22*	.21*	.14	.29**	.24*	.09	.25*	.22*	.15	.22*	.22*	-.26*	-.16	-.12	-.18	-.08	-10	-.09	-.10	-.16
2. Verbal WM	.65**	1	.15	.07	.16	.18	.14	.19	.19	.29**	.20	.21*	.00	.24*	.23*	.10	.13	.24*	-.13	-.15	.00	-.09	-.10	-.14	-.02	-.02	-.13
3. VS STM	.18	.15	1	.46**	.10	.11	.04	.23*	.23*	.20	.13	.09	.06	.22*	.32**	.21	.19	.33*	-.26*	-.28**	-.32**	-.18	-.24*	-.04	-.05	.01	.03
4. VS WM	.16	.07	.46**	1	.10	.09	.07	.13	.23*	.07	.27*	.25*	.05	.16	.18	.22*	.22*	.23*	-.22*	-.30**	-.19	-.16	-.24*	-.24*	.01	-.00	.04
5. WCST Perseverative Responses	.14	.16	.10	.10	1	.98**	-.09	-.04	-.15	.04	.09	.10	-.09	-.01	.01	-.05	.05	-.04	.07	-.13	-.10	.16	-.12	-.01	.11	-.01	.19
6. WCST Perseverative Errors	.17	.18	.11	.09	.98**	1	-.11	-.05	-.17	.06	.06	.08	-.09	-.02	.00	-.05	.05	-.04	.06	-.12	-.09	.14	-.12	-.01	.05	-.01	.15
7. T1 Switching (Time)	.16	.14	.04	.07	-.09	-.11	1	.60**	.57**	.26**	.23*	.05	.44**	.42**	.42**	.59**	.49**	.41**	-.38**	-.23*	-.13	-.28*	-.16	-.30*	.17	.10	.15
8. T2 Switching (Time)	.22*	.19	.23*	.13	-.04	-.05	.60**	1	.55**	.36**	.23*	.04	.22*	.52**	.39**	.31**	.55**	.45**	-.37*	-.34**	-.18	-.24*	-.24*	-.22*	-.05	.05	.02
9. T3 Switching (Time)	.21*	.19	.23*	.23*	-.15	-.17	.57**	.55**	1	.25**	.16	.18	.34**	.50**	.56**	.41**	.40**	.56**	-.17	-.19*	-.09	-.11	-.12	-.17	-.02	.01	.02
10. T1 Switching (Accuracy)	.14	.29**	.20	.07	.04	.06	.26*	.36**	.25**	1	.27**	.03	.05	.17	.32	.16	.28**	.35**	-.27**	-.24*	-.18	-.19*	-.24*	-.04	.08	-.06	-.13
11. T2 Switching (Accuracy)	-.29**	.20	.13	.27*	.09	.06	.23*	.23*	.16	.27*	1	.20*	.04	.12	.08	.22*	.21*	.14	-.15	-.10	-.06	-.13	-.04	-.07	-.02	-.16	-.02
12. T3 Switching (Accuracy)	.24*	.21*	.09	.25*	.10	.08	.05	.04	.18	.03	.20*	1	.06	.14	.14	-.10	.04	.22*	-.10	-.08	0.2	-.08	.06	-.16	-.01	-.06	-.07
13. T1 SW	.09	.00	.06	.05	-.09	-.09	.44**	.22*	.34**	.05	.04	.06	1	.51**	.52**	.66**	.49**	.53**	.28**	-.17	-.20*	-.19*	-.08	-.16	.02	-.10	.12
14. T2 SW	.25*	.24*	.22**	.16	-.01	-.02	.42**	.52**	.50**	.17	.12	.14	.51**	1	.65**	.47**	.62**	.67**	-.35**	-.34**	-.31**	-.27**	-.24*	-.20*	.03	-.02	.10
15. T3 SW	.22**	.23*	.32**	.18	.01	.00	.42**	.39**	.56**	.32**	.08	.14	.52**	.65**	1	.53**	.56**	.73**	-.23*	-.37**	-.28**	-.11	-.23*	-.20*	.15	.08	.10
16. T1 OW	.15	.10	.21	.22*	-.05	-.05	.59**	.31**	.41**	.16	.22*	-.01	.66**	.47**	.53**	1	.57**	.60**	-.40	-.38**	-.42**	-.26**	-.18	-.25**	.01	-.00	.05
17. T2 OW	.22*	.13	.19	.23*	.05	.05	.49*	.55*	.40**	.28**	.21*	.04	.49**	.62**	.56*	.57*	1	.61*	-.44**	-.37**	-.36**	-.37**	-.22*	-.34**	.07	-.11	.11
18. T3 OW	.22*	.24*	.33*	.23*	-.04	-.04	.41*	.45**	.56**	.35**	.14	.22*	.53**	.67**	.73**	.60**	.61**	1	-.36**	-.41**	-.30**	-.23*	-.24*	-.35**	.09	-.09	.00

19. T1 Trails B	-.26*	-.13	-.26*	-.22*	.07	.06	-.38**	-.37**	-.17	-.27*	-.15	-.10	-.28**	-.35**	-.23*	-.40**	-.44**	-.36**	1	.63**	.44**	.81**	.39**	.42**	-.03	.14	.06
20. T2 Trails B	-.16	-.15	-.28**	-.30**	-.13	-.12	-.23*	-.34*	-.19*	-.24*	-.10	-.08	-.17	-.34**	-.37**	-.38**	-.37**	-.41**	.63*	1	.54*	.37**	.73**	.49**	-.05	-.01	-.12
21. T3 Trails B	-.12	.00	-.32**	-.19	-.10	-.09	0.13	-.18	-.09	-.18	-.06	.02	-.20*	-.31**	-.28**	-.42**	-.36**	-.30**	.44**	.54**	1	.23**	.27**	-.02	-.08	-.09	-.03
22. T1 Trails Difference	-.18	-.09	-.18	-.16	.16	.14	-.28**	-.24*	-.11	-.19*	-.13	-.08	-.19*	-.27**	-.11	-.26**	-.37**	-.23*	.81**	.37**	.23**	1	.27**	.36**	.-.08	.18	.87
23. T2 Trails Difference	-.08	-.10	-.24*	-.24*	-.12	-.12	-.16	-.24*	-.12	-.24*	-.04	.06	-.08	.24*	-.23*	-.18	-.22*	-.24*	.39**	.73**	.27**	.27**	1	.33**	-.12	.07	-.15
24. T3 Trails Difference	-.10	-.14	-.04	-.24*	-.01	-.01	-.30*	-.22*	-.17	-.04	-.07	-.16	-.16	-.20*	-.20*	-.25**	-.34**	-.35**	.42**	.49**	-.02	.36**	.33**	1	-.06	.09	-.01
25. T1 Stroop IC	-.09	-.02	-.05	.01	.11	.05	.17	-.05	-.02	.08	-.02	-.01	.02	.03	.15	.01	.07	.09	-.03	-.05	-.08	-.08	-.12	-.06	1	.26**	.19
26. T2 Stroop IC	-.10	-.02	.01	-.00	-.01	-.01	.10	.05	.01	-.06	-.16	-.06	-.10	-.02	.08	-.00	-.11	-.10	.14	-.01	-.09	.18	.07	.09	.26**	1	.18
27. T3 Stroop IC	-.16	-.13	.03	.04	.19	.15	.15	.02	.05	-.13	-.02	-.07	.12	.10	.10	.05	.11	.00	.06	-.12	-.03	.02	-.15	-.01	.19	.18	1

*Note: * p < .05, ** p < .01 (2-tailed)*

5.4. DISCUSSION

5.4.1. Summary of research aims

Pulling together topics discussed within the introductory chapters this study had a number of aims. For a full list of hypotheses see section 5.1.3.

Researchers continue to debate the effects of bilingualism and immersion education (IE) on children's executive function (EF) development (see Chapter 3). Therefore, the first aim of this study was to compare the EF development of two distinct language groups (LGs). 'Bilingual' children educated through early total immersion programmes in Ireland were compared with monolingual peers educated through English and matched for age, gender, non-verbal IQ and socioeconomic status (SES). As LG differences have been reported for some but not all EF components, a battery of tasks was employed to assess participants' specific as well as unified EFs (see section 2.3.). Furthermore, as the IE environment involves the sequential acquisition of an L2 rather than early simultaneous bilingualism (see section 1.2.4) this study assessed whether this language environment could produce the cognitive benefits found for early bilinguals on tasks of EF.

As EF continues to develop through to late adolescence and early adulthood (see section 2.5.) it was expected that children's overall EF performance should continue to improve during primary school or the middle childhood period. Subsequently, the second aim was to examine the developmental nature of EF. Using a longitudinal design, participants were assessed over a three-year period to monitor specific and unified EF performance over time.

As SES has been shown to affect aspects of cognitive, linguistic and EF development (see section 5.1.1.), the third aim of the study was to investigate the effect of SES on children's EF development by comparing children from mid and low-SES groups. The study also examined LG differences between low- and mid-SES groups as bilingualism research has mostly controlled for rather than examined SES differences between LGs.

Finally, the study aimed to understand the relationship between certain EF tasks in order to examine whether findings were in line with the dominant model of EF proposed by Miyake and colleagues (2000; see also Miyake & Friedman, 2012) which postulated a unified yet diverse nature of EF made up of specific components, which can also work together during complex tasks.

5.4.2. Summary of research findings and discussion

5.4.2.1. Control measures

Previous research has found that bilingual children often perform more poorly on tasks of verbal skill compared with monolingual peers (e.g. Gollan et al., 2004; Gollan et al., 2008; see section 3.3.). Although monolinguals in this study had higher means on the PPVT than bilinguals, the difference between groups was not significant. This indicated that the expected verbal disadvantage for bilinguals was not present in this group. However, this was not the first study to show a lack of difference between monolinguals' and bilinguals' PPVT scores (e.g. Allman, 2002; Barac & Bialystok, 2012; Bialystok, 2011; Cromdal, 1999; Yan & Nicoladis, 2009). Bialystok and colleagues (2004) reported no LG differences between younger and older adults' PPVT and RSPM scores, concluding that participants were comparable on verbal and spatial intelligence. Using similar samples to the ones used in this study (see section 5.3.), Parsons and Lyddy (2009) found that by 4th class (approximately 9-10 years of age) children within English-medium, Irish IE and Gaeltacht education did not differ in their levels of English reading, English non-word and English vocabulary skills. They concluded that Irish language skills can be achieved at no cost to English reading development.

One hypothesis for the lack of difference between LGs in this sample may be that IE bilinguals had more exposure to English than in other studies of bilingualism which regularly use simultaneous bilinguals. Furthermore, studies of bilingualism often use participants who have acquired English as an L2 (e.g. Portocarrero et al., 2007) while in this study, English was the L1 for all children. Furthermore, early bilinguals have shown improved PPVT over late bilinguals (Luk, de Sa, & Bialystok, 2011). Therefore, using IE bilinguals who had been exposed to their L2 before 11 years of age, may have prevented a bilingual disadvantage for the PPVT. These findings are also in line with IE research (e.g. Cummins, 1979, 2001; Genesee, 1994, 2004; Parsons & Lyddy, 2009) which has argued that although bilinguals' English skills may lag behind those of monolinguals during the early years of IE programmes, children often catch up with monolingual peers. As the children tested in this study had already been in IE for a minimum of 5 years, there did not appear to be any English deficit by this age although the study cannot comment on children's vocabulary skills during the early years of their IE experience or on their levels of L2 (Irish) vocabulary skills. Furthermore, as the PPVT was only assessed at one time point the study cannot comment on the developmental trajectory of children's English vocabulary skills.

No significant difference between LGs was found on the test of non-verbal intelligence (*g*; Raven et al., 2004). If a significant effect of LG had been found, EF differences may have been mediated by children's general IQ (*g*) rather than their language experience. Previous studies which have failed to control for this variable have been heavily criticised for their lack of adequate controls (e.g. Macnamara, 1966; Saer, 1923; see section 1.3.). Furthermore, most current bilingual studies utilise the RSPM (Raven et al., 2004) as a common control for children's non-verbal IQ or reasoning skills and report no LG differences (e.g. Bialystok, Peets, & Moreno, 2012; De Abreu, 2011).

The SES finding that low-SES groups performed significantly worse than mid-SES groups for both English vocabulary and non-verbal IQ skills is in line with previous research. A common finding has been that SES has a significant impact on children's general cognitive and language skills. Language skills have been shown to differ sharply as a function of SES (Engel et al., 2011; Hackman & Farah, 2009; Hoff, 2003). For instance, vocabulary has been shown to be twice as large in 3-year-olds from professional families compared to families on welfare (Hart & Risley, 1995). Similarly on tasks of general IQ and academic skills, low-SES participants have shown delays compared with mid-SES peers (Alexander et al., 1993; Hanscombe et al., 2012; Hackman & Farah, 2009). Bradley and Corwyn (2002) provided a review of the developmental outcomes affected by low-SES backgrounds and cited studies in which low parental education is associated with lower levels of school achievement and IQ later in childhood (see also, Noble et al., 2007). They also noted that combined SES measures are more predictive of SES than any single indicator. The potential mechanisms for SES differences in IQ and performance will be discussed in section 7.4. Despite there being SES group effects, there was no significant difference between LGs within the mid- or low-SES groups for these measures. As a result, it is concluded that English linguistic skills and non-verbal IQ skills were equivalent between LGs and that SES significantly impacted on these factors.

5.4.2.2. Inhibitory control

Two EF tasks were employed to assess participants' inhibitory control (IC) executive function: the OW (section 4.5.3.2.) and Colour-word Stroop tasks (section 4.5.4.). The benefit of these tasks was that a measure of additional cognitive control demands could be evaluated to examine the additional attentional control required to perform the more difficult, incongruent trials over congruent trials and was calculated using participants' reaction time (RT) differences.

No effect of language or SES group was found for the congruent (Same World) or incongruent (Opposite World) conditions of the Opposite Worlds task across time. However in both conditions a significant interaction was present between LG, SES group and time. The MSB group significantly improved their age-scaled, standardised scores from Time 1 to Time 3 and were the highest scoring group in both the Same and Opposite World conditions at Time 3 of testing. However, the LSB group did not improve their results at the same rate. Furthermore, in the low-SES group, monolingual performance significantly improved from Time 1 to Time 3. In fact, at Time 1, monolinguals in the low-SES group performed significantly worse than bilinguals on the incongruent, Opposite World condition but performed equivalently by Time 2 and Time 3. One issue that may have confounded the high improvement in the LSMs scores was that the novelty in this EF task was lost by Time 2 and Time 3. Therefore, practice effects (PE) may have mediated the differences between monolingual and bilingual groups. Low-SES monolinguals may have had a higher recall of the task and may have been able to utilise certain test-taking strategies to improve their scores at a faster rate than bilinguals. However, what is not known is why such effects were not be found for the bilingual group. The issue of practice effects (PE) was also found when age-scaled scores were converted to percentile bands. On both the Same and Opposite World conditions, participants at Time 1 scored slightly below population average. However, at Time 2 this score improved to the population average and by Time 3 to slightly above population average. As this task was a standardised assessment, if children's development was expected to progress at a steady rate, scores would be predicted at remaining between the 30.9th – 43rd percentile bands as standardisations include age. In fact, the only group to not improve by 1 age-scaled score per year was the LSB group. This may indicate that this group did not have the same level of recall as the other three groups. Although the LSB group did not improve their scores at the same rate as monolingual peers, the group maintained a higher mean score on the Opposite Worlds condition. This shows that the groups had no EF deficit for this task, that EF performance did improve over time and that IC development can be fostered by the practice of having to repeat a task (see section 7.2. for discussion of PE).

The third, 'Worlds Difference' condition examined the difference in time taken to complete the incongruent (OW) over congruent (SW) conditions. In this respect, the Worlds Difference RT scores were similar to the classic Stroop effect paradigm (Stroop, 1935). In other words, this variable looked at the level of extra EF skill needed to complete the more difficult incongruent trials. As raw timing scores were used, no standardised scores were available. There was no main effect of time or SES on this condition. However, there was a significant main effect of LG with bilinguals outperforming monolinguals, showing a reduced timing delay. On closer inspection, the main effect of LG was heavily influenced by Time 3 results and no LG differences were found at Time 1 or Time 2. In other words, the significant effect in favour of bilinguals emerged over time. This is an important finding as cross-sectional analysis may not have picked up group differences at Time 1 or Time 2 and subsequently, researchers may have concluded that bilinguals performed equivalently to monolinguals on tasks of IC. In fact at Time 1, MSMs had slightly lower EF cost than bilinguals. When analysis looked within groups, it appeared that the LG effect was heavily influenced by the low-SES groups, where bilinguals significantly outperformed monolinguals. Therefore, although the LSBs RTs on the congruent and incongruent conditions did not display a significant improvement over time, the degree of EF cost did significantly decrease over time.

Unlike the Opposite Worlds task which used ANOVAs to analyse group differences, the Colour-word Stroop task used individual growth curve analysis (IGC; see section 6.2.3.) to assess children’s performance on the neutral, congruent and incongruent conditions. The facilitation and inhibition (Stroop effect) conditions were analysed using mixed-ANOVAs as they did not meet the criteria for growth curve modelling. A significant linear change in neutral, congruent and incongruent react times (RTs) was found across time. For all conditions mean RTs reduced with age. LG was not a significant predictor of the linear change in any of the Stroop conditions. SES group was a significant predictor in the rate of change for the neutral and congruent conditions of the task but not for the more difficult, incongruent condition. The mid-SES group were faster in their ability to process colours, colour-words and in giving verbal responses than the low-SES group, shown through their reduced RTs. However, SES was not a predictor of change for the incongruent trials indicating no difference between SES groups in processing more complex, incongruent information.

Furthermore, when the facilitation and inhibition conditions were explored using mixed-ANOVAs, no language or SES group differences were present across time or at each specific time point. In other words, groups did not differ in the degree of difficulty associated with inhibiting the unintended response (colour word) or the degree of facilitation experienced on congruent trials. When the inhibition RTs were explored within each language and SES group, a significant difference was found between monolinguals and bilinguals in the mid-SES group at Time 1. At Time 1, bilinguals had significantly slower inhibition RTs than monolinguals. In other words, they had greater difficulty than monolinguals in performing incongruent over congruent trials or using their IC skills. However, by Time 2 and Time 3 this difference had gone with bilinguals performing equivalently to monolingual peers. This indicated that, with time, bilinguals had overcome their IC delays at Time 1 and show an equal level of IC ability to monolinguals at Time 2 and Time 3. Again, this valuable information regarding IC development in bilinguals may have been lost using cross-sectional analysis. Perhaps as a result of their increased exposure in the L2 with increased time spent in IE, bilinguals had become more practiced at inhibiting the internal conflict between their dually-activated L1 and L2 (e.g. Green, 1998).

One repeated finding within the literature is that bilinguals show an advantage on both congruent and incongruent trials using Stroop-like task paradigms (e.g. Bialystok et al., 2004; Costa et al., 2008). However, this result was not replicated in this study as monolinguals and bilinguals showed equivalent performance in both congruent and incongruent trials on the Opposite Worlds and colour-word Stroop tasks. Recently, a number of studies have failed to replicate the advantage for bilinguals on tasks of IC (e.g. Bialystok, 2011; Carlson & Meltzoff, 2008; Costa et al., 2009; Martin-Rhee & Bialystok, 2008) and researchers have begun to move away from explanations which suggest that superior bilingual performance is the result of specific EF component advantages (e.g. Colzato et al., 2008; Bialystok et al., 2012; Kroll & Bialystok, 2013; see section 7.3. for further discussion).

One hypothesis for the lack of differences between LGs relates to how IC has been defined. A distinction has been made between response inhibition and interference suppression (Bunge et al., 2002; Martin-Rhee & Bialystok, 2008; see section 3.4.1.1.). Response inhibition involves participants inhibiting prepotent responses relating to a univalent stimulus, for example say “two” when you see the number ‘1’ (Opposite Worlds task) or say “day” to a picture of night (Day-Night task; Bialystok et al., 2012). Interference suppression requires participants to selectively focus and choose the relevant cue from a bivalent stimulus containing two cues, for example, the word “green” written in red ink (Stroop task). The bilingual advantage is more prevalent in tasks of interference suppression (Bialystok et al., 2008; Bialystok & Senman, 2004; Bialystok & Shapero, 2005) than response inhibition (Bialystok et al., 2006; Carlson & Meltzoff, 2008; Martin-Rhee & Bialystok, 2008) and some researchers have failed to find a LG difference in either type of IC. As the Opposite Worlds task tapped response inhibition, the lack of difference between LGs may not be unexpected in this case. However, no effect was found on the Stroop task, where previously bilinguals had shown superior performance. Reasons for this are discussed below. Furthermore, on the Worlds Difference condition of the Opposite Worlds, a LG effect did emerge over time indicating that bilinguals performed better when extra attentional demands were necessary, particularly at Time 3 of testing.

As interference suppression has been shown to be advantaged in bilinguals, as well as bilinguals often showing reduced Stroop effects, the lack of a bilingual advantage in this study was surprising. However, a study by Nicolay and Poncelet (2013a) also reported a lack of group differences between immersion bilinguals (aged 8 years) and monolinguals on all conditions of IC tasks similar to the Stroop-paradigm (ANT; Fan et al., 2002 and flanker task; Eriksen & Eriksen, 1974). They concluded that previous findings of a bilingual advantage for interference suppression have used samples of highly proficient or simultaneous bilinguals. Children in IE on the other hand are not thought to have the same level of proficiency in their L2 as their L1, particularly in productive skills (see section 1.4.2.). It is the practice of suppressing the unintended language that has led highly proficient bilinguals to become attuned in responding to conflicting cues (Costa & Santesteban, 2004; Green, 1998; Meuter & Allport, 1999), therefore the reduced practice that IE bilinguals had compared to simultaneous bilinguals may explain why no effects were found on this task of interference suppression. However, as the levels of L2 proficiency are not reported within this group, it is not known whether IE participants with high L2 proficiency compared with low L2 proficiency would perform better on the IC tasks. For instance, Inurritegui (2009) found a bilingual effect in favour of simultaneous but not sequential bilinguals on the Stroop task. Issues of LP and Stroop performance will be examined further in Chapter 6 (6.3.2.).

A second hypothesis for a lack of LG effects relates to the nature of the Stroop task itself. As the task involved ignoring colour-words while attending to the ink colour during congruent and incongruent trials, reading levels may have mediated any effects, with higher levels of reading skills adding to task difficulty (creating a higher level of interference). It is for this reason that the facilitation and inhibition conditions are important in the analysis of Stroop findings. However, research has found that young bilinguals display slower reading, reduced receptive vocabulary, and poorer linguistic processing skills compared with monolingual peers (Bialystok, 2009; see section 3.3.). Therefore, it is difficult to determine whether previous effects in favour of bilinguals are the result of poorer reading skills or enhanced IC skills. Furthermore, the use of the colour-word Stroop task in studies of bilingualism has recently received criticism due its use of linguistic stimuli (e.g. Hernández et al., 2010). Green's IC model (1998) of language control maintains that bilinguals must manage two active language representations at all times. As a result, bilinguals display disadvantages for tasks of linguistic processing due to the added difficulty of managing this conflict during linguistic tasks. However, for tasks using non-linguistic information, bilinguals may show an advantage due to their practice in inhibiting the non-intended language. As no effect of bilingualism was found on any condition of the Stroop task here it is difficult to ascertain whether the dual-language activation problem masked any bilingual effects or whether bilinguals simply had equivalent IC skills to their monolingual peers. However, if the IC model is correct and task difficulty was increased for bilinguals who had to select between language lexicons while inhibiting distracting colour-words (during incongruent trials) this did not hinder their RTs over time as no LG effect was found for the inhibition or facilitation conditions. Furthermore, despite mid-SES bilinguals showing a disadvantage for the Stroop effect at Time 1, they overcame this disadvantage by Time 2 and Time 3.

The Opposite Worlds task did not elicit such linguistic competition for bilingual participants. As a result, a bilingual effect did emerge on the Worlds Difference condition of this task but only at Time 3 of testing. This finding suggested that a bilingual advantage for the IC function may emerge over time, perhaps as a function of children's increased LP and time in IE. Although a LG difference was found at Time 3 for the Worlds Difference condition, no effect was found on the standardised Opposite World or incongruent condition of this task.

5.4.2.3. Switching

The task used to assess participants' switching abilities was the Creature Count (CC) task (Manly et al., 1999; see section 4.5.3.1.) which provided standardised scores for children's accuracy and timing across time.

For timing scores, no interactions or main effects were found. Results indicated that, like in the Opposite Worlds task, switching performance improved above participants' expected standardisations. In other words, it would be expected that participants' switching performance would improve relative to their age with children continuing to score in the 30.9^{th}- 43.4^{rd} percentile band. However, participants improved their switching beyond their predicted trajectory with performance improving to within the 56.6th-69.2nd percentile by Time 3. This may show that PE or experience with this task improved children's switching performance over time, so much so that a group who performed below average could improve their standard scores to above population average, possibly as a result of implementing strategies they have remembered from previous testing phases to improve performance.

Of all the groups tested, the group that showed the greatest improvement in scores was the LSM group. Their scores significantly improved from Time 1 to Time 2 and from Time 2 to Time 3. Although the LSBs had higher standardised timing scores than the LSM group at Time 1 and Time 2, their standardised scores remained within the same percentile band across time and by Time 3, the LSM group had continued to improve their scores to above the LSB group (although the difference was not significant).

For the accuracy scores, no main effect of LG was present across time. There was a significant three-way interaction in participants' accuracy scores. There was also an overall effect of time with accuracy scores increasing between Time 1 and Time 2 but not between Time 2 and Time 3 or between Time 1 and Time 3. The MSM and LSB group displayed an increase in their standardised accuracy scores from Time 1 to Time 2 but conversely displayed a decrease from Time 2 to Time 3.

It is important to note that although the changes in accuracy scores were significant in places, these differences were relatively small and were mostly within one-percentile band range. For instance, the MSM and LSB whose scores decreased from Time 2 to Time 3 still maintained accuracy scores within the population average (43.4th-56.6th percentile band). One hypothesis for this discrepancy between groups' accuracy scores is that there are only a small number of trials within the CC task (7). As each trial is marked either correct or incorrect, participants who make only one error within a trail or trial (section 4.5.3.1.) are given no opportunity to correct their mistake and instead the trial is marked as incorrect. This makes it difficult for participants to gain 7/7 trials correct or full marks on this task. Furthermore, if participants make a number of errors (e.g. 5/7) within the task this will often result in them receiving a low standardised score as there are so few trials.

A second hypothesis for why this decrease in standardised scores occurred within the mid- and low-monolingual groups between Time 2 and Time 3 was that there was a speed-accuracy trade-off. At Time 3 participants may have attempted to perform the trials at a fast pace, subsequently reducing their accuracy scores. Results from the timing condition suggested that the LSM group, for example, improved their timing scores significantly across time but this improvement in time may have been at a detriment to their accuracy scores which did not display such a linear increase. This point may also be shown by the accuracy improvement shown by LSBs. LSBs improved their accuracy means from within the 20.2^{nd}-30.9^{th} percentile at Time 1 to the 43.4^{th}-56.6^{th} percentile at Time 2. However, the standardised timing scores of this group did not improve as much from Time 1 to Time 2. In other words, LSBs may have taken more time to complete the CC trials at Time 2 which subsequently saw their accuracy scores improve.
Following a one-way ANOVA which included participants excluded from the mixed-ANOVA due to listwise deletion, a language group effect did emerge at Time 3 within the mid-SES group. Bilinguals at Time 3 had significantly higher (56.6^{th}-69.2^{nd} percentile band) standardised accuracy scores compared with monolinguals (43.4^{th}-56.6^{th} percentile band), although again these differences were small and differed by only one percentile band range.

Meuter and Allport (1999) argued that switching or cognitive flexibility was the mechanism responsible the bilingual advantage. Although studies have found effects with adult participants (Bialystok et al., 2004; Bialystok et al., 2006; Costa et al., 2008), few studies have replicated effects in tasks with children (e.g. Bialystok & Viswanathan, 2009; Kroll & Bialystok, 2013). Furthermore, the mechanisms for any bilingual advantage for switching are still poorly understood and are posulated as relating to the level of difficulty in the task, with higher levels of monitoring and attentional demands often resulting in superior bilingual performance (Bialystok et al., 2012; Costa et al., 2009). Recent research has suggested that rather than bilinguals being enhanced for their switching EF in particular, complex switching tasks which require a higher level of cognitive demand are where bilingual advantages lie and that studies should focus on the unified EF more than switching in particular (see section 5.4.2.5.).

5.4.2.4. Working memory

Using the Working Memory Test Battery-for Children, visuo-spatial (VS) short-term memory (STM) and working memory (WM) differences were examined across groups but at only one time point (Time 2). Results showed no significant difference between LGs on any of the STM or WM tasks. However, as these results are only taken at one time point, the study cannot comment on the developmental trajectories of children's STM and WM skills.

Performance on STM tasks has previously been shown to be equivalent in monolingual and bilingual participants (e.g. Bialystok, 2009; Danahy, Windsor, & Kohnert, 2007; De Abreu, 2011). Therefore it is concluded that bilinguals show no advantage over monolinguals on tasks of verbal or visuo-spatial (VS) memory.
The lack of WM differences between LGs is not an unlikely finding as a number of researchers have reported the equivalent performance of monolinguals and bilinguals on tasks of WM (e.g. Bajo et al., 2000; Bialystok & Viswanathan, 2009; Bonaffici et al., 2011; Da Fontoura & Siegel, 1995; Namazi & Thordardottir, 2010). Bialystok and colleagues (2008) found no difference in the performance of younger and older language groups on a task of self-ordered pointing (Petrides and Milner, 1982), however young adult bilinguals performed better than monolinguals on the backward version of the corsi-blocks task (Milner, 1971). To investigate this WM difference further Morales and colleagues (2012) developed a more complex version of the corsi-blocks task. With this they found no LG differences for WM although a LG effect was found for conditions which included an additionally stringent level of control and IC demands. Engel De Abreau (2011) used a longitudinal analysis over a three year period and found no LG differences in performance on a complex WM span task and on a digit recall task, concluding that the non-significant effect was not as a result of bilinguals' reduced verbal abilities and that bilingual advantages for WM did not emerge with increased exposure to the L2. This study is in line with recent research which has found no bilingual advantage for simple tasks of WM skill (e.g. Bialystok & Feng, 2010; Bonifacci et al., 2011; De Abreu, 2011), however when more EF demands are added, bilinguals may display superior performance (Morales et al., 2013). For instance, Morales and colleagues found an advantage for bilinguals on a modified version of the Simon task. However, their poor

conceptualisation of WM by using a classic IC task did not adequately unpick the WM mechanism for bilingual performance. Furthermore, as their bilingual sample was far from homogenous, incorporating 14 bilingual language pairs, results may have been confounded by the range of languages spoken by participants. Therefore this study concluded that IE bilingual children, in line with current bilingualism research should not display an advantage for STM or WM skills.

Previous tasks using linguistic stimuli to test participants STM (e.g. animal words) have found that bilinguals may perform poorly compared to monolingual peers as a result of their verbal disadvantage (e.g. Fernandes et al., 2007). Although no bilingual disadvantage emerged here, LSBs did show a difference between their verbal and VS WM performance, scoring lower on tasks using verbal information. Similar results have been found with older adult bilinguals who performed significantly worse than monolinguals in their verbal WM yet outperformed monolinguals in VS WM (Luo, Lin, Craik, Fergus, Moreno, Sylvain, & Bialystok, 2013). Although an effect of task stimuli was found in the LSB group, no such difference was obtained in the MSB group. This may indicate that the bilingual disadvantage for linguistic WM tasks may only be present for low-SES groups as language skills have been shown to be highly sensitive to SES factors (see 5.4.2.1.).

Although SES groups did not differ on their VS or verbal STM performance, they differed significantly on the standardised verbal-WM task as well as the number of trials correct on the VS WM task with mid-SES groups performing significantly better than low-SES groups. Although poorer cognitive abilities have been associated with low-SES, few studies have examined the effect of SES on the WM function in particular (e.g. Bradley & Corwyn, 2002; Guo & Harris, 2000; Hackman, Farah, & Meaney, 2010). A recent study found that poor WM and IC but not switching performance was mediated by children’s home environment and by three domains in particular: parental responsivity to the child, home enrichment and family companionship (Bradley & Corwyn, 1977; Sarsour et al., 2010). General IQ (*g*) has also been strongly associated with levels of SES (e.g. Brooks-Gunn & Duncan, 1997; Guo & Mullan-Harris, 2000; National Institute of Child Health and Human Development Early Child Care Research Network, 2005) and recent research suggests a significant and strong link between WM and IQ (Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Alloway & Alloway, 2010). Therefore the association between IQ and WM may explain the poorer performance of low-SES groups on WM tasks tested here.

5.4.2.5. Unified executive function

Two tasks were used to assess participants' unified EF component or the ability to perform more complex tasks, requiring the co-ordination of EF skills: the Trail Making Test (TMT; Armitage, 1946) and the Wisconsin Card Sort Test (Heaton et al., 1981).

Participants' Trails A performance showed a linear trajectory with RTs improving as a function of age but neither LG nor SES predicted this improvement in performance across time. In other words, all groups improved their response times at a similar rate for the less demanding, Trails A condition.

The Trails B response times showed a significant quadratic change in scores across time. Results showed that participants' significantly improved from Time 1 to Time 2 but that the rate of improvement slowed between Time 2 and Time 3. LG predicted the linear decrease but not the quadratic rate of change in response times. Bilinguals showed a faster linear decrease in their Trails B RTs while monolinguals' rate of change appeared to slow from Time 2 to Time 3.

SES significantly predicted the initial status but not in the rate of linear or quadratic change in Trails B response times. In other words, the low-SES group were significantly slower overall than the mid-SES group but both groups' improved their response times at a similar rate across time. Therefore, although low-SES backgrounds may impact negatively on this measure of unified EF, the deficit did not appear to be aggravated as children got older.

There was a significant effect of time, LG and SES on the Trails Difference condition. RTs reduced across time, indicating a greater ability to deal with the additional attentional demand during Trails B performance relative to Trails A as children got older. The significant LG effect in favour of bilinguals was not present at Time 1 but was significant by Time 2 and Time 3. In other words, the advantage in favour of bilinguals emerged over time. The SES effect was significant across all three time-points with low-SES participants performing more poorly than mid-SES.

A number of standardised outcome variables are provided by the WCST and there was a significant difference between LGs on a number of these outcomes including: total errors, perseverative responses and perseverative errors. There was also a significant difference between LGs on the unstandardized outcomes: categories complete and trials to complete first category. In all cases, bilinguals scored significantly higher than monolinguals, although when examined more closely, significant effects were heavily influenced by the mid-SES group and no LG differences were present within the low-SES group. Furthermore, on all outcome variables, no significant difference between SES groups was found.

Much of the recent research which compares the EF skills of LGs has suggested that the mechanism for the bilingual advantage lies in the ability to co-ordinate and manage a number of EF components when performing more complex EF tasks (e.g. Bialystok et al., 2012; Costa et al., 2009; Hilchey & Klein, 2011; Kroll & Bialystok, 2013; Morales et al., 2012). These findings are in line with the updated model of EF by Miyake and Friedman (2012; see section 2.3.4) who have placed more emphasis on the unitary nature of EF components than in their previous model (Miyake et al., 2000, see section 2.3.). Furthermore, the mixed results found in many studies of bilingualism assessing specific EF components (see Hilchey & Klein, 2011) may have been moderated by the level of EF demand required, as the harder the task the more likely the chance of obtaining a LG effect and the better the bilinguals perform (Bialystok et al., 2012; Kroll & Bialystok, 2013; Costa et al., 2009). Costa and colleagues (2009) described this more unified component as “monitoring” while Bialystok (2011) described it as “coordination”. Here it is described it as the “unified EF component”. Kroll and Bialystok (2013) also argued against approaching the bilingualism and EF question from a component-specific perspective as bilingualism is an experience which profoundly affects a number of cognitive networks and changes how language is processed generally. Bialystok even criticised her earlier studies which explained the bilingual advantage through specific EF mechanisms alone and used task paradigms which aimed to tap these specific components (e.g. IC on the Stroop task; Bialystok et al., 2004; Bialystok, 2009). Interestingly, the finding of an enhanced unified EF component for bilinguals was first cited in some of the early studies of bilingualism suggesting research may have come around full circle (Kroll & Bialystok, 2013; Peal & Lambert, 1962). For instance, Ben Zeev (1977b) noted that bilinguals appeared to approach cognitive tasks in a more analytic fashion and appeared more attentive to both structure and details of task instructions and administration, as well as being more sensitive to feedback. These attentional control skills are necessary when performing complex tasks of EF such as the WCST.

In line with this current research, findings from this study suggest that in tasks requiring a number of EF components (or complex EF tasks), bilinguals outperform monolinguals. Furthermore as the TMT showed significant results at Time 2 and Time 3 only and as the WCST was assessed at Time 3 only, the effect of bilingualism on these tasks may have emerged as bilinguals increased their experience with the L2 (or LP) and time spent within Irish-medium IE. In a previous study using the TMT, Bialystok (2011) also found that 6-year old bilinguals outperformed monolinguals on the TMT although effects were found on both Trails A and B. Similarly, previous research has shown a bilingual advantage on the simpler version of the WCST (Bialystok, 1999; Bialystok & Martin, 2004), the Dimensional Change Card Sort Test (DCCS; Zelazo et al., 1995) with pre-schoolers. In a study of LP and WCST performance, simultaneous adult bilingual interpreters who had extensive experience with the L1 and L2 as a result of negotiating both languages in their daily lives were compared with typical adult bilinguals and monolinguals (Yudes, Macizo, & Bajo, 2011). Typical and interpreter bilinguals did not differ in their levels of L2 proficiency ratings although interpreters had more practice and training with their L2 and used it as part of their profession. Bilingual interpreters showed significantly enhanced performance compared with monolinguals and bilinguals in their trials to complete first category, number of errors and in their percentage of perseverative errors made. However, no significant difference was found between groups on the Simon, WM or RSPM tasks. These results indicated that with the extensive practice of switching, monitoring and inhibiting both active languages in the brain, bilinguals may show enhanced performance in their overall EF skills. Although this study may suggest the same conclusion, no measure of LP or L2 (Irish) skill was assessed. The effect of LP on IE children’s EF performance will be examined further in Chapter 6.

One finding that was not predicted was that the low-SES participants would perform significantly worse on the TMT but equivalently on the WCST than mid-SES participants. One hypothesis for this difference relates to the unified and diverse nature of the EF. Although both these tasks have been used as general measures of EF, each has specific characteristics which may correlate more highly with certain EF components.
Sanchez and colleagues (2009) carried out regression analysis to assess the cognitive correlates between Trails conditions and a range of other cognitive assessments. They found that WM (measured by the WAIS-III Digit backward) was the function most correlated with the Trails B than all other functions assessed and after controlling for perceptual skills (visual search measured by Digit Symbol). These findings are also in line with Crowe (1998) who suggested that WM contributed more to performance on the Trails B than task-switching. Findings from this study revealed a significant main effect of SES on the Trails B and Trails Difference but not Trails A condition. SES differences were also obtained in the WM tasks (section 5.4.2.4.). Therefore, the WM deficits for low-SES groups may have mediated their poorer performance on TMT conditions. When correlation analyses were run to explore the relationship between tasks, the Trails B only correlated with the VS WM task and not the verbal WM or STM tasks.

While a significant effect of SES was found for the TMT, no such difference was shown on the WCST as groups performed equivalently. Furthermore, no LG effect was present in the low-SES group. The lack of SES difference may relate to the specific EF components most relied on in the performance of the WCST. Although the WCST is used as a task of unified EF skills, some have associated performance with switching function skills as it requires switching between sorting categories after 10 consecutive correct trials (e.g. Miyake et al., 2000; Vaderploeg, Schinka, & Retzlaff, 1994). In this study, no SES difference was found for the CC switching task at Time 3 of testing (timing and accuracy) yet a LG effect was found in favour of bilinguals (accuracy and Mid-SES only). Therefore, the reliance of the WCST on the switching EF may be the reason for the LG but not SES group effect. However, a study by Gamboz, Borella and Brandimonte (2009) did not replicate the correlation between WCST outcomes and task switching abilities (local switch costs and perseverative errors) of young adults. Similarly no significant correlation was obtained between the switching outcomes at Time 3 and the WCST in this study (see section 5.4.2.6.) as perseverative responses and errors correlated with the RSPM only. Although this does not explain the SES, LG discrepancy, the correlation results do implicate both the TMT and WCST in testing more complex or unified EF skills as the TMT related to a number of EF tasks while the WCST related to general functioning.

A second hypothesis for the contradictory SES findings on unified EF tasks relates to speed of information processing (SIP). SIP is basic component of intellectual functioning (Jensen, 1998) and refers to the rate at which sensory information is passed into the nervous system to be operated on. As a result SIP can affect domain-general cognitive processes (*g*) rather than domain specific-processes (EFs; Bonifacci et al., 2011). While SIP has been shown to be equivalent in monolinguals and bilinguals samples using RT tasks (e.g. neutral conditions of Stroop task; Bialystok et al., 2005; Bonifacci et al., 2011; Costa et al., 2009; Martin-Rhee & Bialystok, 2008), SIP differences may have been what caused low-SES groups to perform poorly on the Trails B and Trails Difference conditions compared with mid-SES groups, as well as the low-SES groups' RT delays discussed earlier for the neutral and congruent conditions of the Stroop task (see section 6.5.2.2.). In a study by Bosco (1972) low-SES, disadvantaged children took significantly longer to process visual information than mid-SES children but performance became more equivalent with age. Furthermore, low-SES groups performed significantly worse on the non-verbal IQ (*g*) test and IQ has been acknowledged as the primary dimension correlated with SIP (Jensen, 1998; Nettlebeck, 1987; Neisser et al., 1996).

5.4.3. Summary

This study provided valuable insights regarding the developmental nature of the EF as well as language and SES group differences. By implementing a battery of EF tasks the developmental trajectories of language and SES groups was compared over a three year period.

Children's EF performance showed improvement from 9-11 years on most EF tasks across time as development displayed a linear trajectory with performance improving as a function of age. These results are in line with current EF literature and studies with middle childhood participants (Best et al., 2011; Diamond & Lee, 2011). One finding that was not predicted was that participants' standardised IC and switching scores improved significantly over time. If children were to improve scores as a function of age, standardised scores would be predicted to remain in the same percentile band across time. However, as this was not the case here it might be argued that children's EF skills for these tasks significantly improved and developed across time. One hypothesis for this improvement was that PE resulted in children recalling the appropriate test-taking strategies for performing the tasks in their LTM, therefore improving performance. Issues of PE are discussed in section 7.2.

Not all EF tasks displayed a linear improvement as accuracy on the switching task decreased slightly from Time 2 to Time 3. These findings support the hypothesis of Best and colleagues (2011) who suggested that accuracy and response times should be reported separately to provide a more in-depth analysis of EF development and processes. Furthermore, children's EF development may be marked by a speed-accuracy trade-off. Davidson (2006) demonstrated this effect when he compared the developmental trajectories of children's EF skills from middle childhood to early adulthood. As children grew older they adjusted their RTs to maintain high levels of accuracy suggesting the influence of metacognition on the development of mature task performance. However, this study demonstrated a speed-accuracy trade-off in the opposite direction. The significant improvement in children's RTs over time may have resulted in lower accuracy scores as children attempted to perform as quickly as possible during this timed task, subsequently resulting in more errors being made.

In line with recent studies of bilingualism (e.g. Kroll & Bialystok, 2013; Bialystok et al., 2012; Barac & Bialystok, 2012; Costa et al., 2009; Colzato et al., 2008; Nicolay & Poncelet, 2013a), this study found that bilinguals showed significantly enhanced performance on unified EF or complex EF tasks (TMT, WCST, and Worlds Difference tasks) but equivalent performance on tasks of IC, WM and switching compared with monolingual peers. The longitudinal nature of this study also showed that the bilingual advantage on certain EF tasks emerged over time. For the Trails B and Trails Difference conditions, bilingual performance was equivalent to monolinguals at Time 1 but significantly better by Time 2 and Time 3. Furthermore, Worlds Difference RTs and switching accuracy scores in the mid-SES group were equivalent between monolinguals and bilinguals at Time 1 and Time 2 yet bilingual performance was superior at Time 3. This information may have been lost if cross-sectional data only had been used. Furthermore, in line with the conclusions of Kroll and Bialystok (2013) these findings suggest that researchers shift their explanations of specific EF mechanisms enhancing bilingual performance and instead consider that bilinguals have an ability co-ordinate and manage EF skills during task performance as a result of managing and controlling two languages in the brain at all times.

Results also showed that SES plays an important role in the development of children's language, non-verbal, and speed of processing information (SIP) skills in particular. These findings are in line with research which has shown that SES strongly impacts on children's language and IQ skills (e.g. Turkheimer et al., 2003). Low-SES group delays were also found for certain EF measures (WM and TMT) although performance in these groups did not deteriorate relative to mid-SES participants over time and instead, showed a similar improvement as children grew older. The rate of improvement for certain RT tasks (Stroop neutral, congruent) was better for mid-SES compared to low-SES groups, however EFs were not thought to be the causal mechanism for these differences. Instead, SIP which is highly correlated with non-verbal IQ was thought to be delayed in low-SES groups, resulting in slower RTs on Stroop conditions and Trails A across time. The SES effect found on the TMT but not the WCST is also thought be the result of this SIP delay as both tasks were postulated to assess unified EF skills. Furthermore, WM skills were poorer for low-SES groups and WM has previously shown a higher correlation with the TMT than other specific EF components (see 5.4.2.5.).

LG differences were present in some but not all measures of EF in the low-SES group. For instance, LSBs performed significantly better than LSMs on the Trails Difference and Worlds Difference RT conditions but equivalently on all other tasks. This demonstrated that bilinguals in the low-SES group may have had enhanced abilities in overcoming the additional cognitive demand necessary to perform the more complex conditions of these tasks yet no LG differences were present on the complex EF task, WCST. Issues of LG and SES performance are discussed further in section 7.4.

5.4.4. Limitations and future directions

A number of limitations for this study are discussed below:

- Some critics of this study may question why parental income was not used as a measure of participants' SES. However, levels of family poverty have not been found to predict children's intellectual development as home environmental factors appear to be more indicative of children's intellectual capabilities (Guo & Harris, 2000). Current research also proposes that to adequately assess SES, individual and aggregate measures should be used in combination with one another as was done here (see section 5.1.2.).

- Some researchers have suggested that the validity of EF measures requires that they be novel and adequately complex to truly assess EF capabilities, as individuals must formulate strategies and monitor their effectiveness (Anderson, 2002). Consequently, if a task is no longer novel then it cannot be guaranteed that participants have utilised the same strategies or EF skills as they did the first time round, decreasing the tasks validity. However, if a longitudinal study, such as this one was to employ different tasks at each time point to overcome this issue of novelty, issues of task impurity could have potentially confounded results (Huizinga et al., 2006). Furthermore, as there was at least 9 months between each testing phase it was suggested that this decreased the opportunity for automaticity of strategies to occur and where possible, timed tasks were used to avoid issues of ceiling effects. The standardised assessments used (e.g. the TEA-Ch) also reported moderate to high test-retest reliabilities for EF tasks and the correlations performed (section 5.3.6.) also found moderate to high test-retest reliabilities for all EF tasks, except for the colour-word Stroop task. The methodological issue of task impurity should not be overlooked and may complicate findings as there is no guarantee that participants performing a task proposed to tap a specific function (e.g. IC) will not utilise functions beyond the one of interest (Best & Miller. 2010; Lehto et al., 2003; Van der Ven, 2011). However, findings indicated that unique EF tasks were assessed as LG and SES effects were found on certain but not all EF tasks and correlation coefficients differed across EF tasks. Furthermore, Van der Ven (2011) noted the lack of longitudinal studies examining the developmental nature of EF skills and those that have prove inconclusive. One method suggested to overcome issues of task impurity is by using a combination of EF tasks, to tap a number of EFs, as was done here.

- Certain measures in this study were only assessed at single time points e.g. PPVT, WM, WCST (see 4.4.) therefore the developmental nature of these tasks and their associated EF skills (e.g. WM) could not be assessed. However, care was taken prior to testing and due to the high level of novelty required in the WCST it has not been recommended that this task be used in longitudinal studies; therefore the TMT was employed to assess unified EF skills over time. Furthermore, it must be noted that participants and their schools were voluntarily giving up their time to allow this study to take place. As a result, issues related to timing and feasibility meant that all tasks could not be performed at each time point (WM) as initially intended.

- Recent studies have suggested that language proficiency (LP) may play an important role in cognitive outcomes related to bilingualism (Kroll & Bialystok, 2013; Kroll & Dussias, 2013; Kross, Dussias et al., 2012; Tao et al., 2011) and may even confound certain results. For instance, Cummins' threshold hypothesis (TH; 1979) suggested that a minimum level of L2 proficiency is needed if bilinguals are to gain any positive cognitive outcomes through their language experience. As LP was not assessed as part of this study it is not known whether bilingual effects could have been mediated or even masked by issues of LP. As a result, the following chapter (Chapter 6) examines issues of LP and language experience further by assessing within-group differences (IE group) as well as comparing IE results with children from Gaeltacht or Irish speaking regions of Ireland (section 1.4.3) who had increased exposure to the L2 outside of the school environment.

CHAPTER SIX
THE ROLE OF LANGUAGE PROFICIENCY IN CHILDREN'S COGNITIVE DEVELOPMENT

6.1. INTRODUCTION

This chapter describes the final study, which aimed to investigate the effect of language proficiency (LP) on children's executive function (EF) performance. The chapter begins with an introduction to issues of LP and methods of LP assessment. Results examined LP in four ways: the first analysis compared the effect of LP on bilinguals within immersion education (IE) by categorising children into high, moderate and low proficiency groups to compare EF skills. The second set of analysis used the same EF tasks to compare monolingual, IE and high proficiency or native-speaking bilinguals from Gaeltacht regions of Ireland (see section 1.4.3) to examine whether the native groups performed better on EF tasks as a result of their higher levels of L2 exposure. A third set of analysis compared the bilingual groups' (immersion and native) performance on English and Irish versions of the inhibitory control (IC) EF tasks to examine whether language of testing impacted on EF performance. Finally, attitudes towards the L2 were explored using thematic analysis and participant self-rating questionnaires. Using these analyses the aim of this study was to gain a more in depth understanding of the impact of L2 proficiency and exposure on children's cognitive skills.

Given that LG effects were found for certain but not all EF tasks in the previous chapter, might the children's levels of L2 proficiency have masked potential bilingual effects? Many have argued that bilingualism should be considered a continuous rather than a categorical variable and recently LP has been considered as a potential mediator of the cognitive effects of bilingualism (e.g. Colzato et al., 2008; Butler & Hakuta, 2004; Kroll & Bialystok, 2013). Furthermore, a number of researchers have argued that the level of bilingualism must be sufficiently high to confer detectable advantages in EF tasks (Bialystok & Majumder, 1998; Carlson & Meltzoff, 2008; Nicolay & Poncelet, 2013a; Ricciardelli, 1992). In a study by Barac and Bialystok (2012), length of time within IE was shown to be directly related to children's EF performance, with performance improving as a function of L2 exposure. Similarly, Tao and colleagues (2011) found that, in a complex task of EF (lateralized attention network test), late bilinguals (L2 acquired after 16 years) showed superior performance on measures of cognitive conflict while early bilinguals showed enhanced monitoring processes and overall reaction time speeds. They concluded that different language experiences may influence mechanisms of cognitive control in different ways despite both bilingual groups' outperforming monolingual controls.

The first aim of the study was to examine effects of LP on EF performance using proficiency ratings from parent, teacher and child questionnaires (see Appendices I, II, and III). Using these ratings IE children were grouped into *low*, *moderate* or *high proficiency* groups in the L2 and were was compared to examine within-group performance differences on EF tasks. It is noted that to prevent any experimenter bias, questionnaires were only read and children ranked into proficiency groups following the completion of data collection.

The second aim of this study was to compare IE bilingual participants at Time 3 of the longitudinal study (Chapter 5) with children from Irish-speaking (Gaeltacht) areas of the Republic of Ireland. As most of the Gaeltacht children are raised with some degree of Irish exposure from birth (see 1.4.3. for details), they are presumed to have higher levels of Irish proficiency and L2 exposure than children within Irish-medium IE programmes. Gaeltacht children were labelled as '*native bilinguals*' and also attended Irish IE schools. Therefore, most were expected to have Irish input from school and home. As the effect of socioeconomic status (SES) was also being investigated as part of the previous longitudinal study, two schools were recruited to make up the native bilingual group: one designated *Deis* or low-SES school (see Section 4.3.2.) and one typical or mid-SES school. Children were assessed using the same battery of EF tasks used in the previous longitudinal study (see 4.5.).

6.1.1. The influence of language proficiency on cognitive outcome measures

The level of proficiency a bilingual child has in their first (L1) and second (L2) language may play an important role in how their bilingualism impacts on cognitive development (e.g. Barac & Bialystok, 2012; Butler & Hakuta, 2004; Cummins, 1978a, 1978b, 1979, 2000; Costa et al., 2009; Kroll & Bialystok, 2013; Tao et al., 2011). Peal and Lambert (1962; see section 1.3.2) recognised the importance of LP and were some of the first to include it as a methodological consideration in their research, subsequently paving the way for and changing the outcomes of future studies of bilingualism. Following in the footsteps of Peal and Lambert, a number of researchers including Hakuta (e.g. Butler & Hakuta, 2004; Hakuta & Diaz, 1985; Hakuta & Garcia, 1989) and Cummins (1976, 1978a, 1978b, 1979, 2000) placed great emphasis on adequately controlling bilingual LP in the L1 and L2 and have criticised those who conduct research using minority language or low L2 proficiency groups. Despite these notes of caution, researchers have noted the continuing lack of emphasis placed on LP in recent studies of bilingualism, which generally categorise participants into either monolingual or bilingual groups (e.g. Hilchey & Klein, 2011, Kroll & Bialystok, 2013; Paap & Greenberg, 2013). Bialystok and Luk (2013) and Colzato et al. (2008) discussed how bilingualism is a continuous variable and that individuals vary in their levels of exposure to the L1 and L2. As a result, contradictory findings may be mediated by issues of LP, such as older but not younger adults demonstrating bilingual effects (e.g. Gold et al., 2013) and different language contexts producing contrasting results (e.g. Linck, Hoshino, & Kroll, 2008). Furthermore, a number of studies have reported that the EF advantages shown for bilinguals may be mediated by LP (e.g. Festman, Rodriguez-Fornells, & Munte, 2010; Singh & Mishra, 2012; Tao et al., 2011). These issues are summarised by Kroll and Bialystok (2013, pp. 5) who argued that:

"…bilingualism needs to be studied in the context of a dynamically changing system of linguistic and cognitive performance, an approach that extends beyond categorical assignment to groups".

Taking these issues into account, this study aimed to explore levels of L2 LP on participants' EF performance as well as examining the impact of different linguistic environments (monolingual, immersion, native) on children's overall cognitive performance. Studying the role of proficiency is all the more important in this study as IE has been shown to elicit varying levels of LP (see section 1.4.2.). '*Additive immersion'* (Cummins, 1998) programmes aim to develop children's L2 language skills at no cost to their L1 and are the programmes utilised in Ireland and within this study. Crucially, additive immersion programmes have been shown to lead to significantly more advanced levels of L2 proficiency than conventional programmes of L2 instruction or L2 learning programmes that are restricted to limited periods of time (Genesee, 2004).

Cummins (1976, 1978a, 1978b, 1983, 1991, 1996) examined effects of IE and found that differing IE environments elicited varying levels of first and second LP. His '*threshold hypothesis'* (TH) postulated that for a child to develop adequate proficiency in the L2, L1 skills must first be well developed to ensure a child's academic success (Cummins, 1976, 1978a; 1978b). Therefore additive IE rather than submersion education is the ideal in terms of developing children's L1 and L2 skills as children's L1 (English) is recognised and promoted alongside the L2 (Irish). This TH later predicted that there is also a L2 threshold necessary if children's bilingualism is to enhance their cognitive and academic performance (Bournot-Trites & Tellowitz, 2002; Cummins, 1979). The TH predicts that low levels of competence in both the L1 and L2 or 'semilingualism' will result in negative cognitive and academic outcomes while high levels of proficiency in both languages or 'additive bilingualism' will result in positive cognitive effects (Cummins, 1979; Toukomaa & Skutnabb-Kangas, 1977). No defined threshold is provided by Cummins but he does argue that the threshold can vary depending on the child's stage of development.

LP has generally been assessed through measures of speaking, listening, reading and writing although methods of assessment can vary across settings (Genesee, 2004). Research has shown that, following IE, children's L2 receptive skills (e.g. listening, comprehension) may be more advanced than L2 productive skills (e.g. speaking, writing; e.g. Cummins, 2001; Lazaruk, 2007). For instance, Genesee (1987) found that children in early French IE programmes developed native-like receptive language skills at around 11 years but that they made linguistic errors in their phonology, vocabulary and grammar. Furthermore, Swain (1996) identified speaking as the weakest skill area for students following IE. However, L1 skills in majority language children are developed equally as well within IE programmes to peers within monolingual programmes (Cummins, 2001). Although this equivalence in L1 skills following IE has widely been cited and used as an argument for the benefits of IE programmes, immersion children often lag behind in their L1 literacy skills during the initial "*early total immersion*" stages where emphasis on the L2 is in place. However, this literacy lag has been shown to decrease after at least one year of L1 instruction, with the development of L1 skills following a typical trajectory (Genesee, 1987, 2004).

It is evident that LP is an important consideration for bilingual studies of linguistic and cognitive development. Therefore, this study aimed to examine the effects of LP within immersion and native bilingual samples. It is also clear that children's L1 (English) development should not be hindered as a result of being taught through their L2 (Irish) and that EF performance should not be delayed as a result of learning through the L2 (Cummins, 1998), a result confirmed from in the previous chapter. LG differences in EF performance is examined further in this study which compares children growing up within Irish-speaking, Gaeltacht regions of Ireland.

6.1.2. Measuring language proficiency

The previous section highlighted how LP levels are an important consideration when exploring the relationship between language and cognitive development. Despite its apparent importance, there have been large discrepancies in how researchers measure L1 and L2 LP. For instance, Butler and Hakuta (2004) described a study by Brunner (1929) where children's proficiencies were categorised using the foreignness of their parents. Children whose parents were both born abroad were assumed to have the highest level of language competence compared with children whose parents were both in their country of residence. Such poor grouping methods create difficulties when interpreting bilingual findings prior to 1962 (Peal and Lambert's study) as LG studies may have been comparing monolinguals with L1 minority learners. Furthermore, although Cummins' TH (1979, 2000) has been widely cited and embraced by bilingual researchers and educationalists alike, it has been criticised for its lack of clarity and arbitrary nature regarding the within-groups L1 and L2 'thresholds' or levels necessary for categorising LGs (e.g. Takakuwa, 2005).

Recent studies of LP have used self-rating, language background questionnaires administered to parents, teachers and participants themselves as measures of language usage and proficiency (e.g. Barac & Bialystok, 2012; Sing & Mishra, 2012; Tao et al., 2011). The type of information gathered usually relates to the age of L2 acquisition, language usage while growing up, current language use, language preference and method of L2 acquisition (e.g. Rosselli et al., 2012). Rosselli et al., (2000; Roselli et al., 2002) also included five-point rating scales for participants to indicate how well they spoke and understood each of their two languages and correlations between self-ratings and language background questionnaires have been reported as high (Hakuta, Bialystok, & Wiley, 2003). Although the term, “balanced bilingual” has been cited as the gold-standard in terms of comparison bilingual groups (Baker, 2011), there is still no accepted definition or assessment for classifying bilinguals as balanced in their L1 and L2 (Ardila, 2007; Baker, 2011). Thomas and Roberts (2011) noted while assessing LP in children (within IE) that the amount of L2 exposure in the classroom is not an adequate measure of LP as children spend most of their day interacting with other children in social scenarios. As a result, children’s social use of language is just as important for their linguistic achievements yet is underestimated and rarely assessed in many studies of LP with children. Their LP questionnaires were administered to children only, between 8 and 11 years of age and included items assessing language usage within the home, community, while in school with friends and teachers, as well as children’s attitudes towards their L1 and L2. Their findings suggested that for children who spoke English within the home, the tendency to use English (L1) over Welsh (L2) within the IE environment was higher than for children who spoke Welsh within the home, and that peer-peer classroom interactions included high levels of English use. They concluded that their statistics were worrying in light of current movements by educationalists and policy makers to encourage the growth of L2 proficiency through IE programmes (Thomas & Roberts, 2011). They also noted that despite Welsh being a strong language input during the early academic years, it does not guarantee uptake of the language, particularly when that language is in the minority (Gathercole & Thomas, 2009).

As this study aimed to assess the effects of children’s LP on measures of executive function (EF), language background and proficiency questionnaires were administered to immersion and native bilingual participants (Appendix III), their parents (Appendix I) and their teachers (Appendix II). Background items requested information such as: language(s) spoken within the home, years within IE, years within English (monolingual) education, experience with L1 and L2, languages spoken with parents, siblings, other family members, within the community, friends in and outside of school. Proficiency rating scales (mark with an 'X' along 10cm line) were also used as measures of English and Irish receptive and productive language skills. Finally, attitudes towards the L1 and L2 were assessed by asking participants how they felt about their L1 and L2 as well as a final, open-ended question regarding what they liked/didn't like about speaking in their L2. Instructions for children's questionnaires were provided in English and Irish and delivered in an age-appropriate manner.

6.1.3. Attitudes towards Irish

Section 1.4.3. highlighted some of the issues facing the Irish Gaeltacht in particular as a result of changing attitudes towards the Irish language and section 1.2.3.5. outlined the importance of attitudes towards the L2 if bilingualism is to be fostered. Kennedy (2012) found that language experiences rather than attitudes towards the Irish language affected children's EF performance on a range of tasks. The attitudes of children in this study towards their L2 will be explored further in this study to investigate whether children in the urban IE and rural Gaeltacht differed in their attitudes towards the L2 (see section 6.4.5.).

6.1.4. Research questions/hypothesis

As the first set of analysis grouped IE participants into low, moderate and high proficiency groups, it was expected that high proficiency participants would outperform moderate and low proficiency groups on measures of EF skill as a result of having more language balance with their L1 and L2.

The second set of analysis compared monolingual, IE bilingual and native bilingual groups on measures of EF skills. As native bilingual groups had increased exposure and experience with the L2 it was predicted that they would outperform monolinguals on certain EF skills such as the unified EF function. However, as the previous chapter found a limited bilingual effect for specific EF skills such as IC and WM, it was not known how the Gaeltacht groups would compare on these tasks; therefore no definitive predictions were made for these assessments.

6.2. METHOD

6.2.1. Participants

In order to investigate the effects of LP, Time 3 data from the longitudinal study was compared with those of children tested in a Gaeltacht region of Ireland. Exclusion criteria were the same as in Chapter 5 (see 4.3.1.). A total of 19 participants were recruited from Gaeltacht areas at Time 3 and were labelled as *native bilinguals*. Following exclusions, 18 participants were included for analysis alongside the 112 participants at Time 3 of the longitudinal study. As in the previous chapter, all bilingual children were educated through a full Irish-medium IE model. The mean age of participants in the native bilingual group was 11 years and 7 months (*SD* = 4.3months; range 10 years 10 months – 12 years 1 month). The sample was made up of 58% males and 42% females. The Time 3 group demographics for all other groups can be found in section 5.2.1.
The native bilingual group was recruited from two schools. One of these schools had been classified as Deis or disadvantaged (*n* = 11) in order to compare results with low-SES groups from the previous study. The second school (*n* = 8) was a typical primary school, comparable to mid-SES groups. Individual SES demographics within the native groups was also explored (Table 36). It is noted that participant numbers in the native bilingual groups were small and may not have been the result of the decreasing numbers of children educated within Gaeltacht areas and the small class sizes within Gaeltacht schools (see section 1.4.3).

6.2.1.1. SES demographics

As in the longitudinal study (see Chapter 5), individual SES demographics were explored within the native bilingual schools using the ISCO-08 occupation ratings and the ISCED-97 highest years' in education ratings (see 5.1.2.). The aggregate measure of school designation to the *Deis* scheme was also used. Both schools were rural and situated in the Co. Donegal Gaeltacht area of the Republic of Ireland. Response rates for the background questionnaires within the non-Deis school

(School 2) was higher at 100% compared to the Deis school (School 1) at 58%. Therefore, lower response rates in School 1 may have biased individual SES results within this group.

Table 36 displays the descriptive statistics for the native group's highest achieving parent/caregiver's average years in education and percentage of individuals within each occupation code. There was no significant difference in parents' highest average years in education, t (12) = .34, p = .75, d = .40 between School 1 and School 2. Results showed that, on average, the highest level of education for parents in both schools was 14 years and had achieved the Leaving Certificate (A-Level) qualification as their highest level of education. Chi-squared results indicated no significant association between Schools and occupation within the native bilingual group, χ^2 (5, N = 14) = 4.20, p = .52. Although groups did not differ significantly in their individual SES demographics as School 1 had been assigned as a disadvantaged or *Deis* by the Irish government, this group is described as *low-SES native bilingual* (LSN) and School 2 is described as *mid-SES native bilingual* (MSN). Immersion bilingual children within the longitudinal study were re-labelled here as low-SES immersion bilinguals (LSI) and mid-SES immersion bilinguals (MSI) in order to avoid confusion. Monolingual terminology remained the same: low-SES monolinguals (LSM) and mid-SES monolinguals (MSM).

Table 36 *Socioeconomic status demographics within the native bilingual groups*

		Native Bilingual Groups		
		Deis School 1 *n* = 12	School 2 *n* = 7	Total *n* = 19
Response Rate		58%	100%	74%
Mean Years in Education		14years	14years	13.5years
(SD)		(3.13)	(1.25)	(2.45)
Occupation Code				
	Occupation Category			
1	Managers	-	-	-
2	Professionals	29%	43%	36%
3	Technicians and associate professionals	14%	14%	14%
4	Clerical support workers	-	14%	7%
5	Service and sales workers	-	-	-
6	Skilled agricultural, forestry and fishery workers	-	14%	7%
7	Craft and related trades workers	-	14%	7%
8	Plant machine operators, and assemblers	14%	-	16%
9	Elementary occupations	-	-	16%
10	Unemployed	43%	14%	29%

6.2.1.2. Language demographics

To assess each child's language background and experience, parental/caregiver (see Appendix I) and children's background and proficiency questionnaires (see Appendix III) were analysed. LP of monolingual children was not assessed as they were assumed to have a high level of proficiency in their L1.

Table 37 shows descriptive statistics for the number of years bilingual children had lived in Ireland, their country of birth, experience with Irish and English and years within immersion and/or monolingual education. The level of Irish and English spoken within and outside the home, with family and friends was also assessed. Parents and children rated the degree of ability and languages spoken by marking with an 'X' along a 10cm line where 0cm = *English only*, 5cm = *English and Irish equally* and 10cm = *Irish only*. Ratings were made of items: language spoken with each parent/caregiver (parent 1/parent 2), with brothers or sisters, friends, other family members, and when involved in community activities (e.g. sports/social clubs). There was a high correlation ($p < .01$) between all parent/caregiver and child ratings on all language background items therefore, for all items, the parent/caregiver ratings were used to compare groups. Two items, the '*language(s) spoken with friends in school'*, r (76) = 04; p = .71, and '*language(s) spoken with friends and home'*, r (76) = 20; p = .08, did not correlate between parents and children. As it was felt that children would provide a more accurate rating of the language spoken with friends and in line with Thomas and Roberts (2011), children's ratings were used to compare these items. A further two items were asked to the children only: what language(s) do they *think in at home* and *at school*.
A comparison of language demographics and experiences between immersion and native bilinguals will be carried out in section 6.4.1.

Table 37 *Means (and standard deviations) for language demographic items across bilingual groups*

	Low-SES Group		Mid-SES Group	
	Native *n* = 9	Bilingual *n* = 12	Native *n* = 7	Bilingual *n* = 41
Response Rate	78%	67%	100%	91%
Country Born	Ireland 83% USA 11% - Scotland 6%	Ireland 92% USA 8% - -	Ireland 86% USA 14% - -	Ireland 98% - China 2% -
Years in Ireland	10.5 (1.34)	11.12 (.39)	11.0 (.82)	10.72 (.57)
Years in Immersion	7.17 (1.95)	8.08 (1.38)	7.57 (.79)	7.54 (.82)
Years in English Education	0 -	.58 (1.24)	0 -	.82 (1.27)
Experience with Irish (Years)	9.56 (1.62)	8.75 (1.96)	10.29 (1.70)	7.84 (1.45)
Experience with English (Years)	10.72 (1.27)	11.17 (.39)	10.14 (1.95)	10.72 (1.27)
Parent 1 home language	*More English than Irish* 1.94 (2.15)	*Mostly English* 1.45 (1.73)	*More Irish than English* 6.46 (2.20)	*Mostly English* 1.17 (1.43)
Parent 2 home language	*More English than Irish* 2.03 (2.03)	*Mostly English* 1.19 (1.86)	*More Irish than English* 6.36 (2.86)	*Almost English only* .69 (1.12)
Language (s) spoken with Brothers/Sisters	*Mostly English* 1.34 (1.72)	*Mostly English* 1.15 (1.64)	*Irish/English equally* 5.46 (3.62)	*Mostly English* 1.52 (1.80)
Other family members	*More English than Irish* 2.40 (2.28)	*Almost English only* .36 (.47)	*Irish/English equally* 5.64 (3.44)	*Almost English only* .84 (1.80)
Friends in school	*Irish/English equally* 5.88 (3.03)	*Mostly Irish* 8.57 (1.53)	*Mostly Irish* 7.70 (1.88)	*Almost Irish only* 9.51 (1.21)
Friends outside of school	*Almost English only* .57 (.98)	*More English than Irish* 2.26 (2.19)	*Irish/English equally* 5.06 (2.86)	*More English than Irish* 2.39 (3.06)
Social activities outside of school	*Mostly English* 1.84 (2.12)	*Mostly English* 1.36 (2.83)	*Mostly Irish* 6.37 (2.68)	*Almost English only* 1.02 (1.44)
Language(s) of thought at home	*More English than Irish* 1.90 (2.44)	*Mostly English* 1.14 (1.83)	*More Irish than English* 5.99 (2.90)	*Mostly English* 2.01 (2.91)
Language(s) of thought in school	*Irish/English almost equally* 4.88 (2.98)	*Irish/English almost equally* 5.71 (3.09)	*Irish/English almost equally* 5.43 (3.80)	*Irish/English almost equally* 5.28 (3.54)

6.2.1.3. Proficiency ratings

To evaluate immersion and bilingual groups' English and Irish proficiency skills, parent/caregiver, teacher and children's responses on proficiency questionnaire items (Appendix I, Appendix II, and Appendix II) were compared and analysed at Time 3 of testing. As with the language demographic items, participants' language skills were rated along a 10cm line where 0cm = *no ability at all* and 10cm = *exceptionally high ability*. The sum of items assessing receptive and productive language skills in either language (English and Irish) were collated for comparison. Receptive skills were the sum of three items (understanding written Irish/English, understanding spoken English/Irish, ability to read in English/Irish) with scores ranging from 0-30 (0 = *no ability at all*; 30 = *exceptionally high ability*) while productive language skills were the sum of two items (ability to speak in Irish/English, ability to write in English/Irish), ranging from 0-20 (0 = *no ability at all*; 20 = *exceptionally high ability*). As receptive skills were marked out of 30 and productive skills were marked out of 20, percentage proficiency (%) was calculated for each skill set (receptive and productive) so that scores could more easily be compared. Therefore, proficiency ratings for each language skill ranged from 0 – 100 (%) where 0% = *no ability at all*, 50% = *moderate ability* and 100% = *exceptionally high ability* (see Table 38).

Descriptive statistics for parent/caregiver, teacher and children's % proficiency ratings are displayed in Table 39 and group comparisons are made in section 6.4.1.

Table 38 *Method of collating percentage proficiency language skills*

English Receptive Skill	Irish Receptive Skill	English Productive Skill	Irish Productive Skills
Ability to read in English + Understanding of spoken English + Understanding of written English = Sum divided by 30 and multiplied by 100	Ability to read in Irish + Understanding of spoken Irish + Understanding of written Irish = Sum divided by 30 and multiplied by 100	Ability to write in English + Ability to speak in English = Sum divided by 20 and multiplied by 100	Ability to write in Irish + Ability to speak in Irish = Sum divided by 20 and multiplied by 100

Table 39 *Means (and standard deviations) for parent/caregiver, teacher and child % ratings of language proficiency*

		Low-SES Group		Mid-SES Group		Total
Language Skill		Native	Immersion	Native	Immersion	
English Receptive Skills	Parent	89% (.05) *n* = 8	93% (.08) *n* = 12	90% (.05) *n* = 7	89% (1.0) *n* = 40	90% (.09) *n* = 67
	Teacher	87% (.15) *n* = 12	89% (.11) *n* = 17	80% (.11) *n* = 7	75% (.17) *n* = 42	80% (.16) *n* = 78
	Child	90% (.12) *n* = 11	81% (.15) *n* = 18	74% (.14) *n* = 7	83% (.14) *n* = 43	83% (.14) *n* = 79
Irish Receptive Skills	Parent	82% (.11)	83% (.14)	86% (.09)	84% (.10)	84% (.11)
	Teacher	80% (.21)	81% (.14)	75% (.17)	75% (.17)	77% (.17)
	Child	76% (.11)	68% (.21)	69% (.14)	66% (.17)	68% (.17)
English Productive Skills	Parent	88% (.08)	92% (.08)	87% (.13)	87% (.10)	88% (1.0)
	Teacher	74% (.23)	78% (.20)	79% (.13)	72% (.18)	74% (.19)
	Child	87% (.11)	81% (.16)	74% (.14)	83% (.13)	82% (.13)
Irish Productive Skills	Parent	76% (.16)	82% (.14)	87% (.07)	83% (.10)	82% (.12)
	Teacher	60% (.26)	69% (.20)	78% (.16)	71% (.17)	69% (.19)
	Child	86% (.14)	70% (.22)	87% (.14)	75% (.24)	76% (.22)

Relationships between ratings

To examine the relationships between parent/caregiver and teacher ratings of participants' language skills, a number of Pearson's correlation tests were run. Table 40 shows that there was a strong, positive, statistically significant correlation between parent and teacher ratings of children's language skills. Scatterplots show the relationship between parent and teacher ratings for Irish receptive (Figures 41) and productive (Figure 42) skills.

Table 40 *Relationship between parent/caregiver and teacher ratings of language proficiency*

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. English Receptive Skill (Parent)	1	.45**	.68**	.37**	.84**	.42**	.55**	.30**
2. English Receptive Skill (Teacher)	.45**	1	.23*	.89**	.36**	.79**	.13	.63**
3. Irish Receptive Skill (Parent)	.68**	.23**	1	.27**	.49**	.34**	.77**	.35**
4. Irish Receptive Skill (Teacher)	.37**	.89**	.27**	1	.23*	.82**	.19	.79**
5. English Productive Skill (Parent)	.84**	.36**	.49**	.23*	1	.33**	.49**	.20
6. English Productive Skill (Teacher)	.42**	.79**	.34**	.82**	.33**	1	.28*	.88**
7. Irish Productive Skill (Parent)	.55**	.13	.77**	.19	.49**	.28**	1	.44**
8. Irish Productive Skill (Teacher)	.30**	.63**	.35**	.79**	.20	.88**	.44**	1

*Note: * p < .05, ** p < .001 (2-tailed)*

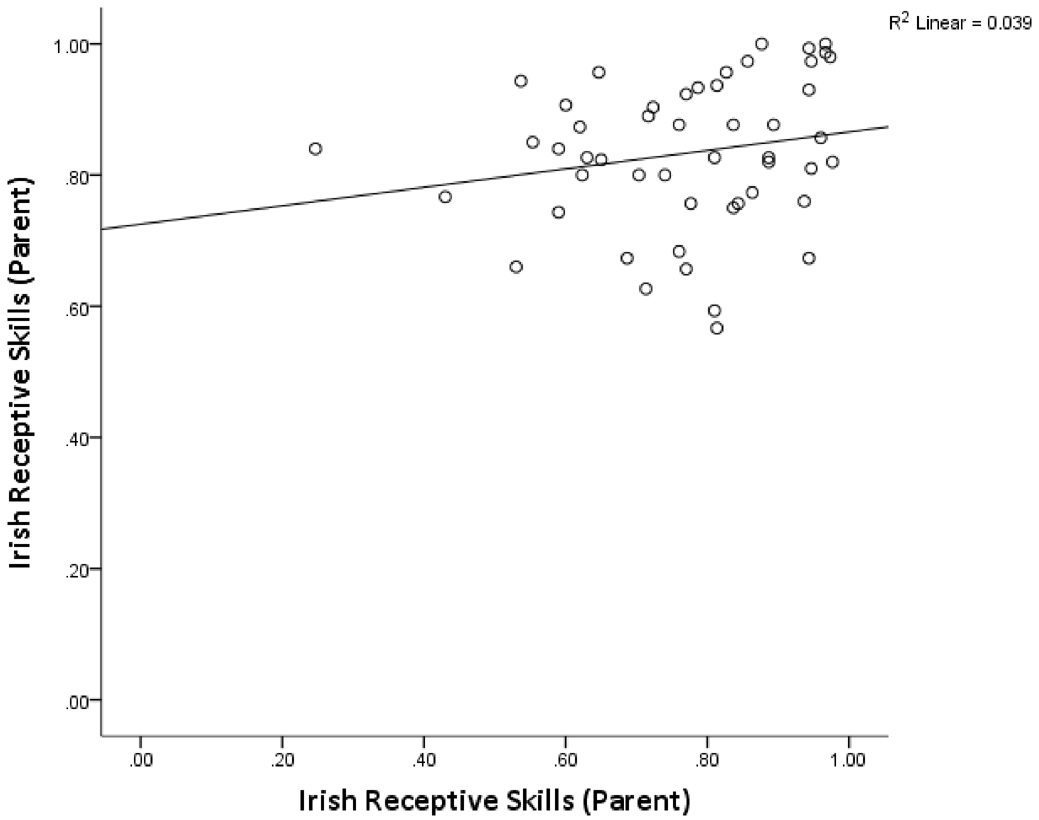


Figure 41 *Parent/caregiver and teacher ratings of Irish receptive language skills*

Figure 42 *Parents/caregiver and teacher ratings of Irish productive language skills*

As there was a high correlation between parent and teacher ratings (see Table 40), parent ratings only were used to represent adult ratings as it was felt that parents would provide a more accurate representation of their child's language abilities overall (Gutiérrez-Clellan & Kreiter, 2003). Table 41 shows a positive, statistically significant correlation between some, but not all of the parent and child ratings of language skills. There was no correlation between parent and child ratings of *Irish receptive skills*, r (52) = .08, p = .56, *English productive skills*, r (53) = - .02, p = 91, or *Irish productive skills*, r (52) = .24, p = .08. The means from Table 39 show that for these skills, parents rated their children as having higher language abilities than children rated themselves. The relationship between parent and child ratings of Irish language skills are shown in Figures 43 and 44. Although there were no significant correlations between certain items, parent/caregivers and children still rated their skills in Irish receptive and productive skills as high (ranging from 66% - 86%).

Table 41 *Relationship between parent and child rating of language proficiency*

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. English Receptive Skill (Parent)	1	.29*	.71**	.14	.91**	.02	.71**	.18
2. English Receptive Skill (Child)	.29*	1	.21	.63**	.25	.58**	.23	.27*
3. Irish Receptive Skill (Parent)	.71**	.21	1	.08	.59**	-.04	.87**	.35*
4. Irish Receptive Skill (Child)	.14	.63**	.08	1	.13	.64**	.16	.46**
5. English Productive Skill (Parent)	.91**	.25	.59**	.13	1	-.02	.73**	.06
6. English Productive Skill (Child)	.02	.58**	-.04	.64**	-.02	1	-.04	.35**
7. Irish Productive Skill (Parent)	.71**	.23	.87**	.16	.73**	-.04	1	.24
8. Irish Productive Skill (Child)	.18	.27*	.35*	.46**	.06	.35**	.24	1

*Note: * $p < .05$, ** $p < .001$ (2-tailed)*

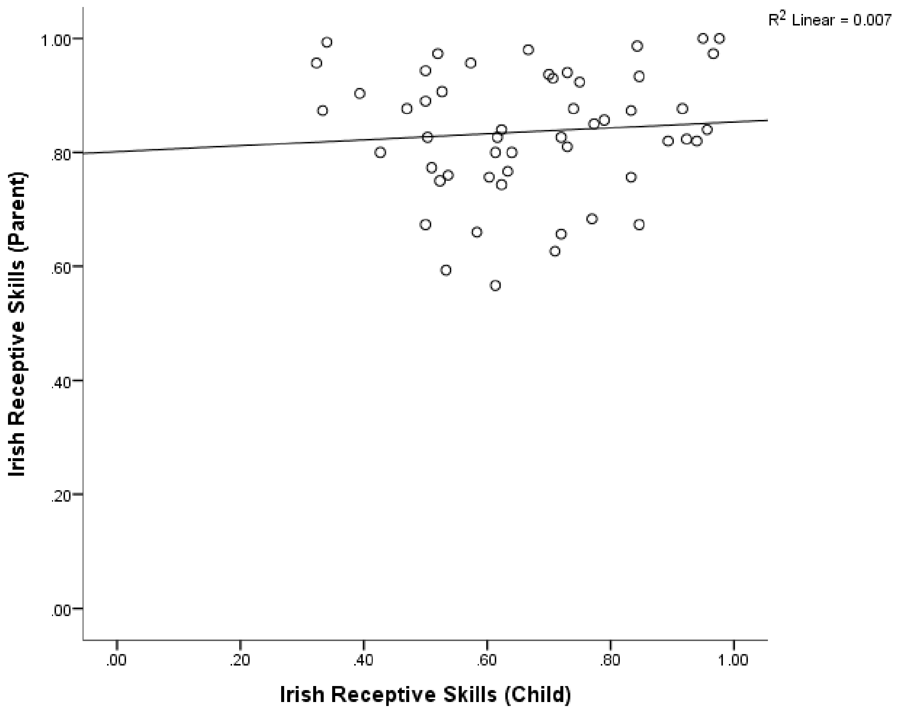


Figure 43 *Parent and child ratings of Irish receptive language skills*

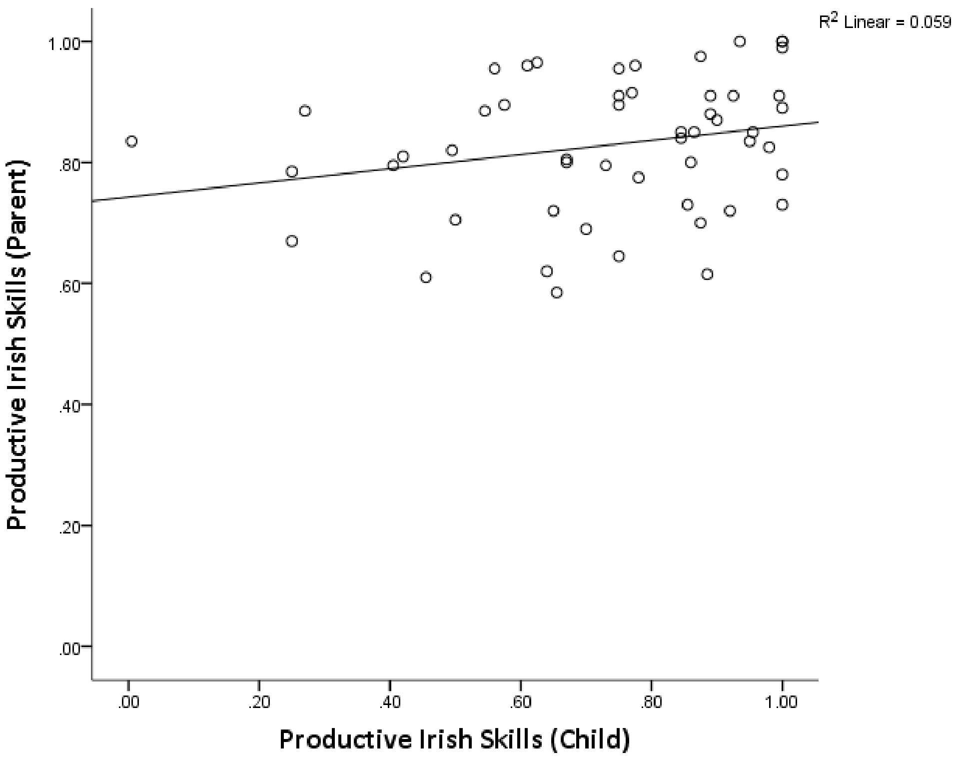


Figure 44 *Parent and child ratings of Irish productive language skills*

6.2.2. Apparatus and procedure

For a list of tasks and a description of the apparatus and procedures used in this study see Chapter 4.

6.2.3. Analysis

As this study compared results from Time 3 of the longitudinal study with cross-sectional data from two native bilingual groups, the analysis used was between-groups ANOVA (see below for layout of results section).

Section 6.3.1. This section examined group differences in the language demographic and proficiency items from parent/caregiver, teacher and children's proficiency questionnaires (sections 6.2.1.2 and 6.2.1.3.). This section also looked at the relationships between language demographic and proficiency questionnaire items within each language group. Descriptive statistics for these items are presented in Tables 37 - 39.

Section 6.3.2. To explore the effects of LP within the immersion bilingual groups, this section compared the EF performance of participants classified as low, moderate of high proficiency with the aim of assessing LP groups differed in their EF performance.

Section 6.3.3. To explore issues of LP and experience further this section compared the EF performance of monolingual, immersion and native bilinguals. Native bilinguals were raised and educated within Gaeltacht or Irish-speaking pockets of Ireland and therefore had higher levels of exposure to Irish than the immersion bilingual group (see section 1.4.3.). This unique groups' performance was compared with participants tested at Time 3 of the longitudinal study.

Section 6.3.4. Immersion and native bilingual groups performed the inhibitory control tasks (Opposite Worlds and colour-Word Stroop) in English and Irish. This section compared results from the English and Irish versions to assess whether language of testing affected children's EF performance.

Section 6.3.5. The final results examined children's opinions of their experience with English and Irish (L1 and L2) to explore how IE has affected children's attitudes towards their L2 (see section 1.2.3.5.).

6.3. RESULTS

6.3.1. Language demographics and proficiency results

Language Demographics

ANOVAs revealed a significant main effect of group for years in English medium education, F (3, 77) = 3.14, p = .03; η^2 = .11, and years' experience with the Irish language, F (3, 77) = 7.88, p < .01; η^2 = .24. Due to the unequal sample sizes, Hochberg's GT2 post hoc analysis was used to unpick main effects and showed a significant difference (p = .05) between the MSI and LSN in their years in monolingual education. There was also a significant difference between the MSI and both the LSN (p < .01) and MSN groups' (p < .01) in their years of experience with Irish. Language demographic means (see Table 36) indicated that the native bilingual groups had significantly more experience with the Irish language than immersion bilinguals as well as having spent no time in English medium education, unlike the immersion groups.

There was a significant main effect of group on language(s) spoken with parent/caregiver 1, F (3, 68) = 20.41, p < .01; η^2 = .49, parent caregiver 2, F (3, 64) = 19.12, p < .01; η^2 = .48, brothers/sisters, F (3, 65) = 8.48, p < .01; η^2 = .29, other family members, F (3, 68) = 14.19, p < .01; η^2 = .40, friends in school, F (3, 78) = 14.31, p < .01; η^2 = .36, friends outside of school, F (3, 78) = 4.04, p < .01; η^2 = .14, during social activities, F (3, 68) = 15.01, p < .01; η^2 = .41 and language of thought at home, F (3, 78) = 5.90, p < .01; η^2 = .19. No other main effects of group were found. Post hoc analysis for the items *language (s) spoken with parent* 1, *parent 2*, *brother/sisters*, *other family members*, *social activities* and *language(s) of thought at home* indicated that the MSN reported significantly higher (p < .01) levels of Irish language use than all other groups (LSI, MSI, LSN, see Table 36).
For the item *languages spoken with friends in school*, the LSN spoke significantly more English than both the LSI and MSI groups who both reported speaking mostly

Irish with friends during school. For the *language(s) spoken with friends at home* item, the MSN group spoke significantly more Irish than the LSN group. No other group differences were reported.

Proficiency Ratings

No significant main effect of group was found for Irish productive or receptive language skill ratings. However, mean % ratings showed (see Table 39) that for Irish receptive skills, parents rated the MSN group as having the highest levels of Irish receptive skills and the LSN as having the lowest level of skills. Teacher ratings indicated that the LSI group had the highest levels Irish receptive skills and the MSI and MSN with the lowest levels.

For Irish productive skill ratings parents, teachers and children all rated the MSN group as having the highest levels of Irish productive skills. However, parents and teachers rated the LSN as having the lowest levels of Irish productive skills. It is important to note that each group made ratings separately from each other and that scores were not standardised according to national statistics therefore ratings were unique within each group and caution must be taken when making comparisons between groups (see section 6.5.3.).

For English language skills, no main effect of group was found for receptive and productive skills apart from the teacher ratings of English receptive skills where a significant effect, F (3, 77) = 3.82, p = .01; η^2 = .13, and post hoc revealed a significant difference between LSI and MSI groups (p = .02) with the LSI group rated as having higher English receptive skills than the MSI (for means see Table 39).

Language Demographics and Proficiency Rating Correlations

To explore the relationships between language demographic items (Table 37) and proficiency ratings (Table 39), a set of Pearson's correlations were run to explore the relationships between items within each language group. Results are shown in Table 42 – 45.

Table 42 *Relationships between language demographic and proficiency items for the LSI group*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	12.	13.	14.
1. Years in IE	1	.78**	.29	.23	.30	.13	-.23	.34	.01	.03	.20	-.30	-.25
2. Years' experience with Irish	.78**	1	.15	-.00	.20	.13	.06	.24	-.05	.22	.36	.02	.00
3. Language spoken with parent 1	.29	.15	1	.90**	.23	.82**	-.30	-.30	-.31	-.46	-.30	-.68*	-.55
4. Language spoken with parent 2	.23	-.00	.90**	1	.21	74**	-.30	-.17	-.24	-.40	-.41	-.80	-.66*
5. Language spoken with siblings	.30	.20	.23	.21	1	.35	.29	-.09	.68*	.46	.54	.20	.32
6. Language spoken with other family members	.13	.13	.82**	.74**	.35	1	-.04	-.32	-.23	-.20	-.08	-.37	-.40
7. Language spoken w/ friends in school	-.23	.06	.-.30	-.30	.29	-.04	1	.32	.43	.50	.52	.65*	.72*
8. Language spoken w/ friends outside school	.34	.24	-.30	-.17	-.09	-.32	.32	1	.18	.21	.31	.24	.28
9. Language spoken in community	.01	-.05	-.31	-.24	.68*	-.23	.43	.18	1	.56	.54	.44	.52
10. Irish receptive skills	.03	.22	-.46	.40	.46	-.20	.50	.21	.56	1	.88*	.64*	.68*
11. Irish productive skills	.20	.36	-.30	-.41	.54	-.08	.52	.31	.54	.88*	1	.72*	.78**
12. English receptive skills	-.30	.02	-.68	-.80**	.20	-.37	.65*	.24	.44	.64*	.72*	1	.91**
13. English productive skills	-.25	.00	-.55	-.66*	.32	-.40	.72*	.28	.52	.68*	.78**	.91**	1

*Note: * p < .05, ** p < .01 (2-tailed)*

Table 43 *Relationships between language demographic and proficiency items for the MSI group*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Years in IE	1	.82**	.30	.06	.20	.20	-.10	.31*	.28	.12	.07	.02	-.10
2. Years' experience with Irish	.82**	1	.49*	.03	.42*	.32*	.01	.39*	-.39*	.27	.24	.18	.09
3. Language spoken with parent 1	.30	.49*	1	.20	.56*	.66*	.09	.03	.19	.27	.29	.23	.17
4. Language spoken with parent 2	.06	.03	.20	1	.01	-.0	-.31	-.00	.17	-.16	-.15	-.04	-.08
5. Language spoken with siblings	.20	.42**	.56**	.01	1	.38*	.06	.12	.16	.25	.20	.18	.10
6. Language spoken with other family members	.20	.32*	.66**	-.08	.38*	1	.15	.16	.17	.19	.29	.10	.18
7. Language spoken w/ friends in school	-.10	.01	.09	-.31	.06	.15	1	-.07	.23	.12	.14	.08	.11
8. Language spoken w/ friends outside school	.31*	.39*	.03	-.00	.12	.16	-.07	1	.14	-.09	-.10	-.18	-.10
9. Language spoken in community	.28	.39*	.19	.17	.16	.17	.23	.14	1	.36*	.31*	.26	.22
10. Irish receptive skills	.12	.27	.27	-.16	.25	.19	.12	-.09	.36*	1	.93**	.93**	.87**
11. Irish productive skills	.07	.24	.29	-.15	.20	.29	.14	-.10	.31*	.93**	1	.91**	.94**
12. English receptive skills	.02	.18	.23	-.04	.18	.10	.08	-.18	.26	.93**	.91**	1	.94**
13. English productive skills	-.10	.09	.17	-.08	.10	.18	.11	-.10	.22	.87**	.94**	.94**	1

*Note: * p < .05, ** p < .01 (2-tailed)*

Table 44 *Relationships between language demographic and proficiency items for the LSN group*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Years in IE	1	.55	.34	.26	.29	.39	.13	-.08	.26	.02	.37	-.04	.15
2. Years' experience with Irish	.55	1	.55	.52	.21	.74*	.40	.33	-.08	.15	.06	.13	-.17
3. Language spoken with parent 1	.34	.55	1	-.02	.69*	.76*	.00	.45	.37	.53	.45	.43	.47
4. Language spoken with parent 2	.26	.52	-.02	1	.16	.68	.27	-.28	-.08	.10	.15	-.11	-.15
5. Language spoken with siblings	.29	.21	.69*	.16	1	.53	-.46	-.30	.69*	.58	.30	.58	.70
6. Language spoken with other family members	.39	.74*	.76*	.68	.53	1	.23	.27	.16	.41	.43	.16	.15
7. Language spoken w/ friends in school	.13	.40	.00	.27	-.46	.23	1	.21	-.46	-.52	-.02	-.68	-.59
8. Language spoken w/ friends outside school	-.08	.33	.45	-.28	-.30	.27	.21	1	-.40	.19	.17	.14	-.02
9. Language spoken in community	.26	-.08	.37	-.08	.69*	.16	-.46	-.40	1	.64	.67	.59	.85**
10. Irish receptive skills	.02	.15	.53	.10	.58	.41	-.52	.19	.64	1	.65	.91**	.84**
11. Irish productive skills	.37	.06	.45	.15	.30	.43	-.02	.17	.67	.65	1	.35	.75*
12. English receptive skills	-.04	.13	.43	-.11	.58	.16	-.68	.14	.59	.91**	.35	1	.74*
13. English productive skills	.15	-.17	.47	-.15	.70	.15	-.59	-.02	.85**	.84**	.75*	.74*	1

*Note: * p < .05, ** p < .01 (2-tailed)*

Table 45 *Relationships between language demographic and proficiency items for the MSN group*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Years in IE	1	-.39	.41	-.62	-.61	-.10	-.36	-.70	-.05	-.20	-.37	.22	.59
2. Years' experience with Irish	-.39	1	.04	.60	.74	.84*	.34	.55	-.66	.53	.73	.45	-.00
3. Language spoken with parent 1	.41	.04	1	.32	.06	.26	-.44	-.49	-.15	.67	.32	.53	.21
4. Language spoken with parent 2	-.62	.60	.32	1	.88*	.59	-.08	.41	-.06	.75	.77*	.20	-.35
5. Language spoken with siblings	-.61	.74	.06	.88*	1	.75	-.11	.70	-.28	.66	.79*	.28	-.22
6. Language spoken with other family members	-.10	.84*	.26	.59	.75	1	-.19	.33	-.50	.70	.87*	.68	.38
7. Language spoken w/ friends in school	-.36	.34	-.44	-.08	-.11	-.19	1	.25	-.22	-.42	-.26	-.42	-.49
8. Language spoken w/ friends outside school	-.70	.55	-.49	.41	.70	.33	.25	1	-.44	.13	.29	-.04	-.43
9. Language spoken in community	-.05	-.66	-.15	-.06	-.28	-.50	-.22	-.44	1	-.36	-.29	-.61	-.15
10. Irish receptive skills	-.20	.53	.67	.75	.66	.70	-.42	.13	-.36	1	.87*	.77*	.20
11. Irish productive skills	-.37	.73	.32	.77*	.79*	.87*	-.26	.29	-.29	.87*	1	.69	.29
12. English receptive skills	.22	.45	.53	.20	.28	.68	-.42	-.04	-.61	.77*	.69	1	.70
13. English productive skills	.59	-.00	.21	-.35	-.22	.38	-.49	-.43	-.15	.20	.29	.70	1

*Note: * $p < .05$, ** $p < .01$ (2-tailed)*

6.3.2. Comparison of low, moderate and high proficiency immersion bilinguals

To examine the effect of LP on EF performance within the immersion bilingual group, participants were categorised according to their levels of Irish (L2) proficiency. The native bilingual groups (LSN and MSN) were not included in this within-groups analysis due to the complexity of their SES and language experience. Furthermore, results from the mean proficiency ratings (section 6.3.1.) revealed unusual discrepancies between native and immersion groups, therefore the LSN and MSN groups' data was compared in a separate analysis (section 6.3.3.). To categorise immersion bilinguals a '*proficiency*' variable was constructed based on their proficiency questionnaire ratings.

As productive language skills in the L2 are often used as an indicatory for IE success, children's Irish productive skills were used to categorise participants into LP groups. Furthermore, research has shown that IE programmes are often successful in developing children's receptive language skills but that children's productive skills may not reach the same level of competency (Genesee, 2004; Ó Duibhir, 2009).
To obtain proficiency ratings for each participant, the sum of parent and teacher % ratings (see Table 39) for Irish productive skills was divided by 2, giving each individual an overall Irish productive ability % rating. Using these % ability scores children were categorised as *low*, *moderate* or *high* proficiency. Participant % proficiency ratings in the bottom 20% of scores were grouped as low proficiency, participants with the top 20% of scores were grouped as high proficiency and participants with scores between 21-79% were grouped as moderate proficiency (see Table 46). The percentage of participants classified as low, moderate or high proficiency within each immersion group (LSI and MSI) is shown in Table 47.

Table 46 *Proficiency questionnaire rankings for grouping participants*

Language Skills Classification	
< 20%	*Low Proficiency*
21-79%	*Moderate Proficiency*
> 80%	*High Proficiency*

Table 47 *Within-immersion group percentages of immersion bilinguals classified as low, moderate or high proficient % Irish Productive Skill*

Proficiency Rating	Low-SES *n* = 11	Mid-SES *n* = 39	Total *n* = 50
Low	18% *n* = 2	23% *n* = 9	22% *n* = 11
Moderate	64% *n* = 7	53% *n* = 21	55% *n* = 28
High	18% *n* = 2	25% *n* = 10	24% *n* = 12

Table 47 shows that the MSI group had a higher percentage (23%) of participants classified as high proficiency than the LSI group (18%) although they also had a higher percentage of participants classified as low proficiency (23%) compared with the LSI group (18%). It is worth noting that sample sizes differed between the MSI (n = 29) and LSI (n = 11) group and chi squared results found no association between SES groups and proficiency classifications.

Within-group performance on executive function tasks

The following results compared low, moderate and high proficiency immersion bilinguals on the EF tasks employed at Time 3 of the longitudinal study and across SES groups.

Table 48 displays the ANOVA results and effect sizes for the variables in which a significant main effect of LP was found: PPVT, Creature Count Accuracy, Irish Opposite Worlds RTs, Irish Incongruent Stroop RTs, and Irish Stroop Inhibition RTs. No other main effects of LP group were found.

Post hoc analysis (Hochberg's GT2) revealed that for the PPVT, the high ($p < .01$) and moderate ($p = .05$) proficiency groups had significantly higher standardised scores than the low proficiency group. For the Creature Count Accuracy variable, the low-proficiency group had significantly higher age-scaled scores ($p = .05$) than the moderate-proficiency group.

For the Irish version of the Opposite Worlds task, the high-proficiency group had significantly faster Opposite Worlds (incongruent) reaction times (sec; $p = .01$) than the low-proficiency group.

Finally for the Irish version of the Colour-word Stroop task, the high-proficiency bilinguals had significantly faster incongruent reaction times (ms; $p = .01$) than the low-proficiency group. The high proficiency bilinguals also had reduced inhibition timing score (sec) than that moderate ($p = .02$) and low proficiency groups ($p = .01$).

Table 48 *Language proficiency group effects*

Task	Cognitive function: Method of Assessment	LP Group Effect	Effect size	Means
PPVT	English vocabulary: Standardised Score	{F (2, 50) = 5.81, p = .01}	$\eta^2 = .19$	Low = 90.82 (7.18) Moderate = 101.68 (14.95) High = 108.42 (13.64)
Creature Count Accuracy	Switching: Age-Scaled Score	{F (2, 50) = 3.19, p = .05}	$\eta^2 = .12$	Low = 11.73 (2.01) Moderate = 9.82 (2.48) High = 10.75 (1.36)
Irish Opposite Worlds	Inhibitory Control: Reaction Time (sec)	{F (2, 50) = 5.24, p = .01}	$\eta^2 = .18$	Low = 26.75sec (2.14) Moderate = 24.64sec (2.82) High = 23.01sec (3.14)
Irish Stroop Incongruent score	Cognitive Conflict Reaction Time (ms)	{F (2, 45) = 5.71, p = .01}	$\eta^2 = .21$	Low = 1067.88ms (131.41) Moderate = 935.41ms (154.93) High = 937.26ms (166.54)
Irish Stroop Inhibition score	Inhibitory Control Reaction Time (ms)	{F (2, 45) = 5.79, p = .01}	$\eta^2 = .21$	Low = 201.83ms (114.58) Moderate = 164.78ms (79.18) High = 76.75ms (92.15)

Socioeconomic group differences

To examine the LP effects within each SES group, a separate set of analyses was carried out within the low- and mid-SES immersion bilingual groups (LSI and MSI).

Within the LSI group, there was a significant main effect of LP on the standardised forward digit WM task, F (2, 10) = 4.73, p = .04; η^2 = .54. Figure 45 and post hoc analysis revealed that the moderate proficiency group scores significantly higher than the low-proficiency group (p = .05).
There was also a significant main effect LP on the Trails Difference response times (sec), F (2, 10) = 5.62, p = .03; η^2 = .58, with low-proficiency bilinguals having a significantly higher timing difference (p = .03) than the moderate-proficiency group (see Figure 46).

Within the MSI group, there was a significant main effect of LP on the PPVT standardised scores, F (2, 38) = 4.45, p = .02; η^2 = .19 with the high proficiency group scoring significantly higher (p = .02) than the low proficiency group. There was also a main effect of LP on the English Stroop neutral, F (2, 36) = 4.23, p = .02; η^2 = .20, congruent, F (2, 36) = 3.44, p = .04; η^2 = .17, and incongruent, F (2, 36) = 3.76, p = .03; η^2 = .19, trials. In all three conditions, the high proficiency group had significantly faster RTs than the low proficiency group (p = .02; p = .02; p = .02) (see Figure 46). Finally, there was a significant main effect of LP for the incongruent RT, F (2, 36) = 5.93, p = .01; η^2 = .26, and inhibition RT, F (2, 36) = 4.50, p = .02; η^2 = .21, conditions of the Irish Stroop task with the high proficiency group recording significantly faster RTs (p = .01; p = .02) than the low proficiency group.

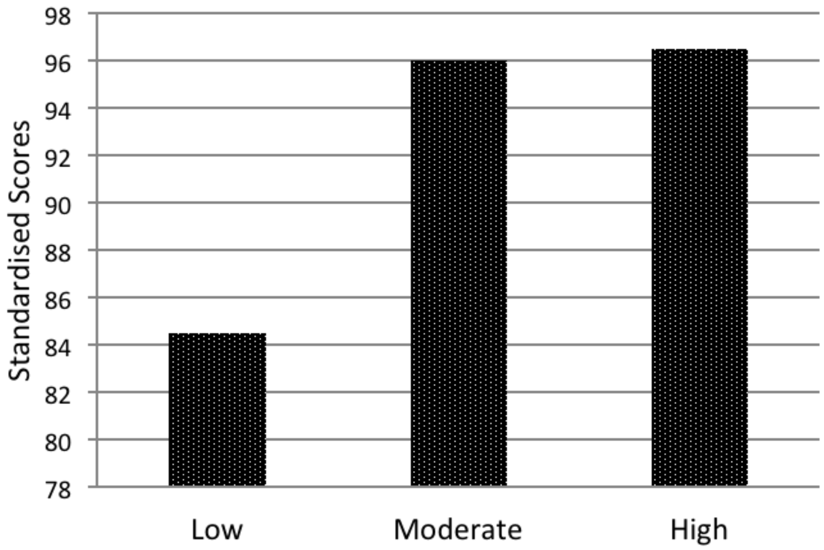


Figure 45 *Forward digit recall performance of low-SES proficiency groups*

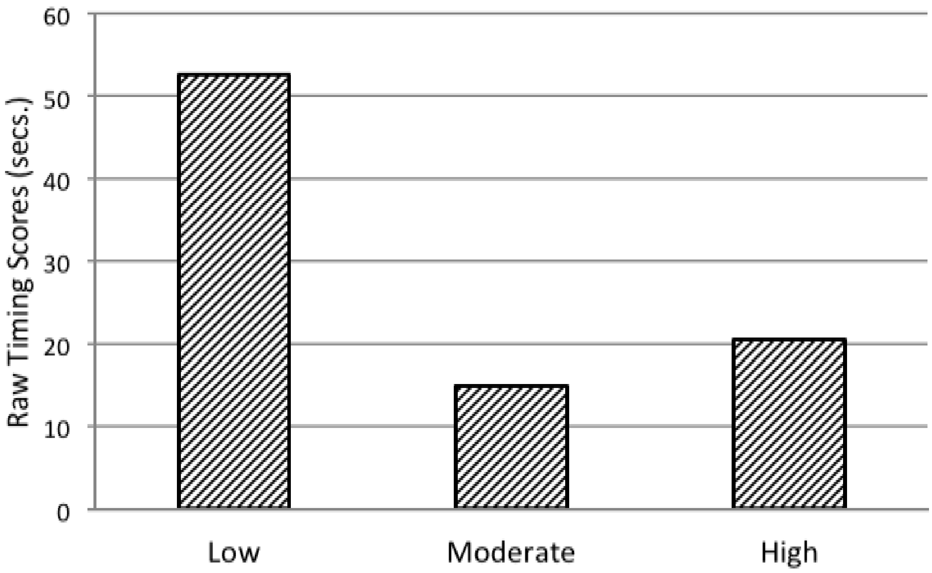


Figure 46 *Trails Difference performance of low-SES proficiency groups*

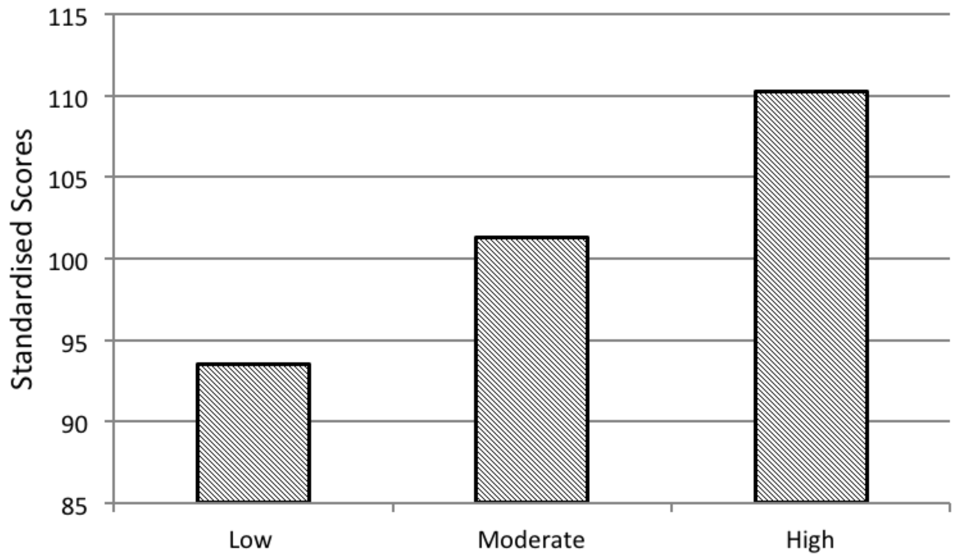


Figure 47 *PPVT of low-SES proficiency groups*

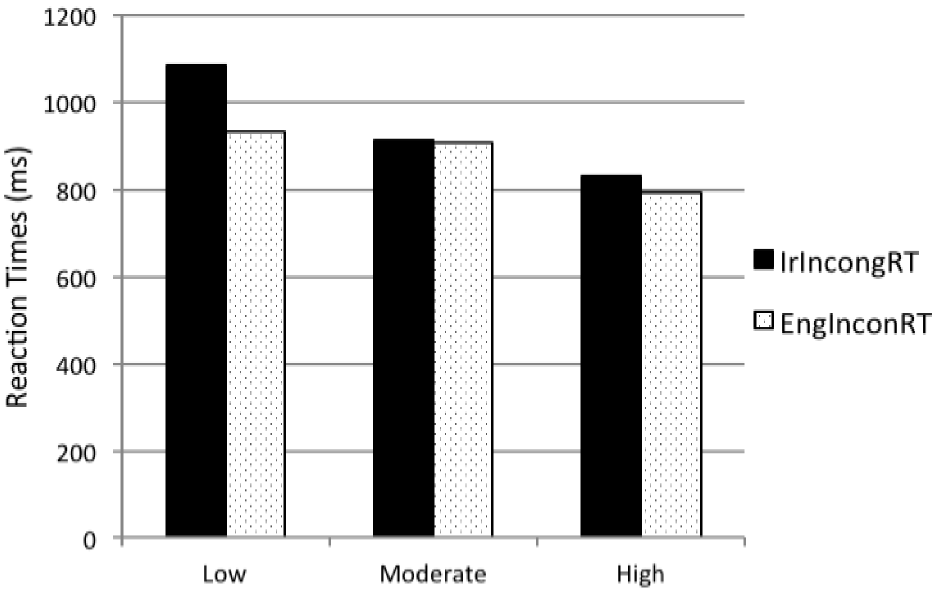


Figure 48 *Stroop incongruent RT performance of low-SES proficiency group*

6.3.3. Comparison of monolingual, immersion and native bilinguals' executive function skills

This section assessed how the native bilingual groups (LSN and MSN) compared with monolingual (LSM and MSM) and immersion bilingual (LSI and MSI) groups on tasks of EF. Between-groups ANOVAs were used to compare results and, due to uneven sample sizes, Hochberg's GT2 was used for post hoc comparisons.

6.3.3.1. Control Measures

There was a significant main effect of group on the Raven's Standard Progressive Matrices (RSPM) {F (5, 134) = 3.55, $p < .01$; $\eta^2 = .12$} and the Peabody Picture Vocabulary (PPVT) test {F (5, 134) = 3.55, $p < .01$; $\eta^2 = .12$}.

Post hoc analyses revealed that the MSI group had significantly higher ($p = .03$) RSPM scores than the LSM group (see Fig. 49). For PPVT the LSM group had significantly lower scores ($p = .02$) than the MSM group. No other group differences were observed (see Fig. 50).

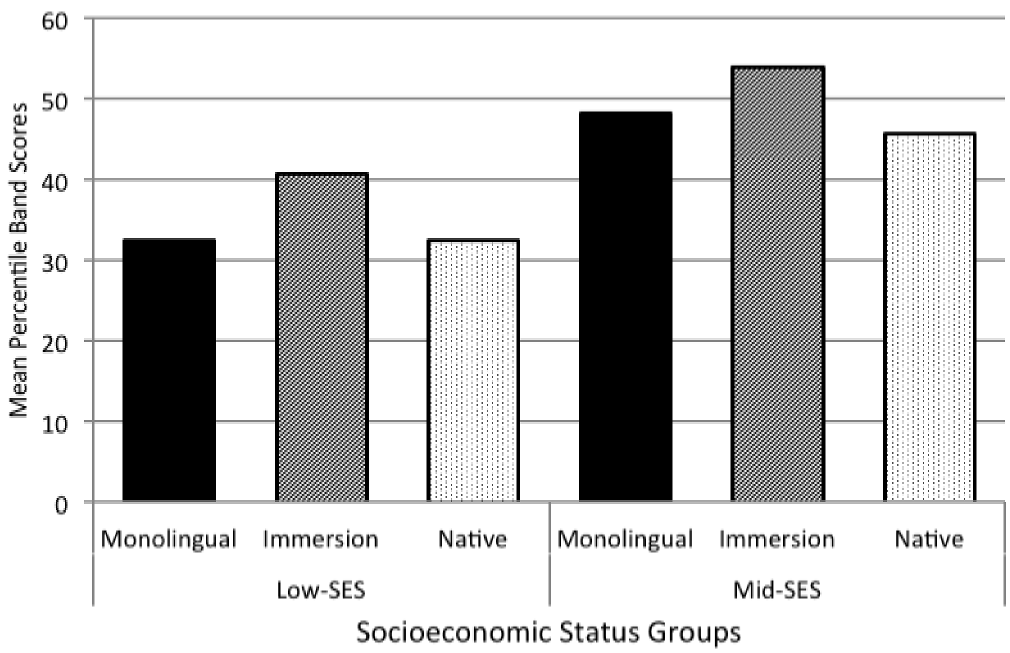


Figure 49 *Mean RSPM scores for monolingual, immersion and native bilingual groups*

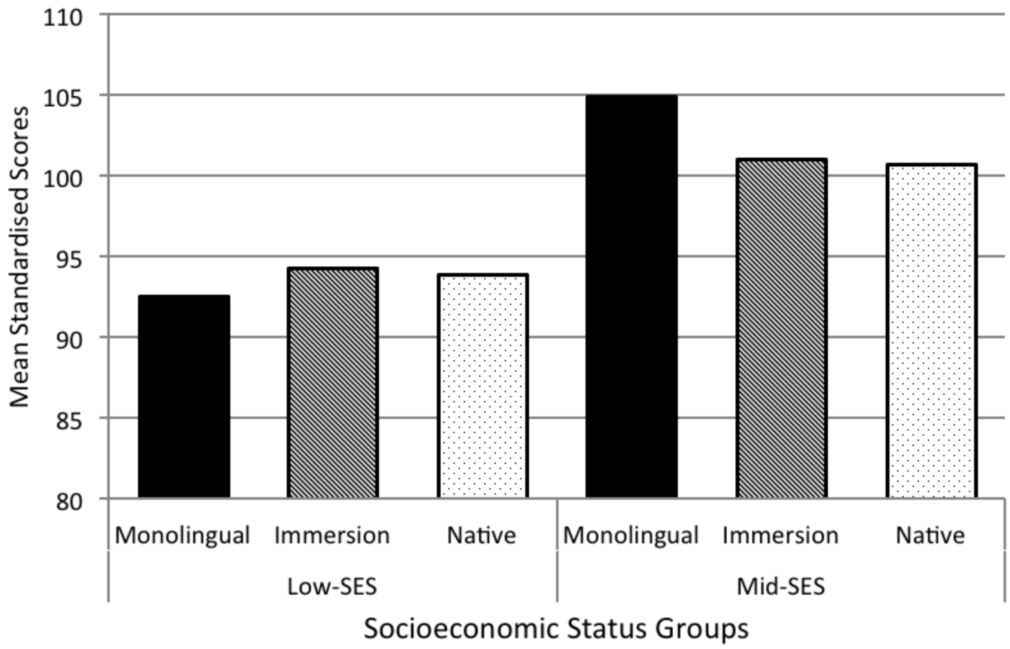


Figure 50 *Mean PPVT scores for monolingual, immersion and native bilingual groups*

6.3.3.2. Inhibitory control tasks

Opposite Worlds

A significant main effect of group was found for the Same World (congruent), *F* (5, 128) = 4.47, $p < .01$; $\eta^2 = .15$, standardised Opposite Worlds (incongruent), *F* (5, 127) = 6.08, $p < .01$; $\eta^2 = .20$, and raw Worlds Difference (sec.), *F* (5, 128) = 3.54, $p < .01$; $\eta^2 = .13$ conditions.

Post hoc analyses and revealed that the MSN group scored significantly lower than both the MSI ($p < .01$) and LSM ($p = .03$) groups on the standardised Same Worlds condition (see Fig. 51). For the Opposite Worlds condition, the LSN and the MSN groups also obtained significantly lower standardised scores ($p < .01$) than the MSI group (see Fig. 52). The MSN also had significantly lower standardised scores than the MSM group ($p = .04$). Finally, for the Worlds Difference condition, the LSI ($p = .04$) and the MSI ($p = .05$) groups had significantly lower raw timing scores (sec) compared with the LSM group. No other group differences were present (see Fig. 51).

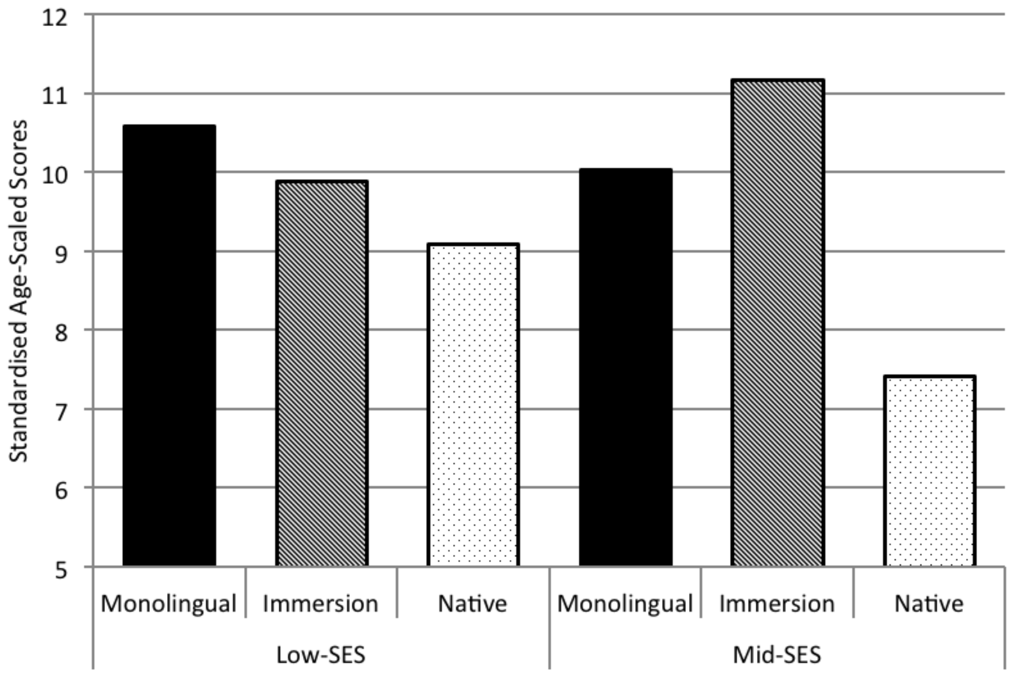


Figure 51 *Age-scaled Same Worlds scores for monolingual, immersion and native bilingual groups*

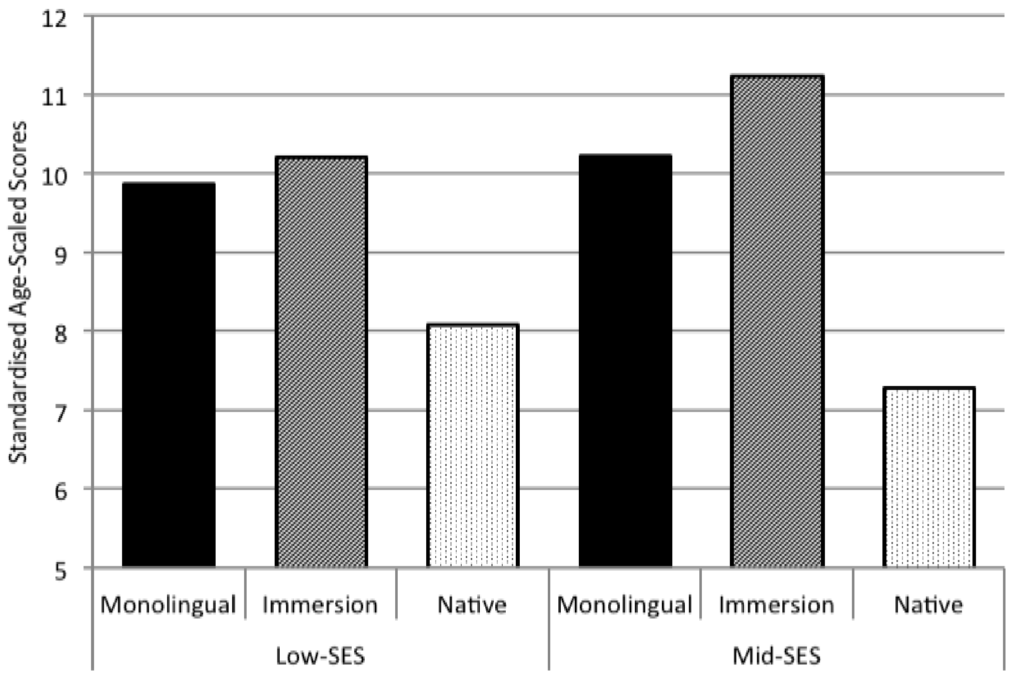


Figure 52 *Age-scaled Opposite Worlds scores for monolingual, immersion and nativebilingual groups*

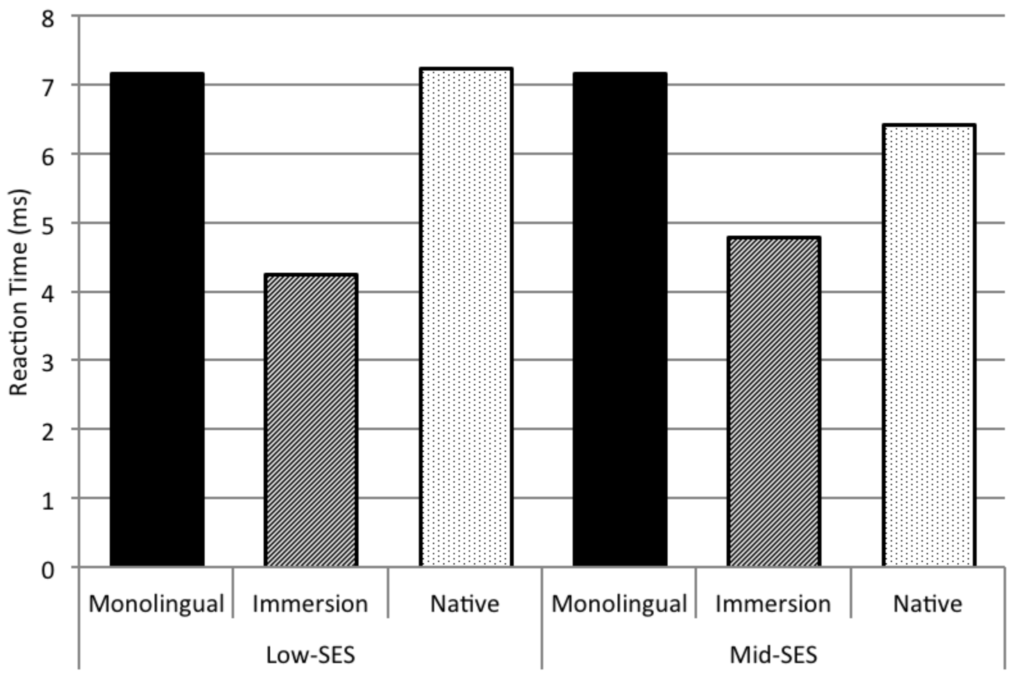


Figure 53 *Raw Worlds Difference timing scores (seconds) for monolingual, immersion and bilingual groups*

Colour-Word Stroop Task

There was a significant main effect of group on the neutral {*F* (5, 114) = 3.35, p = .01; η^2 = .13}, congruent {*F* (5, 114) = 4.12, $p < .01$; η^2 = .16} and incongruent {*F* (5, 114) = 5.65, $p < .01$; η^2 = .21} conditions of the Colour-word Stroop task but not for the inhibition or facilitation conditions.

Post hoc tests revealed that for the neutral (p = .04) and congruent (p = .01) trials, the LSN group took significantly longer than the MSI group on both conditions (see Figures 54 and 55). For incongruent trials, the LSN group took significantly longer than both the MSI ($p < .01$) and MSM (p = .01) groups (see Figure 54). The MSN also took significantly longer on the incongruent trials than the MSI and MSM groups.

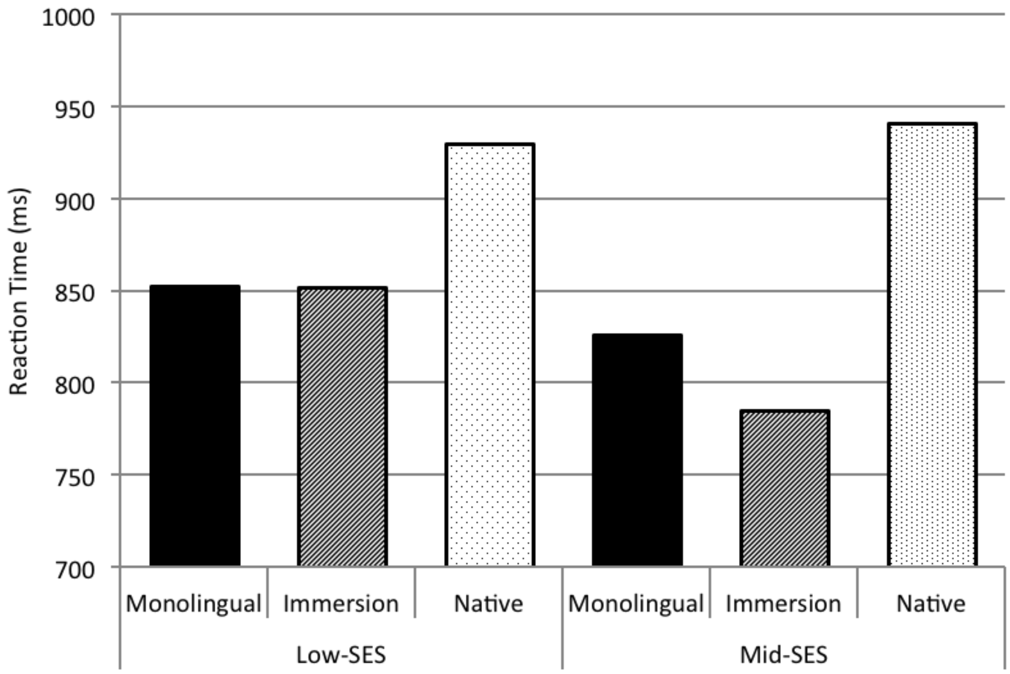


Figure 54 *Mean neutral reaction times (ms) on the Colour-word Stroop task for monolingual, immersion and native bilingual groups*

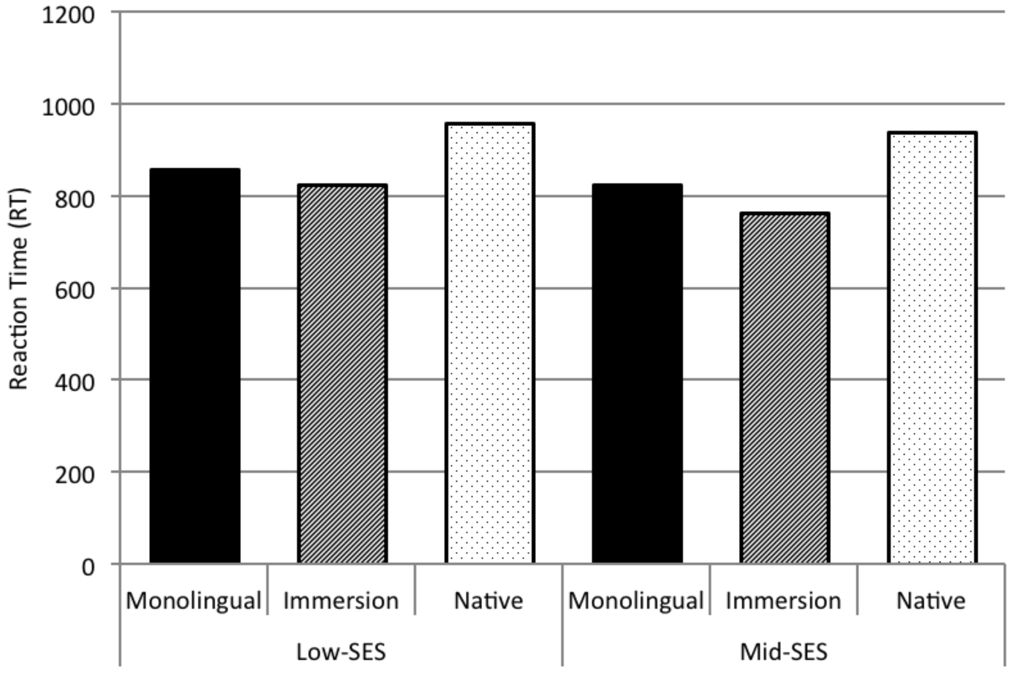


Figure 55 *Mean congruent reaction times (ms) on the Colour-word Stroop task for monolingual, immersion and native bilingual groups*

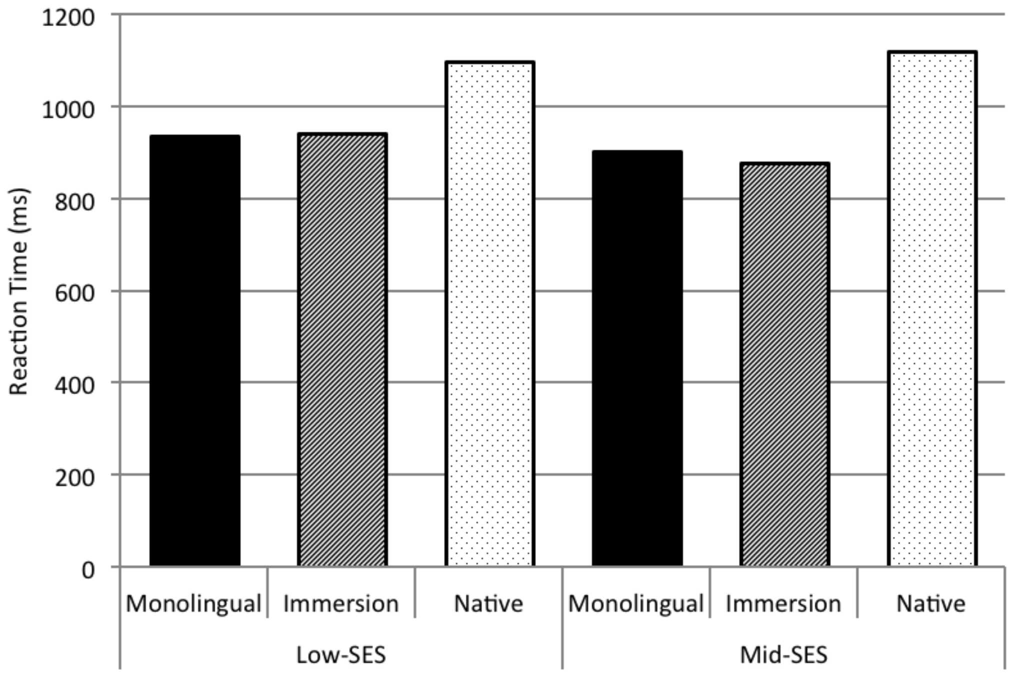


Figure 56 *Mean incongruent reaction times (ms) on the Colour-word Stroop task for monolingual, immersion and native bilingual groups*

6.3.3.3. Switching task

Creature Count Task

There was a significant main effect of group in both the standardised timing {F (5, 129) = 5.12, $p < .01$; $\eta^2 = .17$} and standardised accuracy conditions {F (5, 129) = 2.25, $p = .05$; $\eta^2 = .08$}.

Post hoc analyses indicated that the LSN group had significantly lower standardised timing scores than the LSM ($p = .03$), MSM ($p = .03$) and MSI ($p = .01$) groups. The MSN group also had significantly lower standardised timing scores than the LSM ($p = .02$), MSM ($p = .02$) and MSI ($p = .01$; see Figure 57).
Although the main effect of group was significant for the standardised Creature Count Accuracy scores, post hoc revealed no significant differences between groups for this variable.

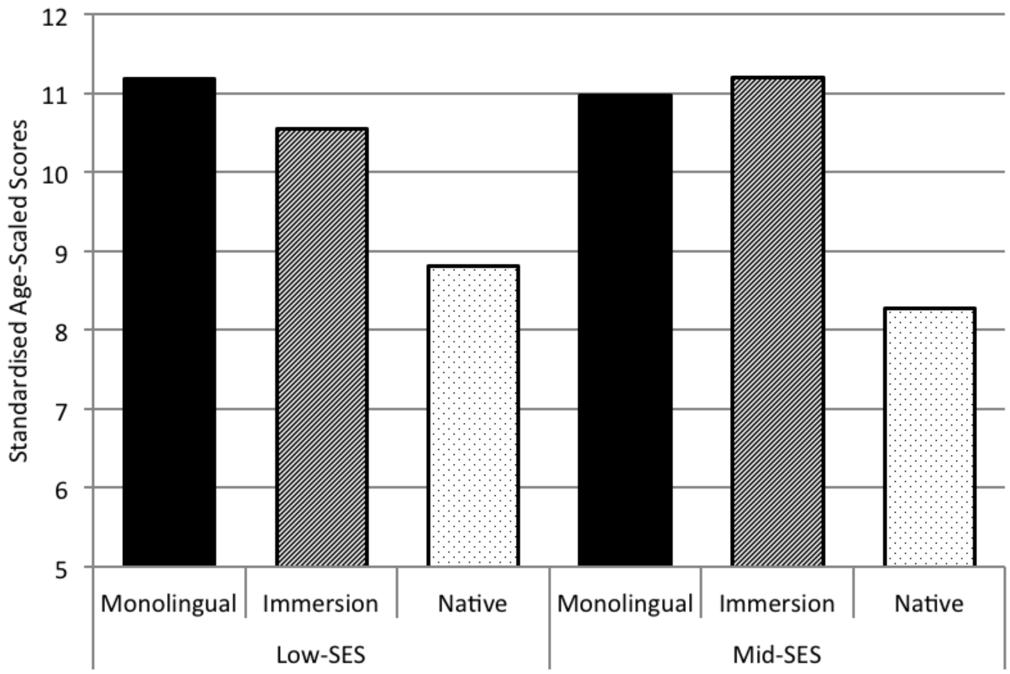


Figure 57 *Mean standardised Creature Count Timing scores for monolingual, immersion and native bilingual groups*

6.3.3.4. Working Memory tasks

Working Memory Test Battery for Children

There was no significant main effect of group for the standardised STM or WM tasks. However, there was a significant main effect of group for the unstandardized backward corsi-blocks assessment, measured by number of items correct, F (5, 107) = 2.45, p = .04; η^2 = .12. Although the post hoc analyses showed no significant differences between groups, mean scores showed that the group with the highest number of items correct on this task were the MSN group and the group with the lowest number of items correct were the LSN group (see Figure 58).

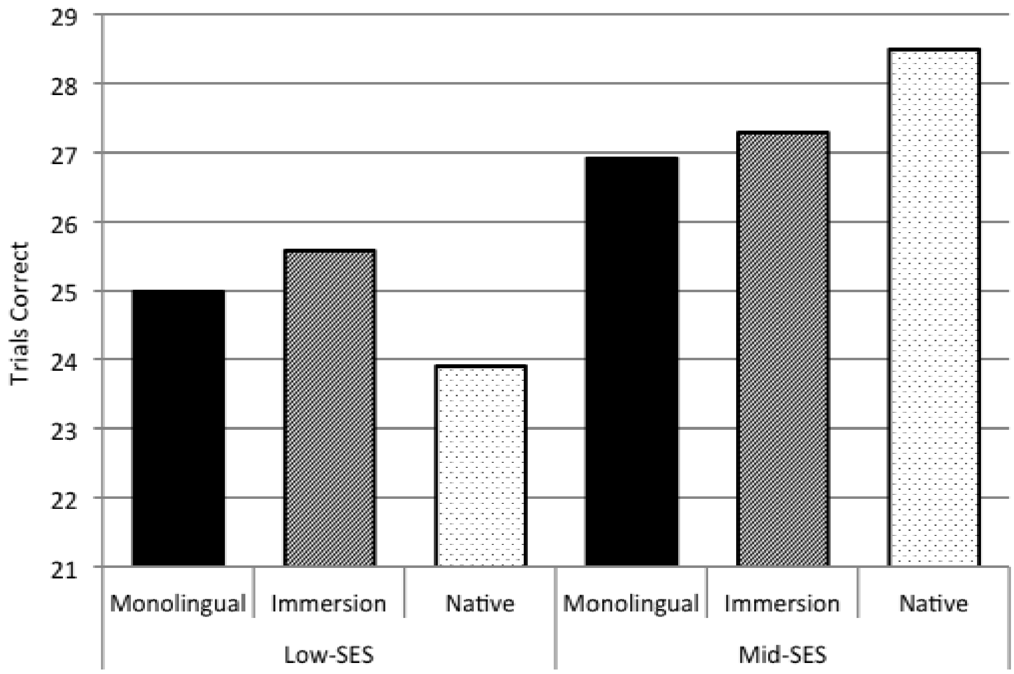


Figure 58 *Mean trials correct for the backward corsi-blocks task for monolingual, immersion and bilingual groups*

6.3.3.5. Unified executive function tasks

Trail Making Test (TMT)

There was a significant main effect of group for the Trails B, F (5, 126) = 4.16, $p <$ .01; η^2 = .15 and Trails Difference, F (5, 126) = 4.06, $p < .01$; η^2 = .14 conditions of the TMT but no main effect for the Trails A condition.

Post hoc analyses indicated that the LSM group had significantly longer response times on the Trails B condition than the LSI (p = .02), MSI (p < .01) and MSM (p = .03) groups (see Figure 59). There was no significant difference between either of the native bilinguals (LSN and MSN) and other groups.
For the Trails Difference, the LSI group had significantly longer response times than the MSI group (p < .01). There were no other differences between groups (see Figure 60).

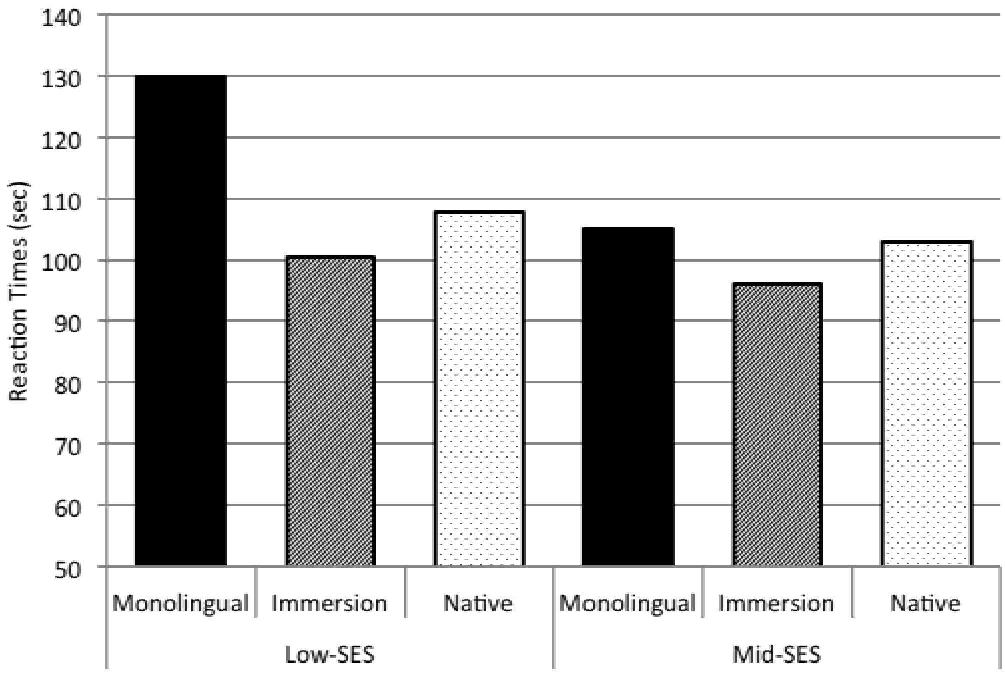


Figure 59*Mean Trails B response times for monolingual, immersion and native bilingual groups*

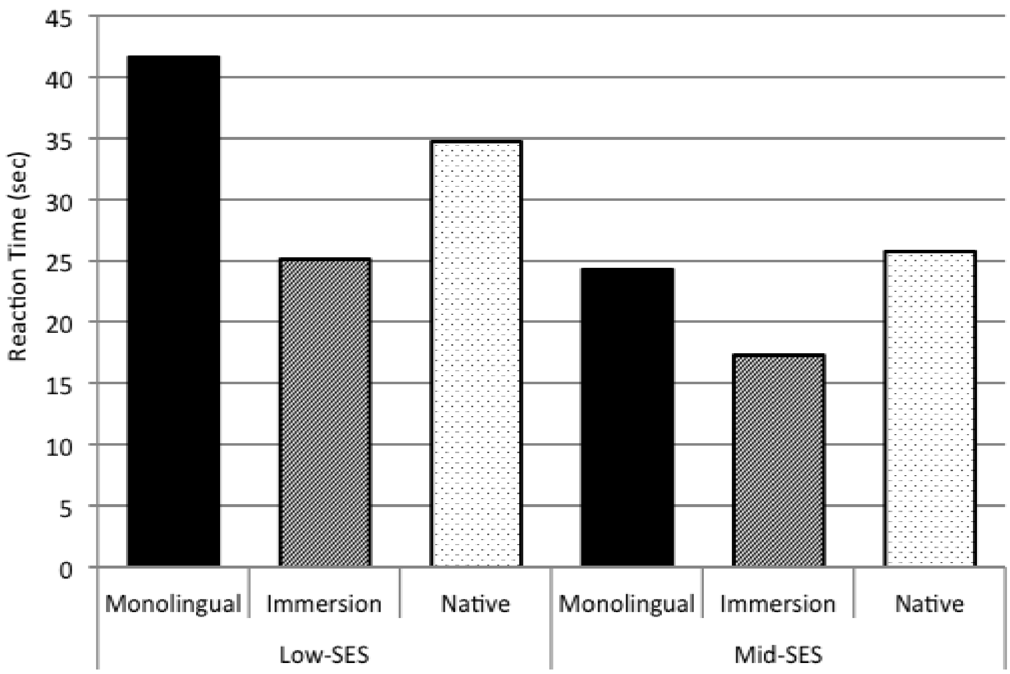


Figure 60 *Mean Trails Difference response times (sec) for monolingual, immersion and native bilingual groups*

Wisconsin Cart Sort Task (WCST)

ANOVAs revealed a significant main effect of group for the following standardised conditions: total errors, F (5, 115) = 2.37, p = .05, η^2 = .10, perseverative responses, F (5, 115) = 3.98, p < .01, η^2 = .15, perseverative errors, F (5, 115) = 3.58, p = .01, η^2 = .14 and failure to maintain set, F (5, 115) = 2.55, p = .03, η^2 = .10. No main effect of nonperseverative errors and trial to complete first category was found.

Post hoc analyses revealed no significant difference between groups for total errors although mean scores indicated that the MSN group had the highest standardised score and that the LSN group had the lowest standardised scores. Descriptive statistics for each group displayed in Table 49.

For perseverative responses, there was a significant difference between the MSN groups standardised scores and the LSM (p = .02), MSM (p < .01) and LSN (p = .02) groups. However, there was no significant difference between the MSN and immersion bilingual groups.

The same effect was found for perseverative errors with the MSN group having the highest mean score which was significantly higher than the LSM (p = .03), MSM (p = .02) and LSN (p = .01) groups.

For the number of categorise complete the MSI group completed significantly more categories than the MSM group (p < .01). The LSI group also completed more categories than the MSM group (p < .05) and the LSN group made fewer failure to maintain set errors than the MSM group (p = .05).

Table 49*Means (and standard deviations) of WCST outcomes for monolingual, immersion and native bilingual groups*

Language Skill	**Low-SES Group**			**Mid-SES Group**			**Total**
	Monolingual *n* = 12	Immersion *n* = 17	Native *n* = 10	Monolingual *n* = 28	Immersion *n* = 43	Native *n* = 6	Total *n* = 116
Total Errors	95.17 (13.60)	103.71 (11.96)	94.50 (13.34)	96.61 (14.73)	102.77 (10.72)	108.50 (16.25)	100.22 (13.13)
Perseverative Responses	97.17 (15.06)	102.82 (12.70)	96.60 (11.86)	97.07 (13.19)	104.37 (10.40)	117.83 (13.93)	101.66 (13.04)
Perseverative Errors	97.25 (15.86)	102.82 (12.81)	94.70 (12.12)	97.96 (13.80)	104.37 (10.85)	117.50 (16.16)	101.71 (13.53)
Nonperseverative Errors	94.50 (12.01)	103.12 (11.21)	93.90 (16.32)	95.29 (13.70)	99.84 (12.66)	99.00 (14.52)	98.11 (13.18)
Categories Complete	5.25 (1.22)	5.65 (.61)	5.60 (.70)	4.71 (1.61)	5.72 (.59)	5.67 (.52)	5.41 (.52)
Failure to Maintain Set	.92 (1.08)	1.0 (.87)	.20 (.42)	1.43 (1.67)	.72 (.85)	1.50 (1.23)	.95 (1.64)
Trials to Complete First Category	16.08 (11.73)	12.94 (3.68)	11.50 (1.51)	16.57 (11.89)	12.33 (3.80)	13.50 (2.88)	13. 50 (2.88)

6.3.4. Task performance through English and Irish

As well as performing tasks in English the immersion and native bilingual groups performed the IC tasks (Opposite Worlds and Colour-word Stroop) in Irish. Numbers in the Opposite Worlds and words in the Colour-word Stroop were translated to Irish to test the effects trial language on bilingual groups' performance. The immersion bilingual groups (LSI and MSI) completed the Irish and English version across time as part of the longitudinal study (Time 1, Time 2 and Time 3; see Chapter 5) and the native bilingual group completed the tasks at Time 3 only. Therefore, the immersion groups' performance was assessed using mixed-ANOVAs to assess performance across time while the native bilingual groups' performance was assessed and compared with immersion bilinguals using between-groups ANOVA and paired t-tests were used to compare their performance on the Irish and English version within native groups.
Immersion bilinguals completed the Opposite Worlds and colour-word Stroop tasks in their L1, English and L2, Irish at three time points. A 3 x 2 x 2 mixed ANOVA compared immersion children's performance on each condition of both tasks in English and Irish. The aim was to understand the influence of the L1 and L2 on task performance. The between-groups variable was socioeconomic group (mid- and low-SES) and the two within-groups variables were language (English and Irish) and Time (Time 1, Time 2 and Time 3). Order of administration was counterbalanced for both IC tasks.

6.3.4.1. Opposite Worlds

Same Worlds Condition

There was a significant interaction between time and trial language {F (2, 120) = 5.83, p = .01; η_p^2 = .09}. Mean scores (Table 50) and Figure 61 show that at Time 1, the Irish version of the task took slightly longer (1 second) than the English version although this difference decreased by Time 2 and Time 3. In the LSI group RTs were faster in English at Time 1 but faster in Irish at Time 2 and Time 3. There was a main effect of time {F (2, 120) = 63.57, p < .01; η_p^2 = .51} and of trial language {F (1,

120) = 4.87, p = .03; η_p^2 = .08}. Mean timing scores decreased across time and the English version of the task was performed more quickly overall (*M* = 22.30sec, *SE* = .39) than the Irish version (*M* = 22.68sec, *SE* = .41). There were no other main effects or interactions.

Table 50 *Means (and standard deviations) of bilinguals' English and Irish raw timing scores on the Opposite Worlds task*

Task Condition	SES Group	Trial Language	Time 1	Time 2	Time 3
Same Worlds	Mid-SES (*n* = 44)	English	24.36 (3.91)	22.13 (3.0)	19.45 (2.54)
		Irish	25.14 (3.93)	22.20 (3.33)	20.11 (3.02)
	Low-SES (*n* = 18)	English	23.78 (5.29)	23.45 (2.55)	20.65 (2.35)
		Irish	25.28 (5.33)	22.83 (2.08)	20.56 (2.17)
	Total (*n* = 62)	English	24.19 (4.32)	22.51 (2.92)	19.80 (2.53)
		Irish	25.18 (4.34)	22.38 (3.01)	20.24 (2.79)
Opposite Worlds					
	Mid-SES (*n* = 44)	English	31.50 (5.11)	27.93 (4.87)	24.25 (3.36)
		Irish	33.07 (5.81)	27.83 (3.95)	24.52 (2.94)
	Low-SES (*n* = 18)	English	31.44 (7.0)	29.65 (3.93)	24.91 (2.73)
		Irish	33.94 (8.01)	29.50 (4.05)	24.99 (2.91)
	Total (*n* =62)	English	31.48 (5.66)	28.43 (4.65)	24.44 (3.18)
		Irish	33.32 (6.47)	28.31 (4.02)	24.66 (2.91)
Worlds Difference					
	Mid-SES (*n* =44)	English	7.14 (3.96)	5.80 (3.73)	4.80 (2.20)
		Irish	7.93 (4.15)	5.63 (2.83)	4.42 (2.23)
	Low-SES (*n* = 18)	English	7.67 (6.82)	6.20 (3.35)	4.26 (2.13)
		Irish	8.67 (5.08)	6.67 (3.23)	4.43 (2.67)
	Total (*n* = 62)	English	7.29 (4.91)	5.92 (3.60)	4.65 (2.18)
		Irish	8.15 (4.41)	5.93 (2.97)	4.42 (2.27)

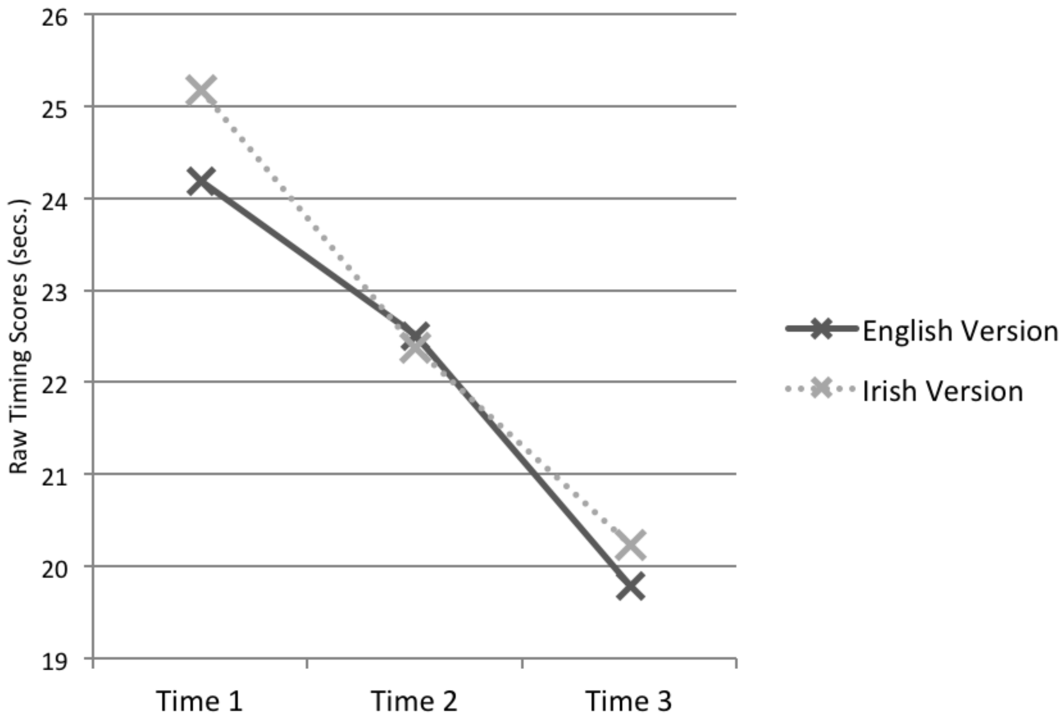


Figure 61 *Mean Same World raw timing scores for the two language versions*

Opposite Worlds Condition

There was a significant interaction between time and trial language {*F* (2, 120) = 3.84, *p* = .03; η_p^2 = .06}. At Time 1, participants has slower RTs in Irish than in English although this difference decreased by Time 2 and Time 3. There was a main effect of time {*F* (2, 120) = 124.03, *p* < .01; η_p^2 = .67} and of trial language {*F* (1, 120) = 4.73, *p* = .03; η_p^2 = .07}. Mean timing scores decreased across time and the English version of the task was performed more quickly overall (*M* = 28.28sec, *SE* = .54) than the Irish version (*M* = 28.98sec, *SE* = .53) although this difference was very small. No other main effects or interactions were found (see Fig. 62).

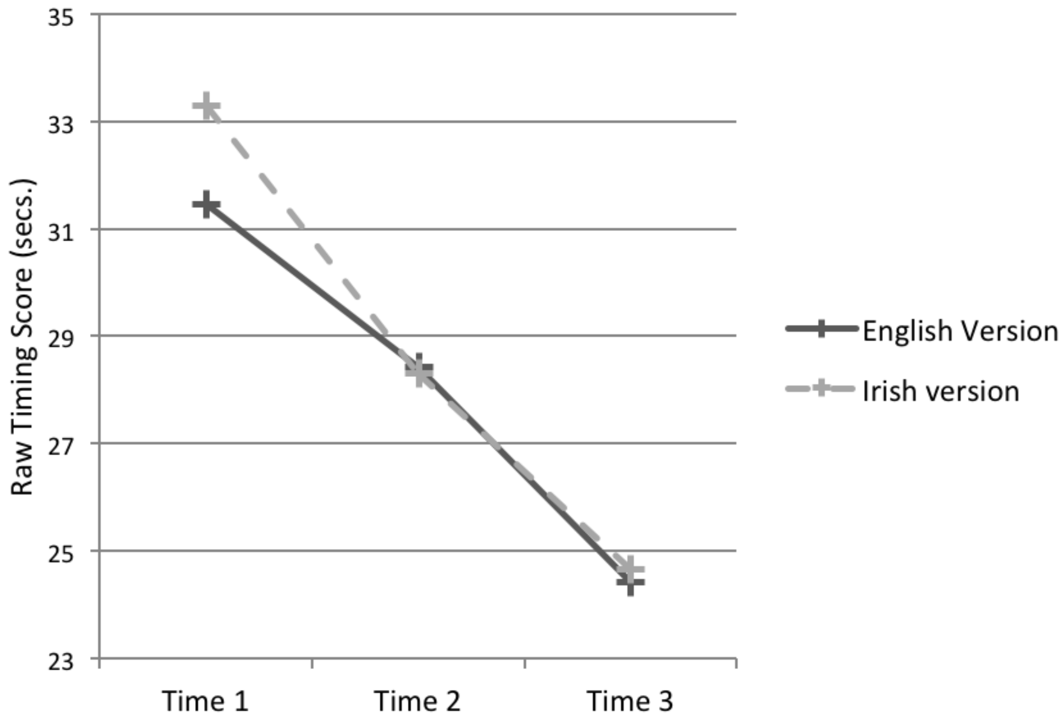


Figure 62 *Mean Opposite Worlds raw timing scores for the two language versions*

Worlds Difference

A similar pattern was observed in the Worlds Difference condition with a higher level of interference experienced in the Irish version of the task than in the English at Time 1, with differences reducing at Time 2 and Time 3. At Time 3, the Irish version had a lower inhibitory cost than the English version. However the mixed-ANOVA revealed no significant interactions or main effects, other than for time {F (2, 120) = 30.70, p < .01; η_p^2 = .34}, with inhibitory cost RTs reducing across time (see Table 50 and Figure 63).

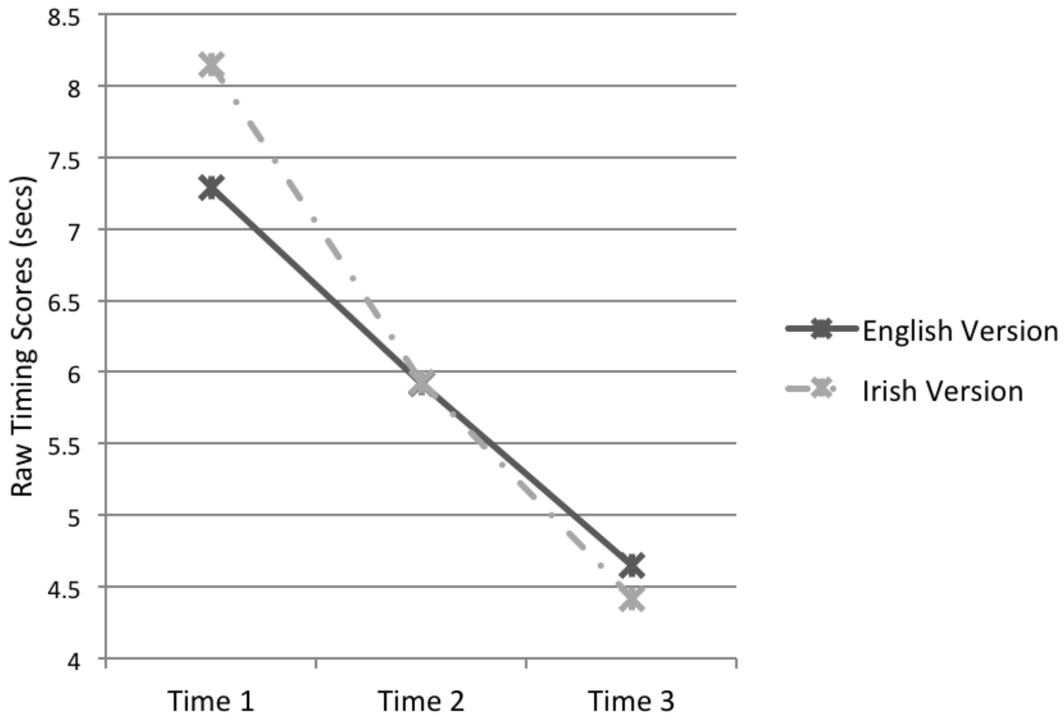


Figure 63 *Mean Worlds Difference raw timing scores for the two language versions*

6.3.4.2. Colour-word Stroop

Neutral condition

There was a main effect of time {F (2, 96) = 27.94, p < .01; η_p^2 = .37} and of trial language {F (1, 48) = 13.78, p < .01; η_p^2 = .22} on this condition. Mean RTs decreased across time and the English version of the task was performed more quickly overall (M = 872.38ms, SE = 16.51) than the Irish version (M = 921.19, SE = 16.91). There were no other interactions or main effects (see Table 51).

Congruent condition

There was a significant interaction between trial language and time {F (2, 96) = 3.68, p = .03; η_p^2 = .07}. At Time 1, participants took longer to perform congruent trials in Irish (M = 964.33ms, SD = 201.27) than in English (M = 908.23ms, SD = 167.91) but by Time 2 and Time 3 were performing almost equivalently (see Table 51). There was a main effect of time {F (2, 96) = 22.74, $p < .01$; $\eta_p^2 = .32$} as RTs decreased across time (see Table 51). There were no other interactions or main effects.

Table 51 *Means (and standard deviations) RTs for trial languages in each condition of the colour-word Stroop*

Task Condition	SES Group	Trial Language	Time 1	Time 2	Time 3
Neutral	Mid-SES (n = 33)	English	949.27 (179.54)	850.37 (119.21)	776.09 (113.02)
		Irish	979.92 (175.30)	860.53 (132.13)	802.29 (132.93)
	Low-SES (n = 15)	English	903.13 (86.74)	855.37 (130.63)	851.67 (126.06)
		Irish	979.92 (175.30)	944.33 (102.31)	884.23 (134.82)
	Total (n = 48)	English	934.85 (157.02)	852.20 (121.51)	799.71 (121.17)
		Irish	998.34 (175.30)	886.72 (128.63)	827.90 (137.54)
Congruent	Mid-SES (n = 33)	English	914.97 (174.19)	855.92 (127.66)	762.55 (118.22)
		Irish	954.08 (198.36)	855.33 (148.89)	767.97 (143.73)
	Low-SES (n = 15)	English	888.17 (165.52)	879.60 (169.08)	821.60 (134.65)
		Irish	975.37 (222.97)	856.40 (120.21)	824.00 (121.59)
	Total (n = 48)	English	906.59 (170.22)	863.32 (140.48)	781.00 (125.23)
		Irish	960.73 (204.20)	855.67 (139.28)	785.41 (138.41)
Incongruent	Mid-SES (n = 33)	English	1094.29 (224.14)	970.24 (164.79)	869.06 (131.65)
		Irish	1150.17 (205.36)	1009.12 (157.20)	902.41 (174.62)
	Low-SES (n = 15)	English	1078.43 (172.90)	965.97 (126.05)	941.53 (109.17)
		Irish	1238.97 (195.79)	1044.37 (125.30)	983.07 (124.04)
	Total (n = 48)	English	1089.33 (207.76)	968.91 (152.40)	891.71 (128.46)
		Irish	1177.92 (204.60)	1020.14 (147.55)	927.61 (153.62)

Task Condition	SES Group	Trial Language	Time 1	Time 2	Time 3
Facilitation	Mid-SES (*n* = 33)	English	33.18 (136.90)	1.37 (82.93)	16.21 (64.15)
		Irish	68.72 (164.14)	10.01 (91.58)	45.99 (101.78)
	Low-SES (*n* = 15)	English	53.03 (198.96)	-27.53 (77.19)	26.47 (61.50)
		Irish	143.16 (147.81)	80.19 (95.16)	49.53 (85.09)
	Total (*n* = 48)	English	39.53 (157.57)	-7.88 (81.49)	19.49 (62.87)
		Irish	92.54 (161.44)	32.47 (97.54)	47.12 (95.90)
Inhibition	Mid-SES (*n* = 33)	English	145.93 (140.40)	114.29 (89.77)	89.32 (87.01)
		Irish	170.56 (132.30)	146.00 (106.79)	90.59 (122.84)
	Low-SES (*n* = 15)	English	146.84 (179.27)	120.25 (94.53)	90.25 (78.46)
		Irish	188.28 (157.30)	101.59 (94.53)	110.09 (103.28)
	Total (*n* = 48)	English	146.22 (152.03)	116.20 (94.76)	89.62 (83.57)
		Irish	176.23 (139.40)	131.79 (104.18)	96.83 (116.25)

Incongruent trials

There was a main effect of trial language {F (1, 48) = 23.04, $p < .01$; $\eta_p^2 = .32$} and of time {F (2, 96) = 55.63, $p < .01$; $\eta_p^2 = .54$}. Reaction times decreased across time (see Table 51) and the incongruent trials were performed more quickly through English overall (*M* = 990.19ms, *SE* = 20.39) than Irish (*M* = 1055.71ms, *SE* = 20.67). There were no other interactions or main effect.

Facilitation

No significant main effects or interactions.

Inhibition

There was a main effect of trial language {F (1, 48) = 16.65, $p < .01$; $\eta_p^2 = .26$} and of time {F (2, 96) = 7.72, $p < .01$; $\eta_p^2 = .14$}. Reaction times decreased across time (see Table 52) and inhibition delays were less in English version (*M* = 117.82ms, *SE* = 11.25) than the Irish version (*M* = 179.37ms, *SE* = 14.11). There were no other interactions or main effect.

6.3.4.3. Comparison of immersion and native bilingual performance on Irish versions of inhibitory control tasks

Bilingual groups were compared to examine if native bilinguals performed the Irish IC tasks at Time 3 more quickly than the immersion bilinguals as a result of their higher levels of LP.

Opposite Worlds

There was a main effect of group for the Same World (congruent) {F (3, 80) = 7.06, $p < .01$; $\eta^2 = .22$} and Opposite World (incongruent) {F (3, 80) = 9.77, $p < .01$; $\eta^2 = .28$} Irish conditions of the task, but not for the Worlds Difference condition {F (3, 80) = .92, $p = .44$; $\eta^2 = .03$}. Post hoc tests (Heichberg's GT2) revealed that the MSI group performed significantly more quickly than both the LSN ($p = .02$; $p < .01$) and MSN ($p < .01$; $p < .01$) groups on the Same and Opposite conditions of the task respectively. The LSI group ($p = .03$) also performed the Opposite World condition more quickly than the native bilingual groups.

Colour-word Stroop

Although there was no significant main effect of group for Irish Stroop accuracy scores, there was a significant main effect of group on the Stroop reaction time (RT) conditions including neutral RTs {F (3, 80) = 7.06, $p < .01$; $\eta^2 = .19$}, congruent RTs {F (3, 80) = 7.06, $p < .01$; $\eta^2 = .28$} and incongruent RTs {F (3, 80) = 7.06, $p < .01$; $\eta^2 = .45$} at Time 3. Mean scores for each of the Stroop task conditions are shown in Figure 64. For neutral RTs, the MSI group had significantly lower timing scores than both the LSN ($p = .05$) and MSN ($p = .02$) groups. For the congruent RTs with MSI scoring significantly lower timing scores than the LSN ($p < .01$) and MSN ($p < .01$). The LSI also recorded significantly lower RTs for congruent trials than the LSN (p =

.05) and MSN groups (p = .04). Finally in the incongruent trials, MSI and LSI both had significantly lower RTs than the LSN ($p < .01$; $p < .01$) and MSN ($p < .01$; $p < .01$) groups respectively.

No differences were found between groups for facilitation RTs but there was a main effect of group for inhibition RTs {F (3, 66) = 4.07, p = .01; η^2 = .16}. Although there was a main effect of group, post hoc analysis revealed no significant differences between groups although mean scores showed that the native bilingual groups had higher RTs than immersion groups.

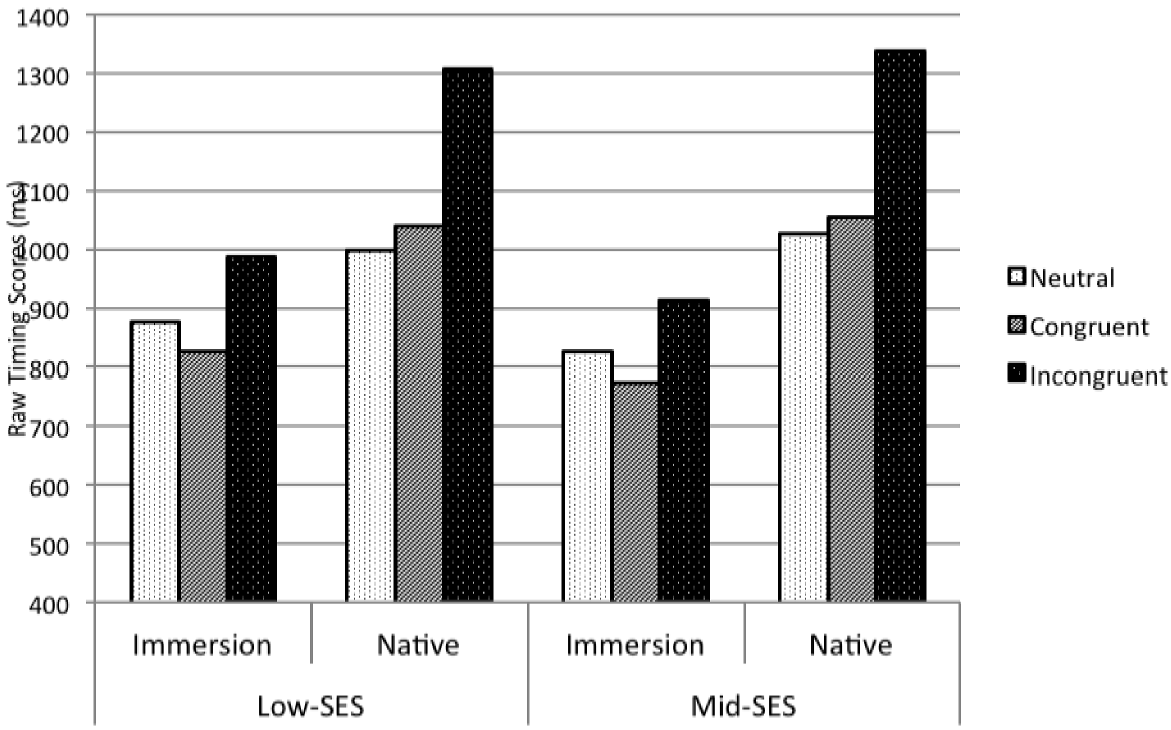


Figure 64 *Mean RTs for immersion and native bilingual Irish Stroop conditions*

6.3.4.4. Native bilinguals' performance in English and Irish

To compare whether the native bilingual groups' performed the IC tasks better in English or in Irish paired-samples t-tests were used to compare participants' raw timing scores on the IC tasks.

Paired t-tests revealed no difference in the time taken for the LSN or MSN to perform the English and Irish versions of the Opposite Worlds task conditions. The only significant difference found for the colour-word Stroop task was in the MSN group where participants performed the task significantly faster in English (*M* = 1119.08ms; *SD* = 116.67) than in Irish (*M* = 1337.08ms; *SD* = 234.17; *t* (6) = 1.78, *p* = .04).

6.3.5. Attitudes towards English and Irish

At Time 3, children's attitudes towards English and Irish was assessed by asking children to rate (along a 10cm line) their opinions of a number of different language-related activities e.g. reading in English/Irish, watching television in English/Irish. Children rated their 'like' or 'dislike' of items using a 10cm rating scale where 0 = *I really don't like* and 10cm = *I really like*. Questions were presented in a child-friendly manner with sad, smiley and neutral faces (see Figure 65). Table 52 displays the descriptive statistics for participants' responses.

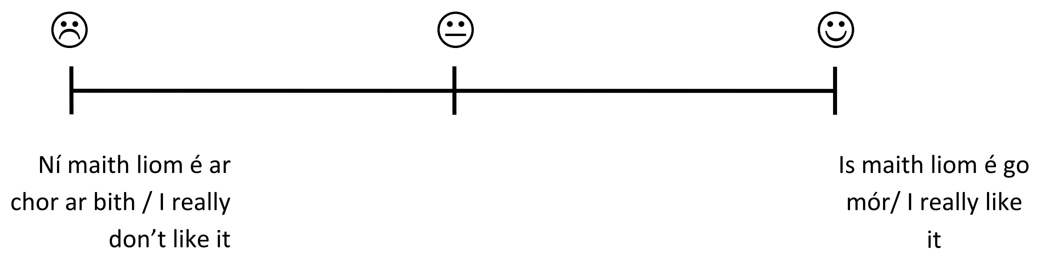


Figure 65 *Example of children's attitude rating item*

There was a significant main effect of group on the items, attitude towards reading in English {F (3, 78) = 3.02, p = .04; η^2 = .11}, reading in Irish {*F* (3, 78) = 3.40, *p* = .02; η^2 = .12} and writing in English {*F* (3, 78) = 4.29, *p* = .01; η^2 = .15}. Post hoc analyses showed that the LSI group liked reading in English less than the MSI group although the difference was only approaching significance (*p* = .06). For the reading in Irish item, the LSN reported liking to read in Irish significantly more than the LSI group (*p* = .03). Finally the MSI group reported liking to write in English significantly more than the LSI group (*p* = .01; see Table 52 for means). All groups gave high ratings in the *speaking in Irish* item and the MSN means showed that they preferred to speak in Irish than in English.

Paired samples t-tests showed that all groups preferred to read, write and watch televisions in English then in Irish but that they equally liked speaking in English then in Irish.

Table 52 *Means (and standard deviation) of attitudes towards English and Irish for the immersion and native bilingual groups*

	Low-SES Group		**Mid-SES Group**	
	Native *n* = 11	Bilingual *n* = 18	Native *n* = 7	Bilingual *n* = 43
Reading in English	9.10 (1.72)	6.64 (3.62)	7.76 (1.77)	8.55 (2.27)
Reading in Irish	8.10 (2.12)	4.72 (3.54)	7.29 (1.94)	5.63 (3.16)
Writing in English	7.78 (2.21)	6.47 (2.80)	8.70 (1.48)	8.77 (2.28)
Writing in Irish	7.93 (2.07)	5.99 (2.82)	8.73 (1.28)	6.86 (2.75)
Speaking in English	9.66 (.77)	9.08 (2.22)	7.50 (2.15)	8.33 (2.25)
Speaking in Irish	9.25 (1.33)	8.07 (2.03)	8.63 (1.60)	8.04 (2.53)
Watching TV in English	9.84 (.39)	9.11 (1.54)	8.20 (1.58)	9.02 (1.97)
Watching TV in Irish	3.33 (4.07)	1.93 (2.35)	5.26 (2.67)	3.26 (3.56)

At the end of the proficiency questionnaire children were asked if they would like to say anything about why they liked or didn't like using Irish in school. As this was optional, response rates were low (native bilinguals: n = 4, immersion bilinguals: n = 12). Children were instructed that they could answer in English or in Irish but more chose to answer in English. From the children's answers two themes emerged: that *speaking Irish is unique and different* and that *Irish is part of our identity and heritage*.

Speaking Irish is unique and different

A number of children mentioned how, in many other countries English is spoken, yet only people in Ireland have the opportunity to speak in Irish and that even they (IE students) have to speak in English at home:

"I like having a new language because I speak in English at home all the time"

"I'm from Ireland. Everyone has English…in America, in England. My family all have Irish that nobody else does."

"We're (immersion education pupils) able to speak a different language".

"Some don't understand it so it's like a secret language".

Irish is part of our identity and heritage

A number of children linked their ability to speak in Irish with their identity and where they come from and that speaking the language is part of what makes them Irish:

"Is é ar dteanga féin": (Traslated)…It's our own language".

"It's (Irish) our traditional language".

"It's (Irish) part of our country".

"The Irish language is our language".

6.4. SUMMARY OF RESULTS

Language demographics and proficiency results

Most children within the sample were born in Ireland and there was no difference in time spent within IE between the immersion and native bilingual groups. However, bilinguals in the immersion bilingual groups (LSI and MSI) had more experience in English medium education than the native bilinguals (LSN and MSN) as well as less experience with the Irish language and more experience with the English language. The MSN group had the highest level of exposure to Irish within the home (speaking with parents/caregivers, siblings and other family members) and which was positively correlated with their levels of Irish productive proficiency skills (see Table 44). The only other correlations between Irish proficiency and language demographic items were in the MSI group where levels of Irish receptive and productive skills showed a significant positive correlation with the amount of Irish spoken during social activities (in the community; see Table 42). Proficiency skills in Irish did not differ between LGs although the MSN group had the highest levels of Irish productive skills shown by results from parents, teacher and children's ratings scales. An interesting finding was that the immersion bilingual groups spoke more Irish with friends in school than the LSN and MSN groups (see Table 36) and that the MSN group spoke more Irish with friends outside of school than all other groups.

Comparison of low, moderate and high proficiency results

When immersion bilinguals were categorised into low, moderate and high proficiency groups it appeared that PPVT scores differed between groups with high and moderate LP groups having significantly higher mean scores than the low LP group. However, low LP groups performed significantly better on the accuracy condition of the switching EF task (Create Count). Irish versions of the IC tasks (Opposite Worlds and colour-word Stroop) were also performed more quickly and with lesser attentional delays for IC conditions by high LP groups compared with the low LP groups although findings differed within SES groups (see section 6.3.2.).

Comparison of monolingual, bilingual and native bilinguals EF skills

The native bilingual groups differed from immersion groups on certain tasks of EF. For the Opposite Worlds tasks both the LSN and MSN groups had significantly lower standardised age-scaled scores at Time 3 than the MSI group with mean scores between the 12.2nd – 20.2nd (MSN) and 30.9th – 43.4th (LSN) percentile band. Similarly on the Stroop IC task LSN and MSN groups' had significantly slower RTs than the MSM and MSI immersion groups but did not differ from the LSM or LSI groups. The LSN group in particular had significantly slower RTs on all conditions of the Stroop task compared with the MSI group. For switching skills, the native bilingual groups had significantly lower timing age-scaled scores than the monolingual and MSI groups. However the native bilinguals' standardised scores were only slightly below the population average (30.9th – 43.4th percentile band). There was a significant main effect group for visuo-spatial WM skills, and mean scores revealed that the MSN group had the highest mean of all six groups. For the unified EF function, native bilinguals did not differ in their performance on the TMT. However, on the WCST, the MSN group performed significantly better than all other groups and on the perseverative outcome variables in particular, performed far above population norms (in the 87th percentile). Conversely, the LSN group scored significantly lower than all other groups for WCST outcome variables although they still performed at approximately population average.

Task performance through English and Irish

Immersion bilingual performed the English versions of the IC tasks more quickly than the Irish versions of the task. However, performance delays on the Irish versions at Time 1 reduced by Time 2 and Time 3 and were performed equivalently in many cases, demonstrating the increased abilities of immersion bilinguals to perform well on Irish EF tasks. Although MSI bilinguals performed the incongruent conditions of the Opposite Worlds and colour-word Stroop tasks more quickly than native bilinguals, there was no difference between groups on the Worlds Difference, facilitation or inhibition conditions of the IC tasks. Furthermore, there was no difference between the native bilinguals and LSI group on the Irish versions of the EF tasks.

Attitudes towards English and Irish

The attitudes questions asked to children revealed that overall children preferred to read and write through English although they liked speaking in Irish as much as they liked speaking in English. Overall, the LSI group enjoyed reading in English and Irish and writing in English significantly less than the other groups which may have been indicative of their SES. Overall, the native bilinguals recorded higher ratings of liking Irish than immersion groups including television watching, reading, writing and speaking in Irish. Thematic analysis showed that the Irish language was closely linked to the children's sense of identity and that the language offers them something unique and different.

6.5. DISCUSSION

The aim of this study was to examine the effects of language proficiency (LP) on children's executive function (EF) performance. To assess the effects of LP, the EF performance of the immersion bilinguals categorised as low, moderate and high proficiency was compared. A second set of analysis also looked as LP differences by comparing the EF skills of children from Gaeltacht or Irish speaking areas of Ireland with monolingual and immersion bilingual groups. Performance in Irish and English versions of the IC tasks as well as children's attitudes towards the Irish language was also assessed. As increased proficiency in the L2 (Irish) was thought to increase IE children's experience with controlling and monitoring two active languages in the mind (see sections 3.4.1. and 6.1.1.), the high proficiency group was predicted to show enhanced EF performance compared to the low proficiency group and the Gaeltacht bilingual groups were predicted to perform better on tasks of EF than the monolingual and immersion groups (see section 6.1.4. hypotheses).

6.5.1. Language Proficiency and executive function performance

Children were approximately 11 years of age at time of testing and in both the immersion and native bilingual samples had received approximately 8 years of formal instruction through Irish immersion education (IE). It can be argued. therefore, that participants had received adequate time to develop a high level of L2 skill as IE students have been shown to develop 'native-like' receptive school-based language skills by approximately 11 years (Cummins, 2001; Genesee, 1987; Lazaruk, 2007). However, the groups differed in the amount of exposure they received outside of the school environment. Only one parent in the LSI and MSI group reported speaking more Irish than English with their child and most of the parents/caregivers of immersion bilinguals spoke English to their children at home. Furthermore, immersion children spoke mostly English with other family members and English or a mixture of English and Irish with siblings, often as a result of their siblings being IE pupils.

As the language demographics and skills of immersion bilinguals varied little between individuals, it was decided that Irish LP ratings should be used to classify immersion bilinguals into LP groups. Irish productive language skills were used to assess bilinguals L2 skill and children in the top 20% of scores were grouped as high proficiency and those in the bottom 20% were grouped as low proficiency (for detail on group classification see section 6.2.1.3.). Using LP classifications, immersion bilinguals were compared to examine the effects of LP on children's EF skills.

LP groups did differ in their English receptive vocabulary and EF switching accuracy scores as well as their performance on Irish versions of the IC tasks. PPVT scores increased as a function of LP with high proficiency bilinguals having significantly higher PPVT scores than low proficiency bilinguals. These findings are in line with Cummins' interdependence hypothesis (Cummins, 1979) which stated that children's L1 skills may be positively influenced by the development of L2 skill in IE and that L1 knowledge can be instrumental in the development of corresponding L2 skills. In other words, bilinguals' language systems may complement each other in terms of their development; provided that the L1 is sufficiently developed prior to intensive exposure to the L2 (see also, Rabia, 2001). Here, it seemed that children's Irish and English skills complimented each other as high proficiency skills in productive Irish skills were closely associated with English receptive vocabulary skills.

The high proficiency bilinguals appeared to perform better on the Irish IC tasks although low proficiency participants had significantly higher accuracy scores for the switching task than moderate proficiency bilinguals. The discrepancy between proficiency groups' IC RTs and switching task accuracy scores may relate to a recent study by Tao and colleagues (2011). Their study found that early bilinguals (simultaneous bilinguals) showed reduced cognitive costs when performing a complex version of the ANT task (Fan et al., 2002) compared to late bilinguals (L2 acquired after age 16). Contrastingly, late bilinguals displayed an advantage for conflict resolution in terms of RTs and error rates compared with early bilinguals. Similarly, this study found that low proficiency bilinguals displayed enhanced performance over moderate proficiency bilinguals on the accuracy condition of the Creature Count switching task. Although early and late bilinguals do not equate to low and high immersion bilinguals, both late and low proficiency bilinguals may have similarities in that L2 proficiency levels are generally lower than for early and high proficiency bilinguals who have had more experience and exposure to the L2. A hypothesis for the advantage of low proficiency bilinguals in the switching task was as a result of a speed-accuracy trade-off. Low proficiency bilinguals may have taken more time in performing the switching task which subsequently resulted in their higher accuracy rates. In other words participants may have placed more value on answering correctly than responding quickly which subsequently increased their accuracy rates. The decreased RTs on the Irish versions of the IC is in line with current research which suggests that bilingual performance may improve as a function of increased LP (e.g. Bialystok et al., 2012; Kroll & Bialystok, 2013). Furthermore, it makes sense that participants who have stronger Irish productive skills would find conditions of the Irish Stroop and Opposite Worlds less taxing then low proficiency groups as they may more practiced in controlling their dual-language interference from the L1 and L2.

After the groups were split according to SES, further differences emerged between LP groups. In the LSI group, low proficiency group had lower scores on the forward digit recall task and a higher timing delay on the Trails Difference complex EF task. These findings may relate to the finding from the previous study (5.4.2.3.) that children in the Low-SES IE groups had difficulty compared with other groups on the verbal-WM task. As bilinguals have shown disadvantages on tasks of verbal ability as well as SES being closely linked to language skills, non-verbal tasks seem more appropriate for use with bilingual participants to explore the effects of bilingualism on children’s WM skill and research has shown that SES is closely linked to language skills (6.2.1.1.). For the MSI group, LP groups demonstrated a similar pattern to the overall trend for LP in that performance on the PPVT increased as a function of LP and that the Stroop task conditions were performed more quickly by high proficiency bilinguals than any other LP group.

6.5.2. Executive function performance of Gaeltacht children

Although native bilinguals had more exposure to Irish than immersion bilinguals (see Table 36), the two native schools tested differed in how much Irish exposure they had outside the school environment which may have been related to the schools’ SES

classifications. Although SES individual differences were not found between schools, School 1 had been designated as disadvantaged or *Deis* by the Irish Education Board. The MSN group had a higher number of parents who spoke only Irish or mostly Irish with their child within the home (85%) than in the LSN group (30%). Although two sets of parents in the MSN group reported speaking mostly Irish with their children, they also reported that their child had had 10 years experience with English. This made it difficult to decipher whether these children were L1 English or L1 Irish as proficiency ratings for both languages were high (although in one case Irish skills were rated as slightly above English skills). Despite a smaller number of parents speaking only Irish in the home with LSN bilinguals, this group still maintained higher levels of exposure and experience with Irish than the immersion groups. Because of these discrepancies between native bilinguals, EF performance was examined separately as SES and language demographics were not thought to be similar enough to treat them as one group.
Despite the language exposure and SES differences between native bilinguals, the general attitudes of these groups towards Irish were similar and both reported high levels of like towards speaking, reading and writing in Irish compared with immersion bilingual groups. Therefore, the complexity and nature of the native bilingual group highlighted the variation in factors of LP and the importance of understanding the backgrounds of bilingual groups before drawing conclusions from the data (e.g. Tao et al., 2011).

One important finding from standardised assessments of native bilinguals' performance was that monolinguals, immersion, and native bilinguals were comparable on their levels of English receptive (PPVT) and non-verbal intelligence skills (RSPM). This finding is in line with studies with Welsh IE students. Children who spoke Welsh-only at home and in school achieved the same levels of English skill as children who spoke Welsh only in school (Gathercole & Thomas, 2009). This study concluded that for children growing up in bilingual communities where there is a dominant language (English) alongside a minority language (Irish), children will still attain equivalent levels of the dominant language to monolinguals, regardless of home language patterns. Similar results have been found in French and Irish IE programmes where, following an initial few years of intense L2 exposure, children will catch up to their monolingual peers in terms of their English language skills (Cummins, 2000; Parsons & Lyddy, 2009).

In terms of the IC tasks, native bilinguals displayed a disadvantage compared to monolinguals and immersion bilinguals in both response inhibition and interference suppression. Previous research has found a bilingual advantage on tasks of interference suppression but not response inhibition therefore it was not expected that these groups should perform worse than their peers (Martin-Rhee & Bialystok, 2008; Morton & Harper, 2007; 2009; Nicolay & Poncelet, 2013). For the response inhibition task of IC (Opposite Worlds) native bilinguals displayed a deficit compared to MSI and MSN groups and MSN bilinguals also performed more poorly than the MSM group. For the interference suppression task of IC (Stroop) the LSN group did not differ from LSM or LSI groups. However, the MSN group were slower on incongruent trials than both the MSI and MSM groups.

The first hypothesis for this finding relates to the issue of PE. As monolinguals and immersion bilinguals had already performed these tasks twice, approximately 10 and 12 months previously, they had received practice in performing the task and may have performed better than would be expected if they were being tested for the first time at Time 3. Indeed when we compared the means of the monolingual and immersion groups at Time 1 with the native bilingual groups at Time 3 there appeared to be no significant difference between groups although statistical analysis of these differences was not performed. Due to PE positively biasing monolinguals and immersion bilinguals, results of the response inhibition task are not surprising as recent research has found no difference in monolingual and bilingual performance on such IC tasks (e.g. Barac & Bialystok, 2012; Carlson & Meltzoff, 2008; Martin-Rhee & Bialystok, 2008). Despite the issue of PE and task novelty in IE samples (see section 5.4.4.), test-retest reliabilities of the TEA-Ch were reported as moderate to high (.65-.85; Manly et al., 1999) and test-retest were also reported as moderate to high in correlational analysis carried out in Chapter 5 (5.3.6.).

The second hypothesis considers the higher levels of exposure and experience that native bilinguals had to their L2. As a result of high levels of Irish exposure, native bilinguals and MSNs in particular, may have experienced higher levels of interference from the Irish than immersion and monolingual participants, resulting in a greater delay or interference from Irish during performance of the incongruent trials. A similar finding was reported by Linck and colleagues (2009) who found that young adult bilinguals attending IE had difficulty accessing their L1 as a result of switching into the L2 mental set during school hours, making it more difficult to switch back into the dominant L1 than the less dominant L2 (Meuter & Allport, 1999; see section 3.4.1.2.). However, this language switching theory does not account for the difference between immersion and native bilinguals on the Irish version of the Stroop task in which LSI and MSI outperformed MSNs on the incongruent trials. Although a difference in RTs was found between groups on conditions of the Stroop task, no group differences were found for the Irish and English inhibition and facilitation RTs. In other words although overall RTs were slower for native bilinguals, the level of delay or interference (Stroop effect) was equivalent to other groups. Therefore, this shows that although accessing English for production may have been more difficult for native bilinguals, they did not display cognitive impairment for IC and furthermore, no IC advantage was found for this group. These results were in spite of performance advantages favouring monolingual and immersion groups who had more practice performing the Stroop task (at Time 1 and Time 2).

The hypothesis of native MSN bilinguals having higher levels of Irish active within the mind than immersion bilinguals may have also explained their poor performance relative to the MSM and MSI groups on the switching EF task (Creature Count). However, again issues of PE may have biased these results and comparing natives' performance against the monolingual and bilingual groups at Time 1 may have been more accurate although it would not take into account the increased proficiency levels of language groups at Time 3.

Contrary to tasks which used RTs to measure performance, the two tasks which were not speeded tasks showed an advantage in favour of the MSN group. On both the visuo-spatial WM (backward corsi-blocks) and unified EF task (WCST), the MSN group showed enhanced performance compared with all other groups. WM advantages were not found for the MSN on the verbal-WM task (backward digit recall). These results are in line with current research which suggests that with increased proficiency and use of L2, bilinguals are able to monitor, control and flexibly switch between rules in their mind. Although native bilinguals showed a delay in language switching and IC tasks, when timing scores were not used they appeared to perform equivalently or better then immersion and monolingual peers. Furthermore, the WCST findings are in line with the findings from Chapter 5 which found that immersion bilinguals had a greater ability to overcome perseverative responses and errors. Not only that but the MSN group in this study scored significantly above population average for the perseverative responses and errors outcomes of the WCST, scoring in the 87th percentile band of the population. These results indicate that despite native bilinguals showing delayed performance on specific EF tasks, their performance on the WCST, a task of unified EF was greatly enhanced compared with other groups. In line with current research (e.g. Bialystok et al., 2012; Kroll & Bialystok, 2013; Nicolay & Poncelet, 2013a) it is suggested that as a result of having to monitor, update, inhibit and control two active languages in the mind, native bilinguals have been practiced in using their EF skills and subsequently outperform monolingual peers on tasks of unified EF function.

An important result from this study was that the pattern found for MSN on WM and unified measures of EF was not found for the LSN group. At Time 3, the LSN group performed significantly worse than all other groups for these measures demonstrating the substantial difference between them and MSN group. However, as in Chapter 5, the LSN group did not significantly differ in their WM or WCST performance to LSM or LSI groups. While it could be proposed that reduced levels of Irish exposure led to this difference between MSN and LSN groups, it is more likely that school 1ow-SES and school designation as Deis resulted in the LSN deficit. What this finding shows is that SES factors may be more important than language experience or exposure to the L2 in predicting children's success on tasks of EF.

6.5.3. Limitations and future research

- Some may suggest that the reason bilingual effects were found for some but not other EF tasks was due to the level of linguistic stimuli required for tasks. As the WM and WCST did not utilise linguistic stimuli, this may have resulted in the MSN advantage for these tasks (e.g. Costa et al., 2009).

However, no such effects were found in the LSN group therefore linguistic stimuli alone could not explain these findings. It is proposed that future research endeavour to use tasks requiring minimal linguistic skills to prevent this acting as a confounding variable for results.

- The lack of standardised assessments for participants' Irish skills in this study made it difficult to accurately assess children's levels of L2 skill. The nature of the language proficiency questionnaires meant that they were somewhat subjective and parents and teachers were rating children against their fellow classmates rather than the population as a whole. Future research should try to assess children's LP in the L1 and L2 using standardised assessments to prevent these issues of subjectivity. However, the lack of such assessments in Ireland makes this a difficult task in this case.

- The small sample size in the Gaeltacht group (*n* = 19) made it difficult to generalise these findings to more global bilingual populations. Furthermore, as Gaeltacht participants were not present in the longitudinal study (Chapter 5), it is not known how their EF skills developed as a function of their time within IE. Furthermore, it is evident that the Gaeltacht case is a unique linguistic environment where two schools, only 10 minutes apart can vary greatly in their linguistic and SES experiences. These findings suggest that careful consideration of each bilingual sample is important and can dramatically alter the nature of conclusions.

6.5.4. Summary

The findings from this study highlight the complexity of children's language experience and its potential impact on EF skills. While IC skills showed a disadvantage in native bilinguals, the MSN group outperformed all groups on the backward corsi-blocks and perseverative skills on the WCST. However, for the LSN group no such advantages for WM or unified EF tasks were found, implicating factors beyond language exposure, such as SES on children's EF development.

CHAPTER SEVEN
SUMMARY AND GENERAL DISCUSSION

7.1. SUMMARY OF FINDINGS AND DISCUSSION

This thesis examined the impact of bilingualism on children's executive function (EF) development. The developmental trajectories of the EF skills were examined using a longitudinal design which tracked children's performance on a battery of EF tasks including tasks of inhibitory control (IC), switching, working memory (WM) and unified EF skills. This research also explored the relationship between EF skills, language proficiency (LP) and language experience by comparing the EF performance of three distinct language groups: monolinguals, bilinguals who had acquired their L2 (Irish) from immersion education (IE) programmes and high proficiency or native bilinguals who had exposure to the L1 (English) and L2 (Irish) prior to attending Irish IE programmes. A second facet to this research was an investigation of the language group (LG) effects on children from distinct socioeconomic status (SES) backgrounds. The EF skills of children from low- and mid-SES groups were compared to assess whether bilingualism had a similar impact on children from typical and disadvantaged backgrounds.

As the areas of the brain controlling speech and language production have been shown to overlap with the areas associated with EF, it was predicted that IE children would display enhanced EF performance on certain EF tasks as a result of maintaining and controlling two active languages in the mind (section 3.4.1.). However, it was not known if any findings would change as children developed and spent increased time in IE. Furthermore, the mechanisms eliciting any EF advantages for bilinguals are poorly understood and it was hoped that insights into the potential causal mechanisms for bilingual gains would be obtained through this research, by employing a longitudinal design. To examine issues of bilingualism and EF development further, a group of high proficiency bilinguals with increased exposure to the L2 (Irish) than IE bilinguals were assessed at Time 3 of the longitudinal study and EF performance was compared with monolingual and IE bilinguals. It was predicted that children with higher levels of L2 proficiency or with more balance between their L1 and L2 would subsequently have more practice in dual-language control and display enhanced EF performance. A full list of the aims and hypotheses for each study (Chapter 5 and 6) can be seen in section 5.1.3 and 6.2.4.

What made this thesis unique was that few studies in current bilingualism research have utilised longitudinal designs to explore the developmental nature of bilingual effects. Although some studies have examined the effects of bilingualism on specific EF skills, most include samples of early simultaneous bilinguals and few have examined the impact of the IE experience or sequential bilingualism on children's general EF performance. Research has shown bilingual effects on a number of EF outcomes (e.g. Bialystok, 2009; Carlson & Meltzoff, 2008; Colzato et al., 2008; see sections 3.4.1.), yet there is little agreement over the mechanisms causing such effects as well as contribution of L2 proficiency to such results (Cummins et al., 2001; Kroll & Bialystok, 2013). Cross-sectional studies have struggled to disentangle the cause-and-effect relationship between bilingualism and potential EF advantages. The question of cause and effect still has to resolve whether enhanced EF skills improves children's level of bilingualism or whether bilingualism enhances children's EF skills? Using longitudinal analysis this thesis hopes to shed light upon some of these issues.

Irish IE has been a topic of interest for many researchers since its surge in popularity during the last two decades (Harris, 2007; see 1.4.3.). However, most studies of Irish IE have investigated the impact of IE from a socio-cultural, pedagogical and linguistic perspective (e.g. Hickey, 1999; Ó Duibhir, 2009; Ó hIfearnáin, 2007). These studies are extremely valuable as they provide a linguistic context to the Irish immersion sector as well helping to improve the effectiveness Irish IE teaching practices (e.g. Harris et al., 2006; Hickey, 1997; 1999; 2001, Ó Duibhir, 2009; Ó Muircheartaigh & Hickey, 2008; Parsons & Lyddy, 2009). However, few researchers have examined how Irish IE affects children's cognitive development, and EF development in particular (e.g. Cummins, 1978; Hickey, 1997; Kennedy, 2012; Macnamara, 1966). The literature available suggests that this is the first longitudinal study to compare monolingual and IE children's EF development within Ireland. Understanding the EF trajectories of this sample is important for our understanding of how Irish IE is affecting children's cognitive progress as well contributing to the general IE debate within Ireland. Furthermore, few studies have shown how the IE experience can impact on the cognitive development of children from low-SES or disadvantaged backgrounds, particularly in Ireland.

The key findings of this research are summarised below:

- Monolingual children, children in immersion education (IE) and children from Gaeltacht areas of Ireland showed equivalent skills in their levels of English vocabulary and non-verbal intelligence (*g*).

- The developmental trajectory of children's EF skills showed improvements as a function of age across a three year period. The EF switching task (accuracy) was the only measure to not show a linear improvement with age although scores remained at around population average and reaction times (RTs) on this task did improve as predicted.

- The EF development and performance of IE bilinguals and monolinguals on tasks of IC, switching and WM were equivalent. However, mid-SES bilinguals showed significantly greater EF improvements relative to monolingual peers for the response inhibition (Worlds Difference RTs), switching (CC accuracy) and unified EF (TMT) tasks.

- For tasks of unified EF, bilinguals demonstrated significantly better performance than monolinguals. For the Trail Making Test (TMT) the LG effect was not present at Time 1 but emerged at Time 2 and Time 3 where bilinguals outperformed monolinguals. At Time 3 bilinguals also outperformed monolinguals on the Wisconsin Card Sort Test (WCST) of general EF skills although the LG effect was heavily influenced by mid-SES group's scores. Alternatively, the emergent LG effect on the TMT was heavily influenced by the low-SES group's scores.

- Socioeconomic status (SES) had a significant impact on children's non-verbal IQ and English vocabulary skills. In terms of EF development low-SES groups showed delays on the switching, WM and timed unified EF task

(TMT). It also appeared that the low-SES groups seemed to struggle on certain timed tasks requiring fast sensory information processing (SIP, e.g. Stroop conditions and TMT). However, low-SES groups showed no disadvantages in their IC abilities and although SES put children at an initial disadvantage, it did not impede on the developmental trajectories of their EF skills, which improved as a function of age.

- At Time 3, bilinguals from Gaeltacht areas of Ireland (native bilinguals) demonstrated poorer performance on IC and switching EF tasks compared with monolingual and IE bilingual groups (although these groups had less experience with these tasks than participants who had also performed the tasks at Time 1 and Time 2). However, for non-linguistic tasks of WM and unified EF skills, the mid-SES native group showed higher performance levels than all other groups while the low-SES native group had the poorest results. The language demographics and backgrounds of the two Gaeltacht schools tested were qualitatively different and highlighted the importance of factors beyond language of education for predicting children's EF abilities, such as SES.

7.2. EXECUTIVE FUNCTION DEVELOPMENT

Developmentally, children's EF skills improved as a function of their age between 9 -11 years. This was shown through improved RTs on the IC, switching and unified EF tasks as well as the maintenance of standardised scores at approximately population average on IC, switching, WM and unified EF tasks. The linear developmental trajectory of specific and unified EF skills across time has been replicated in other studies of EF (Diamond, 2013; Diamond & Lee, 2011; Garon et al., 2008; see section 2.5.) and was the expected trajectory of children in middle childhood (Anderson et al., 2001; Best et al., 2011; Huizinga & Smidt, 2010; Levin et al., 1991; Sarsour et al., 2011; Welsh, Pennington, & Groisser, 1991). An important implication of this finding is that any improvement in children's EF skills during this period may shift the entire distribution of outcomes and yield improvements in lifelong EF skills (Diamond & Lee, 2011; Moffitt et al., 2011). In line with the opinions of some researchers, this thesis suggests that bilingualism may be an experience which has an ability to improve children's EF skills through mechanisms of neuroplasticity (Diamond, 2013; Bialystok, 2011a; see section 3.4.1.).

Standardised timing scores on the IC and switching (timing only) tasks improved significantly across time, a result that had not been predicted. At Time 1 participants' recorded mean scores below the population average (see section 5.4.2.). However, by Time 3 participants were performing above population average. If children were to improve scores as a function of age, standardised scores would be expected to remain within the same percentile band across time. As scores increased, it is argued that children's EF skills significantly improved across time for these tasks as a result of practice. Practice effects (PE) may have resulted in children recalling the tasks in their long-term memory (LTM). Subsequently, children may have recalled certain strategies to improve their task performance, resulting in improvements beyond their expected trajectories. A suggested strategy may be that as a result of children recalling the EF task, children may have anticipated stimuli switches (e.g. on switching and IC tasks) resulting in improved performance.

Furthermore, PE may have resulted in native bilingual children performing significantly worse than monolinguals and IE bilinguals at Time 3 of testing. In a study by Basso and colleagues (1999), PE over a 12 month period were evident in some EF tasks (WCST) but not others (TMT). It was suggested that during performance on the WCST, participants retained information relating to the task which they could apply to test-taking strategies, resulting in enhanced performance (e.g. that the order of sorting principles followed a particular pattern). They also suggested that the TMT task was not subject to significant PE due to task complexity issues as the TMT is not thought to require strategies as complex as the WCST. Furthermore, the WCST has also been shown to require a higher degree of novelty and subsequently this thesis considered issues of PE prior to testing, employing the WCST only once during the longitudinal study (Miyake et al., 2000; Huizinga & Smidt, 2003; Huizinga & van der Molen, 2007). The improved performance found for the IC and switching task may also have been the result of children's improved test-taking strategies such as planning and anticipatory skills which are shown to improve with age (Czernochowski et al., 2010; Munakata et al., 2011).

Although issues of PE were present in this thesis, the improved results shown by monolingual and IE bilinguals across time may not be a limitation as researchers use practice, along with other techniques to help enhance children's EF skills (e.g. aerobic exercise, noncomputer games, see Diamond & Lee, 2011; see section 2.6.). For instance, Thorell and colleagues (2009) employed computerised training (CogMed, Pearson Education) games to improve the WM skills of typically developing children. Practice was shown to improve EF skills as well as transferring to other WM tasks. These improvements were also maintained by participants in a 6 month follow-up assessment (Holmes et al., 2010; Holmes, Gathercole, & Dunning, 2009). In line with this thesis (CC task) research has also shown that children with initially poor EF skills or from low-SES backgrounds may show the highest level of gains following EF training, demonstrating the strength of practice as a method to reduce the achievement gap (Moffitt et al., 2011; O'Shaughnessy et al., 2003) and EF improvements have been most salient in complex rather than specific EF tasks (Diamond & Lee, 2011). These findings are in line with our research which found that with practice, children may improve their EF skills from below to above population average across a three-year period and significant improvements were also shown for children's unified EF skill (TMT task). One issue with using practice as a method to improve children's EF skills is that benefits do not often transfer to unpractised EF skills and for those that do, the transfer is narrow (Bergman-Nutley et al., 2011; Diamond & Lee, 2011). However, most of the training techniques used to improve children's EF skills focus on the WM component, while this thesis found improvements on IC, switching and unified EF component tasks. Therefore the findings from this thesis may indicate that skills beyond WM may transfer to other EF skills, although to test this theory, it is recommended that future research employ further tasks of each EF skill to examine the transfer effect more closely.

Issues of training and experience as modifiers of EF skill is also shown by the language and SES group effects found in this thesis. Although children were tested at similar ages and time points across a three-year period, the unified EF skill of LGs showed different developmental trajectories, with performance improving at a faster rate for bilinguals than monolingual (see section 5.3.5.1.). Results showed that bilinguals improved their RT performance on the Trails B, Trails Difference and Worlds Difference tasks, all of which required higher levels of cognitive demand than other tasks of EF (e.g. switching). Similarly, low-SES groups showed a slower rate of change in the neutral and congruent conditions of the Stroop task, possibly related to their lower levels of English linguistic skills (PPVT) or their delays in sensory information processing (SIP) which correlate highly with non-verbal IQ as well as lower performance on the switching (accuracy) task across time. Although research has suggested that the bilingual experience may enhance children's EF skills, there have been issues regarding which of the functions are most affected by the bilingual experience. The longitudinal results from this thesis proposed that in line with current research, the enhanced EF improvement in bilinguals is most salient in complex EF tasks, suggesting a bilingual advantage for the unified rather than specific EF skills (e.g. Bialystok, 2011b; Carlson & Meltzoff, 2008; Costa et al., 2009; Kroll & Bialystok, 2013; Prior & MacWhinney, 2010).

This thesis also found that in line with current models (Miyake et al., 2000; Miyake & Friedman, 2012) the EF is both unified and diverse in nature as some but not all EFs were affected by language and SES (see section 7.3. and 7.4.). For example, low-SES groups demonstrated poorer performance on tasks requiring WM but not IC skills. Similarly, LG effects in favour of bilinguals were found for complex rather than more specific EF tasks. A bilingual advantage for tasks assessing the unified EF or more complex EF skills, e.g. TMT, WCST, Worlds Difference task was found. However, on tasks assessing specific EF skills e.g. IC, switching and WM, no significant LG effect was obtained. These results support the EF model of Miyake and Friedman (2012; Miyake et al., 2000) which proposed that EF contains both unified and specific components and that experience may modify each of these components in unique ways. Potential mechanisms for the SES and LG effects are discussed in sections 7.3. and 7.4. respectively.

Correlations between EF tasks (section 5.3.6.) indicated that as expected the unified EF task, TMT correlated with a number of specific EF tasks, e.g. IC and WM, implicating its' ability to assess co-ordinated rather than specific EF skills. Although the WCST outcome variables did not significantly correlate with specific skills it did correlate with the RSPM measure of non-verbal IQ, again implicating its use for more general EF skills. Test-retest reliability coefficients were moderate to high for all EF tasks tested over time, indicating the validity of longitudinal analysis using a battery of EF measures.

7.3. BILINGUALISM AND COGNITIVE DEVELOPMENT

LG findings revealed that bilinguals had superior performance to monolinguals on some but not all EF tasks. The advantages found in favour of bilingual groups were particularly evident in more complex tasks of EF including the TMT and WCST. In the Gaeltacht group, mid-SES participants outperformed all other groups on WM and WCST although findings were most prevalent on the WCST. For this measure the MSN group not only performed better than other the groups tested, but mean perseverative responses and errors were in the 87th percentile band, far above population average. Furthermore, this group performed worse compared with IE bilinguals and monolinguals on tasks of IC and switching and maintained equivalent performance on the TMT tasks of unified EF (for a discussion of these results see section 6.6.2.). In the longitudinal study, the LG effect in favour of bilinguals was not present at Time 1 but emerged over time where by Time 3, bilinguals outperformed monolinguals on complex tasks of EF in particular e.g. TMT, WCST and Worlds Difference task.

These findings are significant as recent research has begun to shift its explanation of bilingual advantages for EF tasks towards mechanisms utilising a unified EF component (e.g. Adesope et al., 2010; Bialystok, 2011; Kroll & Bialystok, 2013; Paap & Greenberg, 2013). Initially, researchers postulated that specific EFs such as IC and switching were the areas advantaged by the bilingual experience (e.g. Bialystok, 2001; Bialystok, 2009; see section 3.4.1.). Mechanisms for these effects included the dual-language activation hypothesis which suggested that proficient bilinguals must constantly monitor two languages in the mind which are both active regardless of linguistic context (Bialystok, 2011; Kroll & deGroot, 1997; Kroll et al., 2012; Marian et al., 2003). To adequately control both languages, it has been shown (through behavioural, clinical populations and fMRI studies) that bilinguals utilise areas of the brain known to include EF processes, specifically the frontal lobes (Abutalebi et al., 2011; Abutalebi & Green, 2008; Bialystok et al., 2012; Luk et al., 2012). As a result, children become more practiced using specific EF functions such as IC as they must rely on this mechanism to inhibit the non-intended language during speech production and task performance (Green, 1998). Similarly, bilinguals must utilise the switching EF to switch between languages being encoded and produced depending on linguistic context (Costa, 2005; Costa et al., 1999; Meuter & Allport, 1999). Through the practice of using IC and switching to control the L1 and L2, bilinguals demonstrate an advantage for these skills during non-linguistic EF tasks compared to monolingual peers who have not had to utilise these skills as much during language production. Recent studies have found that rather than such specific functions being advantaged, bilinguals utilise a number of EFs during dual-language control. Furthermore, issues of EF task impurity have raised questions regarding the specific nature of the bilingual advantage (e.g. Anderson, 2002; Burgess et al., 1998; see section 2.7.). As bilinguals must monitor, inhibit, attend to and switch between both languages in day-to-day life, it makes sense that unified rather than specific EFs should be enhanced by the bilingual experience. This is also in line with Miyake and colleagues' updated model of EF (Miyake & Friedman, 2012) which suggests a common core that is shared between EF subcomponents. In other words, no function is used in isolation when controlling two languages, therefore bilinguals may be better able to co-ordinate their EFs during complex tasks such as the WCST (Kroll & Bialystok, 2013). This theory of a bilingual advantage in co-ordinating EFs during task performance is in line with this thesis which found that all bilinguals, bar the low-SES Gaeltacht group, outperformed monolinguals on the WCST and that an advantage for the IE bilinguals emerged over time on the TMT.

As the bilingual advantage emerged over time with IE bilinguals improving certain EF skills at a faster rate than their monolingual peers, the developmental trend in favour of bilinguals may be the result of their increased experience with the L2, which in turn, increased their practice of using EFs to control and monitor both languages. The concept of bilingualism has been considered a matter of degree rather than a categorical variable (Carlson & Meltzoff, 2008). Therefore, it would be expected that with increased exposure to the L2 during IE, children's LP in the L2 will improve with age. Furthermore, the improved results in favour of IE bilinguals across time are not unsurprising as increased bilingualism and L2 proficiency have previously been shown to modify cognitive outcomes (Bialystok et al., 2012; Kroll & Bialystok, 2013). To follow up on issues of LP, Chapter 6 split the IE group into low, moderate and high proficiency groups. Results found a minimal difference between groups although the high proficiency participants differed in their levels of English vocabulary skills. As proficiency increased, so too did children's English receptive skills. This finding was in line with Cummins' interdependence hypothesis (1979) which stated that L1 skills can benefit from exposure to the L2 as an interplay between languages resulting in one language supporting the development of the other (so long as the L1 is well developed prior to L2 exposure). Furthermore, Perani and colleagues (1999) found that proficiency rather than age of acquisition was more important for brain development. Although English skills were enhanced in high proficiency bilinguals, LP differences were only observed in the Irish versions of EF tasks. This indicates that the time spent within IE may be more predictive of EF skills than the levels of Irish productive skill.

Furthermore, Gaeltacht or native bilingual groups did not perform better on tasks of IC or switching. As results from Chapter 5 showed no bilingual advantage on these tasks results may be unsurprising, as native bilinguals had less experience in performing tasks compared to participants in the longitudinal study. As native bilinguals also had more exposure to Irish than IE bilinguals, they may have experienced more interference from the L2 than IE bilinguals, delaying their RTs. However, what was surprising was that mid-SES native bilinguals outperformed all groups on non-verbal WM and unified EF tasks. One hypothesis for this result was that these tasks had no linguistic-stimuli compared with the other EF tasks, resulting in a LG effect in favour of native bilinguals. However, if this was the case and linguistic stimuli were affecting native bilingual performance, then a higher result would have been expected from both the IE and low-SES native groups. Instead, it is proposed that through their enhanced experience and exposure to Irish, mid-SES native bilinguals and IE bilinguals (in Chapter 5) developed unified EF skills through having to control and monitor two active languages in the minds, subsequently improving their performance.

The poor performance of low-SES native bilinguals highlighted the importance of factors beyond language and IE which may have affected EF performance. As this school was situated in a rural area of Ireland and was designated as Deis, it automatically placed this group at a disadvantage relative to all other groups (see

4.3.2.). Furthermore, language background questionnaires revealed that this group had reduced exposure to Irish within the home and school environments which may have resulted in poorer L2 proficiency relative to IE and mid-SES native bilingual groups. Kroll and Bialystok (2013) made the point that bilingualism is a multidimensional experience depending on many factors such as linguistic, social, educational and SES environments. They suggested that there should be careful consideration of language context and interactions between environmental influences prior to making conclusions regarding the nature of bilingualism. This thesis also concluded that children’s language and environmental contexts are crucial for any study of bilingualism and especially this sample. The differences between IE and Gaeltacht bilingual groups are discussed further in section 7.5.

7.4. SOCIOECONOMIC STATUS AND COGNITIVE DEVELOPMENT

Findings showed that SES plays an important role in the development of children’s EF skills, however it appears that certain EFs more than others are affected by SES backgrounds. Furthermore, the two variables which demonstrated the strongest effect of SES were the control variables English language skills (PPVT) and non-verbal IQ (RSPM). These findings are in line with studies which have shown that SES strongly affects children’s language and IQ skills (Mezzacappa, 2004; Hackman & Farah, 2009; Noble et al., 2007).

A positive outcome from this research was that although low-SES children demonstrated delays on a number of EF measures including WM and unified EF (TMT), their developmental trajectories did not differ from mid-SES groups. Instead, it appeared that the low-SES children’s development is delayed rather than being disadvantaged and improves at the same rate as children from mid-SES backgrounds. On the switching and IC tasks, low-SES monolingual children in particular significantly improved their performance across time. However, on speeded tasks (congruent and neutral Stroop RTs and Trails RTs) mid-SES groups improved at a significantly faster rate than low-SES groups. RTs may have been slower in low-SES groups as a result of speed of information processing (SIP) which has also shown a close link with general IQ skills (e.g. Bonifacci et al., 2011) and therefore disadvantaged in the low-SES groups. WM performance in particular was reduced in low-SES groups with poorer performance on verbal and visuo-spatial WM tasks as well as on the TMT which has been shown to correlate highly with WM skills (e.g. Alloway & Alloway, 2010). The WM difference may relate to the close relationship between WM and IQ, a theory in line with these results (Alloway & Alloway, 2010). SES groups did not differ on the complex WCST task. This lack of difference may have been the result of the task being untimed, providing low-SES children with an opportunity to consider potential strategies during performance. One implication of this finding is that future researchers should consider using untimed tasks when assessing the EF skills of children from low-SES backgrounds, as timed tasks may place these children at an automatic disadvantage. Furthermore, the skills most likely to show a deficit in low-SES samples require language and WM skills and can be related to general IQ. EF skills such as switching, IC and unified EF, do not display such disadvantages, which indicates that the EF potential children from low- and mid-SES backgrounds may be similar.

LG differences were present in some but not all EF tasks in the low-SES group. For instance, LSI bilinguals performed significantly better than LSMs on the Trails Difference and Worlds Difference conditions but equivalently on all other tasks. Participants in the LSN group performed worse on IC, switching and unified EF tasks compared to monolinguals and IE bilinguals. This demonstrated that although IE bilinguals had enhanced abilities in overcoming the additional cognitive demand necessary to perform the more complex conditions, differences may depend on factors beyond language experience, such as home environments. A current issue within many Gaeltacht schools is that more and more English is being spoken during school hours as a result of societal changes, not all of which are purely linguistic (Mercator European Research Centre on Multilingualism and Language Learning, 2008). The report by the Mercator group discussed how the linguistic background of pupils within the Gaeltacht is varied and changing, particularly with more pupils coming from returned Gaeltacht emigrant families brought up abroad. This finding is line with this research which found that in the LSN group, children were speaking a mixture of English and Irish with their friends in school compared with other groups who spoke mostly Irish with friends during school. Furthermore, a number of the children in the LSN group were children of immigrant families (see Table 37). These factors must be taken into consideration when evaluating the cognitive impact of IE in Gaeltacht areas and are discussed in more detail in the following section.

7.5. IMMERSION EDUCATION AND BILINGUALISM

Unlike early studies of bilingualism which found that the BE experience may negatively impact on children's academic and linguistic performance in the L1, this thesis found that children within Irish-medium IE show equivalent non-verbal IQ and English skills to their monolingual counterparts. Previous studies with Irish-IE children have also found no delays in IE children's academic and language abilities, therefore IE should be considered as a viable academic environment for fostering bilingual children (Harris et al., 2006; Nicolay & Poncelet, 2013; Ó Duibhir, 2009; Parsons & Lyddy, 2009).

For the Irish IE case this thesis found that IE in general should, at least, not delay the EF development of children. Furthermore, children who had been attending IE for a minimum of 7 years displayed enhanced performance on complex EF tasks, although advantages may depend on the nature of the task itself. This thesis argues that the use of participants within IE programmes in urban and majority L1 environments may be more generalizable to IE populations globally, particularly if the groups are selected from mid-SES backgrounds. A similar study by Nicolay and Poncelet (2013a) found that EF advantages can be obtained through IE despite children's L2 exposure being sequential rather than simultaneous and despite the L1 being the dominant language. They also argued that a lack of balance between the L1 and L2, as in the IE case, requires bilinguals to rely more on their EF skills, resulting in advantages for complex EF tasks, as was shown in this thesis. Although participants with higher levels of LP within the IE groups of this thesis also showed higher levels of L1 skills, the time spent in IE was more important in predicting children's EF performance than LP categorisation. Furthermore, indicators such as SES and linguistic environments are influential on EF outcomes and should not be overlooked when assessing the success of IE, as differences were found between children in urban IE programmes and rural Gaeltacht IE programmes.

In the Gaeltacht, factors beyond language experience may impact on children's EF development, as children in the low-SES Gaeltacht group performed worse on a number of EF tasks compared with other groups. The Gaeltacht case is a unique linguistic context and therefore this thesis suggests it should be treated as such by future researchers. Recently, the Gaeltacht has been struggling to maintain the level of Irish traditionally spoken in these communities as well as in school, as English becomes the language of work in these areas. English is also increasingly being used in the homes of Gaeltacht families and therefore levels of Irish competency and experience varies between children entering primary school, creating challenges for teachers who are supposed to teach through the medium of Irish only (Ó hIfearnáin, 2007). Hickey (1999) discussed the demands on Gaeltacht teachers to maintain the native-Irish abilities in some children while increasing the level of Irish in children with minimum exposure to Irish. Reports have also shown that Gaeltacht schools vary in their practice of language instruction, depending on the proportion of pupils fluent in Irish on entry and the policies of the schools themselves towards Irish (COGG; Mac Donnacha et al., 2005). These findings are in line with this thesis as in Gaeltacht school 1 (mid-SES native), Irish only was spoken during school hours and anecdotally, a higher degree of Irish was also used within the community. However, in school 2, only a 10 minute drive from school 1, a higher level of English was heard and spoken in the school yard (low-SES native) and participants in this group reported speaking more English with their friends in school and a lower frequency of Irish within the home (see Thomas & Roberts, 2011 for similar issues of school-yard language mixing). These groups also displayed different outcomes on EF tasks and the unified EF advantage in particular was found for school 1 but not school 2. This thesis would argue that the more a school adopts a single language of instruction (Irish), the greater the opportunity for linguistic and non-linguistic gains, particularly if English is spoken at a higher frequency within the home and community. However, group differences may have occurred as a result of aggregate-SES differences between these schools, although no such differences were found at an individual-level. Baker (1993) commented that there is a danger in placing too much reliance on the bilingual education environment as the salvation for heritage languages, such as Irish. He also argued that influences within and outside of school such as television, 'pop' culture and the information technology revolution provide children with additional skills in the majority language (English) and that IE alone cannot reverse this trend. Baker argued that the L2 (Irish) should be promoted within the community and as part of the children's daily lives. In line with his arguments, this thesis suggests that for the Gaeltacht in particular, communities should try to encourage their children to use Irish as a means of communication outside of the classroom (e.g. within the school yard, while with friends at home) if linguistic and cognitive advantages are to be gained.

7.6. IMPLICATIONS OF RESEARCH

In terms of its contribution to EF research, the longitudinal study in this thesis found that EF skills may be independent of one another but that a unified EF component does exist and has a unique developmental trajectory to other EF skills. These findings are in line with the EF model proposed by Miyake and Friedman (2012), adding to the growing body of research suggesting the existence of a unified or more complex EF component.

In line with developmental studies, this thesis also found that EF develops as a function of age from 9 to 11 years and that the EF demonstrates a protracted development relative to other cognitive skills. Sarsour and colleagues (2011) pointed out that middle childhood is a period of increased intellectual challenges, increased independence and relationships with peers. As a result, the lack of research during middle childhood is unfounded and researchers should cease their disregard for this period of cognitive development as it may provide important insights into children and young adults' cognitive development.

SES and LG effects were found for some but not all EFs and the bilingual effect was significant on both unified EF tasks. The LG findings were in line with some of the most recent papers on bilingualism suggesting that EF advantages for bilinguals involve a number of EF processes rather than being the result of enhanced skills for one specific EF component (e.g. Bialystok, 2011a; Bialystok et al., 2012; Colzato et al., 2008; Nicolay & Poncelet, 2013; Sarsour et al., 2011). This thesis proposes that future bilingualism research examines the EF skills of simultaneous and sequential bilinguals using complex, non-linguistic tasks to examine the limits of the bilinguals' enhanced unified EF skills. Furthermore, research must consider the independent effects of SES and LG. A study by Calvo (2011; cited in Kroll & Bialystok, 2013) also showed the independent influence of bilingualism and SES on a group of simultaneous (5 year old) bilinguals whose language and attention abilities were primarily affected by SES and whose EFs were primarily affected by bilingualism.

The thesis found that LP ratings were indicative of L1 skills but not EF skills. These results highlight the need for standardised assessments of Irish productive and receptive skills as the nature of LP questionnaires used may not have given accurate indicators of children's LP skill. Gaeltacht teachers in particular have been petitioning for more resources in Irish and particularly for standardised Irish assessments for their pupils (Coady & Ó Laoire, 2002; Mercator, 2008; Ó hIfearnáin, 2007; Parsons & Lyddy, 2009). The LP findings did give support to Cummins' interdependence hypothesis (1979) which postulated that L1 abilities can be improved with increased exposure to the L2. Another implication of this thesis is the importance of considering the Gaeltacht and IE linguistic experiences independently when drawing conclusions, as a wide range of socio-political, linguistic and environmental factors appear to vary between these groups and not simply children's levels of LP (e.g. Ó Duibhir, 2009). While the findings from IE research may be generalizable to the global IE picture, results from Gaeltacht group were drawn from a significantly smaller sample size. Furthermore, even though the Gaeltacht schools tested were only 10 minutes apart from one another, EF results differed between schools.

The impact of EF on children's lifelong skills is currently a topical area of research (e.g. Diamond, 2013; Diamond et al., 2011). EF researchers have been working to find methods and experiences to improve children's EF skills in the hope of improving long-term outcomes. This thesis would suggest that bilingualism is one of these experiences. Recently, correlations have found a significant relationship between complex EF and academic achievement including maths and reading in children between 5 to 17 years (Best et al., 2011). Furthermore, practice with EF tasks results in improvement over time, although the strategies employed by participants to improve their results are unclear. Although the PE on EF tasks have previously been shown to have a low level transfer to skills lifelong skills (e.g. Diamond & Lee, 2011) bilingualism may be an experience which enhances a number of EF skills and subsequently may transfer to lifelong skills such as academic ability as this language experience appears to enhance children's unified EF in particular.

7.6. LIMITATIONS AND FUTURE DIRECTIONS

- One of the main limitation of this thesis was that certain EF tasks, and the colour-word Stroop task in particular, used linguistic stimuli (colour words), which may have complicated LG effects. For instance, the Gaeltacht group showed poor performance on tasks with linguistic stimuli (numbers, colour-words) but showed equivalence or an advantage on tasks requiring no linguistic stimuli (e.g. WM, WCST). Although numerous bilingualism studies have utilised tasks such as the Stroop (e.g. Bialystok, 2001, 2009), Costa and colleagues (2009) have argued against its use with bilingual samples. Although it is suggested that future research utilise non-linguistic tasks, bilingual advantages still emerged for tasks requiring letters and numbers, e.g. TMT and Worlds Difference and low-SES Gaeltacht participants did not perform better on assessments using non-linguistic stimuli.

- Another limitation of this study was the nature of certain task paradigms. For instance in the Creature Count task of switching, accuracy scores were gathered from 7 trials only. As a result, any error made within a trial was discounted and marked incorrect. Subsequently, participants who may have performed well on the task overall, still received low accuracy scores. Similarly the nature of the Stroop task made it difficult to draw conclusions from results. Previous research has found an advantage for bilinguals on all conditions of the Stroop task (e.g. Bialystok, 2001; Bialystok, 2009; Bialystok et al., 2008; Hernández, 2010). However, performance on the Stroop task is heavily influenced by children’s reading and automaticity skills (Stroop, 1935). Therefore, if a child’s reading skills are poor, than performance on the Stroop task may be increased as children have less difficulty ignoring colour-words and naming the ink colour. On the other hand, if children’s ability to inhibit a dominant response is enhanced, they too will show an advantage on incongruent trials. In other words, previous findings implicating a bilingual advantage on the Stroop task does not tell us whether the children had poorer reading skills or were better at inhibiting responses. Similarly, as children in this study showed no effect of LG over time on the Stroop, this result may be a result of their higher reading levels (equivalent on PPVT with monolinguals) or a lack of an IC advantage. Furthermore, as bilingualism has been shown to disadvantage children's lexical retrieval skills, the lack of an advantage on the Stroop may have been the result of the delay caused by children having to select between the L1 and L2 lexicons. This thesis suggests that future research utilises IC tasks which do not involve linguistic material (e.g. flanker, ANT) to decipher LG effects and due to the range of factors required, it is recommended that the Stroop task paradigm not be used as a measures of IC in bilinguals (see also Tao et al., 2011; Verleger et al., 2009).

- The subjective nature of the parent/caregiver, teacher and participants' LP questionnaires meant that accurately assessing children's Irish and English skills was difficult. Furthermore, parents/caregivers and teachers were comparing children against other children in their class rather than the population as a whole. This made it difficult to objectively decipher the LP skills of individuals, particularly in the Gaeltacht groups as parents and teachers were not considering the levels of Irish/English spoken by the population as a whole. This issue requires standardised assessment to assess children's Irish skills and such resources are in limited supply within Ireland.

7.8. CONCLUDING REMARKS

This research has shown that certain language experiences, and in particular, the immersion education (IE) experience has the ability to enhance children's executive function (EF) skills. However, contrary to widely held opinion that bilingualism impacts positively on specific EFs such as IC, switching and WM, this thesis found that, in line with recent literature, the bilingual advantage may lie in their ability to utilise a number of EF skills in collaboration with one another. This potential advantage in general EF skills for IE bilinguals may not be present during the primary stages of schooling but instead, emerges over time. Furthermore, children attending disadvantaged or low-SES IE schools are not hindered by their bilingual experience. Finally, children within Gaeltacht or Irish speaking areas of Ireland may also witness these unified EF advantages although this linguistic environment is unique and therefore factors beyond education must be considered. In conclusion, this research suggests that the experience of immersion education and using two languages can foster the development of children's unified executive functioning skills.

REFERENCES

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242-275.

Abutalebi, J., Brambati, S. M., Annoni, J. M., Moro, A., Cappa, S. F., & Perani, D. (2007). The neural cost of the auditory perception of language switches: An event-related functional magnetic resonance imaging study in bilinguals. *The Journal of Neuroscience*, *27*(50), 13762-13769.

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*(2), 207-245.

Adi-Japha, E., Berberich-Artzi, J., & Libnawi, A. (2010). Cognitive flexibility in drawings of bilingual children. *Child Development*, *81*(5), 1356-1366.

Adler, M. K. (1977). *Collective and individual bilingualism: a sociolinguistic study*. Hamburg: Helmet Buske Verlag.

Albert, M.L & Obler, L.K. (1978). *The bilingual brain*. New York: Academic Press.

Allman, T. M. (2002). Patterns of spelling in young deaf and hard of hearing students. *American Annals of the Deaf*, *147*(1), 46-64.

Alloway, T. P. (2007). *Automated Working: Memory Assessment: Manual*. Pearson.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20-29.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and Visuospatial Short-Term and Working Memory in Children: Are They Separable?. *Child Development*, *77*(6), 1698-1716.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, 87, 85–106.

Amso D, Casey BJ. (2006). Beyond what develops when: Neuroimaging may inform how cognition changes with development. *Current Directions in Psychological Science*, *15*, 24–29.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological Review*, *102*(1), 68.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71-82.

Anderson, V. (1998). Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychological Rehabilitation*, *8*(3), 319-349.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental neuropsychology*, *20*(1), 385-406.

Anisfeld, E. A. (1964). A comparison of the cognitive functioning of monolingual and bilinguals. Unpublished doctoral dissertation, McGill University.

Anokhin, A. P., Golosheykin, S., Grant, J. D., & Heath, A. C. (2010). Developmental and genetic influences on prefrontal function in adolescents: a longitudinal twin study of WCST performance. *Neuroscience letters*, *472*(2), 119-122.

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of clinical and experimental neuropsychology*, *22*(4), 518-528.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*(2), 115-129.

Ardila, A. (2007). Bllingualism in the contemporary world. *Speech and language disorders in bilinguals*, 1.

Argyri, E., & Sorace, A. (2007). Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, *10*(01), 79-99.

Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs*, *60*(1).

Asato, M. R., Sweeney, J. A., & Luna, B. (2006). Cognitive processes in the development of TOL performance. *Neuropsychologia*, *44*(12), 2259-2269.

August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language-minority children: A research agenda*. National Academies Press.

Awh E, Anllo-Vento L, Hillyard SA (2000) The role of spatial selective attention in working memory for locations: evidence from event related potentials. *Journal of Cognitive Neuroscience 12*(5):840–847.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in cognitive sciences*, *5*(3), 119-126.

Axelrod, B. N., Greve, K. W., & Goldman, R. S. (1994). Comparison of four Wisconsin Card Sorting Test scoring guides with novice raters. *Assessment*,*1*(2), 115-121.

Axelrod, B. N., Jiron, C. C., & Henry, R. R. (1993). Performance of adults ages 20 to 90 on the abbreviated Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*, *7*(2), 205-209.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 5-28.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1996). Exploring the central executive. Quarterly Journal of Experimental Psychology, 49A, 5–28.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485.

Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. *Models of working memory: Mechanisms of active maintenance and executive control*, 28-61.

Baddeley, A.D. (1986). *Working memory.* Oxford: Oxford University Press

Baddeley, A.D. and Hitch, G.J. (1974) Working memory. In The *Psychology of Learning and Motivation* (Bower. G.A., ed.), pp. 47-89, Academic Press

Baetens-Beardsmore, H. (ed.) (1991) *European Models of Bilingual Education*. Clevedon: Multilingual Matters.

Bain, B. (1974). Bilingualism and cognition: Towards a general theory. In *Bilingualism, biculturalism, and education Proceedings from the Conference at College Universitaire Saint Jean* (pp. 119-128). Edmonton: The University of Alberta.

Bajo, M. T., Padilla, F., & Padilla, P. (2000). Comprehension processes in simultaneous interpreting. *Benjamins Translation Library*, *39*, 127-142.

Baker, C. (1988). *Key issues in bilingualism and bilingual education* (Vol. 35). Multilingual matters.

Baker, C. (1997). Deaf children: Educating for bilingualism. *Deafness and Education, 21*, 3-9.

Baker, C. (2003). 6. Education as a Site of Language Contact. *Annual Review of Applied Linguistics*, *23*, 95-112.

Baker, C. (2011). *Foundations of bilingual education and bilingualism* (Vol. 79). Multilingual matters.

Baker, C. R., & Jones, S. P. (Eds.). (1998). *Encyclopedia of bilingualism and bilingual education*. Multilingual Matters.

Baker, C. R., & Sienkewicz, A. (2000). *The care and education of young bilinguals: An introduction for professionals*. Multilingual Matters.

Baker, C., & Jones, M. P. (2000). Welsh language education: A strategy for revitalization. *Language revitalization: Policy and planning in Wales*, 116-137.

Balkan, L. (1970). *Les effets du bilinguisme Français-Anglais sur les aptitudes inteleectuelles*. Bruxelles: AIMAV.

Bamford, K. W., & Mizokawa, D. T. (1989). Cognitive and Attitudinal Outcomes of an Additive-Bilingual Program.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child development*, *83*(2), 413-422.

Barceló, F. (2003). The Madrid card sorting test (MCST): a task switching paradigm to study executive attention with event-related potentials. *Brain Research Protocols*, *11*(1), 27-37.

Barceló, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia*, *40*(3), 349-356.

Barceló,F. (1999). Electrophysiological evidence of two different types of error in the Wisconsin Card Sorting Test. *Neuroreport* 10, 1299–1303.

Barik, H. C., & Swain, M. (1976). A longitudinal study of bilingual and cognitive development. *International Journal of Psychology*, *11*(4), 251-263.

Barke, E. M. (1933). A study of the comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *British Journal of Educational Psychology*, *3*(3), 237-250.

Barke, E. M., & Williams, D. P. (1938). A further study of the comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *British Journal of Educational Psychology*, *8*(1), 63-77.

Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., & Metevia, L. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *Journal of abnormal child psychology*, *29*(6), 541-556.

Bastarache, M. (1987). Bilingualism and the Judicial System. *Language Rights in Canada*, 123-174.

Baumeister, R.F., & Vohs, K.D. (2004). *Handbook of self-regulation: Research, theory, and applications*. New York: Guilford Press.

Becker, M., Isaac, W., & Hynd, G. (1987). Neuropsychological development of nonverbal behaviors attributed to ‘‘Frontal Lobe’’ functioning. *Developmental Neuropsychology, 3,* 275-298.

Bedard AC, Nichols S, Barbosa JA, Schachar R, Logan GD, Tannock R. (2002) The development of selective inhibitory control across the life span. *Developmental Neuropsychology, 2,* 93-111.

Ben-Zeev, S. (1977a). The influence of bilingualism on cognitive strategy and cognitive development. *Child development*, 1009-1018.

Ben-Zeev, S. (1977*b*). The effect of Spanish-English bilingualism in children from less privileged neighborhoods on cognitive development and cognitive strategy. *Working Papers on Bilingualism, 14,* 83-122.

Berko, J. (1958). *The child's learning of English morphology* (Doctoral dissertation, Radcliffe College).

Berlin, L., Bohlin, G., & Rydell, A. M. (2004). Relations between inhibition, executive functioning, and ADHD symptoms: A longitudinal study from age 5 to 8½ years. *Child Neuropsychology*, *9*(4), 255-266.

Best JR, Miller PH, Jones LL. Executive functions after age 5: Changes and correlates. Developmental Review. 2009; 29:180–200.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641-1660.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327-336.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 57,* 498–510.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, *24*, 560–567.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology*, 60, 68–79.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(01), 3-11.

Bialystok, E. (2010). Global–local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental psychology*, *46*(1), 93.

Bialystok, E. (2011). Reshaping the mind: the benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *65*(4), 229.

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*(1), 67-73.

Bialystok, E., & Craik, F. I. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science*, *19*(1), 19-23.

Bialystok, E. & Feng, X. (2011). Language proficiency and its implications for monolingual and bilingual children. *Language and literacy development in bilingual settings*, 121-138.

Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics*, *19*(01), 69-85.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental science*, *7*(3), 325-339.

Bialystok, E., & Senman, L. (2004). Executive processes in appearance–reality tasks: the role of inhibition of attention and symbolic representation. *Child development*, *75*(2), 562-579.

Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science, 8*(6), 595-604.

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*(3), 494-500.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*, 89–129.

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*(2), 459-464.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, *16*(4), 240-250.

Bialystok, E., Craik, F. I., & Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: The effects of bilingualism and ageing. *The Quarterly Journal of Experimental Psychology*, *59*(11), 1968-1983.

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1341.

Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, *24*(1), 40-49.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging*, *19*(2), 290.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 859.

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, *24*(01), 27-44.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*(1), 103-119.

Bialystok, E., Peets, K., & Moreno, S. (2012). Producing bilinguals through immersion education: Development of metalinguistic awareness. *Applied Psycholinguistics*, *1*(1), 1-15.

Blair, C., Razza, R.P. (2007). Relating effortful control, executive function, and false-belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 782*, 647—663.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology*, *28*(2), 561-571.

Bloomfield, L. (1933) *Language*. New York: Holt, Rhinehart and Winston.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition*, *118*(2), 245-257.

Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental science*, *14*(2), 256-269.

Boone, K. B., Miller, B. L., Lesser, I. M., Hill, E., & D'Elia, L. (1990). Performance on frontal lobe tests in healthy, older individuals. *Developmental Neuropsychology*, *6*(3), 215-223.

Bornstein, M. H., Hahn, C. S., Suwalsky, J. T., & Haynes, O. M. (2003). The Hollingshead four-factor index of social status and the socioeconomic index of occupations. *Socioeconomic status, parenting, and child development*, 29-81.

Bosco, J. (1972). The visual information processing speed of lower-and middle-class children. *Child Development*, 1418-1422.

Bournot-Trites, M., & Tellowitz, U. (2002). *Report of current research on the effects of second language learning on first language literacy skills*. Printing House.

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*, *8*(3), 243-254.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual review of psychology*, *53*(1), 371-399.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, *39*(4), 713-726.

Brennan-Wilson, A. (2013). Cognitive development of low- and typically- achieving monolingual and bilingual children. Unpublished PhD Dissertation. Queen's University, Belfast.

Brisk, M. E. (1998). *Bilingual education: From compensatory to quality schooling.* Mahwah, NJ: Lawrence Erlbaum.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental neuropsychology*, *26*(2), 571-593.

Brooks-Gunn, J., Denner, J., & Klebanov, P. K. (1995). Families and neighborhoods as contexts for education. *Changing populations, changing schools: Ninety-fourth yearbook of the National Society for the Study of Education, Part II*, 233-252.

Brooks-Gunn, J., Duncan, G., & Aber, J. L. (Eds.). (1997). *Neighborhood poverty. 1. Context and consequences for children* (Vol. 1). Russell Sage Foundation.

Brown, V. J., & Bowman, E. M. (2002). Rodent models of prefrontal cortical function. *Trends in neurosciences*, *25*(7), 340-343.

Bruck, M. (1978). The suitability of early French immersion programs for the language disabled child. *Canadian Journal of Education, 3,* 51-72.

Bruck, M. (1982). Language disabled children: Performance in an additive bilingual education program. *Applied Psycholinguistics, 3,* 45-60.

Bruck, M., & Genesee, F. (1995). Phonological awareness in young second language learners. *Journal of child language*, *22*, 307-307.

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of the performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45*, 743–754.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental neuropsychology*, *33*(3), 205-228.

Bunge, S. A., & Crone, E. A. (2009). Neural correlates of the development of cognitive control. *Neuroimaging in developmental clinical neuroscience*, 22-37.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J. & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*, 301–311.

Burgess, P. W. (1997). Theory and methodology in executive function research. *Methodology of frontal and executive function*, 81-116.

Burke, H. R. (1972). Raven's Progressive Matrices: Validity, reliability and norms. *Journal of Psychology, 82*, 253–257

Butler, Y. G., & Hakuta, K. (2004). Bilingualism and second language acquisition. *The handbook of bilingualism*, 114-144.

Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *The Journal of Educational Research*, *90*(5), 269-277.

Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, *13*(1), 61-68.

Carlson, S. M. (2003). Executive function in context: Development, measurement, theory, and experience. *Monographs of the Society for Research in Child Development 2003, 68*(3):138–151.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental neuropsychology*, *28*(2), 595-616.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental science*, *11*(2), 282-298.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Developmental psychology*, *40*(6), 1105.

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,* 404– 431.

Carrasquillo, A. L., & Rodr̂aiguez, V. (2002). *Language minority students in the mainstream classroom* (Vol. 33). Multilingual Matters.

Carringer, D. C. (1974). Creative Thinking Abilities of Mexican Youth The Relationship of Bilingualism. *Journal of Cross-Cultural Psychology*, *5*(4), 492-504.

Carrow, W. M. A. (1957). Linguistic functioning of bilingual and monolingual children. *Developmental Psychology, 5,* 398-405.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of experimental child psychology*, *33*(3), 386-404.

Casey, B. J., Tottenham, N., Liston, C., Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Science, 9*(3), 104–110.

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., et al. (1997). A developmental functional MRI study of prefrontal activation during performance of a go-no-go task. *Journal of Cognitive Neuroscience. 9,* 835–847.

Cazden, C. B. (1974). Play and metalinguistic awareness: One dimension of language experience. *The Urban Review*, *7*(1), 28-39.

Cenoz, J. 2009. *Towards Multilingual Education: Basque Educational Research in International Perspective.* Clevedon: Multilingual Matters.

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: examination of task-switching performance. *Developmental psychology*, *37*(5), 715.

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201-216.

Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: FMRI and meta-analytic evidence of a domain-general control network for learning. *Brain Research: Cognitive Brain Research, 25,* 607–623.

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting test. *Journal of Clinical and Experimental Neuropsychology*, *8*(3), 219-228.

Chomsky, N. (1965). *Aspects of the Theory of Syntax* (Vol. 11). The MIT press.

Chomsky, N. (1995). Language and nature. *Mind*, *104*(413), 1-61.

Christ, S. E., White, D. A., Mandernach, T., & Keys, B. A. (2001). Inhibitory control across the life span. *Developmental Neuropsychology, 20,* 653–669.

Chugani, H. T., Phelps, M. E., Mazziotta, J. C. (1987). Positron emission tomography study of human brain functional development. *Ann Neurol*. **22**, 487–497.

Cicerone, K. D., & Azulay, J. (2002). Diagnostic utility of attention measures in postconcussion syndrome. *The Clinical Neuropsychologist*, *16*(3), 280-289.

Clearfield, M. W., Diedrich, F. J., Smith, L. B., Thelen, E. (2006). Young infants reach correctly in A-not-B tasks: on the development of stability and perseveration. *Infant Behavior and Development*, *29,* 435–444.

Coady, M., & Laoire, M. Ó. (2002). Mismatches in language policy and practice in education: The case of Gaelscoileanna in the Republic of Ireland. *Language Policy*, *1*(2), 143-158.

Cloud, N., Genesee, F. & Hamayan, E.V. (2000). *Dual language Instruction: A Handbook for Enriched Education.* Boston: Heinle & Heinle.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by< i> g</i>. *Intelligence*, *32*(3), 277-296.

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of memory and language*, *45*(4), 721-736.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 302.

Connor, A., Franzen, M. D., & Sharp, B. (1988). Effects of practice and differential instructions on Stroop performance. *International Journal of Clinical Neuropsychology*.

Corrigan, J. D., & Hinkeldey, N. S. (1987). Relationships between parts A and B of the Trail Making Test. *Journal of clinical psychology, 43*(4), 402-409.

Costa, A. (2005). Lexical access in bilingual production. *Handbook of Bilingualism*, 308.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition, 2*(3), 231–244.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491-511.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59-86.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135-149.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, *41*(3), 365-397.

Cragg, L., & Nation, K. (2008). Go or no-go? Developmental improvements in the efficiency of response inhibition in mid-childhood. *Developmental Science*, *11*(6), 819-827.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in cognitive sciences*, *10*(3), 131-138.

Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: Analysis and control in young Swedish-English bilinguals. *Applied Psycholinguistics*, *20*(1), 1-20.

Crone, E. A., Bunge, S. A., Van Der Molen, M. W., & Ridderinkhof, K. R. (2006). Switching between tasks and responses: A developmental study. *Developmental Science*, *9*(3), 278-287.

Crone, E. A., Ridderinkhof, R. K., Worm, M., Somsen,R .J. & VanDer Molen, M. W. (2004). Switching between spatial stimulus–response mappings: a developmental study of cognitive flexibility. *Developmental Science, 7,* 443–455.

Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L. & Bunge, S. A. (2006): Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences, 103,* 9315–9320.

Crowe, S. F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. *Journal of clinical psychology, 54*(5), 585-591.

Crystal, D. (1997). *The Cambridge Encyclopedia of Language, 2nd edition.* Cambridge: Cambridge University Press.

Csikszentmihali, M. (1999). Implications of a systems perspective for the study of creativity. In R. J. Sternberg (Ed.), Handbook of Creativity (pp. 313-338). New York: Cambridge University Press. Srivastava, B. (1991). Creativity and linguistic proficiency. *Psycho-Lingua, 21,* 105–109.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism*, 9, 1-43.

Cummins, J. (1977). Cognitive factors associated with the attainment of intermediate levels of bilingual skills. *The Modern Language Journal*, *61*(1-2), 3-12.

Cummins, J. (1978a). Bilingualism and the development of metalinguistic awareness. *Journal of cross-cultural psychology*, *9*(2), 131-149.

Cummins, J. (1978b). Educational Implications of Mother Tongue Maintenance in Minority Language Groups. *Canadian Modern Language Review*, *34*(3), 395-416.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49, 222-51. (Reprinted in the National Dissemination and Assessment Center, Bilingual Education Paper Series, September, 1979).

Cummins, J. (1983). *Heritage Language Education: A Literature Review*. Publication Sales, Ontario Institute for Studies in Education, 252 Bloor St. West, Toronto, Ontario M5S 1V6 Canada..

Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. *Language processing in bilingual children*, 70-89.

Cummins, J. (1996). *Negotiating identities: Education for empowerment in a diverse society* (pp. 1-368). Ontario, CA: California Association for Bilingual Education.

Cummins, J. (1997). Educational attainment of minority students: A framework for intervention based on the constructs of identity and empowerment. In A. Sjogren (ed.) Language and Environment (pp. 89-101). Stockholm: Mangkulturellt Centrum.

Cummins, J. (1998). Beyond adversarial discourse: Searching for common ground in the education of bilingual students. *Presentation to the California State Board of Education. Sacramento, California*.

Cummins, J. (2000). Immersion education for the millennium: What we have learned from 30 years of research on second language immersion. *Retrieved October*, *5*, 2000.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire* (Vol. 23). Multilingual Matters.

Cummins, J. (2001). Empowering minority students: A framework for intervention.

Cummins, J. (2001). Instructional conditions for trilingual development. *International Journal of Bilingual Education and Bilingualism*, *4*(1), 61-75.

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal*, 585-592.

Cummins, J. (2009). 10 Bilingual and Immersion Programs. *The handbook of language teaching*, 161.

Cummins, J., & Gulutsan, M. (1974). Some effects of bilingualism on cognitive functioning. *Bilingualism, biculturalism and education. Edmonton: University of Alberta*.

Cummins, J., Baker, C., & Hornberger, N. H. (Eds.). (2001). *An introductory reader to the writings of Jim Cummins* (Vol. 29). Multilingual Matters.

Cushen, P. J., & Wiley, J. (2011). Aha! Voila! Eureka! Bilingualism and insightful problem solving. *Learning and Individual Differences*, *21*(4), 458-462.

Czernochowksi, D., Mecklinger, A., & Johansson, M. (2009). Age-related changes in the control of episodic retrieval: An ERP study of recognition memory in children and adults. *Developmental Science, 12,* 1026–1040.

D'Esposito, M., Postle, B. R., Jonides, J., & Smith, E. E. (1999). The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proceedings of the National Academy of Sciences*, *96*(13), 7514-7519.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Darcy, N. T. (1953). A review of the literature on the effects of bilingualism upon the measurement of intelligence. *The Pedagogical Seminary and Journal of Genetic Psychology*, *82*(1), 21-57.

David, A., & Wei, L. (2008). Individual differences in the lexical development of French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, *11*(5), 598-618.

Davidson, M. C., Amso, D., Anderson, L. C. & Diamond A. (2006). Development of cognitive control and executive functions from 4–13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia 4,*:2037–78

Dayan, E., & Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron*, *72*(3), 443-454.

De Abreu, E. P. M. (2011). Working memory in multilingual children: Is there a bilingual effect?. *Memory*, *19*(5), 529-537.

De Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism Enriches the Poor Enhanced Cognitive Control in Low-Income Minority Children. *Psychological science*, *23*(11), 1364-1371.

De Bot, K. (2008). Introduction: Second language development as a dynamic process. *The Modern Language Journal*, *92*(2), 166-178.

De Courcy, M., Warren, J., & Burston, M. (2002). Children from diverse backgrounds in an immersion programme. *Language and Education*, *16*(2), 112-127.

De Houwer, A. (2009). *Bilingual first language acquisition*. Multilingual Matters.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the CANTAB. I: development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, *25*(2), 242-254.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. Psychological Corporation.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12,* 45–75.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193-222.

D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain and cognition*, *41*(1), 66-86.

Diamond A. 1985. Development of the ability to use recall to guide action, as indicated by infants' performance on A-not-B. *Child Development, 56,* 868–83

Diamond, A., Carlson, S. M. & Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. *Devlopmental Neuropsychology, 28,* 689–729

Diamond, A, & Taylor C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "do as I say, not as I do. *Developmental Psychobiology, 29,* 315–334.

Diamond, A. & Lee K. (2011). Interventions and programs demonstrated to aid executive function development in children 4–12 years of age. *Science 333,* 959–64

Diamond, A. & Goldman-Rakic, P. S. (1989). Comparison of human infants and rhesus monkeys on Piaget's Anot-B task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research, 74,* 24–40.

Diamond, A. (1991). Developmental time course in human infants and infant monkeys, and the neural bases of inhibitory control in reaching. Annals of the New York Academy of Sciences: pt. VII. Inhibition and executive control: 608: 637–704.

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). London: Oxford University Press.

Diamond, A. (2006). The early development of executive functions. *Lifespan cognition: Mechanisms of change*, 70-95.

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135-168.

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the society for research in child development*, i-206.

Diamond, J. (2010). The benefits of multilingualism. *Science*, *330*(6002), 332-333.

Dias, R., Robbins, T. W. & Roberts, A.C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience*, *17*, 9285–9297.

Dias, R., Roberts, A. & Robbins, T.W. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature 380,* 69 –72.

Diaz, R. M. (1983). Thought and two languages: The impact of bilingualism on cognitive development. *Review of Research in Education*, *10*, 23-54.

Diebold, A. R. (1964). *Incipient bilingualism.* In Hymes, D. ed., Language in Culture and Society. New York: Harper & Row, pp. 495–511.

Dijkstra, T. (2005). Bilingual visualword recognition and lexical access. In Kroll & de Groot (eds.), pp. 179-201.

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test–retest reliability and practice effects of expanded Halstead–Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society*, *5*(04), 346-356.

Dimyan, M. A., & Cohen, L. G. (2011). Neuroplasticity in the context of motor rehabilitation after stroke. *Nature Reviews Neurology*, *7*(2), 76-85.

Döpke, S. (1992). *One parent–one language: an interactional approach* (Vol. 3). John Benjamins Publishing.

Dorian, N. C. (1982). Defining the speech community to include its working margins. In Romaine, S. ed., Sociolinguistic Variation in Speech Communities. London: Edward Arnold, pp. 25-33.

Dornic, S. (1980). Language dominance, spare capacity and perceived effort in bilinguals. *Ergonomics*, *23*(4), 369-377.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Changes in grey matter induced by training. *Nature*, *427*, 311–312.

Drewe, E. A. (1974). The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. *Cortex*, *10*(2), 159-170.

Duncan, J. (1986) . Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, *3*, 271-290.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in neurosciences*, *23*(10), 475-483.

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.

Duncan, O. D., Featherman, D. L., & Duncan, B. (1972). Socioeconomic background and achievement. New York: Seminar Press

Duncan, S. E. (2005). Child bilingualism and cognitive functioning: A study of four Hispanic groups. *Dissertation Abstracts International, A: The Humanities and Social Sciences, 65*(8), 2895–A.

Duncan, S. E., & De Avila, E. A. (1979). Bilingualism and cognition: Some recent findings. *NABE journal*, *4*(1), 15-50.

Dunn, L. & Dunn, D. (2007). *Peabody Picture Vocabulary Test*- Fourth Edition (PPVT-4), Bloomington, MN: NCS Pearson, Inc.

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., et al. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science*, *9*, 1–20.

Echevarria, J., & Graves, A. (1998). *Sheltered content instruction: Teaching English-language learners with diverse abilities.* Boston, MA: Allyn and Bacon.

Echevarria, J., Vogt, M. E., & Short, D. (2004). *Making content comprehensible for English learners: The SIOP model* (2nd ed.). Boston: Pearson Allyn & Bacon.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature neuroscience*, *8*(12), 1784-1790.

Elley, W. B., & MacArthur, R. S. (1962). The Standard Progressive Matrices as a culture-reduced measure of general intellectual ability. *Alberta Journal of Educational Research*.

Emmett, W. G. (1949). Factor analysis by Lawley's method of maximum likelihood. *British Journal of Statistical Psychology*, *2*(2), 90-97.

Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The Source of Enhanced Cognitive Control in Bilinguals Evidence From Bimodal Bilinguals. *Psychological Science*, *19*(12), 1201-1206.

Engel de Abreu, P. M. (2011). Working memory in multilingual children: Is there a bilingual effect?. *Memory*, *19*(5), 529-537.

Engel de Abreu, P. M. J., Gathercole, S. E., & Martin, R. (2011). Disentangling the relationship between working memory and language: The roles of short-term storage and cognitive control. *Learning and Individual Differences*, *21*(5), 569-574.

Engel de Abreu, P. M., Conway, A. R., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence*, *38*(6), 552-561.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. *Models of working memory: Mechanisms of active maintenance and executive control*, 102-134.

Engle, R.W., & Kane, M.J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In Ross, B. (ed.). The psychology of learning and motivation (pp. 145–199). New York: Academic Press.

Erickson, K. I., Kim, J. S., Suever, B. L., Voss, M. W., Francis, B. M., & Kramer, A. F. (2008). Genetic contributions to age-related decline in executive function: a 10-year longitudinal study of COMT and BDNF polymorphisms. *Frontiers in human neuroscience*, *2*(11), 1-9.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143-149.

Eson, M. E., Yen, J. K., & Bourke, R. S. (1978). Assessment of recovery from serious head injury. *Journal of Neurology, Neurosurgery & Psychiatry*, *41*(11), 1036-1042.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., and Glisky, M. L. (1999). Executive functioning in preschool children: Performance on A-not-B and other delayed response format tasks. *Brain and Cognition, 41*, 178–199.

Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages after frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery, and Psychiatry, 68*, 650–652.

Faltis, C. (1997). Case study methods in researching language and education. In *Encyclopedia of language and education* (pp. 145-152). Springer Netherlands.

Fan, J., McCandliss, B.D., Sommer, T., Raz, A., Posner, M.I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14* (3), 340–347.

Farah, M. J., Shera, D. M., Savage, J. H., Betancourt, L., Giannetta, J. M., Brodsky, N. L., & Hurt, H. (2006). Childhood poverty: Specific associations with neurocognitive development. *Brain research*, *1110*(1), 166-174.

Feinstein, A., Brown, R., & Ron, M. (1994). Effects of practice of serial tests of attention in healthy subjects. *Journal of clinical and experimental neuropsychology*, *16*(3), 436-447.

Fenson, L., Dale, P. S., Reznick, J. S., Thal. D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual.* San Diego, CA: Singular Publishing Group.

Fernandes, M. A., Craik, F., Bialystok, E., & Kreuger, S. (2007). Effects of bilingualism, aging, and semantic relatedness on memory under divided attention. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *61*(2), 128.

Fernandez-Duque, D., & Posner, M. I. (1997). Relating the mechanisms of orienting and alerting. *Neuropsychologia*, *35*(4), 477-486.

Festman, J., Rodriguez-Fornells, A., & Münte, T. F. (2010). Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behav. Brain Funct*, *6*(5).

Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1075.

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition*, *9*(02), 153-166.

Fischer, K. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review, *87*, 477-531.

Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages* (Vol. 76). Multilingual matters.

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 874-890.

Fox, M. T., Barense, M. D., & Baxter, M. G. (2003). Perceptual attentional set-shifting is impaired in rats with neurotoxic lesions of posterior parietal cortex. *The Journal of neuroscience*, *23*(2), 676-681.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101.

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.

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

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

Frye, D., Zelazo, P. D. & Palfai T. (1995) Theory of mind and rule-based reasoning. *Cognitive Development, 10,* 483–527.

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at head start. *Developmental psychology*, *47*(2), 404.

Fuster, J. M. (1997). The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe. New York and Philadelphia: Lippincott-Raven.

Fuster, J. M. (2008). *The Prefrontal Cortex, Fourth Edition* (London: Academic Press)

Galambos, S. J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition*, *34*(1), 1-56.

Gamboz, N., Borella, E., & Brandimonte, M. A. (2009). The role of switching, inhibition and working memory in older adults' performance in the Wisconsin Card Sorting Test. *Aging, Neuropsychology, and Cognition*, *16*(3), 260-284.

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., ... & Ávila, C. (2010). Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. *Neuroimage*, *53*(4), 1272-1278.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological bulletin*, *134*(1), 31.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Lawrence Erlbaum Associates, Inc.

Gathercole, S. E., & Baddeley, A. D. (1994). *Working Memory and Language*. Hove: Lawrence Erlbaum Associates, Ltd.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177-190.

Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, *12*(2), 213-237.

Gathercole, V. C. M. (2002). Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. *Language and literacy in bilingual children*.

Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, *12*(2), 213-237.

Gathercole, V. C., Thomas, E. M., Jones, L., Guasch, N. V., Young, N., & Hughes, E. K. (2010). Cognitive effects of bilingualism: Digging deeper for the contributions of language dominance, linguistic knowledge, socio-economic status and cognitive abilities. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 617-664.

Gaudino, E. A., Geisler, M. W., & Squires, N. K. (1995). Construct validity in the Trail Making Test: what makes Part B harder?. *Journal of clinical and experimental neuropsychology*, *17*(4), 529-535.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in cognitive sciences*, *16*(2), 129-135.

Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education* (Vol. 163). Cambridge, MA: Newbury House.

Genesee, F. (1989). Early bilingual development: one language or two? *Journal of Child Language* 16, 161-79.

Genesee, F. (1994). *Educating second language children: The whole child, the whole curriculum, the whole community*. Cambridge University Press.

Genesee, F. (2004). What do we know about bilingual education for majority language students. *Handbook of bilingualism and multiculturalism*, 547-576.

Genesee, F., Paradis, J., & Crago, M. B. (2004). *Dual language development & disorders: A handbook on bilingualism & second language learning, vol 11*. Paul H Brookes Publishing.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 245.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. Cognition, *53*(2) 129-153.

Gersten, R., & Woodward, J. (1995). A longitudinal study of transitional and immersion bilingual education programs in one district. *The Elementary School Journal*, 223-239.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., . . . Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2,* 861–863.

Gittins, J. (1952). *Approved School Boys;*. HM Stationery Office.

Gnys, J. A., & Willis, W. G. (1991). Validation of executive function tasks with young children. *Developmental Neuropsychology*, *7*(4), 487-501.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Science, 101*, 8174–8179.

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging.*The Journal of Neuroscience*, *33*(2), 387-396.

Golden, C.J. (1978). Stroop colour and word test: a manual for clinical and experimental uses. Chicago (IL): Stoelting.

Golden, Z. L., & Golden, C. J. (2002). Patterns of performance on the Stroop Color and Word Test in children with learning, attentional, and psychiatric disabilities. *Psychology in the Schools*, *39*(5), 489-495.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. *Comprehensive Physiology*.

Goldstein, F. C., & Green, R. C. (1995). Assessment of problem solving and executive functions. In *Clinical neuropsychological assessment* (pp. 49-81). Springer US.

Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246.

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, *13*(02), 197-208.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787-814.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220-1234.

Gopnik, A., & Rosati, A. (2001). Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. *Developmental Science*, *4*(2), 175–183.

Gort, M. (2006). Strategic codeswitching, interliteracy, and other phenomena of emergent bilingual writing: Lessons from first grade dual language classrooms. *Journal of Early Childhood Literacy*, *6*(3), 323-354.

Graf, P., Uttl, B., & Tuokko, H. (1995). Color-and picture-word Stroop tests: performance changes in old age. *Journal of Clinical and Experimental Neuropsychology*, *17*(3), 390-415.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of experimental psychology*, *38*(4), 404.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and language*, *27*(2), 210-223.

Green, D. W. (1993). Towards a model of L2 comprehension and production. In R. Schreuder & B. Weltens (eds.), The Bilingual Lexicon, pp. 249±277. Amsterdam/Philadelphia: John Benjamins.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(02), 67-81.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes* 3:233-274.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and language, 36*(1), 3-15.

Grosjean, F. (1996). Living with two languages and two cultures. *Cultural and language diversity and the deaf experience*, 3-19.

Grosjean, F. (1997). The bilingual individual. *Interpreting*, *2*(1-2), 1-2.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1:131-149.

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition*, *1*(03), 175-176.

Grosjean, F. (1999). The right of the deaf child to grow up bilingual. *Deaf Worlds, 15*(2): 29-31.

Grosjean, F. (2001). The bilingual's language modes. *One mind, two languages: Bilingual language processing*, 1-22.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). Bilingualism, biculturalism, and deafness. *International Journal of Bilingual Education and Bilingualism*, *13*(2), 133-145.

Guo, G., & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography*, *37*(4), 431-447.

Gutiérrez-Clellen C, Kreiter J. (2003) Understanding child bilingual acquisition using parent and teachers reports. *Applied* Psycholinguistics, 24, 267–288

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in cognitive sciences*, *13*(2), 65-73.

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651-659.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9*, 416–423.

Hakuta, K. (1984). The causal relationship between the development of bilingualism, cognitive flexibility, and social-cognitive skills in his Hispanic elementary school children. Final Report, National Institute of Education Grant-81-012. Available from the National Clearinghouse for Bilingual Education, Arlington, VA.

Hakuta, K. (1986). *Mirror of language: The debate on bilingualism*. New York: Basic Books.

Hakuta, K. (2001). The education of language-minority students. Testimony to the U.S. Commission on Civil Rights. Washington, DC, April 13.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence a test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31-38.

Hale, S., Bronik, M. D., & Fry, A. F. (1997). Verbal and spatial working memory in school-age children: Developmental differences in susceptibility to interference. *Developmental Psychology*, *33,* 364–371.

Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: A critical discussion and some new longitudinal data. *Children's language*, *5*, 319-344.

Hakuta, K., & Garcia, E. E. (1989). Bilingualism and education. *American Psychologist*, *44*(2), 374.

Hamers, J. F., & Blanc, M. H. (2000). *Bilinguality and bilingualism*. Cambridge University Press.

Hammill, D. D., Pearson, N. A., & Wiederholt, J. L. (1997). *Comprehensive test of nonverbal intelligence (CTONI)*. Pro-ed.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: towards a unified developmental approach. *Developmental Science*, *13*(4), 622-635.

Harris, C.L., Gleason, J. B., & Aycicegi, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. *Bilingual education and bilingualism*, *56*, 257.

Harris, J., & Murtagh, L. (1999). *Teaching and Learning Irish in Primary School: A Review of Research and Development*.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of experimental psychology: Learning, memory, and cognition*, *17*(1), 163.

Haugen, E. I. (1956). *Bilingualism in the Americas: A bibliography and research guide*. Publications of the American Dialect Society 26.

Hauser, R. M., & Warren, J. R. (1997). Socioeconomic indexes for occupations: A review, update, and critique. *Sociological methodology*, *27*(1), 177-298.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*,*109*(2), 340.

Heaton, R. K., & Pendleton, M. G. (1981). Use of neuropsychological tests to predict adult patients' everyday functioning. *Journal of consulting and clinical psychology*, *49*(6), 807.

Heaton, R. K., Chelune, G. J. Talley, J. L., Kay, G., Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual, revised edition*, Psychological Assessment Resources, Odessa, FL. .

Henry, L. A., & Bettenay, C. (2010). The assessment of executive functioning in children. *Child and adolescent mental health*, *15*(2), 110-119.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of child psychology and psychiatry*, *53*(1), 37-45.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language?. *Bilingualism: Language and Cognition*, *1*(03), 213-229.

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: how bilingual?. *International journal of bilingual education and bilingualism*, *15*(2), 131-145.

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish–English bilinguals: An fMRI study. *NeuroImage*, *14*(2), 510-520.

Hickey, T. (1991) Leisure reading in a second language: An experiment with audio tapes. *Language, Culture and Curriculum* 4 (2), 119–133.

Hickey, T. (1997). *Early Immersion Education in Ireland: Na Naíonraí*. Dublin: Linguistics Institute of Ireland.

Hickey, T. (1998). Early immersion in Ireland: The *naíonra* experience. In V. Regan (Ed.), *Contemporary approaches to second language acquisition* (pp. 46-61). Dublin: UCD Academic Press.

Hickey, T. (1999). *Luathoideachas triGhaeilge sa Ghaeltacht.* Gaillimh: Udaras na Gaeltachta & Intitiuid Teangeolafochta Eireann

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin & review*, *18*(4), 625-658.

Hillman, C. H., Erickson, K. I. & Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *National Review of Neuroscience*, *9,* 58–65

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child development*, *74*(5), 1368-1378.

Hoff, A. L., Svetina, C., Shields, G., Stewart, J., & DeLisi, L. E. (2005). Ten year longitudinal study of neuropsychological functioning subsequent to a first episode of schizophrenia. *Schizophrenia research*, *78*(1), 27-34.

Holmes, J., Gathercole, S.E., & Dunning, D.L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12,* 9–15.

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*(6), 725-743.

Huagen, E. (1953). *The Norwegian language in America: a study in bilingual behavior.* Philadelphia: University of Pennsylvania Press.

Hudspeth, W. & Pribram, K. (1990). Stages of brain and cognitive maturation. *Journal of Educational Psychology, 82,* 882-884.

Huettner, M. I. (1989). Neurological basis of language and reading. *Learning and Individual Differences*, *1*(4), 407-421.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233-253.

Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair?. *Developmental neuropsychology*, *28*(2), 645-668.

Hughes, C., & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions?. *Child and adolescent mental health*, *7*(3), 131-142.

Huizinga, M., & Smidts, D. P. (2010). Age-related changes in executive function: A normative study with the Dutch version of the Behavior Rating Inventory of Executive Function (BRIEF). *Child neuropsychology*, *17*(1), 51-66.

Huizinga, M., & van der Molen, M. W. (2007). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Developmental Neuropsychology*, *31*(2), 193-215.

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

Humes, G. E., Welsh, M. C., Retzlaff, P., & Cookson, N. (1997). Towers of Hanoi and London: Reliability and validity of two executive function tests. *Assessment,* 4, 249–257

Huttenlocher, P. R. (2013, June). Plasticity in Human Cerebral Cortex. In *Threats To Optimal Development: Integrating Biological, Psychological, and Social Risk Factors: the Minnesota Symposia on Child Psychology* (Vol. 27, p. 35). Psychology Press.

Ianco-Worrall, A. D. (1972). Bilingualism and cognitive development. *Child development*, 1390-1400.

Ikkai, A., & Curtis, C. E. (2011). Common neural mechanisms supporting spatial working memory, attention and motor intention. *Neuropsychologia*, *49*(6), 1428-1434.

ILO. Internation Standard Classification of Occupations (ISCO-08). (2008). http://www.ilo.org/public/english/bureau/stat/isco/isco08/

Inurritegui, S. (2009). Bilingualism and the Development of Cognitive Control. Ph.D. Dissertation, Katholieke Universiteit of Leuven.

Iseman, S. J., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD and LD: A randomized controlled study. *Journal of Learning Disabilities*, *44*, 184–195.

Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive function in preschool children: Examination through everyday behavior. *Developmental neuropsychology*, *26*(1), 403-422.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production?. *Acta psychologica*, *127*(2), 277-288.

Jahanshahi, M., Profice, P., Brown, R. G., Ridding, M. C., Dirnberger, G., & Rothwell, J. C. (1998). The effects of transcranial magnetic stimulation over the dorsolateral prefrontal cortex on suppression of habitual counting during random number generation. *Brain*, *121*(8), 1533-1544.

James, L. E. & Burke, D. M. (2000) Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1378–1391.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jespersen, O. (1922). *Language: Its nature, development, and origin.* London: Allen and Unwin.

John, V. (1970). Cognitive development in the bilingual child. In J. E. Alatis (Ed.), Monograph Series on Languages and Linguistics. Georgetown University School of Languages and Linguistics, 23, 59-67.

Johnson, R. K., & Swain, M. (Eds.). (1997). *Immersion education: International perspectives*. Cambridge University Press.

Johnstone, R. (1999). Research on language learning and teaching: 1997-98. *Language Teaching*, *32*(3), 137-156.

Johnstone, R. (2007). Characteristics of immersion programmes. *Bilingual education and bilingualism*, *61*, 19.

Jones, W. R. (1933). Tests for the Examination of The Effect of Bilingualism on Intelligence. *Unpublished MA Thesis, University of Wales*.

Jones, W. R. (1959). *Bilingualism and Intelligence*. Cardiff: University of Wales Press.

Junker, D. A., & Stockman, I. J. (2002). Expressive vocabulary of German-English bilingual toddlers. *American Journal of Speech-Language Pathology*, *11*(4), 381.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology review*, *17*(3), 213-233.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological review*, *99*, 122-149.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic bulletin & review*, *9*(4), 637-671.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169.

Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning*, *57*(1), 119-163.

Kennedy, I. (2012). Irish medium education: cognitive skills, linguistic skills, and attitutudes towards Irish. Unpublished PhD. Bangor University, Wales.

Kennedy, W. A., Van De Riet, V., & White, J. C. (1963). A normative sample of intelligence and achievement of Negro elementary school children in the southeastern United States. *Monographs of the Society for Research in Child Development*, *28*(6), 1-112.

Kiefer, M., Marzinzik, F., Weisbrod, M., Scherg, M., & Spitzer, M. (1998). The time course of brain activations during response inhibition: evidence from event-related potentials in a go/no go task. *Neuroreport*, *9*(4), 765-770.

Kim, K. H., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *388*(6638), 171-174.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: the role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General*, *122*(4), 411.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. *Proceedings of the National Academy of Sciences*, *92*(7), 2899-2903.

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the dimensional change card-sorting task. *Developmental Science*, *8*(1), 44-56.

Knight, R. T., & Stuss, D. T. (2002). Prefrontal cortex: the present and the future. *Principles of frontal lobe function*, 573-597.

Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The Development of Self-Regulation in the First Four Years of Life. *Child development*, *72*(4), 1091-1111.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child development*, *67*(2), 490-507.

Kohlberg, L. (1963). *Moral Development and Chicago*: Uni. of Chicago Press.

Konaka, K. (1997). *The relationship between degree of bilingualism and gender to divergent thinking ability among native Japanese-speaking children in the New York area* (Doctoral dissertation, New York University, School of Education).

Kortte, K. B., Horner, M. D., & Windham, W. K. (2002). The trail making test, part B: cognitive flexibility or ability to maintain set?. *Applied Neuropsychology*,*9*(2), 106-109.

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*, *12*(1), 48-54.

Kovács, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*, 6556–6560.

Kovelman, I., Baker, S. A., & Petitto, L. A. (2008). Bilingual and monolingual brains compared: a functional magnetic resonance imaging investigation of syntactic processing and a possible “neural signature” of bilingualism. *Journal of cognitive neuroscience*, *20*(1), 153-169.

Kovelman, I.; Petitto, (2003) LA. Stages of language development in bilingual children exposed to their other language at different ages. Poster presented at the annual meeting of the International Symposium on Bilingualism; Phoenix, AZ. 2003.

Kovelman, I.; Petitto, LA. (2002). Bilingual babies' maturational and linguistic milestones as a function of their age of first exposure to two languages. Poster presented at the conference for the Society for Neuroscience; Orlando, FL. 2002.

Kray, J. (2006). Task-set switching under cue-based versus memory-based switching conditions in younger and older adults. *Brain research*, *1105*(1), 83-92.

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica*, *115*(2), 143-165.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, (ahead-of-print), 1-18.

Kroll, J. F., & Dussias, P. E. (2013). *The comprehension of words and sentences in two languages. In T. Bhatia & W. Ritchie (Eds.), The handbook of bilingualism and multilingualism* (2nd ed., pp. 216-243). Malden, MA: Wiley-Blackwell.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. *The handbook of second language acquisition*, 104-129.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta psychologica*, *128*(3), 416-430.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes- Kroff, J. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), The psychology of learning and motivation (Vol. 56, pp. 229_262). San Diego, CA: Academic Press.

Kusche, C. A., Cook, E. T., & Greenberg, M. T. (1993). Neuropsychological and cognitive functioning in children with anxiety, externalizing, and comorbid psychopathology. *Journal of Clinical Child Psychology, 22,* 172-195.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. *Handbook of bilingualism*, 289.

Lakes, K. D., & Hoyt, W. T. (2004). Promoting self-regulation through school-based martial arts training. *Applied Developmental Psychology, 25*, 283–302.

Lakes, K. D., Bryars, T., Sirisinahal, S., Salim, N., Arastoo, S., Emmerson, N., ... & Kang, C. J. (2013). The Healthy for Life Taekwondo Pilot Study: A Preliminary Evaluation of Effects on Executive Function and BMI, Feasibility, and Acceptability. *Mental Health and Physical Activity*.

Lambert, W. E. (1977). The effects of bilingualism on the individual: Cognitive and sociocultural consequences. *Bilingualism: Psychological, social and educational implications*, *9*.

Landry, R. G. (1974). A comparison of second language learners and monolinguals on divergent thinking tasks at the elementary school level. *The modern language journal*, *58*(1-2), 10-15.

Lazaruk, W. (2007). Linguistic, academic, and cognitive benefits of French immersion. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *63*(5), 605-627.

Lee, H., & Kim, K. H. (2011). Can speaking more languages enhance your creativity? Relationship between bilingualism and creative potential among Korean American students with multicultural link. *Personality and Individual Differences*, *50*(8), 1186-1190.

Lehto, J. E., Juujärvi, P., Kooistra, L. & Pulkkinen L. (2003) Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59–80.

Leopold, W. F. (1939-1949). *Speech development of a bilingual child: A linguist's record (Volumes I - IV).* Evanstron, IL: Northwestern University Press.

Leung, A. K. Y., Maddux, W. W., Galinsky, A. D., & Chiu, C. Y. (2008). Multicultural experience enhances creativity: the when and how. *American Psychologist*, *63*(3), 169.

Levin, H. S., Boldstein, F. C.,Williams, D. H., & Eisenberg, H. M. (1991). The contribution of frontal lobe lesions to the neurobehavioural outcome of closed head injury. In H. S.

Levin, H. M. Eisenberg & A. L. Benton (Eds.), Frontal lobe function and dysfunction (pp. 318-338). New York: Oxford University Press.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting Your Native Language The Role of Retrieval-Induced Forgetting During Second-Language Acquisition. *Psychological Science*, *18*(1), 29-34.

Lewis, E. G. (1977). Bilingualism and bilingual education: The ancient world to the Renaissance. In Spolsky, B. & Cooper, R. L. (eds.) *Frontiers of Bilingual Education. Rowley, MA: Newbury House*.

Lewis, W.G. (2008). Current challenges in bilingual education in Wales. In Cenoz, J. & Gorter, D. (eds) Multilingualism and Minority Languages: Achievements and Challenges in Education. AILA Review, 21. Amsterdam: Benjamins.

Lezak, M. (1995). *Neuropsychological assessment.* Vol. 3. New York: Oxford University Press.

Lezak, M. D. (1983). *Neuropsychological Assessment.* 2nd ed. Oxford University Press, New York.

Lezak, M. D. (Ed.). (2004). *Neuropsychological asessment 4 Ed*. Oxford university press.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). Executive functions and motor performance. *Neuropsychological assessment*, 611-646.

Liedtke, W. W., & Nelson, L. D. (1968). Concept formation and bilingualism. *Alberta Journal of Educational Research*, *14*(4), 225-232.

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The mental lexicon*, *3*(3), 349.

Lindholm-Leary,K. J.,&Borsato, G. (2001). *Impact of two-way bilingual elementary programs on students'attitudes toward school and college*. Santa Cruz, CA: Center for Research on Education, Diversity & Excellence.

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*, 273–293.

Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function in children using the Cambridge Neuropsychological Testing Automated Battery (CANTAB): Performance in 4 to 12 year-olds. Developmental Neuropsychology, 22, 595 – 623.

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005).The development of nonverbal working memory and executive control processes in adolescents. *Child Development, 76,* 697–712.

Luk G, et al. Is there a relation between onset age of bilingualism and enhancement of cognitive control? Bilingualism: Language and Cognition. 2011; 14:588–595.

Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of Neuroscience*, *31*(46), 16808-16813.

Luk, G., De Sa, E. R. I. C., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control. *Bilingualism: Language and cognition*, *14*(4), 588-595.

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, *27*(10), 1479-1488.

Luna B. 2009. The maturation of cognitive control and the adolescent brain. In: Aboitiz F, et al. From attention to goal-directed behavior: neurodynamical, methodological and clinical trends. Heidelberg (Germany): Springer-Verlag. p. 249--274.

Luna, B., Garver, K.E., Urban, T.A., Lazar, N.A., & Sweeney, J.A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75*, 1357–1372.

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and cognition*, *72*(1), 101-113.

Lupien, S. J., King, S., Meaney, M. J., & McEwen, B. S. (2000). Child's stress hormone levels correlate with mother's socioeconomic status and depressive state. *Biological psychiatry*, *48*(10), 976 980.

Lupien, S. J., King, S., Meaney, M. J., & McEwen, B. S. (2001). Can poverty get under your skin? Basal cortisol levels and cognitive function in children from low and high socioeconomic status. *Development and psychopathology*, *13*(3), 653-676.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior.* Oxford: Pergamon.

Luria, A. R. (1966). *Higher Cortical Functions in Man*. Tavistock. London.

Luria, A. R. (1973). The frontal lobes and the regulation of behavior.

Mackey, W. (1967). *Bilingualism as a World Problem*. Montreal: Harvest House.

Mackey, W. F. (1965). *Language Teaching Analysis.* Bloomington: Indiana University Press.

Mackey, W. F. (1968). The Description of Bilingualism. In Fishman ed., pp. 554-84.

Mackey, W. F. (1970). A typology of bilingual education. *Foreign Language Annals*, *3*(4), 596-606.

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

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

Macnamara, J. (1966). *Bilingualism and primary education: A study of Irish experience*. [Edinburgh] University Press.

Macnamara, J. (1967). The effects of instruction in a weaker language. *Journal of Social Issues*, *23*(2), 121-135.

Macnamara, J., & Kushnir, S. L. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, *10*(5), 480-487.

Macnamara, J., Svarc, J. & Horner, S. (1976). Attending a primary school of the other language in Montreal. In A. Simoes Jr (ed.) The Bilingual Child. New York: Academic Press.

Maddux, W. W., Adam, H., & Galinsky, A. D. (2010). When in Rome.... learn why the Romans do what they do: How multicultural learning experiences enhance creativity. *Personality and Social Psychology Bulletin*, *36*, 731–741.

Mahon, M., & Crutchley, A. (2006). Performance of typically-developing school-age children with English as an additional language on the British Picture Vocabulary Scales II. *Child Language Teaching and Therapy*, *22*(3), 333-351.

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., & Robertson, I. H. (2001). The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *Journal of Child Psychology and Psychiatry*, *42*, 1065–1081.

Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith, I. (1999). *TEA-Ch: the test of everyday attention for children.* Bury St. Edmunds, England: Thames Valley Test Company.

Manly, T., Robertson, I. H., Galloway, M,, & Hawkins, K.(1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37,* 66 1-670.

Manuel, H. T. (1935). *Spanish and English editions of the Stanford-Binet in relation to the abilities of Mexican children* (No. 3532). The University.

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals*. *Journal of child language*, *37*(4), 817.

Marinove-Todd, S. H., Marshall, D. B., & Snow, C. E. (2000). Three misconceptions about age and L2 learning. *Tesol Quarterly*, *34* (1), 9-34.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism Language and Cognition*, *11*(1), 81.

Mayberry, R. I., & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and language*, *87*(3), 369-384.

Mazzocco, M. M., & Kover, S. T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology*, *13*(1), 18-45.

McCall, R. B. (1981). Nature-nurture and the two realms of development: A proposed integration with respect to mental development. *Child Development*, 1-12.

McLaughlin, B. (1978). *Second language acquisition in childhood.* Hillsdale, N.L: Erlbaum.

McLaughlin, B. (1984). *Second language acquisition in childhood: Vol. 1. Preschool children* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

McLaurin, W. A., & Farrar, W. E. (1973). Validities of the Progressive Matrices Tests against IQ and grade point average. *Psychological Reports, 32*, 803-806.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American psychologist*, *53*(2), 185.

McVeigh, C. (2013). Language and cognition in developing bilinguals: examining working memory and executive control. Unpublished PhD. Queen's University, Belfast.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757-757.

Meiran, N. (1996) Reconfiguration of processing mode prior to task performance. *The Journal of Experimental Psychology: Learning, Memory, and Cognition 22:* 1423–1442.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive psychology*, *41*(3), 211-253.

Meltzer, L. (Ed.). (2011). *Executive function in education: From theory to practice*. Guilford press.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language*, *40*(1), 25-40.

Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child development*, *75*(5), 1373-1386.

Michael, E.B., & Gollan, T.H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J.F. Kroll & A.M.B. de Groot (Eds.), Handbook of bilingualism: Psycholinguistic approaches (pp. 389–407). New York: Oxford University Press.

Middleton, F. A., & Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *The Journal of Neuroscience*, *21*(2), 700-712.

Middleton, F. A., & Strick, P. L. (2002). Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cerebral Cortex*, *12*(9), 926-935.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*(1), 167-202.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27,* 272–277.

Milte, C. M., Parletta, N., Buckley, J. D., Coates, A. M., Young, R. M., & Howe, P. R. (2012). Eicosapentaenoic and docosahexaenoic acids, cognition, and behavior in children with attention-deficit/hyperactivity disorder: A randomized controlled trial. *Nutrition*, *28*(6), 670-677.

Mitrushina, M. (Ed.). (2005). *Handbook of normative data for neuropsychological assessment*. Oxford University Press.

Mitrushina, M. N., Boone, K. L., & D'Elia, L. (1999). *Handbook of normative data for neuropsychological assessment.* New York: Oxford University Press.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8-14.

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.

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93-148). Hove, UK: Erlbaum.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134-140.

Montessori, M. (1949). In Tong, F., & Rodriguez, L. (2013). Montessori Bilingual Education and Brain-Compatible Learning. *Education for the Human Brain: A Road Map to Natural Learning in Schools*, 83.

Morales, J., Calvo, A., & Bialystok, E. (2012). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187-202.

Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of neurolinguistics*, *23*(6), 564-579.

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.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science*, *10*(6), 719-726.

Morton, J., Munakata, Y., (2002). Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology, 40,* 255–265.

Moulden, D. J. A., Picton, T. W., Meiran, N., Stuss, D. T., Riera, J. J., & Valdes-Sosa, P. (1998). Event-related potentials when switching attention between task sets. *Brain and Cognition, 37,* 186–190.

Mueller Gathercole, V. C., Thomas, E. M., Jones, L., Guasch, N. V., Young, N., & Hughes, E. K. (2010). Cognitive effects of bilingualism: Digging deeper for the contributions of language dominance, linguistic knowledge, socio-economic status and cognitive abilities. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 617-664.

Mullane, J. C., Corkum, P. V., Klein, R. M., McLaughlin, E. N., & Lawrence, M. A. (2009). Interference control in children with and without ADHD: A systematic review of flanker and simon task performance. *Child Neuropsychology*, *15*, 321-342.

Munakata, Y. (2001). Graded representations in behavioural dissociations. *Trends in Cognitive Sciences*, *5*, 309–315.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences, 15,* 453–459.

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*(3), 218-228.

Muraven, M. (2010). Building self-control strength: practicing self-control leads to improved self-control performance. *Journal of Experimental Social Psychology, 46,* 465–68

Murtagh, L., & van der Slik, F. (2004). Retention of Irish skills: A longitudinal study of a school-acquired second language. *International Journal of Bilingualism*, *8*(3), 279-302.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing* (Vol. 11). Cambridge University Press.

Myers-Scotton, C. (2002). Frequency and intentionality in (un) marked choices in codeswitching:"This is a 24-hour country". *International Journal of Bilingualism*, *6*(2), 205-219.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2012). Planning, Attention, Simultaneous, Successive. *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, 178.

Namazi, M., & Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 597-616.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive psychology*, *9*(3), 353-383.

Nebel, K., Wiese, H., Stude, P., de Greiff, A., Diener, H. C., & Keidel, M. (2005). On the neural basis of focused and divided attention. *Cognitive Brain Research*, *25*(3), 760-776.

Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American psychologist*, *51*(2), 77.

Nelson, C.A., & Bloom, F.E. (1997). Child development and neuroscience. *Child Development, 68,* 970-987.

Nettelbeck, T. (1987). Inspection time and intelligence. In Phillip, V. (Ed.). Speed of information-processing and intelligence. (pp. 295-346). Westport, CT, US: Ablex Publishing.

Nicolay, A. C., & Poncelet, M. (2013a). Cognitive advantage in children enrolled in a second-language immersion elementary school program for three years. *Bilingualism: Language and Cognition*, *1*(1), 1-11.

Nicolay, A. C., & Poncelet, M. (2013b). Cognitive abilities underlying second-language vocabulary acquisition in an early second-language immersion education context: A longitudinal study. *Journal of experimental child psychology*, *115*(4), 655-671.

Nieuwenhuis, S., & Yeung, N. (2005). Neural mechanisms of attention and control: losing our inhibitions?. *Nature Neuroscience*, *8*(12), 1631-1633.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological bulletin*, *126*(2), 220.

Nisbet, R. A. (1953). *The quest for community: A study in the ethics of order & freedom*. San Francisco: ICS Press.

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental science*, *10*(4), 464-480.

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental science*, *8*(1), 74-87.

Nobre, A. C., & Stokes, M. G. (2011). Attention and short-term memory: Crossroads. *Neuropsychologia*, *49*(6), 1391-1392.

Norman, D. A., & Shallice, T. (1980). *Attention to action: Willed and automatic control of behavior* (No. CHIP-99). California University San Diego La Jolla center for human information processing.

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*, *71*(3), 437-451.

Ó hAinifein, P. (2008). *Mionteangacha Sna Meáin Craoltóireachta: Ról Na Gaeilge Agus Na Catalóinise* (Doctoral dissertation).

Ó hIfearnáin, T. Ó. (2007). Raising children to be bilingual in the Gaeltacht: Language preference and practice. *International Journal of Bilingual Education and Bilingualism*, *10*(4), 510-528.

Ó Duibhir, P. (2009). *The spoken Irish of sixth-class pupils in Irish immersion schools*. Unpublished PhD. Trinity College: Dublin, Ireland.

Ó Duibhir, P. (2011). 'I Thought That We Had Good Irish': Irish Immersion Students' Insights into Their Target Language Use, in Immersion Education: Practices, Policies, Possibilities, Tedick, D.J., Christian, D. & Fortune, T.W. (2011). Multilingual Matters

Ó Murchú , H. (2001). Irish: The Irish language in education in the Republic of Ireland. Ljouwert, Leeuwarden: Regional Dossier Mercator Education.

Ó Murchú, H. (2003). Gaelic-medium education in the Republic of Ireland. Gaelic-medium Education Provision: Northern Ireland, the Republic of Ireland, Scotland and the Isle of Man. Belfast: Ollscoil na Banríona.

Ó Murchú, H. (2003). Retention and attrition of Irish as a second language. A longitudinal study of general and communicative proficiency in Irish among second level school leavers and the influence of the instructional background, language use and attitude/motivation variables. PhD Dissertation, University of Groningen.

O'Donnell, J. P., MacGregor, L. A., Dabrowski, J. J., & Oestreicher, J. M. (1994). Construct validity of neuropsychological tests of conceptual and attentional abilities. *Journal of Clinical Psychology*.

Official Languages Act: Acht na dTeangacha Oifigiúla (2003). Number 23 of 2003. Office of the Attorney General.

Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children* (Vol. 2). Multilingual Matters.

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, *28*(02), 191-230.

Ozonoff, S. (1997). Components of executive function in autism and other disorders. In J. Russell (Ed.), Autism as an executive disorder (pp. 179–211). New York, USA: Oxford University Press.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive psychology*, *66*(2), 232-258.

Paolo, A. M., Tröster, A. I., Axelrod, B. N., & Koller, W. C. (1995). Construct validity of the WCST in normal elderly and persons with Parkinson's disease.*Archives of Clinical Neuropsychology*, *10*(5), 463-473.

Parsons, C. E., & Lyddy, F. (2009). Early reading strategies in Irish and English: Evidence from error types. *Reading in a Foreign Language*, *21*(1), 22-36.

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children.*Developmental Neuropsychology*, *1*(4), 349-370.

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: general and applied*, *76*(27), 1-23.

Pearson, B. Z., Fernandez, S. C. & Oiler, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, *43*, 93-120.

Pearson, B. Z., Fernandez, S., Lewedag, V., & Oller, D. K. (1997). Input factors in lexical learning of bilingual infants (ages 10 to 30 months). *Applied Psycholinguistics*, *18*, 41–58.

Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525-531.

Penfield, W., and Roberts, L. (1959). Speech and Brain-mechanisms. Princeton University Press, Princeton, N.J.

Pennington, B. F. (2002): *The Development of Psychopathology*. New York: Guilford Press.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., ... & Mehler, J. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain*, *121*(10), 1841-1852.

Petitto, L. A., & Dunbar, K. (2004). New findings from educational neuroscience on bilingual brains, scientific brains, and the educated mind. In *Conference on Building Usable Knowledge in Mind, Brain, & Education, Havard Graduate School of Education*.

Petitto, L. A., & Kovelman, I. (2003). The bilingual paradox: How signing-speaking bilingual children help us to resolve it and teach us about the brain's mechanisms underlying all language acquisition. *Learning Languages*, *8*(3), 5-19.

Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*, Delachaux & Niestlé, Paris.

Piaget, J. (1954). The construction of reality in the child (M. Cook, Trans.). New York, NY: Basic Books.

Pickering, S. J., & Gathercole, S. E. (2001). *Working Memory Test Battery for Children.* London: Psychological Corporation.

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C).* Psychological Corporation.

Pinker, S. (1984). *Language learnability and language development.* Harvard University Press.

Pinker, S. (1994). *The language instinct.* New York: Morrow.

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415-422.

Posner, M. I., & DiGirolamo, G. J. (1998). Executive attention: conflict, target detection, and cognitive control.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25–42

Posner, M. I., & Rothbart, M. K. (1991). Attentional mechanisms and conscious experience. In M. Rugg & A.D. Milner (Eds.), The neuropsychology of consciousness (pp. 91-112). London: Academic Press.

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, *17*(1), 57-70.

Poulisse, N. (1997). Language production in bilinguals.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production* (Vol. 20). John Benjamins Publishing.

Price, C. J. (2000). The anatomy of language: contributions from functional neuroimaging. Journal of Anatomy, 197, 335-359.

Price, C.J., Green D. & von Studnitz, R. A. (1999). Functional imaging study of translation and language switching. *Brain, 122,* 2221–2236.

Prior, A., & Gollan, T. H. (2011). Good Language-Switchers are Good Task-Switchers: Evidence from Spanish??? English and Mandarin??? English Bilinguals. *Journal of the International Neuropsychological Society*, *17*(4), 682.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(02), 253-262.

Rabia, S. A. (2001). Testing the interdependence hypothesis among native adult bilingual Russian-English students. *Journal of Psycholinguistic Research*, *30*(4), 437-455.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language*, *26*(4), 392-405.

Raven, J. C., Court, JH, & Raven, J.(1986). *Raven's progressive matrices and Raven's coloured matrices*.

Raven, J., Court, J. H. et al. (1995). Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No.4, A Compendium of Normative, Reliability and Validity Studies. Oxford, England: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation

Redlinger, W. E., & Park, T. Z. (1980). Language mixing in young bilinguals. *Journal of child language*, *7*(2), 337-352.

Reitan, R. M. (1971) Trail Making Test results for normal and brain-damaged children. *Perceptual and Motor Skills*, 33, 575–581

Reitan, R. M. (1979). *Manual for administration of neuropsychological test batteries for adults and children*. Neuropsychology Laboratory, Indiana University medical Center.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.

Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review*, *4*(3), 161-198.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of psycholinguistic research*, *21*(4), 301-316.

Ricciardelli, L. A. (1993). An investigation of the cognitive development of Italian- English bilinguals and Italian monolinguals from Rome. *Journal of Multilingual and Multicultural Development, 14*, 345–346.

Riggs, N. R., Blair, C. B., & Greenberg, M. T. (2004). Concurrent and 2-year longitudinal relations between executive function and the behavior of 1st and 2nd grade children. *Child Neuropsychology*, *9*(4), 267-276.

Robbins, T. W. (1998). Dissociating executive functions of the prefrontal cortex. In: The prefrontal cortex. Executive and cognitive functions (Roberts AC, Robbins TW, Weiskrantz L, eds), pp. 117–130. Oxford: Oxford University Press.

Roberts Jr, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental neuropsychology*, *12*(1), 105-126.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, *16*(4-6), 635-645.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention., *Journal-International Neuropsychological Society*, *2*, 525-534.

Robinson, A. L., Heaton, R. K., Lehman, R. A., & Stilson, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of consulting and clinical psychology*, *48*(5), 605.

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Nösselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, *415*(6875), 1026-1029.

Rodriguez-Fornells, A., Van Der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, *17*(3), 422-433.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of experimental psychology: General*, *124*(2), 207.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford, UK: Blackwell.

Romaine, S. (1999). Bilingual language development. The development of language, ed. by Martyn Barrett, 251–75. Hove, UK: Psychology Press.

Romine, C. B. & Reynolds C. R. (2005). A model of the development of frontal lobe function: Findings from a meta-analysis. *Applied Neuropsychology, 12,* 190–201. [PubMed: 16422660]

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*(8), 1027-1041.

Ronjat, J. (1913). *Le developpement du langage observe chez un enfant bilingue.* Paris: Libraire Ancienne H. Champion.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: the role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental science*, *14*(5), 1161-1175.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of memory and language*, *39*(3), 418-436.

Rosselli, M., & Ardila, A. (1993). Developmental norms for the Wisconsin Card Sorting Test in 5 to 12 year-old children. *The Clinical Neuropsychologist*, *7*(2), 145-154.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, *7*(1), 17-24.

Rosselli, M., Ardila, A., Jurado, M. B., & Salvatierra, J. L. (2012). Cognate facilitation effect in balanced and non-balanced Spanish–English bilinguals using the Boston Naming Test. *International Journal of Bilingualism*.

Rosselli, M., Ardila, A., Santisi, M., Arecco, M., Salvatierra, J., Conde, A., et al. (2002) Stroop effect in Spanish-English bilinguals. *Journal of the International Neuropsychological Society,* 8, 819–827.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 763.

Rushworth, M. F., Passingham, R. E., & Nobre, A. C. (2005). Components of attentional set-switching. *Experimental Psychology*, *52*(2), 83.

Sacks, T. L., Clark, C. R., Pols, R. G., & Geffen, L. B. (1991). Comparability and stability of performance of six alternate forms of the dodrill-stroop colour-word test. *The Clinical Neuropsychologist*, *5*(3), 220-225.

Saer, D. J. (1922). "An Inquiry into the Effect of Bilingualism upon the Intelligence of Young Children." *Journal of Experimental Pedagogy* 6:232-40, 266-74.

Saer, D. J. (1924). *Bilingual Problem*. Hughes & Son.

Saer, D.J. (1923). "The Effects of Bilingualism on Intelligence." *British Jourrial of Psychology* 14:25 38.

Saer, H. (1931). Experimental inquiry into the education of bilingual peoples. *Education in a changing commonwealth*, 116-21.

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological psychology*, *54*(1), 35-54.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*(4), 566.

Sanchez-Cubillo, I., Perianez, J. A., Adrover-Roig, D., Rodriguez-Sanchez, J. M., Rios-Lago, M., Tirapu, J., & Barcelo, F. (2009). Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, *15*(3), 438.

Sanchez-Cubillo, I., Perianez, J. A., Adrover-Roig, D., Rodriguez-Sanchez, J. M., Rios-Lago, M., Tirapu, J., & Barcelo, F. (2009). Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, *15*(3), 438.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency: The dual-task analogy. *Bilingualism: Language and Cognition*, *13*, 231–252.

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., & Boyce, W. T. (2010). Family socioeconomic status and child executive functions: the roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society*, *17*(1), 120.

Schneider, S. L., & Kogan, I. (2008). The International Standard Classification of Education 1997: Challenges in the application to national data and the implementation in cross-national surveys. *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for*, *15*, 13-46.

Scotton, C. M. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford University Press.

Segalowitz, S. J., and Davies, P. L. (2004). Charting the maturation of the frontal lobe: an electrophysiological strategy. *Brain Cogn.* 55, 116–133.

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.

Shallice T. (2002). Fractionation of the supervisory system. In: Stuss D, Knight R, editors. Principles of frontal lobe function. Oxford: Oxford University Press. p. 261--277.

Shallice, T. & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114,* 727-741.

Shallice, T. & Burgess, P.W. (1996). Domains of supervisory control and the temporal organisation of behaviour. *Philosophical Transactions of the Royal Society B, 351*, 1405–1412. [Reprinted in A.C. Roberts, T.W. Robbins, & L. Weiskrantz (Eds.), *The prefrontal cortex: Executive and cognitive function*

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*,*298*(1089), 199-209.

Shallice, T. (1988). *From neuropsychology to mental structure*. New York: Cambridge University Press.

Shallice, T., & Burgess, P. (1993). Supervisory control of action and thought selection.

Shallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organisation of behaviour. *Philosophical Transactions of the Royal Society of London, 131*, 1405–1412.

Shallice, T., Burgess, P., 1996. The domain of supervisory processes and temporal organization of behaviour. Philos. Trans. R. Soc. London, B 351, 1405– 1412.

Shallice, T., Stuss, D. T., Picton, T. W., Alexander, M. P., & Gillingham, S. (2008). Mapping task switching in frontal cortex through neuropsychological group studies. *Frontiers in neuroscience*, *2*(1), 79.

Shek, D. T. (2011). Quality of life research: Responses to emerging issues in a changing world. *Social Indicators Research*, *100*(3), 371-374.

Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia*, *40*(6), 605-619.

Shipstead, Z., Redick, T. S. & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138,* 628–54

Shute, G. E., & Huertas, V. (1990). Developmental variability in frontal lobe function. *Developmental Neuropsychology*, *6*(1), 1-11.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology, 7,* 268–288.

Simon, J. R. (1969). Reactions towards the source of stimulation. *Journal of Experimental Psychology, 81, 174−176.*

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(04), 771-781.

Singleton, D. (2003). Critical Period or General Age. *Age and the acquisition of English as a foreign language*, 1.

Singleton, D. D. M., & Ryan, L. (2004). *Language acquisition: The age factor* (Vol. 47). Multilingual matters.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research*, *75*(3), 417-453.

Skutnabb-Kangas, Tove (1987). Are the Finns in Sweden an Ethnic Minority? Finnish Parents Talk about Finland and Sweden. Research Project "The education of the Finnish minority in Sweden", Working Paper nr 1. Roskilde: Roskilde University Centre, Institute VI.

Skutnabb-Kangas, Tove (2000). *Linguistic genocide in education – or worldwide diversity and human rights?* Mahwah, New Jersey & London: Lawrence Erlbaum Associates.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*(5408), 1657-1661.

Smith, F. (1923). Bilingualism and mental development. *British Journal of Psychology. General Section*, *13*(3), 271-282.

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: the development of causal reasoning in toddlers and infants. *Developmental psychology*, *42*(6), 1103.

Somsen, R. J. M. (2007). The development of attention regulation in the Wisconsin Card Sorting Task. *Developmental Science*, *10*, 664-680.

Sorace, A. (2011). Pinning down the concept of interface in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1-33.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E. & Toga A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience*, *24*, 8223--8231.

Spreen, O., Strauss, E. (1998). *A Compendium of neuropsychological tests : Administration, norms, and commentary.* (2nd edition). NY. Oxford University Press

St. Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*(4), 745-759.

Stark, W. (1939). The effect of bilingualism on general intelligence: an investigation carried out in certain Dublin Primary Schools. Theses Presented for Higher Degrees, University of Glasgow.

Statistics Canada. (2003a). Education indicators in Canada: report of the Pan-Canadian education indicators programme. *Catalogue no. 81-582-XPE*. Ottawa, ON: Canada. Authority of the Minister Responsible for Statistics Canada, Toronto.

Stevens, C., Lauinger, B., & Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: an event-related brain potential study. *Developmental science*, *12*(4), 634-646.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Stuss, D. T. (1992). Biological and psychological-development of executive functions. *Brain and cognition*, *20*(1), 8–23.

Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: a conceptual view. *Psychological research*, *63*(3-4), 289-298.

Stuss, D. T., & Benson, D. F. (1986). The Frontal LobesRaven Press. *New York*, 97ff.

Swain, M. (1996). Discovering successful second language teaching strategies and practices: From programme evaluation to classroom experimentation. *Journal of Multilingual and Multicultural Development*, *17*(2-4), 89-104.

Swain, M. K. (1972). *Bilingualism as a first language*. Ph.D. dissertation, University of California.

Swain, M., & Johnson, R. K. (1997). Immersion education. *Immersion education: International perspectives*, 1-16.

Swain, M., & Lapkin, S. (1991). Heritage language children in an English-French bilingual program. *Canadian Modern Language Review*, *47*(4), 635-641.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*(3), 320-337.

Swain, N. & Lapkin, S. (1982). Evaluating bilingual education. Clevedon, England: Multilingual Matters.

Sylvester, C. Y. C., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, *41*(3), 357-370.

Takakuwa, M. (2003). Lessons from a paradoxical hypothesis: A methodological critique of the Threshold Hypothesis. In *El adás: 4th International Symposium on Bilingualism. Arizona State University, April* (Vol. 30).

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in psychology*, *2*.

Thatcher R. (1991).Maturation of the human frontal lobes: Physiological evidence for staging. *Developmental Neuropsychology, 7,* 397–419.

Thatcher RW. Neural coherence and the content of consciousness. Consciousness and Cognition: An International Journal. 1997; 6(1):42–49.

Thomas, E. M., & Roberts, D. B. (2011). Exploring bilinguals' social use of language inside and out of the minority language classroom. *Language and Education*, *25*(2), 89-108.

Thorell LB, Lindqvist S, Bergman N, Bohlin G, Klingberg T. 2009. Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12:106–13

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203-214.

Torrance, E. P. (1998). *The Torrance tests of creative thinking norms—technical manual figural (streamlined) forms A & B.* Bensenville, IL: Scholastic Testing Service, Inc.

Torrance, E. P., Gowarn, J. C., Wu, J. M. & Aliotti, N. C. (1970). Creative functioning of monolingual and bilingual children in Singapore. *Journal of Educational Psychology, 61,* 72-75.

Toukomaa, P., & Skutnabb-Kangas, T. (1977). *The intensive teaching of the mother tongue to migrant children at pre-school age*. Department of sociology, University of Tampere.

Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Clevedon, U.K.: Multilingual Matters.

Tuckman, B. W., & Hinkle, J. S. (1986). An experimental study of the physical and psychological effects of aerobic exercise on schoolchildren. *Health Psychology, 5*, 197–207.

Turian, D., & Altenberg, E. P. (1991). 14 Compensatory strategies of child first language attrition. *First language attrition*, 207.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological science*, *14*(6), 623-628.

UNESCO. International Standard Classification of Education (ISCED-97). (1997). http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm

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*(5), 427-449.

Van der Ven, S. H., Kroesbergen, E. H., Boom, J., & Leseman, P. P. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*(1), 100-119.

Van Heuven, W. J., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex*, *18*(11), 2706-2716.

Van Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review*, *17*(2), 147-177.

Van Valin, R. (1991). Functionalist linguistic theory and language acquisition. *First Language, 11*(31), 7-40.

Vanderploeg, R. D., Schinka, J. A., & Retzlaff, P. (1994). Relationships between measures of auditory verbal learning and executive functioning. *Journal of Clinical and Experimental Neuropsychology*, *16*(2), 243-252.

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, *110*(1), 84-99.

Vihman, M. M. (1985). Language differentiation by the bilingual infant. *Journal of child language*, *12*(2), 297-324.

Vincent, K. R. & Cox, J. A. (1974). A re-evaluation of Raven's Standard Progressive Matrices. *The Journal of Psychology*, *88*, 299-303.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of child language*, *5*(02), 311-326.

Wager, T. D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: a meta-analysis. *Neuroimage*, *22*(4), 1679-1693.

Wais, P. E., & Gazzaley, A. (2011). The impact of auditory distraction on retrieval of visual memories. *Psychonomic bulletin & review*, *18*(6), 1090-1097.

Waldman, I. D. (2005). Statistical approaches to complex phenotypes: evaluating neuropsychological endophenotypes for attention-deficit/hyperactivity disorder. *Biological psychiatry*, *57*(11), 1347-1356.

Walsh, K. W. (1978). *Neuropsychology: a clinical approach.* Edinburgh: Churchill Livingstone.

Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *NeuroImage*, *35*(2), 862-870.

Wass, S. et al. (2011) Training attentional control and working memory: Is younger better? Devlopmental Review

Wechsler, D. (1981). *Wechsler adult intelligence scale-revised*. Psychological Corporation.

Wecker, N. S., Kramer, J. H., Wisniewski, A., Delis, D. C., & Kaplan, E. (2000). Age effects on executive ability. *Neuropsychology*, *14*(3), 409.

Weinreich, U. (1953). Languages in contact. New York: Linguistic Circle of New York.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, *72*(3), 655-684.

Welsh, M. C., & Pennington, B. F. (1988). Assessing frontal lobe functioning in children: views from developmental psychology. *Developmental Neuropsychology, 4*(3), 199-230

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. Developmental Neuropsychology, 7(2), 131_149.

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.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, *108*(3), 436-452.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological psychiatry*, *57*(11), 1336-1346.

Willcutt, E. G., Pennington, B. F., Olson, R. K., & DeFries, J. C. (2007). Understanding comorbidity: A twin study of reading disability and attention-deficit/hyperactivity disorder. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *144*(6), 709-714.

Williams, B., Ponesse, J., Schacher, R., Logan, G., & Tannock, R. (1999). Development of inhibitory control across the life span. Developmental Psychology, 35, 205–213.

Williams, I. W. (2003). Our children’s language: The Welsh-medium schools of Wales, 1939-2000. Talybont (Wales): Y Lolfa.

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2010). The measurement of executive function at age 3 years: psychometric properties and criterion validity of a new battery of tasks. *Psychological assessment*, *22*(2), 306.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition*, *13*, 103–128.

Yan, S., & Nicoladis, E. (2009). Finding le mot juste: Differences between bilingual and monolingual children’s lexical access in comprehension and production. *Bilingualism: Language and Cognition*, *12*(3), 323-335.

Yang, H., Yang, S., Ceci, S. J., & Wang, Q. (2005). Effects of bilinguals’ controlled-attention on working memory and recognition. In *Proceedings of the 4th international symposium on bilingualism*.

Yang, S., Shih, J., & Lust, B. (2005). Exploring cognitive advantages of childhood bilingualism through new tests of executive attention. In *Poster presented at the biennial meeting of the Society for Research in Child Development, Atlanta, GA* (pp. 7-10).

Yang, S., Yang, H., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism: Language and Cognition*, *14*(3), 412-422.

Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience & Biobehavioral Reviews*, *33*(8), 1168-1177.

Yeatman, J. D., Ben-Shachar, M., Glover, G. H., & Feldman, H. M. (2010). Individual differences in auditory sentence comprehension in children: An exploratory event-related functional magnetic resonance imaging investigation. *Brain and language*, *114*(2), 72-79.

Yelland, G. W., Pollard, J., & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. *Applied psycholinguistics*, *14*(04), 423-444.

Yoshida, H. (2008). The cognitive consequences of early bilingualism. *Zero to Three*, *29*(2), 26-30.

Yudes, C., Macizo, P., & Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in psychology*, *2*.

Zagar, R., Arbit, J., & Friedland, J. (1980). Structure of a psychodiagnostic test battery for children. *Journal of Clinical Psychology*, *36*(1), 313-318.

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature neuroscience*, *14*(5), 656-661.

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science*, *7*(4), 121-126.

Zelazo, P. D., & Mu¨ller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), Handbook of childhood cognitive development (pp. 445 – 469). Oxford, England: Blackwell.

Zelazo, P. D., & Müller, U. (2010). Executive function in typical and atypical children. *Blackwell handbook of cognitive development*, *2*, 574-603.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of general psychology*, *1*(2), 198.

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. Acta Psychologica, 115, 167–184.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*(1), 37-63.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood: I. The development of executive function. *Monographs of the society for research in child development*.

Zelazo, P. D., Qu, L., & Müller, U. (2005). Hot and cool aspects of executive function: Relations in early development. *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind*, 71-93.

Zelazo, P. D., Resnick, J. S. & Pinon, D. E. (1995). Response control and the execution of verbal rules. *Developmental Psychology*, 31, 508–517.

Appendices

APPENDIX I – Parent/caregiver socioeconomic status, language demographic and proficiency questionnaire

Queen's University Belfast

Children's Language Proficiency/Background Questionnaire for Parents/Guardians
Ceistiúchán Inniúlacht Teanga an Pháiste do Thuismitheoirí/Caomhnóirí

Dáta/ Date:__________________

Scoil/ School: ___________________

Ainm an páiste/Child's name: ________________________________

Iscne an páiste / Child's Gender (ciorcal/ circle):

Male/Fireann   Female/Baineann

Tír breithe do pháiste / Child's country of birth: ___________________

Dáta breithe do pháiste / Child's date of birth:________________

Aois/Age: (Blianta/Years) _____

Blianta ina c(h)ónaí in Éireann/ Years living in Ireland:______________

Current year in school: _________________(e.g., Rang 1, Rang 5)

Uimhir bhlianta teagaisc foirmiúil (Naíscoil san áireamh)/ Number of years of formal instruction (including Nursery):

Tríd Béarla/ through English: _______________
Tríd Gaeilge/ through Irish: _______________

Uimhir bhlianta táithí le/Number of years' experience with:
Béarla/ English:_________________
Gaeilge/ Irish:_______________

Teanga labhartha (ciorcal)/ **Languages spoken** (circle relevant):

English   /   Irish   / Other(s):   (please name)_______________________

BACKGROUND INFORMATION QUESTIONNAIRE

***Tuismitheoir/Caomhnóir 1/* Parent/Guardian 1:**

Caidreamh len pháiste/Relationship to child (e.g. Father): ___________________________

Fostaíocht faoi láthair/Current Occupation:

____

Stádas Fostaíocht/ Employment Status:

Full-Time ☐ Part-Time ☐ Unemployed ☐

Other (Please Specify Below) ☐

____

Cáilíochtaí/ Qualifications (*Tic gach ceann a bhaineann leat*/ Tick all that apply):

Junior Certificate ☐ Leaving Certificate ☐ NVQ L1 ☐ L2 ☐ L3 ☐ Bachelor's Degree ☐

Master's Degree ☐ Other Postgraduate Qualification (e.g. PhD) ☐ Professional Qualification (Please Specify Below) ☐ Other (Please Specify Below) ☐

____

***Tuismitheoir/Caomhnóir 2 (Roghnach)/* Parent/Guardian 2**

Caidreamh len pháiste/Relationship to child (e.g. Mother): ___________________________

Fostaíocht faoi láthair/Current Occupation:

____

Stádas Fostaíocht/ Employment Status:

Full-Time ☐ Part-Time ☐ Unemployed ☐

Other (Please Specify) ☐

____

Cáilíochtaí/ Qualifications (*Tic gach ceann a bhaineann leat*/ Tick all that apply):

Junior Certificate ☐ Leaving Certificate ☐ NVQ L1 ☐ L3 ☐ ☐ Bachelor's Degree ☐

Master's Degree ☐ Other Postgraduate Qualification (e.g. PhD) ☐ Professional Qualification (Please Specify Below) ☐ Other (Please Specify Below) ☐

NOTE: All the information provided is completely confidential and anonymous. Once filled out if you would like to seal the questionnaires in the envelope provided and sign over the seal to maintain anonymity. A sealed box will be left in the classroom where children can post the sealed envelope. We would also like to take this opportunity to thank you for your patience, time and energy in completing these forms. We are extremely grateful and hope that this information will help to improve the quality of our work.

Instructions / Treoracha

Ar dtús, tá cúpla scála ceistiúcháin le líonadh. Beidh ort marc a chur ar gach scála chun an ceist a fhreagairt.

Below, there are a few questionnaire scales to complete. You need to mark each scale in order to answer the question

Seo sampla chun an stíl freagrach a léiriú:
Here is an example to demonstrate the style of answering:

I do thuaraim, cé comh maith is atá do pháiste ábalta Béarla a léamh?

In your opinion, how well is your child able to read in English?

Níl sé/sí ábalta é a léamh ar chor ar bith/ He/She cannot read at all
Cannot read at all

Ábaltacht an-ard léitheoireacht/ Exceptionally high reading ability

Má cheapann tú gur léitheoir maith é/í do pháiste; ach ó am go ham, go mbíonn deacrachtaí aige/aicí, b'fheidir go gcuirfeá marc ceartingearrach ar an scála mar seo:

If you feel that your child reads well in English, but experiences some difficulties from time to time, you may wish to mark the scale like this:

Níl sé/sí ábalta é a léamh ar chor ar bith/ He/She cannot read at all
Cannot read at all

Ábaltacht an-ard léitheoireacht/ Exceptionally high reading ability
high ability

Do thuaraim féin atá i gcéist, agus níl freagair ceart nó mícheart ann. Tá na freagraí seo rúnda agus usáidtear iad ar mhaithe leis an stáidear seo amháin.

Remember, this is in your opinion, and there are no right and wrong answers. These answers will be treated as confidential and only used for the purposes of this study.

Cuir marc ingearach ar na scála seo thíos, le do thoil

Please vertically mark along the scales below

Ag Léamh/Reading

1. I do thuairim, cé comh maith is atá do pháiste ábalta Béarla a léamh?
In your opinion, how well is your child able to read in English?

Níl sé/sí ábalta é a léamh ar chor ar bith/ He/She cannot read at all Cannot read at all	Ábaltacht an-ard léitheoireacht/ Exceptionally high reading ability

2. I do thuairim, cé comh maith is atá do pháiste ábalta Gaeilge a léamh?
In your opinion, how well is your child able to read in Irish?

Níl sé/sí ábalta é a léamh ar chor ar bith/ He/She cannot read at all Cannot read at all	Ábaltacht an-ard léitheoireacht/ Exceptionally high reading ability

Ag Scríobh/Writing

3. I do thuairim, cé comh maith is atá do pháiste ábalta scríobh as Béarla?
In your opinion, how well is your child able to write in English?

Níl sé/sí ábalta scríobh as Béarla ar chor ar bith/Cannot write in English at all	Ábaltacht an-ard scríbhneoireacht i mBéarla/ Exceptionally high writing ability in English

4. I do thuairim, cé comh maith is atá do pháiste ábalta scríobh as Gaeilge?
In your opinion, how well is your child able to write in Irish?

Níl sé/sí ábalta scríobh as Gaeilge ar chor ar bith/Cannot write in Irish at all

Ábaltacht an-ard scríbhneoireacht i nGaeilge/ Exceptionally high writing ability in Irish

Ag Labhairt/Speaking

5. I do thuairim, cé comh maith is atá do pháiste ábalta Béarla a labhairt?
In your opinion, how well is your child able to speak in English?

Níl sé/sí ábalta Béarla a labhairt ar chor ar bith/Cannot speak English at all

Tá Béarla líofa aige/aici/ Fluent in English

6. I do thuairim, cé comh maith is atá do pháiste ábalta Gaeilge labhairt?
In your opinion, how well is your child able to speak in Irish?

Níl sé/sí ábalta Gaeilge a labhairt ar chor ar bith/ Cannot speak Irish at all

Tá Gaeilge líofa aige/aici / Fluent in Irish

Tuiscint/Understanding

7. I do thuairim, cé comh maith is atá do pháiste ábalta Béarla labhartha a thuiscint?
In your opinion, how well is your child able to understand spoken English?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand any spoken English

Tuiscint den scoith ar Béarla labhartha/ Exceptionally high understanding of spoken English

8. I do thuairim, cé comh maith is atá do pháiste ábalta Gaeilge labharta a thuiscint?
In your opinion, how well is your child able to understand spoken Irish?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand any spoken Irish
Cannot understand any

Tuiscint den scoith ar Béarla labhartha/ Exceptionally high understanding of spoken Irish ability

9. I do thuaraim, cé comh maith is atá do pháiste ábalta Béarla scríofa a thuiscint?
In your opinion, how well is your child able to understand written English?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand any written English
Cannot understand

Tuiscint den scoith ar Béarla labhartha/ Exceptionally high understanding of written English ability

10. I do thuaraim, cé comh maith is atá do pháiste ábalta Gaeilge scríofa a thuiscint?
In your opinion, how well is your child able to understand written Irish?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand any written Irish
Cannot understand any

Tuiscint den scoith ar Béarla labhartha/ Exceptionally high understanding of written Irish ability

<u>Úsáid Teanga/Language Usage</u>

11. Cén teanga is minice a úsáidtear agus tú sa bhaile le do pháiste?
What language do you use most often when you are at home with your child?

Béarla/English | Go Cothrom/Equally | Gaeilge/Irish

12. Má tá tuismitheoir/caomhnóir eile ag an páiste, cén teanga is minice a úsáideann siad iad sa bhaile leo?
If the child has another parent/guardian, what language do they use most often at home with them?

Béarla/English Go Cothrom/Equally Gaeilge/Irish

13. Cén teanga is minice a úsáideann do pháiste agus iad ag labhairt le deartháireacha/deirfiúracha sa bhaile?
What language does your child use most often when speaking to brothers/sisters at home?

Béarla/English Go Cothrom/Equally Gaeilge/Irish

14. Cén teanga is minice a úsáideann do pháiste agus iad ag labhairt le daoine eile i do theaghlach?
What language does your child use most often when speaking to other members of your family?

Béarla/English Go Cothrom/Equally Gaeilge/Irish

15. Cén teanga is minice a úsáideann do pháiste agus iad ag labhairt lena c(h)airde?
What language does your child use most often when speaking to his/her friends?

Béarla/English Go Cothrom/Equally Gaeilge/Irish

16. Cén teanga is minice a úsáideann do pháiste agus iad ag glacadh páirt in imeachtaí sóisialta (m.sh. spóirt, damhsa, seirbhísí reiligiúnacha srl.)?

What language does your child use most often when taking part in social events (e.g. sports, dance, religious services etc.)?

Béarla/English Go Cothrom/Equally Gaeilge/Irish

ANY ADDITIONAL COMMENTS: AON BARÚIL EILE

Go Raibh Míle Maith Agat

Thank You

APPENDIX II – Teacher proficiency questionnaire

CEISTIÚCHÁN MAIDIR LE OILTEACHT: LEAGAN GAEILGE DO MÚINTEOIRÍ
LANGUAGE PROFICIENCY QUESTIONNAIRE: TEACHER'S RATING

Ainm Páiste/Child's name: __________________

Data/Date:__________________

Scoil/School: ______________________________

Cuir marc ceartingearrach ar na scála méadracha seo thíos, le do thoil:
Please vertically mark along the scales below:

Ag Léamh/Reading

I do thuairim, cé comh maith is atá an dalta ábalta Béarla a léamh?
In your opinion, how well is the student able to read in English?

Níl sé/sí ábalta é a léamh ar chor ar bith/ Cannot read at all	Béarla líofa/ Fluent, exceptionally high ability

I do thuairim, cé comh maith is atá an dalta ábalta Gaeilge a léamh?
In your opinion, how well is the student able to read in Irish?

Níl sé/sí ábalta é a léamh ar chor ar bith/ Cannot read at all	Gaeilge líofa/ Fluent, exceptionally high ability

Ag Scríobh/Writing

I do thuairim, cé comh maith is atá an dalta ábalta scríobh as Béarla?
In your opinion, how well is the student able to write in English?

Níl sé/sí ábalta é a scríobh ar chor ar bith/ Cannot write at all	Béarla líofa/ Fluent, exceptionally high ability

I do thuairim, cé comh maith is atá an dalta ábalta scríobh as Gaeilge?
In your opinion, how well is the student able to write in Irish?

Níl sé/sí ábalta é a scríobh ar chor ar bith/ Cannot write at all	Gaeilge líofa/ Fluent, exceptionally high ability

Ag Labhairt/Speaking

I do thuairim, cé comh maith is atá an dalta ábalta Béarla a labhairt?
In your opinon, how well is the student able to speak in English?

Níl sé/sí ábalta é a labhairt ar chor ar bith/ Cannot speak at all	Béarla líofa/ Fluent, exceptionally high ability

I do thuairim, cé comh maith is atá an dalta ábalta Gaeilge a labhairt?
In your opinon, how well is the student able to speak in Irish?

Níl sé/sí abalta é a labhairt ar chor ar bith/ Cannot speak at all	Gaeilge líofa/ Fluent, exceptionally high ability

Tuiscint/Understanding

I do thuairim, cé comh maith is atá an dalta ábalta Béarla labhartha a thuiscint?
In your opinon, how well is the student able to understand spoken English?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand at all — Béarla líofa/ Fluent, exceptionally high ability

I do thuairim, cé comh maith is atá an dalta ábalta Gaeilge labhartha a thuiscint?
In your opinon, how well is the student able to understand spoken Irish?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand at all — Gaeilge líofa/ Fluent, exceptionally high ability

I do thuairim, cé comh maith is atá an dalta ábalta Béarla scríofa a thuiscint?
In your opinon, how well is the student able to understand written English?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand at all — Béarla líofa/ Fluent, exceptionally high ability

I do thuairim, cé comh maith is atá an dalta ábalta Gaeilge scríofa a thuiscint?
In your opinon, how well is the student able to understand written Irish?

Níl sé/sí ábalta é a thuiscint ar chor ar bith/ Cannot understand at all — Gaeilge líofa/ Fluent, exceptionally high ability

Go raibh míle maith agat

APPENDIX III – Children's language demographic and proficiency questionnaire

Your languages questionnaire

Treoireacha/ Instructions

You can take part in this study if you want to. If you don't want to, you can tell the researcher, at any time. We would like to find out how children who go to Irish Medium schools feel about their skills in Irish and English. We would also like to find out how you get to use your Irish and English, and what kinds of things you like about both of your languages. Any questions? OK to continue?

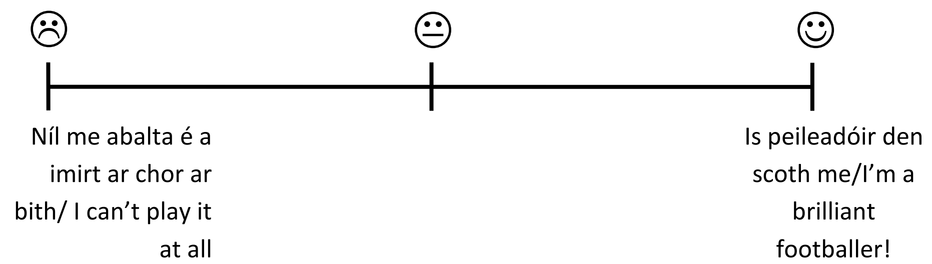
The researcher will help you to mark your answers. Here is an example question.

Tá mé go maith ag imirt péil/ I am good at playing football

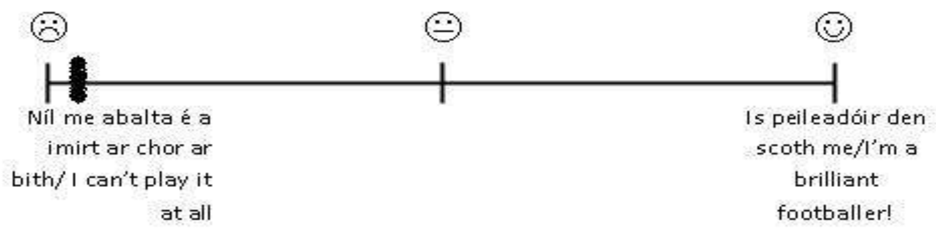
Look at the line below. On the left there is a sad face that says "I can't play at all". On the right is a really happy smiley face who says "I'm a brilliant footballer". You can point anywhere along the line from "can't play it at all" to "I'm a brilliant footballer" to show how good you think you are at football. Where do you think you would be on the line? Point to where you think is right for you and I will help you to mark the place where you pointed to. Do you have any questions about how to do this?

Remember, there are no right and wrong answers, just point to what you think are right.

Tá mé go maith ag imirt péil/ I am good at playing football



Here is my answer as an example. I am not good at football at all, but I can play it. So I would probably mark my answer quite low down, like this.



You don't have to copy my answer, it was just to show you that your answer can be anywhere along the line.

Try your best with this question now. Point to where you think you would be.

Is maith liom ag féachaint ar Blue Peter/ I like watching Blue Peter

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it

Is maith liom é go mór/ I really like it

Maith thú. An dtuigeann tú cad atá i gceist anois? Sásta chun leanadh ar aghaidh?
Good. Do you understand what you have to do now? Ok to continue?

Ag léamh/Reading

Tá mé go maith ag léamh i mBéarla/ I'm good at reading in English

☹ 😐 ☺

Níl mé abalta léamh i mBéarla/ I can't read at all in English

Tá mé ar fheabhas ag léamh i mBéarla /I'm brilliant at English reading

Tá mé go maith ag léamh i nGaeilge/ I'm good at reading in Irish

☹ 😐 ☺

Níl mé abalta léamh i nGaeilge/ I can't read at all in Irish

Tá mé ar fheabhas ag léamh i nGaeilge /I'm brilliant at Irish reading

Tá mé go maith ag scríobh i mBéarla/ I'm good at writing in English

☹ 😐 ☺

Níl mé abalta scríobh i mBéarla/ I can't write at all in English

Tá mé ar fheabhas ag scríobh i mBéarla /I'm brilliant at English writing

Tá mé go maith ag scríobh i nGaeilge/ I'm good at writing in Irish

☹ 😐 ☺

Níl mé abalta scríobh i nGaeilge/ I can't write at all in Irish

Tá mé ar fheabhas ag scríobh i nGaeilge/I'm brilliant at Irish writing

Ag labhairt/Speaking

Tá mé go maith ag labhairt i mBéarla/ I'm good at speaking English

☹ 😐 ☺

Níl mé abalta labhairt i mBéarla/ I can't speak at all in English

Tá mé ar fheabhas ag labhairt i mBéarla /I'm brilliant at speaking English

Tá mé go maith ag labhairt i nGaeilge/ I'm good at speaking Irish

☹ 😐 ☺

Níl mé abalta labhairt i nGaeilge/ I can't speak at all in Irish

Tá mé ar fheabhas ag labhairt i nGaeilge/I'm brilliant at speaking Irish

Ag Tuiscint/Understanding

Tá mé go maith ag tuiscint Béarla labhartha/ I'm good at understanding spoken English

☹ 😐 ☺

Níl mé abalta Béarla labharta a thuiscint/ I can't understand spoken English

Tá mé ar fheabhas ag tuiscint Béarla labharta/I'm brilliant at understanding spoken English

Tá mé go maith ag tuiscint Gaeilge labhartha/ I'm good at understanding spoken Irish

☹ 😐 ☺

Níl mé abalta Gaeilge labharta a thuiscint/ I can't understand spoken Irish

Tá mé ar fheabhas ag tuiscint Gaeilge labharta/I'm brilliant at understanding spoken Irish

Tá mé go maith ag tuiscint Béarla scríofa/ I'm good at understanding written English

☹ 😐 ☺

Níl mé abalta Béarla scríofa a thuiscint/ I can't understand written English

Tá mé ar fheabhas ag tuiscint Béarla scríofa/I'm brilliant at understanding written English

Tá mé go maith ag tuiscint Gaeilge scríofa/ I'm good at understanding written Irish

☹ 😐 ☺

Níl mé abalta Gaeilge scríofa a thuiscint/ I can't understand written Irish

Tá mé ar fheabhas ag tuiscint Gaeilge scríofa/I'm brilliant at understanding written Irish

Language Use/ Usáid teangacha

Usáideann tú cén teanga is minic nuair atá tú sa bhaile le do Mhamaí?

What language do you mostly use when you are at home with your mum?

Bearla amháin/ All English

Gaelige agus Béarla comh minic/ Equally both

Gaeilge amháin/ All Irish

Usáideann tú cén teanga is minic nuair atá tú sa bhaile le do Dhaidí?

What language do you mostly use when you are at home with your dad?

Bearla amháin/ All English | Gaelige agus Béarla comh minic/ Equally both | Gaeilge amháin/ All Irish

…ga is minic nuair … deartharacha/…

What language do you mostly use wher… with brothers/ sisters?

Bearla amháin/ All English | Gaelige agus Béarla comh minic/ Equally both | Gaeilge amháin/ All Irish

Usáideann tú cén teanga is minic nuair atá tu ag labhairt le daoine eile i do chlann?

What language do you mostly use when speaking with other members of your family?

Bearla amháin/ All English | Gaelige agus Béarla comh minic/ Equally both | Gaeilge amháin/ All Irish

Usáideann tú cén teanga is minic nuair … le do chairde ina bhfuil Gaelige acu, nuair atá tú sa scoil?

What language do you mostly use when speaking with Irish speaking friends, when you are in school?

Bearla amháin/ All English | Gaelige agus Béarla comh minic/ Equally both | Gaeilge amháin/ All Irish

Usáideann tú cén teanga is minic nuair atá tú ag labhairt le do chairde ina bhfuil Gaelige acu, taobh amuigh den scoil?

What language do you mostly use when speaking with Irish speaking friends, outside of school?

Bearla amháin/ All English | Gaelige agus Béarla comh minic/ Equally both | Gaeilge amháin/ All Irish

Usáideann tú cén teanga is minic nuair atá tú ag glacadh páirt in imeachtaí sóisialta (m.s., imeachtaí spóirt, drámaíocht, seirbhisí reiligiúin)?

What language do you mostly use when you are taking part in community activities (e.g., sports, drama, religious services)?

Bearla amháin/ All English	Gaelige agus Béarla comh minic/ Equally both	Gaeilge amháin/ All Irish

C*én teanga ina mbíonn tú ag smaoinea tú sa scoil? I rith ranganna, sos agus lón.*

What language do you think in when you are in school? During classes, break and lunch.

Bearla amháin/ All English	Gaelige agus Béarla comh minic/ Equally both	Gaeilge amháin/ All Irish

C*én teanga ina mbíonn tú ag smaoineamh tríd, nuair atá tú sa bhaile?*

What language do you think in when you are at home?

Bearla amháin/ All English	Gaelige agus Béarla comh minic/ Equally both	Gaeilge amháin/ All Irish

Dearcthaí/ Attitudes

Is maith liom ag léamh i mBéarla/ I like reading in English

☹	😐	☺
Níor mhaith liom é ar chor ar bith / I really don't like it		Is maith liom é go mór/ I really like it

Is maith liom ag léamh i nGaeilge/ I like reading in Irish

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it

Is maith liom é go mór/ I really like it

Is maith liom ag scríobh i mBéarla/ I like writing in English

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it

Is maith liom é go mór/ I really like it

Is maith liom ag scríobh i nGaeilge/ I like writing in Irish

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it

Is maith liom é go mór/ I really like it

Is maith liom ag labhairt i mBéarla/ I like speaking in English

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it

Is maith liom é go mór/ I really like it

Is maith liom ag labhairt i nGaeilge/ I like speaking in Irish

☹ 😐 ☺

Is maith liom ag féachaint ar clár teilifíse i nGaeilge/ I like watching T.V. programmes in Irish

Níor mhaith liom é ar chor ar bith / I don't like it — Is maith liom é go mór/ I really like it

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it — Is maith liom é go mór/ I really like it

Is maith liom ag féachaint ar clár teilifíse i nGaeilge/ I like watching T.V. programmes in Irish

☹ 😐 ☺

Níor mhaith liom é ar chor ar bith / I really don't like it — Is maith liom é go mór/ I really like it

How do you feel about being able to speak Irish? Do you like it/not like it? Cad a cheapann tú faoi do chuid Gaeilge. An maith leat/nach maith leat é?

APPENDIX IV – Opposite Worlds Raw Timing Scores (sec)

Condition	Socioeconomic Status	Language Group	Time 1	Time 2	Time 3
	Mid-SES	Monolingual (*n* = 32)	24.56 (3.31)	23.09 (2.55)	20.63 (2.27)
		Bilingual (*n* = 48)	24.75 (4.05)	22.25 (3.15)	19.45 (2.54)
		Total (*n* = 80)	24.67 (3.75)	22.59 (2.93)	19.94 (2.49)
Same World (Congruent)	Low-SES	Monolingual (*n* = 18)	24.17 (2.26)	23.28 (2.47)	20.80 (1.77)
		Bilingual (*n* = 19)	24.21 (5.47)	23.45 (2.55)	20.65 (2.35)
		Total (*n* = 37)	24.19 (4.17)	23.36 (2.48)	20.72 (2.06)
	Total	Monolingual (*n* = 50)	24.60 (4.46)	22.59 (3.02)	19.80 (2.53)
		Bilingual (*n* = 67)	24.42 (2.96)	23.16 (2.50)	20.69 (2.09)
		Total (*n* = 117)	24.52 (3.87)	22.84 (2.81)	20.19 (2.38)
	Mid-SES	Monolingual (*n* = 32)	31.28 (4.79)	29.18 (4.01)	25.89 (3.64)
		Bilingual (*n* = 48)	31.88 (5.34)	27.94 (4.84)	24.25 (3.36)
		Total (*n* = 80)	31.64 (5.11)	28.45 (4.53)	24.94 (3.55)
Opposite World (Ingoncruent	Low-SES	Monolingual (*n* = 18)	33.94 (4.28)	30.01 (2.62)	27.97 (2.70)
		Bilingual (*n* = 19)	32.53 (8.28)	29.65 (3.93)	24.91 (2.73)
		Total (*n* = 37)	33.22 (6.59)	29.83 (3.30)	26.39 (3.09)
	Total	Monolingual (*n* = 50)	32.24 (4.75)	29.48 (3.57)	26.61 (3.46)
		Bilingual (*n* = 67)	32.06 (6.25)	28.42 (4.64)	24.44 (3.18)
		Total (n = 117)	32.14 (5.64)	28.88 (4.22)	25.40 (3.47)

APPENDIX V – Individual Growth Curve Model Fit Statistics

Task		-2 log likelihood	AIC	BIC
Stroop Neutral	Model 1	4050.86	4054.86	4062.37
	Linear	3961.25	2969.25	3984.25
	Predictors	3919.5	3927.5	3942.42
Stroop Congruent	Model 1	2084.77	2088.81	2096.44
	Linear	1872.08	1880.21	1895.4
	Predictors	1866.18	1874.3	1889.42
Stroop Incongruent	Model 1	1875.04	1879.04	1886.7
	Linear	1822.05	1830.05	1845.36
	Predictors	1824.05	1832.05	1847.3
Trails A	Model 1	3154.84	3158.84	3166.48
	Linear	3012.67	3029.67	3044.94
	Predictors	3003.24	3011.24	3026.44
Trails B	Model 1	3401.16	3405.16	3412.8
	Linear	3218.25	3226.25	3241.51
	Quadratic	3207.26	3215.38	3230.52
	Predictors	3155.08	3163.08	3178.23

APPENDIX VI – Example of parent/caregiver consent form for participation

A thuismitheor/ chaomhnóir,

Cláir Nic Stiofáin is ainm dom agus is mac léinn PhD mé ón Scoil Siceolaíochta, Ollscoil na Banríona, i mBéal Feiriste. I gceann cúpla seachtain beidh mé ag tabhairt cuairt ar scoil do pháiste le haghaidh taighde a dhéanamh, ar aird i bpáistí. Tá an taighde seo, atá urraithe ag an Comahirle um Oideachas Gaeltachta agus Gaelscolaíochta (COGG), páirteach de clár taighde in Ollscoil na Banríona ar páistí dátheangach, agus an tumoideachas. Deirtear go bhfuil buntáistí éagsúla ag páistí atá i ngaelscoileanna, go háirithe bainte le fobairt cognaíoch na páistí. Tá muidne ag iarraidh níos mó eolas a fháil faoi na difríocht seo.

Beidh páistí i Rang a 5 ag déanamh cúraimí gearr, páipéar-agus-peann luaidhe, ag tomhais scileanna airde. Beidh na cúraimí curtha i láthair i bhfoirm cluichí agus beidh na páistí curtha ar a shuaimhneas nach scrúdú atá ann. Tabharfaidh mé na cluichí do na páistí go aonarach i seomra ciúin sa scoil. I dtáithí s'againn baineann na páistí an-taitneamh as na cluichí.

Bhéimis an-buíoch dá mbeifeá sásta cead a thabhairt do do pháiste páirt a ghlacadh sa stáidéar. Tá taighde s'againn tughta de na caighdeáin is airde eitice agus proifisiúta, mar a leagtar thíos ag an 'British Psychological Society' agus glactar gach cúram chun spéis agus sábhaltacht gach páiste a chinntiú. Chomh maith le seo, tugadh formheas eiticiúil don taighde seo ón Choiste Eitice Síceolaíochta, Ollscoil na Banríona. *Tá gach píosa eolais a bfhuil baint aige le rannpháirtithe go hiomlán anaithnide, agus thig le páiste tarraingt siar ón stáidéar in am ar bith gan míniú.* Má tá tú sásta cead a thabhairt do do pháiste páirt a ghlacadh sa taighde seo, líon agus sínigh an foirm thíos chomh luath agus is féidir, le do thoil. Agus seo déanta, iarr ar do pháiste é a thatbairt ar ais dá m(h)úinteoir.

Go raibh mile maith agat do do chuid ama agus chabhrach. Má tá aon cheist agat, déan teagmháil le Cláir Nic Stiofáin: cstephens01@qub.ac.uk

Dear Parent/Guardian

My name is Claire Stephens and I am a PhD student at the School of Psychology, Queen's University Belfast. Next week I will be coming along to your child's school to carry out a research study on children's attention. The research study, which is sponsored by *an Comhairle um Oideachas Gaeltachta agus Gaelscolaiochta* (COGG), is part of a programme of research at Queen's on bilingualism and immersion education. It is believed that children who are in immersion schools may display some advantages in terms of their cognitive development and we're hoping to investigate this effect further.

Children in fifth class will be asked to perform short pencil-and-paper based tasks, measuring attentional skills. These tasks will be in a game-like format and the children will be reassured that they are not being tested. I will administer the tasks individually in a quiet room in the school. In our experience children find the tasks enjoyable and stimulating.

We should be most grateful if you would be willing to give permission to allow your child to participate in our study. Our research is subject to the highest standards of ethical and professional conduct as laid down by the British Psychological Society and great care is taken to protect the interests and well-being of all children. Ethical approval for the research has been granted by the Psychology Research Ethics Committee of Queen's University Belfast. *All information concerning a pupil's performance is completely anonymous and confidential, and any child may withdraw from the study at any time without explanation.*
If you are willing to allow your child to take part in the study, please complete and sign the attached slip as soon as possible and ask your child to return it to his/her teacher.

Thank you so much for your time and assistance. If you have any further questions please contact Claire Stephens on: cstephens01@qub.ac.uk

Claire Stephens, BSc

Dr Gerry Mulhern BSc PhD AFBPsS CPsychol, Senior Lecturer in Psychology
Dr Judith Wylie BA PhD, Lecturer in Psychology

Tugaim cead do mo pháiste páirt a ghlacadh sa taighde seo.

I give my child consent to take part in this study.

Tuigim go bhfuil sonraí ar bith atá bailíthe go hiomlán agus anaithnide.

I understand that the data collected are entirely anonymous and confidential.

Tuigim go dtig tarraingt siar ón staidéar in am ar bith gan míniú.

I understand that my child may withdraw from the study at any time without explanation.

AINM AN PHÁISTE
CHILD'S NAME ______________________________

DO SÍNÚ
YOUR SIGNATURE ______________________________ DÁTÁ
DATE ____________________

APPENDIX VII – Ethics Approval

Queen's University Belfast

School of Psychology

Queen's University
Belfast
David Keir Building
18-30 Malone Road
BELFAST BT9 5BP
Tel: 028 90975445
Fax: 028 90975486
psychology@qub.ac.uk
www.psych.qub.ac.uk

31 March 2009

Ms Claire Stephens

Dear Claire

Re: PRE-UG-53-09

I acknowledge receipt of your completed ethics application countersigned by a member of staff and the associated documentation. I can confirm that your application has been approved and that you can now begin collection of data in consultation with your supervisor.

Yours sincerely

pp. [signature]

Dr Ian Sneddon (Chair)
Psychology Research Ethics Committee

cc Dr Martin McPhillips

APPENDIX VIII – Ethics amendment approval

Queen's University Belfast

School of Psychology

Queen's University
Belfast
David Keir Building
18-30 Malone Road
BELFAST BT9 5BP
Tel: 028 90975445
Fax: 028 90975486
psychology@qub.ac.uk
www.psych.qub.ac.uk

25 March 2011

Dear Ms Stephens

Full title of Study: The effects of immersion education on children's cognitive development
PREC reference number: 58-2009

I write to advise you that the application for amendment to the previously approved application PREC No 58-2009, (as detailed in your letter to the Chair of 16 March 2011), has been approved by the School of Psychology Research Ethics Committee, on behalf of Queen's University Belfast.

It is the responsibility of the Chief Investigator to ensure that the research has been recorded on the University's Human Subjects Research Database otherwise it will not be covered by the University's indemnity insurance. This database can be found in the 'My Research' section of Queen's On-line.

Yours sincerely

pp B Devlin

Dr Ian Sneddon (Chair)
Psychology Research Ethics Committee