

From executive behaviors to neurophysiological markers of executive function:
Measuring the bilingual advantage in young adults

by

William Rylie Moore
BA, University of British Columbia – Okanagan, 2010
MSc, University of Victoria, 2012

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

© William Rylie Moore, 2016
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by
photocopy or other means, without the permission of the author.

Supervisory Committee

From executive behaviors to neurophysiological markers of executive function:
Measuring the bilingual advantage in young adults

by

William Rylie Moore
BA, University of British Columbia – Okanagan, 2010
MSc, University of Victoria, 2012

Supervisory Committee

Dr. Mauricio Garcia-Barrera, Department of Psychology
Supervisor

Dr. Clay Holroyd, Department of Psychology
Departmental Member

Dr. Barbara Rutherford, University of British Columbia - Okanagan
Additional Member

Abstract

The ease at which individuals acquire a second language is astounding. Individuals are capable of learning a second language at any point through out their lifespan, although it is easier to learn a second language early in life. With increasing knowledge about linguistic neural processing and the brain's capacity for plasticity, the research on bilingualism has increased substantially. Researchers have become increasingly more interested in the long-term effects of acquiring a second language, especially the enhancement of executive function (EF). This enhancement, also known as *bilingual advantage*, has been studied for a range of EFs, including inhibition, attention, problem solving, and reasoning. Although this effect was first demonstrated in bilingual children, researchers have extended the quest for understanding to young, middle, and older adults; however, the research findings are mixed for young adults. In order clarify these mixed results, the age of second language acquisition has been included as an experimental variable, producing three relevant groups: early bilinguals, late bilinguals, and monolinguals.

There are several ways in which EFs can be measured, including behavioral rating scales, computerized cognitive tasks with behavioral outcomes (i.e., response times and accuracy), and computerized event-related potential cognitive tasks. A novel multi-level approach to measuring the bilingual advantage was developed and used as a framework for the current dissertation; i.e., the bilingual advantage was measured at three levels of measurement. This approach predicts that more complex levels of measurement (i.e., executive behaviors) would produce null findings between the three groups, while differences between early bilinguals and the other two groups would be predicted for less

complex levels of measurement (i.e., neurophysiological markers). This approach predicts mixed results for levels of measurement that involve moderate complexity (e.g., computerized tasks of EF). Early bilinguals, late bilinguals, and monolinguals were compared across three hierarchical levels of measurement: (i) executive behaviors; (ii) information processing (i.e., computerized tasks of EF); and (iii) neurophysiology (i.e., event-related potential paradigm). Findings generally support the multi-level approach: no differences were found at the executive behavior level, limited and mixed differences were found at the information processing level, and differences between groups were found at the neurophysiological level.

Table of Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	v
List of Tables	viii
List of Figures	ix
Acknowledgments	xii
Dedication	xiii
Prologue	xiv
Chapter 1	1
The bilingual advantage on executive function task performance in young adult: A review and suggestions for future research.....	2
Executive function	3
The bilingual brain and the need for an executive control system	8
The bilingual advantage.....	15
Additional variables of interest.....	24
Opponents of the Bilingual Advantage	40
Looking forward: A multi-level approach to the measurement of the bilingual advantage	42
Recommendations for future research	46
Conclusions	49
Chapter 2	52
Examination of the bilingual advantage in young adults using two behavioral rating scale measures of executive function.....	53
The Bilingual Advantage.....	53
Measuring EF	55
Objective.....	61
Hypotheses.....	62
Method	63
Participants and Measures	63
Statistical Analyses.....	66
Results.....	67
Discussion	68
Chapter 3	75
The bilingual advantage in young adults: Examining the importance of age of acquisition when measuring executive function.....	76
Executive function	77
Measuring the bilingual advantage.....	80
The present study	84
Method	89

Participants	89
Materials	90
Stimuli & Procedure	91
Statistical analyses	96
Results	98
Inhibition	98
Switching	99
Updating	99
Complex EF Tasks	99
Analyses with two language groups	99
Correlations	101
Discussion	103
Inhibition	103
Switching	105
Updating	106
Complex EF tasks	108
Limitations	110
Conclusions and Future research	111
Chapter 4	114
Measuring the neurophysiological markers of executive function advantages in bilingual young adults: Inhibition, shifting, and updating working memory	115
Executive function and the bilingual advantage	115
Electroencephalography/Event-related potentials	117
Neurophysiological differences of EF in bilinguals and monolinguals.....	127
Statement of the problem.....	131
Hypotheses.....	133
Method	136
Questionnaires	136
Participants	137
Stimuli & Procedure	139
Data acquisition & ERP analyses	145
Results.....	150
Demographic and language between-group comparisons	150
Behavioral data	151
ERP data	152
Discussion	172
Inhibition	173
Switching	178
Updating	181
Limitations.....	183
Conclusions and future directions	188
References	194
Appendix A	227
Appendix B	242

Appendix C 247

List of Tables

Table 1. <i>Demographic Information</i>	65
Table 2. <i>Dominant and Nondominant Language Spoken by Participants</i>	65
Table 3. <i>Mean and SD Table</i>	67
Table 4. <i>ANOVA Table</i>	67
Table 5. <i>Demographic Information for Chapter 3</i>	90
Table 5. <i>Table of Means, SDs, and ANOVAs Results for Three Groups</i>	98
Table 7. <i>Table of Means, SDs, and ANOVAs for Two Groups</i>	100
Table 8. <i>Correlation Table Between Individual EF Components</i>	101
Table 9. <i>Correlation Table Between Complex EF Tasks</i>	102
Table 10. <i>Participant Demographics</i>	138
Table 11. <i>Dominant and Nondominant Language Spoken by Participants</i>	139
Table 12. <i>ERP Rejection and Correction Table</i>	145
Table 13. <i>Behavioral Data: means, SDs, and one-way ANOVA results</i>	151
Table 14. <i>Midline Electrode Mean Voltages</i>	152
Table 15. <i>Peak Amplitude and Latency Means and SDs</i>	153

List of Figures

- Figure 1. Adapted from Kroll (1993). It represents a cognitive processing model where the top layer (i.e., lexical level) represents the form of the dominant (L1) and nondominant (L2) languages and their relative connections. The conceptual level (i.e., conceptual store) is a common system that houses semantic information and is shared by both languages. Solid lines represent relatively stronger connections; dotted lines represent relatively weaker connections.10
- Figure 2. Adapted from Crosbie et al., 2008 and Garcia-Barrera, Frazer, & Areshenkoff, 2012. This figure represents the relationship between genes, environment, and executive behaviours.44
- Figure 3. This figure represents a proposed developmental trajectory of the development of EF. The solid line represents the theoretical trajectory for monolinguals, while the dotted line represents that of bilinguals. The difference (i.e., shaded area) represents the bilingual advantage.50
- Figure 4 represents a multilevel perspective, where multiple interactions between genes, proteins, and structures interact with the environment during the production of cognitive processing, ultimately resulting in executive behaviors. This model was adapted from Crosbie et al. 2008 (model for endophenotypes) and Garcia-Barrera, Frazer, & Areshenkoff, 2012.71
- Figure 5. A schematic representation of Navon figures, where a “global” triangle is made up of "local" circles on the right and vice versa on the left.93
- Figure 6. Illustration of the IGT, demonstrating feedback after selecting deck 'A.'94
- Figure 7. Illustration of the Tower of London (PEBL version, 0.12) with four disks.95
- Figure 8. An illustration showing stimuli from the Berg Card Sorting Test (PEBL, version 0.12).96
- Figure 9. Schematic example of the computerized go/nogo inhibitory task. The arrow indicates the progression of time. The first three trials are examples of “go” trials, and the last trial is an example of a “nogo” trial.142
- Figure 10. Represents (a) the stimulus-locked go and nogo trial grand average waveforms for all three groups; and (b) the grand average waveforms for only nogo trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.154
- Figure 11. Represents stimulus-locked go and nogo trial grand average waveforms for monolinguals at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.154
- Figure 12. Represents the stimulus-locked go and nogo trial grand average waveforms for early bilinguals at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.155

- Figure 13. Represents the stimulus-locked go and nogo trial grand average waveforms for late bilinguals at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200..... 155
- Figure 14. Represents the go and nogo trial grand average waveforms for monolinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300..... 156
- Figure 15. Represents the go and nogo trial grand average waveforms for early bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300..... 157
- Figure 16. Represents the go and nogo trial grand average waveforms for late bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300..... 157
- Figure 17. Represents (a) the stimulus-locked pre-switch and switch trial grand average waveforms for all three groups; and (b) the stimulus-locked grand average waveforms for Switch trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window. PreSw = pre-switch trials. 159
- Figure 18. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for monolinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window. 159
- Figure 19. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for early bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window. 160
- Figure 20. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for late bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window. 160
- Figure 21. Represents the grand average waveforms for switch trials for monolinguals, early bilinguals, and late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b..... 162
- Figure 22. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for monolinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b..... 163
- Figure 23. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for early bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b..... 163
- Figure 24. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b..... 163
- Figure 25. Represents (a) the 2-back target and nontarget trial grand average waveforms for all three groups at electrode site Pz; and (b) the grand average waveforms for 2-

- back target trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300. NonTarg = Nontarget trials. 165
- Figure 26. Represents the 2-back target and nontarget trial grand average waveforms for monolinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials. 165
- Figure 27. Represents the 2-back target and nontarget trial grand average waveforms for early bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials. 166
- Figure 28. Represents the 2-back target and nontarget trial grand average waveforms for late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials. 166
- Figure 29. Represents (a) the 3-back target and nontarget trial grand average waveforms for all three groups; and (b) the grand average waveforms for 3-back target trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300. NonTarg = Nontarget trials..... 168
- Figure 30. Represents the 3-back target and nontarget trial grand average waveforms for monolinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials. 168
- Figure 31. Represents the 3-back target and nontarget trial grand average waveforms for early bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials. 169
- Figure 32. Represents the 3-back target and nontarget trial grand average waveforms for late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials. 169
- Figure 33. This figure was adapted from Crosbie et al. (2008) and from Garcia-Barrera, Frazer, & Areshenkoff (2012). It represents the relationship between genes, environment, and executive behaviours, and provides a theoretical framework from which one can assess the bilingual advantage effect at varying levels of measurements, as outlined in Chapter 1 of this dissertation. 191

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Dr. Mauricio Garcia-Barrera, whose expertise, support, guidance, and patience made this dissertation a success. I am grateful for his support and encouragement to pursue this line of research, which has expanded my scientific inquiry to include multiple forms of executive function measurement. Further, his support allowed me to return my research passions back to bilingualism. His wisdom and mentorship has set a strong foundation upon which I can build a fruitful career in Clinical Neuropsychology. Thanks to Dr. Garcia-Barrera, I was able to integrate education, research, and clinical work.

I would also like to thank the other members of my committee, Drs. Clay Holroyd and Barbara Rutherford, for going above and beyond their role as committee members. These individuals offered invaluable knowledge, time, and assistance throughout every stage of my dissertation. Their input and feedback helped foster a dissertation that is not only novel in its attempt to understand the complexities in the measurement of the bilingual advantage, but also offers inspiring avenues for future research.

I would like to extend my deepest love and appreciation to my partner, Christopher Shewchuk, who has been my greatest supporter for the last 10 years. I cannot express enough gratitude for his patience, understanding, and encouragement while I achieve my educational and career goals. I will endeavour to support his career and educational goals with the same care! Without his support, this dissertation would not have been such a success, and for that, I am ever grateful.

This dissertation took the corroboration of many researchers, students, collaborators, and clinicians alike, in order to arrive at a successful completion. I must extend my appreciation to Dr. Holroyd and Dr. Tanaka for the use of their lab resources. Furthermore, I would like to thank Corson Areshenkoff for assisting with the computer programming that facilitated the collection of data across all the studies in this dissertation. I would also like to thank Emily Duggan who assisted with data collection and provided endless research and emotional support throughout the process. As well, I would like to thank the honour's students, Karolina Karas and Violeta Nunez, as well as research assistant, Travis Graetz, for their assistance with data collection. I truly appreciate your help! And finally, I would like to thank Dr. Iris Gordon, post-doctoral researcher, whose help in data collection and analysis made this research a success.

In conclusion, I recognize that this research would not have been possible without the financial assistance of CIHR and the Department of Psychology at the University of Victoria (Teaching Assistantships, Graduate Research Scholarships).

Dedication

I dedicate this work to my family, who encouraged and supported my year abroad in France, where my passion for languages all started.

— — —

In loving memory of my mother,
Lana Diane Jeffery
I know you would be so proud.

Prologue

The following dissertation consists of four related, but distinct, manuscripts examining various levels of measurement of the bilingual advantage. Together, they form a cohesive collection of empirical research that meet the following research aims related to the multi-level assessment of EF in bilingual and monolingual individuals: (i) critically review the extant bilingual advantage literature in young adults; (ii) provide a novel conceptual approach to measurement that seeks to shed insight into the elusive nature of the bilingual advantage in young adults and provide a framework for hypotheses in future research; and (iii) conduct novel empirical studies that follow this approach in an attempt to ascertain its usefulness in this field. The autonomous nature of the articles introduces some redundancies within the dissertation as a whole, including the reviewed literature and, to a lesser extent, the individual discussions and conclusions provided; however, the manuscripts are written to complement one another and to contribute to current and future research pertaining to the bilingual advantage in young adults. Further, the autonomous nature of the manuscripts will facilitate submission for publication.

Chapter 1

**The bilingual advantage on executive function task performance in young adult: A
review and suggestions for future research**

The bilingual advantage on executive function task performance in young adult: A review and suggestions for future research

The ‘bilingual advantage,’ which posits that individuals who speak a second language have enhanced executive function (EF), especially for children, has been well established in the literature (e.g., Bialystok, 2011). Simply explained, it is thought that this advantage is the outcome of managing the output of two languages. Further, this advantage is strongly demonstrated in older adults in the form of a cognitive buffer, associated with delayed onset of cognitive impairment, dementia, and other signs of cognitive aging (e.g., Bialystok, 2008, 2011, 2012). However, the literature is mixed for young adults. Although not mutually exclusive, there are several interpretations for these findings, including: (i) regardless the number of languages they manage, individuals in this age range have similar levels of EF that are not fortified by bilingual language processing because they are at their developmental peak (i.e., ceiling) for this set of processes; (ii) the tasks used to assess EF are not sensitive enough to measure subtle differences in EF processes; and (iii) bilingualism does not foster EF enhancement across the lifespan. Given the paucity of longitudinal research that would help to address these hypotheses, it is difficult to assess their accuracy. However, this review paper will: (i) briefly discuss EF; (ii) outline how the EF system may be recruited for language control; (iii) review the empirical support for this advantage in young adult bilinguals using computerized executive function tasks; (iv) discuss the role of measuring additional variables when investigating this advantage in young adults; (v) briefly review arguments proposed by opponents of the bilingual advantage; (vi) posit a novel framework to

conceptualize and measure the bilingual advantage; and (vii) offer suggestions for future research.

Executive function

Executive function is an umbrella term and serves as a general referent to a set of cognitive abilities that are drawn upon when encountering novel problems or goal-oriented activities (Anderson, 2008). Recent research has focused on the developmental course, dimensionality, and malleability of EF; the latter of which is prominent in the bilingual advantage literature. Research findings have indicated that EF follows a protracted developmental course (e.g., Romine & Reynolds, 2005), show evidence of both unity and diversity (e.g., Garon, Bryson, & Smith, 2008; Miyake & Friedman, 2012), and are malleable, showing deficits in response to stress (e.g., Arnsten, 2000) as well as enhancements in response to structured activities (e.g., Bryck & Fisher, 2012) and experience (e.g., Bialystok, Craik, Green, & Gollan, 2009).

A Model of EF. There are many ways to conceptualize and understand EF, whether it is through the investigation of individual components (i.e., diversity), through the interconnections between such components, or through a common, *unitary* EF construct. In fact, many research studies, especially in bilingual research, select individual components, such as inhibitory control (e.g., Bialystok, Craik, et al., 2005), task switching (e.g., Prior & Macwhinney, 2009), working memory (e.g., Soliman, 2014), problem solving (e.g., Cushen & Wiley, 2011), and planning (e.g., Festman, Rodriguez-Fornells, & Münte, 2010) to compare monolingual and bilingual individuals. Although assessing the individual components of EF has its benefits (e.g., more simple and concise), a major shortcoming is the lack of ability to assess the interconnections between

proposed executive components. Miyake and colleagues offer a way to conceptualize EF and serves as a template for the structure of EF that facilitates the investigation of the individual components, the interconnections between those components, and its unity nature. These authors (e.g., Miyake & Friedman, 2012; Miyake et al., 2000) provide compelling factor analytic evidence that EF can be represented by intercorrelations between latent variables (i.e., individual components), while also representing the shared variance between these components, which produces a “common” EF component. These authors stipulate three basic EF components: (i) inhibition (i.e., inhibit a prepotent response); (ii) shifting (i.e., shift between mental sets or tasks); and (iii) updating working memory (i.e., update and revise mental representations in working memory).

In their seminal study that first suggested EF was made up of at least of these three components, Miyake and colleagues (2000) used several computerized tasks for each component and factor analysis to assess the individual components and how those components contribute to a common factor. To assess inhibition, they used a Stroop task¹, an antisaccade task², and a stop-signal task³, all of which require deliberately stopping a response that is relatively automatic; the specific response that needs to be inhibited differs across tasks. To assess updating working memory, the authors used a keep track task⁴, a letter memory task⁵ (similar to the n-back task), and a tone-monitoring task⁶. All

¹ Computerized task; participants were required to verbally name the color of a stimulus (either an asterisk or a word spelling a color) as quickly as possible in each trial, with RTs measured by voice key.

² Computerized task; participants were presented two stimuli, a distracting stimulus (i.e., flashing black box) and an arrow. Participants had to inhibit being distracted by the first stimulus to be able to respond to the direction of the arrow.

³ Computerized task. Consisted of two blocks of trials; during the first block, participants developed a particular response pattern to categorize animals versus nonanimals. In the second block, participants were required to reverse their response pattern.

⁴ Computerized task; participants were required to remember the last word presented in each of the target categories and then write down these words at the end of the trial.

three involve constantly monitoring and updating information in working memory, although the nature of the information that needs to be updated as well as the goals of the tasks are rather different. And finally, to assess switching, they chose a plus–minus task⁷, the number–letter task⁸, and the local–global task⁹, which all required shifting between mental sets; the specific operations that need to be switched back and forth are rather different across tasks.

In another seminal article, Miyake & Friedman (2012) investigated the nature and structure of EF by using a nested-factors model to represent how the individual components relate to each other (i.e., interconnections) and to a common factor (i.e., unity). Similar to their previous study, they used nine EF tasks¹⁰ (three for each component) that measured the three EF components: inhibition, switching, and updating. Instead of supporting their previous three component model, they found that a Common EF factor loads directly on the nine selected tasks, and two specific factors that load on the updating and shifting tasks, respectively, suggesting these two EF components involve abilities beyond what is common to the three previously identified factors. Individual differences in inhibition were entirely explained by what is common amongst

⁵ Computerized task; participants were required to recall the last 4 letters presented in the list of serially presented letters. To ensure that the task required continuous updating, the instructions required the participants to rehearse out loud the last 4 letters by mentally adding the most recent letter and dropping the 5th letter back, and then saying the new string of 4 letters out loud.

⁶ Participants were required to respond when the 4th tone of each particular pitch was presented (e.g., after hearing the 4th low tone, the 4th medium tone, or the 4th high tone), which required participants to monitor and keep track of the number of times each pitch had been presented.

⁷ Paper and pencil task; participants were required to add, subtract, or switch between tasks.

⁸ Computerized task; participants were required to indicate whether the number was odd or even when the number–letter pair was presented in either of the top two quadrants and to indicate whether the letter was a consonant or a vowel when the number–letter pair was presented in either of the bottom two quadrants.

⁹ Computerized task; participants were required to respond to either the global- or local- features of a Navon figure

¹⁰ For inhibition, they used an antisaccade task, a Stroop task, and a stop-signal task. For switching, they used a color-shape switch task, a number-letter switch task, and a category-switch task. For updating, they used a letter memory task, a keep track task, and a spatial 2-back task.

the EF tasks, suggesting it is involved in all of the EF tasks used in the study. That is, once the variance attributed to what is common across the nine EF tasks used in the study, there is no unique variance remaining for the inhibition-specific factor.

Regardless of which of these two models may best describe EF, it is important to note that there are multiple latent factors or individual components (i.e., inhibition/common, shifting, and updating) and these components are related to one another and/or to a common EF factor. Future research in the domain of EF modelling will continue to shed light on its elusive unitary and diverse nature; however, in its current state, it is important to consider the three proposed components in an attempt to capture the diversity of EF in bilingual samples. As such, a more detailed discussion of the three components, the related research findings, and tasks used to assess such findings is reviewed below.

Inhibition. Although both types of inhibition involve voluntary or deliberate control, many authors distinguish between suppression of an over-practiced, automatic, or ‘prepotent’ behavioural response, and the ability to suppress or ignore information that is irrelevant but is currently interfering with or eliciting a conflicting response on the immediate task. The former of these two processes has generally been named response inhibition (e.g., Barkley, 1999; Nigg, 2003) or behavioral inhibition (Barkley, 1997); while the latter process is typically referred to as interference control (Friedman & Miyake, 2004), conflict resolution (Posner & DiGirolamo, 1998), or executive attention (Posner & Rothbart, 2007). Miyake and colleagues (Miyake & Friedman, 2012; Miyake et al., 2000) combine these types of inhibition because of the shared *voluntary* nature, and thus, *inhibition* is defined as the deliberate, controlled suppression of a dominant or

prepotent responses. These authors clearly delineate the difference between the deliberate control over inhibitory processes (i.e., executive) versus inhibition in spreading activation models of neural networks and ‘reactive suppression’ (e.g., inhibition of return and negative priming) because neither of the latter types of inhibition are intended nor deliberate.

Switching. The construct of ‘attention’ is complex and is conceptualized as involving multiple functions and processes, and thus, cannot be described as a single system. As such, there are many models and theories of attention (see Neumann, 1996). For example, Posner and Rothbart (2007) argue for three attention systems, including (i) a system for orienting attention, (ii) one for maintaining a state of system alertness, and finally (iii) a system of attention under the influence of executive control (i.e., executive attention). In addition, the Supervisory Attentional System (SAS) as proposed by Norman and Shallice (1986) involves a number of sources of action control, such as an attentional controller capable of overriding habitual response patterns when a novel schema needs to be initiated for dealing with a novel situation, suggesting an *element* of switching ability. Across both of these models of attention, shifting attention is included and is thought to be an important element of executive control. Miyake and colleagues (2000) included three switching tasks, all with distinctly different stimulus information, which concerns shifting back and forth between multiple tasks, operations, or mental sets; thus, the commonality between these tasks is the *shifting* requirement.

Updating. Working memory (WM) is a cognitive system dedicated to the *temporary* processing, maintenance, and integration of information during the performance of everyday cognitive tasks (Baddeley, 2003b). Given its purpose of

maintaining task-relevant information in mind, WM is critical because it allows individuals to hold information received from the environment or retrieve information stored in long-term memory, and subsequently maintain that information if it is relevant to the task goals (Unsworth & Engle, 2007). Accordingly, if information becomes irrelevant, it is deleted and replaced with new, relevant information (i.e., updating). As a result, an individual can use and organize this updated information to execute goal-directed behaviors. Therefore, it is necessary for any model of EF to include a short-term storage component that maintains current (updated), task-relevant information until it is no longer needed, according to the goals of a given task at a given time. Miyake and colleagues (2000) stressed the importance of the *updating* element of WM in their conceptualization of EF because the essence, or executive nature, of this component is the active manipulation of relevant information in WM, rather than the passive storing of information.

Further, the importance of this system in linguistic processing has been demonstrated for tasks as comprehension of written and spoken text (Gernsbacher & Faust, 1991; Just & Carpenter, 1992) and fluency in language production (Rosen & Engle, 1997). People naturally vary in their working memory capacity, and it is well-known that limitations in working memory capacity are related to language processing difficulty and misinterpretation in both first and second languages (Baddeley, 2003a).

The bilingual brain and the need for an executive control system

Bilingual lexical control involves the *mental lexicon*, which is the mental “dictionary” of words associated with language. This dictionary reflects the *orthography* (i.e., the spelling of the words in a language), the *phonology* (i.e., the sound of the words

in a language), and also the *semantics* (i.e., the meaning of words in a language) of familiar words (Altarriba & Heredia, 2008). A bilingual person has two collections of words, each associated with one particular language. These collections are differentiated by abbreviations where L1 refers to the native or dominant language and L2 refers to the second or less dominant language.

Regardless of when an individual learned his/her second language, there is still the idea of how lexical items are stored within the brain. There are several types of models, including hierarchical, connectionist, and interactive models. Hierarchical models consist of two layers: (i) a lexical layer (i.e., represents orthography and phonology of each language) and (ii) a semantic layer (i.e., the meaning of the lexical item). Even though there are a variety of theoretical hierarchical models using this structure, they differ on the connections between the layers (see Cook, 2002 for a review). The most well-supported and parsimonious of the hierarchical models is the Revised Hierarchical Model (RHM) proposed by Kroll and colleagues (Kroll & Stewart, 1994; Kroll, van Hell, Tokowicz, & Green, 2010), which assumes a direct link between the direct translation of words in L1 and L2 (i.e., lexical layer) and vice versa. Furthermore, it assumes an indirect connection between these representations through the conceptual node shared between L1 and L2 (a connection that includes a direct link between the L1/L2 conceptual node and the L1 and L2 form nodes). This forms a 3-way connection between the conceptual node and the two lexical representations. Even though this model assumes direct connections between all parts, it also posits directional strength differences between nodes. There is a strong link between L1 - conceptual layer and L2 to L1 form nodes, and there is a weak link between L2 - conceptual layer and L1 to L2

(see Figure 1). The directional strength of this model is flexible and can be influenced by auxiliary characteristics of bilingualism (e.g., age of acquisition, relative fluency, context).

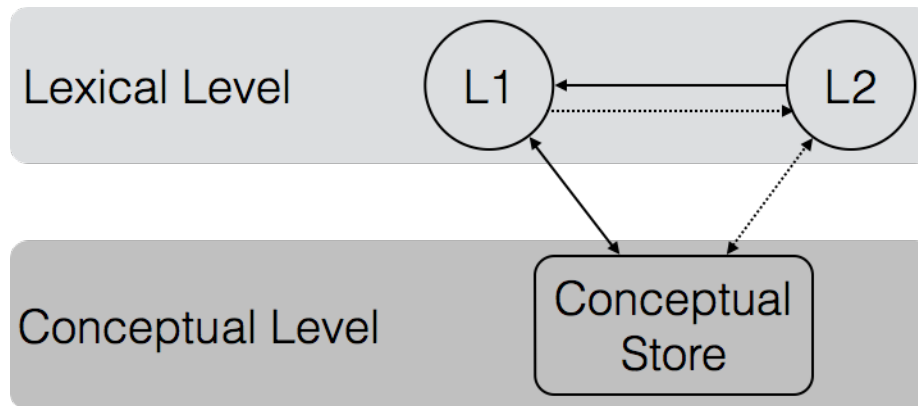


Figure 1. Adapted from Kroll (1993). It represents a cognitive processing model where the top layer (i.e., lexical level) represents the form of the dominant (L1) and nondominant (L2) languages and their relative connections. The conceptual level (i.e., conceptual store) is a common system that houses semantic information and is shared by both languages. Solid lines represent relatively stronger connections; dotted lines represent relatively weaker connections.

The RHM (Kroll & Stewart, 1994; Kroll, 1993) suggests that all nodes within the representation are interconnected. Similar to the RHM, some authors (e.g., Fox, 1996; Schwanenflugel, 1986) have concluded that bilinguals possess a shared semantic representational system and separate (but partially integrated) lexical entries for each language, while others (e.g., (Dijkstra & van Heuven, 2002) have suggested a connectionist model which proposes a single, integrated, word identification system, where lexical items, as well as semantic information, from both languages are stored together and activated in a language nonselective process of word recognition. This model assumes a temporal delay where the orthographical or phonological information is activated first, depending on the stimulus type (e.g., visual versus auditory), and activation of those nodes activates the remaining, related nodes. For example, for word

reading, the orthographical nodes are activated first with activation flowing subsequently into phonological and semantic nodes. Included in this word identification system are *language nodes* (L1 and L2) that are connected to the orthographical and phonological nodes, which assist with correct language identification. In addition to this system, there is a task/decision system, which has several functions, including (i) receiving continuous information from the word identification system; (ii) determining specific processing steps for the task at hand; and (iii) using decision criteria to determine if a response is made, based on relevant codes. In addition to the assumption of interconnectivity within the word identification system, this model assumes a connection between the word identification system and other, higher order language systems.

In both models, it is suggested that there is interconnection between L1 and L2 lexical information, and thus, if the language-separate lexical nodes are interconnected, it begs the question of whether or not the lexical representations are language-selective (i.e., only representations within the target language are activated) or language-nonspecific (i.e., representations in both languages become active). While the former would produce little or no conflict for the language system, the latter of these would produce a significant conflict for the language system, such that two equally probable response options would be available for selection. In a seminal study, Guttentag, Haith, Goodman, & Hauch (1984) presented bilingual participants stimuli consisting of a target word surrounded above and below by two copies of a to-be-ignored flanker word. Depending on the condition, the target word was either in the same language or in a different language. Results revealed similar target response times for both conditions, suggesting that language activation is nonspecific. Corroborating this evidence, reviews of the

literature from other studies using similar cross-language paradigms converge on the same conclusion: when bilinguals read a word in one of their two languages, activation of the orthographically-, phonologically-, and semantically-similar nodes occurs across both languages (e.g., Dijkstra & van Heuven, 2002; Kroll, Sumutka, & Schwartz, 2005). As such, most theories of bilingual language processing incorporate the idea of language nonselectivity (e.g., Dijkstra & van Heuven, 2002; Green, 1998; van Heuven, Schriefers, Dijkstra, & Hagoort, 2008).

Cook (2002) argues that in total nonselectivity, irrespective of contextual factors (e.g., topic of conversation, experimental setting), external input or internally-generated conceptual content (i.e., thinking) always activates lexical representations in both of the bilingual's language forms – and always to the same extent. Ultimately, the correct language form is selected for both proper comprehension and production. If separate language lexica are simultaneously activated, how is it, then, that a bilingual person can accurately access the appropriate lexicon for both perception and production? There are a number of opposing models of a L2 processing (i.e., production and comprehension) that have been the focus of much research; however, given the strong supportive evidence in the literature and its relation to executive function, the Inhibitory Control (IC) model (Green, 1998) will be reviewed briefly.

The IC model (Green, 1998) purports that language control is exerted through a process of active inhibition between the language entries at both the lexical and semantic levels. Upon perception or production of output, lexical representations are initially activated in both languages. In response to contextual information (e.g., current language in use, topic of conversation, laboratory settings), the targeted or selected language

reactively inhibits activation of the other language through the use of language *tags*. Using the RHM (Kroll & Stewart, 1994; Kroll, 1993), this inhibitory effect acts in a feed-forward manner through the opposing lexical representations (i.e., the phonological and orthographical layers), resulting in successive inhibition of those representations; thus, fluency in the target language is more easily obtained. According to Green's (1998) model, inhibition is stronger from L1 to L2 so that switching the target language to L2 must overcome greater levels of inhibition whereas switching from L2 to L1 would be more easily obtained. This is supported by studies of the *slip-of-the-tongue* phenomena where one is speaking fluently in L2 accidentally uses a word from L1 (Poullisse, 1999). This occurs because the relative inhibition from L2 to L1 for that particular concept is very weak or nonexistent and the wrong lexical representation is activated (i.e., the nontarget representation is produced or comprehended) because of the lack of inhibition. In conclusion, the IC model is well-supported (e.g., Bialystok, Craik, Green, & Gollan, 2009; Bialystok, Martin, & Viswanathan, 2005; Bialystok, 2011; Martin-Rhee & Bialystok, 2008; Rodriguez-Fornells, De Diego Balaguer, & Munte, 2006), suggesting that at least, in part, this method of neural control is necessary to manage the production of two languages.

In a recent expansion of this model, Green and colleagues (Bialystok, Craik, Green, & Gollan, 2009) proposed an eloquent hypothesis: that, in addition to inhibitory control, the bilingual brain recruits the executive control network to facilitate the accurate perception and production of two languages. The need for the recruitment of this system arises from the need to select an appropriate response from the target language system in

the context of compelling and active alternatives from the non-target language system.

Green and colleagues state that,

“the response to this conflict is to recruit the executive control system that has evolved to resolve conflict across all domains of perceptual and cognitive processing. The constant use of this executive control system for bilingual language management opens the possibility that the system itself is modified, changing its valence or efficiency for all tasks. That is, the use of a set of executive control procedures to manage attention to language, to avoid interference from the nontarget language, and to monitor two simultaneously active languages may alter the nature or efficiency of those executive control processes more generally” (p.97).

In order to support their argument, the authors review the existing literature for bilinguals across the lifespan. In terms of cognitive control and the bilingual advantage, the authors reviewed relevant literature involving monolingual and bilingual participants for inhibition, switching, and *memory processes*. Only two of the three components studied by Miyake and colleagues (2000) were included in their review, likely as a result of a lack of updating task literature for the comparison between monolinguals and bilinguals.

Before reviewing the existing bilingual advantage literature, the requirement for the EF system to manage the production and perception of two languages in the bilingual brain must be further elucidated. Using conversation as an example, the interlocutors must keep the current topic of speech in mind and formulate responses, which may be facilitated by working memory processes (Juffs & Harrington, 2011); as such, this process would be holding the recent speech in mind and updating response options. If one

imagined that only a single correct response was required for updating, there would be *two* responses simultaneously activated in the bilingual brain, given both languages are activated (e.g., Green, 1998). Therefore, this simultaneous activation would cause a problem *or conflict* for the language system. As such, a bilingual individual's language network must be able to reduce this conflict (i.e., conflict resolution) and intentionally *inhibit* the nontarget language (i.e., a prepotent response). The current environmental and language contexts (e.g., current language being spoken by other speaker, and at home, work, or school) would likely determine the relevant selection of the correct language node. Lastly, language switches may be required by the speaker; for example, an individual's friend may be speaking in L1 at the store when a cashier asks a question to the bilingual individual in L2. At that moment, the bilingual speaker's brain must *update* the language context (from L2 to L1), while inhibitory processes release inhibition of L2 in order to inhibit L1, all while attentional resources are *switched* from L1 to L2; accordingly, this latter process would likely enhance the *updating* and *inhibitory* processes through the interactions of the EF system.

The bilingual advantage

The existing literature suggests that simultaneous activation of two languages occurs when using one language alone, and as such, it has become increasingly clear that the need to manage this nonselectivity in neural activation is required in bilinguals. Therefore, the language networks make use of the EF networks to facilitate this management. Within these integrated network connections, the constant use of the EF system to manage language inhibition, switching, and updating, which is posited to provide a fortifying effect for EF. This bilingual advantage is prominent in childhood,

where bilingual children outperform monolingual children on tasks of inhibition, working memory, and task switching (Bialystok, 2011; Kroll & Bialystok, 2013). Further, it has been suggested that the advantage extends to older adults, where they experience enhanced EF compared to monolingual older adults, delayed onset of dementia (up to five years) and less decline in other cognitive functions (Bialystok, 2011). The current review paper is focused, however, on the fortification of the EF in young adults, where the EF system should be at its peak performance. We are interested in first answering the question of whether or not the bilingual advantage is present across different executive components in this age range, and therefore, the research literature of the bilingual advantage in young adults is reviewed below for the three components.

Inhibition. If both lexica are activated during language processing, and if only one language is to be perceived or produced, then the other must be inhibited. It is posited that bilinguals develop an enhanced inhibitory capacity to facilitate this process, and these advantages transfer throughout the EF system (i.e., to nonlanguage tasks). In the field of bilingualism, authors have used various forms of a go/nogo task, which requires a participant to withhold a response to a specific stimulus in the context of a sustained and frequent response to similar stimuli (e.g., Festman, Rodriguez-Fornells, & Münte, 2010); this task measures the *response inhibition* aspect of inhibition. However, the majority of the research has been conducted using the Simon task (e.g., Bialystok, 2006; Bialystok et al., 2005; Bialystok, Craik, Klein, & Viswanathan, 2004) and the Stroop task (e.g., Bialystok et al., 2008; Bialystok, Poarch, Luo, & Craik, 2014; Blumenfeld & Marian, 2013; Costa, Hernández, & Sebastián-Gallés, 2008; Kousaie & Phillips, 2012; Kousaie et al., 2014), while others have used the lateralized attention network test (LANT; e.g., Tao,

Marzecová, Taft, Asanowicz, & Wodniecka, 2011). The commonality between these tasks is suppressing a more salient element of a stimulus in order to arrive at a correct response. For example, in the Stroop task, participants must state the color of ink a word is printed in, rather than reading the word, which is a more salient element. These tasks measure the *conflict resolution* aspects of inhibition.

A review of the extant literature in young adults¹¹ reveals mixed evidence about whether or not young adult bilinguals experience advanced inhibitory control processes. Many of the studies used a version of the Simon task (Bialystok et al., 2004, 2008; Bialystok & Depape, 2009; Bialystok, 2006b; Bialystok, Craik, et al., 2005; Blumenfeld & Marian, 2013; Gathercole et al., 2014; Kousaie, Sheppard, Lemieux, Monetta, & Taler, 2014; Salvatierra & Rosselli, 2010; Linck, Hoshino, & Kroll, 2008; Mercier, Pivneva, & Titone, 2014; Mor, Yitzhaki-Amsalem, & Prior, 2014), which is based on stimulus–response compatibility and assesses the extent to which the prepotent association to irrelevant spatial information affects participants’ response to task- relevant nonspatial information; this task measures the *conflict resolution* aspect of inhibitory control. Using this task, four empirical studies of young adults found a bilingual advantage (Bialystok et al., 2004; Bialystok & Depape, 2009; Bialystok, Martin, et al., 2005; Linck et al., 2008) whereas several other studies did not find a significant bilingual advantage (Bialystok et al., 2008; Blumenfeld & Marian, 2013; Gathercole et al., 2014; Kousaie et al., 2014; Salvatierra & Rosselli, 2010; Mor et al., 2014). The Stroop task is another widely used measure of inhibitory control that also measures *conflict resolution*. Compared to studies that used the Simon task, there was more positive evidence for a bilingual advantage

¹¹ Please refer to Appendix A for a table of studies included in the review of inhibitory control advantages in young adults. This table includes participant characteristics, proficiency, age of acquisition, task, outcome measures, and statistical outcome.

using the Stroop task: six studies found that bilinguals outperformed monolinguals (Bialystok et al., 2008; Bialystok, Poarch, Luo, & Craik, 2014; Blumenfeld & Marian, 2013; Costa, Hernández, & Sebastián-Gallés, 2008; Kousaie & Phillips, 2012; Kousaie et al., 2014); two studies demonstrated no difference between groups (Blumenfeld & Marian, 2013; Mor et al., 2014); and one study found mixed results (Bialystok & Depape, 2009). Using various other tasks of inhibitory control, no evidence for a bilingual advantage was found for an antisaccade task (Bialystok, 2006b), a stop signal task (Colzato et al., 2008), the Attention Network Test (Costa et al., 2009), the Lateralized Attention Network Test (LANT; conflict component; Tao et al., 2011) and a flanker task (Luk, De Sa, & Bialystok, 2011); however, a bilingual advantage was found using a modified antisaccade task (Bialystok, 2006) and an inhibition of return task (Colzato et al., 2008).

In a review of the literature, Hilchey and Klein (2011) found poor evidence for a bilingual advantage for interference control (what they called the “bilingual inhibitory control advantage” or BICA), but found strong support for a global advantage (termed the “bilingual executive processing advantage” or BEPA), where bilinguals consistently demonstrated faster response times than monolinguals for congruent and incongruent trials. The authors noted that this latter effect was only present when the task itself elicited a sufficient level of response conflict. However, in an update of their original paper to include studies that had been published since 2011, (Hilchey, Saint-Aubin, & Klein, in press) the BEPA effect disappeared. The authors argue that the disappearance of this effect between 2011 and 2016 was related to publication bias, such that fewer studies were published that demonstrated little or null effects prior to 2011.

More research has been conducted using tasks that measure conflict resolution, rather than response inhibition. However, there does not appear to be a relationship between the type of nonlinguistic inhibitory control task and the outcome, except for more studies finding an advantage for bilinguals using the Stroop task. As is evident by the studies identified to measure inhibitory control mechanisms in bilinguals, this EF component is the most widely studied, which may be the result of initial theories of bilingual language control (e.g., IC model; Green, 1998). A more consistent and prominent bilingual advantage for inhibition has been found in children, middle-aged adults, and older adults (Bialystok, Martin, et al., 2005) than young adults. Therefore, it is possible that peak EF development (i.e., a ceiling effect) in young adulthood reduces our ability to identify, via task performance (e.g., accuracy and response time), the bilingual advantage. Following this logic, tasks measuring *information processing* (e.g., computerized EF tasks, paper-and-pencil EF tasks) may not always be sensitive enough to pick up subtle differences in inhibition differences in this age group.

Switching. A person who speaks two languages needs to attend to the language that is appropriate in the particular context and ignore the language that is irrelevant; this kind of experience may lead to development of more effective attentional mechanisms. In the bilingualism literature, researchers often use nonlinguistic switch tasks (e.g., Garbin et al., 2010), global-local tasks (e.g., Bialystok, 2010), trail making tasks (e.g., Bialystok, 2010), dichotic-listening tasks (e.g., Soveri, Laine, Hämäläinen, & Hugdahl, 2010), the lateralized attention network test (e.g., Tao et al., 2011) and other task switching paradigms (e.g., Prior & Macwhinney, 2009) to measure attention switching. In monolingual and bilingual young adults, three studies have investigated the role of

switching mechanism fortification in young adult bilinguals using nonverbal attention switch tasks¹², such as switching between the color or shape of a stimulus on screen, or using a dichotic listening task; all three of these studies found a bilingual advantage (i.e., reduced switch cost for bilingual participants) (Garbin et al., 2010; Prior & Macwhinney, 2009; Soveri et al., 2010). Other tasks that investigated the role of additional attention processes, such as orienting, alerting, and sustained attention, did not find bilingual advantages (Hernández, Costa, Fuentes, Vivas, & Sebastián-Gallés, 2010; Tao et al., 2011), suggesting that the enhancement that bilinguals received may be confined to the process used most by the bilingual language networks: switching.

Updating. The theoretical explanation for the use of working memory in the bilingual brain is gaining support, such that some have argued for the role of working memory during second language acquisition and bilingual language processing (e.g., Szmalec, Brysbaert, & Duyck, 2012). If this is the case, then one could make similar hypotheses for working memory advantages in bilinguals as are made for inhibition and switching. The working memory tasks typically used in the bilingual literature do not involve the *updating* element of working memory posited by Miyake et al. (2000). However, a brief review of the literature follows.

Feng and colleagues (2009) report that they were the first to investigate directly the role of working memory advantages in bilingual children. In the first of their two-experiment study, they used four tasks, which purported to measure: (i) only working memory (i.e., a verbal sequencing span task, where participants had to hold in mind a sequence of numbers and reorder those numbers; and a frog matrix task, where

¹² Please refer to Appendix A for a table of studies included in the review of Attention switching advantages in young adults. This table includes participant characteristics, proficiency, age of acquisition, task, outcome measures, and statistical outcome.

participants had to reproduce a sequence of spatial locations); and (ii) working memory and inhibition (i.e., a faces task, where participants had to press the response key corresponding to the direction the eyes, even if it was not the side where the head and eyes were located; and a pictures task, where participants responded to a right button press for picture A and a left button press for picture B). Results from this task revealed superior performance on all tasks for the bilingual participants compared to the monolingual participants. In their second experiment, which was conducted to replicate and extend their findings related to the frog matrix task. Thus, they use a complex version of the frog matrix task with four levels of difficulty: (i) simple spatial span, requiring participants to simply hold information in mind without manipulation; (ii) spatial memory with distraction, where irrelevant items were present during a delay interval; (iii) temporal order memory, which was the same as the frog matrix in their first experiment; and (iv) spatial memory plus reordering, where participants were shown items in a sequence, but had to reorder that sequence in their response. Results revealed bilingual children outperformed their monolingual counterparts on the latter two conditions of the visual-spatial task, which most taxed working memory; no differences were found between monolingual and bilingual participants for the first two conditions.

In an attempt to consolidate the literature and relationship between working memory capacity and second language acquisition, Linck, Osthus, Koeth, & Bunting, (2014) conducted a meta-analysis involving 79 independent study samples of 3707 young adult participants. These authors define working memory along two task-type continua, including: (i) simple (i.e., measuring storage only) versus complex (i.e., measuring storage and manipulation); and (ii) verbal and nonverbal stimuli. They also defined

language-related variables, including (i) L2 performance measures (i.e., a measure of comprehension, production, or both); and (ii) L2 proficiency, which was further categorized as highly proficient or less proficient. Their results revealed that second language proficiency (comprehension and production) was positively correlated with working memory capacity across the task-type variables. The effect sizes were relatively larger in verbal working memory measures than nonverbal working memory measures and in complex measures versus simple span measures, suggesting the bilingual advantage for working memory tasks may be circumscribed to more complex and verbal working memory tasks, rather than a general fortification of all working memory processes.

The main similarity between the initial study reviewed and the meta-analysis was the use of working memory tasks. However, there are several differences between the initial study reviewed and the meta-analysis, namely: (i) the initial study conducted by Feng and colleagues (Feng et al., 2009) was conducted on children; (ii) the meta-analysis by Linck et al. (2014) was conducted on adults; (iii) the initial study compared monolingual and bilingual participants; (iv) the meta-analysis was conducted on bilinguals only; (v) the initial study found a bilingual advantage for simple auditory and visual working memory tasks, as well as complex visual-spatial tasks; and (vi) the meta-analysis found a positive correlation between task complexity and L2 proficiency, as well as larger effect sizes for verbal working memory tasks than nonverbal working memory tasks. Although Linck et al. (2014) did not compare monolingual and bilingual participant performance, their results did suggest that higher proficiency was related to better working memory capacity, especially in complex tasks. Furthermore, the authors

included the n-back task in their complex, nonverbal category, which is one of the tasks that measure updating as posited by Miyake et al. (2000). However, since their results revealed smaller effect sizes for nonverbal working memory tasks and they did not compare bilingual to monolingual participant performance, it remains unclear whether or not a bilingual advantage would emerge.

To expand on the aforementioned literature, various studies investigating the bilingual advantage have included measures of working memory¹³; these measures are used to control for working memory ability and do not include the updating component posited in Miyake and colleagues' (2000) conceptualization of EF. However, Bialystok et al. (2008) used the Corsi blocks test, which is a visual-spatial working memory task where participants have to point to a series of blocks, and found a significant bilingual advantage for young bilinguals; however, using the same task, Wodnieka and colleagues (2010) did not find a bilingual advantage. In fact, in this latter study, the authors actually found a bilingual *disadvantage* in an auditory-verbal digit span working memory task¹⁴. Two other studies that have included visual-spatial and auditory span tasks have not found significant group differences (Bialystok et al., 2004; Luk et al., 2011).

To our knowledge, only one study (Soveri, Rodriguez-Fornells, & Laine, 2011) has investigated the role of *updating* mechanisms related to working memory in adult bilinguals. The author's aimed at introducing a "complementary analysis approach" by using regression analyses within 38 adult bilinguals to assess the relative contribution of different EF tasks to the prevalence of language switches in daily conversations. In this

¹³ Please refer to Appendix A for a table of studies included in the review of working memory advantages in young adults. This table includes participant characteristics, proficiency, age of acquisition, task, outcome measures, and statistical outcome.

¹⁴ The authors suggested that this finding was linked to lower proficiency in L1, which was the language that the task was given in.

study, the authors used a spatial n-back task, in combination with several other tasks assessing inhibition and switching components of EF. Although a direct comparison between monolingual and bilingual groups was not conducted for the n-back task, the authors' regression analysis of bilinguals-only revealed updating processes were related to age, such that younger bilinguals produced smaller processing costs. Working memory was not significantly related to the language-related variables, which included language switches, contextual switches, and unintended switches¹⁵.

Taken together, the literature on working memory differences between bilingual and monolingual individuals is limited, mixed, and lacking the updating element of working memory posited by Miyake et al. (2000). Further, the literature reveals a similar pattern of results as the empirical literature for inhibition and switching. It is clear that more research is needed on the updating nature of working memory in bilinguals and monolinguals.

Additional variables of interest

Age of acquisition. In bilingualism research, the age at which L2 learning begins or the age at which an individual is first exposed to L2 is referred to as the age of acquisition (AoA). The earlier an individual learns an L2, the more *native*-like that language will be (Costa & Sebastián-Gallés, 2014); however, even if an individual is exposed to an L2 since birth, a dominant language prevails (Sebastian-Galles, Echeverria, & Bosch, 2005). Furthermore, AoA has been suggested to drive a different neural representation of those bilinguals who learned both languages at different points in their

¹⁵ Data for all three of these variables was collected using a 5-point scale where participants rated their perception from never (1) to always (5). Language switches were the tendency to switch between languages when a word was not known and was calculated by taking the average for Finnish-English and English-Finnish switches. Contextual switches were those that occurred in specific situations where switches always occurred. And finally, Unintended switches were those that were difficult to control or unintended

life (Sabourin, Brien, & Burkholder, 2013). To our knowledge, Soares & Grosjean (1984) were the first to methodologically define late bilinguals, which they defined as first coming into contact with an L2 after the age of 12; rationale for their definition was not provided. Looking further into the literature, there are inconsistencies in how AoA is operationally defined. Some studies have considered AoA to be the age of first exposure to a second language (Newman, Tremblay, Nichols, Neville, & Ullman, 2012), the age at which subjects moved to a country that spoke the L2 being studied (Weber-Fox & Neville, 1996), or the age at which an individual started using both languages on a daily basis (Luk, De Sa, et al., 2011). Further, some others have used different age cutoffs, such that early bilinguals were defined as those who learned L2 before the age of three or four and late bilinguals learned L2 after the age of 10 (Isel, Baumgaertner, Thrän, Meisel, & Büchel, 2010; Perani et al., 1998). More recently, authors defined early bilinguals as learning L2 before the age of three and late bilinguals after the age of seven (Klein, Mok, Chen, & Watkins, 2014); whereas other studies have been more vague in their definition (Kaushanskaya & Marian, 2007). It is apparent that a common operational definition for AoA has not been consistently used in the literature; however, despite these methodological differences, striking effects of AoA on language processing and executive control have been found.

One meta-analytic study that investigated the underlying neural lateralization of monolingual and bilingual speakers developed moderator divisions for different age groups, depending on when a second language was acquired. Hull and Vaid (2007) conducted a meta-analysis with a central aim to systematically examine and quantify outcomes from the behavioral bilingual laterality literature in order to assess the

functional organization of language in the bilingual brain. The authors state the attributes used to define bilingualism and its subtypes varied widely from study to study, and thus, they were required to develop moderator divisions for L2 acquisition. Although their rationale for the moderator divisions was not stated directly, the authors were investigating lateralization differences depending on the age at which an L2 was acquired, which is influenced by and/or influences brain and cognitive development. Further, they use observations from the brain development literature, including (i) the observations that the cerebral cortex nor the corpus callosum is fully developed until the age of 5 or 6 years (or even later) and (ii) that myelination of neural pathways is not established until that time and even later. Taken together, the authors' operational definitions were: (i) early (or infant) bilinguals learned L2 by the age of 6 and (ii) late (or adult) bilinguals learned L2 at or after the age of 13. These authors had a third group of 'childhood' bilinguals whose L2 acquisition fell in the middle.

In their meta-analysis, Hull and Vaid (2007) found support for their operational definitions of AoA from differential lateralization effects for languages depending on AoA. The authors examined bilingual functional lateralization based on studies that directly compared both monolinguals and bilinguals. They found that monolinguals and late bilinguals were reliably left hemisphere dominant across language tasks regardless of proficiency, whereas early bilinguals showed reliable bilateral hemispheric involvement. This finding was supported by an earlier meta-analysis where comparable differences between early and late bilinguals were also reported (Vaid & Hull, 1991). Taken together, this evidence suggests that the primary predictor of functional language lateralization in

adulthood is whether an individual learned one versus two languages early in childhood, and further, that long lasting neural changes occur when one learns an L2 early in life.

Extending and complimenting the neural differences and lateralization findings by Hull and Vaid (2007), other researchers have investigated the cognitive processing differences between monolinguals, early bilinguals, and late bilinguals. For example, in a recent study, Bialystok and colleagues (Luk, De Sa, et al., 2011) found that AoA was an important variable when studying the effects of enhanced cognitive control on young adults. These authors used the age at which the participants started to use both languages on a daily basis, suggesting that the brain-related changes associated with using two languages is important for underlying brain processing (i.e., enhanced EF), similar to the argument by Hull and Vaid (2007). Further, they selected an arbitrary cutoff of age 10 as a ceiling for when a participant could have started using his/her second language on a daily basis. They argued that “previous research reporting support for a critical hypothesis [for brain development] was around the age of puberty. Since the sample in the study was young adults around the age of 20, age 10 was slightly before puberty and divided bilinguals into those who have used two languages essentially for all their lives or half of their lives” (p.590). Despite this definition, they used a traditional flanker task to assess inhibitory control processes in their sample and found that early bilinguals produced the smallest response time (RT) cost for incongruent trials (i.e., flanker effect) with no difference between monolinguals and late bilinguals. Their findings continue to suggest that the earlier one starts to learn and use – *on a daily basis* – a second language, the more enhanced the effects on EF are in early adulthood.

Extending the cognitive findings from Bialystok and colleagues (2011) and matching an operational definition closer to Hull and Vaid (2007), another study investigating the role of cognitive control processes in early and late bilinguals was conducted by Tao and colleagues (2011). In measuring AoA in this study, these authors considered early bilinguals as those who immigrated to another country at or before the age of six, whereas late bilinguals were immigrants who moved at or after the age of 12; they did not control for exposure to L2 before immigrating. For group comparisons, the authors used the lateralized attention network test (LANT), which is a computerized task that measures three components of attention: alerting, orienting, and executive control, as posited by Posner and Rothbart (2007). Comparison between the results in a valid spatial cue condition (which informs participants where the target will occur) and the results in a center cue condition provide information about the efficiency of orienting to the target location. As such, these authors found that early and late Chinese-English bilinguals had a more efficient attentional network than monolinguals. In a comparison between congruent and incongruent flanker LANT conditions, they also found that the late bilinguals demonstrated the greatest advantage in conflict resolution (i.e., smallest RT cost), which was an index of the executive network's efficiency. Taken together, this study demonstrated that AoA has an impact on the development of differential advantages for early and late bilinguals.

Contributing further to the cognitive processing differences for bilinguals who learned an L2 at different time points, two recent empirical studies investigating the role of AoA found different effects of AoA on EF outcomes, as follows. Kalia, Wilbourn, and Ghio (2014) sought to (i) determine whether early and late bilinguals vary from one

another, and (ii) exhibit cognitive advantages in EF relative to monolinguals. They defined early bilinguals as individuals who learned an L2 prior to age 6 and classified late bilinguals as individuals who learned an L2 at and after age 6. Although the authors cite Hull and Vaid (2007) in the rationale for their definition, their operational definition of late bilinguals is different. To measure group differences, they used an auditory cued number numerals task and found no differences between the three groups, which they argue suggested AoA and bilingualism do not produce advantages for the EF system. However, in another study investigating AoA and the bilingual advantage, Pelham and Abrams (2014) measured and compared outcomes in three groups: (i) monolinguals, (ii) early bilinguals (no later than age 7), and (iii) late bilinguals (no earlier than age 13). They used the attentional network test, which is a combination of the Simon task and a flanker task, and results revealed similar performance for early and late bilinguals, where both groups performed better (i.e., less RT cost) than monolinguals. The authors argued that the process of learning a second language, regardless of age, produce a fortification of the EF system.

Expanding on these cognitive processing findings, various studies have investigated the role of AoA on brain size and function. Using structural MRI scans, Klein and colleagues (2014) investigated cortical thickness of various brain regions in: (i) monolinguals, (ii) simultaneous bilinguals (learned both languages before the age of three), (iii) early sequential bilinguals (learned an L2 after developing proficiency in L1 but between the ages of 4-7, and (iv) late childhood sequential bilinguals (learned L2 between ages 8-12). They found thicker grey matter in the left inferior frontal gyrus (IFG) and thinner cortex in the right IFG in late bilinguals. The left IFG is important in

language production and commonly referred to as Broca's area, while the right IFG has been linked to the go/nogo task and inhibitory control (Aron, Robbins, & Poldrack, 2004). No differences were found in simultaneous bilinguals and monolinguals for any measured brain region. Furthermore, within the bilingual groups, significant correlations between age of acquisition of L2 and cortical thickness were seen in the same regions: cortical thickness correlated with age of acquisition positively in the left IFG and negatively in the right IFG. The authors argue that thicker cortex for early and late childhood sequential bilinguals may reflect the idea that acquiring an L2 as a new skill after infancy induces specific structural changes in brain areas demanded by the task. Taken together, they argue that their "results provide structural evidence that AoA is crucial in laying down the structure for language learning" (p. 5).

In conclusion, differential effects, including structural and function brain imaging, as well as cognitive processing outcomes, have been found for early and late bilinguals compared to monolinguals. However, there are several methodological challenges when synthesizing the literature, mostly due to the vast differences in operational definitions used to categorize early and late bilinguals. Despite these challenges, however, differences in outcomes have emerged. Compared to late bilinguals and monolinguals, learning an L2 early in life may produce more fortified enhancements of the EF system while it is developing, resulting in long lasting benefits throughout an individual's life. Furthermore, there are grounds to suggest that late bilinguals also experience fortification of the EF system, but that the neural function and structure of the system may be different than early bilinguals. Included in the aforementioned studies, language proficiency was similar across early and late bilinguals; a high level of proficiency of both languages is

required when comparing groups, as a result of the influence of language proficiency on language networks (e.g., Mechelli et al., 2004; Newman et al., 2012; Nichols, 2013; Perani et al., 1998).

Language proficiency. Language proficiency has an impact on the functioning of the language networks (e.g., Dijkstra & van Heuven, 2002; Kroll & de Groot, 2005; Kroll et al., 2010), including the level of inhibition required to suppress nodes in the nontraget language (Green, 1998), the size of brain regions (Mechelli et al., 2004), lexical access (Sabourin et al., 2013), and neurophysiological differences (Fernandez, Tartar, Padron, & Acosta, 2013). In fact, language proficiency is a factor typically accounted for in models of bilingual language use, such as Kroll's HRM (Kroll & Stewart, 1994), and the BIA+ model (Dijkstra & van Heuven, 2002). Similar to the literature reviewed thus far, the bilingual language proficiency literature is mixed; however, individuals who study the effects of bilingualism employ measures of language proficiency in an attempt to measure the effects of language proficiency on dependent variables or as a control measures (e.g., Bialystok & Viswanathan, 2009; Bialystok, 2007; Carlson & Meltzoff, 2008; Moreno, Wodniecka, Tays, Alain, & Bialystok, 2014). Some authors have used a standardized measure of English proficiency, including various versions of the Peabody Picture Vocabulary Test (PPVT) (e.g., Bialystok & Viswanathan, 2009; Luk et al., 2011; Martin-Rhee & Bialystok, 2008; Prior & Macwhinney, 2009), whereas others have used self-report indices (e.g., Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Costa, Santesteban, & Ivanova, 2006; Emmorey, Luk, Pyers, & Bialystok, 2008). There are pros and cons to employing both of these forms of measurement. For example, standardized measures of language proficiency provide more objective measurement of

proficiency. However, these measures have several shortcomings: (i) they can be time consuming, (ii) they are language specific, which can limit the participant sample to specific languages, and (iii) they can be costly. The benefits of using a self-report rating scale include: (i) they can be administered quickly and (ii) they can be used with a diverse range of L2 languages. However, a shortcoming for self-report measures of proficiency is less objectivity and potential bias in self-report.

There have been studies that investigated the role of language proficiency as an independent variable and found differential effects (e.g., Kalia et al., 2014; Luk & Bialystok, 2013; Luo, Luk, & Bialystok, 2010; Mechelli et al., 2004). For example, Mechelli and colleagues (2004) used magnetic resonance imaging (MRI) and voxel-based morphometry to assess the relationship between proficiency (high versus low) and grey matter density. These authors did not report how they measured language proficiency. Their results revealed greater grey matter density in the inferior parietal cortex for bilinguals compared to monolinguals. Further, they found a positive correlation between proficiency and grey matter density. The authors argue that their results are consistent with growing evidence demonstrating that structural changes in the brain occur in response to environmental demands (e.g., learning a second language and being highly proficient). Similar to structural changes related to proficiency, an event-related potential (ERP) study, which measures neurophysiological or *functional* differences, investigated the role of inhibitory control in bilinguals and monolinguals. To measure language proficiency, the authors used the Oral vocabulary subtest of the Bilingual Verbal Ability Test (BVAT), a standardized measure of proficiency. Using a non-linguistic, auditory Go/NoGo task, a task linked to non-motor, cognitive inhibition, these authors found a

correlation between N200 amplitude (a measure of inhibitory control) and second language proficiency (Fernandez, Tartar, Padron, & Acosta, 2013).

In another study investigating the potential factors associated with the bilingual experience (e.g., language proficiency, daily usage, etc.), Luk and Bialystok (2013) assessed 110 heterogeneous bilinguals using a questionnaire and standardized English proficiency measures (i.e., the Peabody Picture Vocabulary Test-III and the Expressive Vocabulary test). Using exploratory and confirmatory factor analysis, these authors analyzed the factors that quantified the relevant dimensions of language experience for bilinguals. As such, they found two correlating factors: (i) daily bilingual usage, and (ii) English proficiency. These authors argue that their results support the idea that bilingual experience is composed of multiple related dimensions that should be included when testing the impact of bilingualism on other factors, such as EF.

In conclusion, measurement of language proficiency is prevalent in the bilingualism literature, including accounting for language proficiency in models of bilingual language use. Together with the impact of proficiency on outcome measures, the evidence suggests that it is necessary to, at least, control for the level of language proficiency to ensure similarity between groups.

Socioeconomic Status. Socioeconomic status (SES) refers to a measure of an individual's or family's economic and social position in relation to others, based on variables, such as income, work experience, education, and occupation. In a critical review of the relevant bilingual children literature, Barac, Bialystok, Castro, & Sanchez (2014) argued that “the studies included in the review varied greatly in terms of socio-economic status ... [and] information about socio-economic background was not always

reported; when socio-economic status was included it was measured by using parents' education and/or income as a proxy" (pp. 701-702). The authors further identified that although some studies appropriately matched participants on SES, it was not considered in the statistical analyses because groups were equivalent, and when bilingual and monolingual children differed significantly in terms of socio-economic background, this effect was accounted for in the statistical analyses by covarying out differences in SES between the groups.

The research reviewed thus far involves bilingual children from varying SES brackets and suggests that the bilingual advantage extends beyond the impact of SES, at least in children. Turning to the adult literature and specifically to the studies reviewed earlier in adult bilinguals across three EF components, only three studies included criteria for matching SES across participant groups (Fernandez, Acosta, Douglass, Doshi, & Tartar, 2014; Moradzadeh, Blumenthal, & Wiseheart, 2014; Tao et al., 2011), while the remainder did not explicitly measure or account for SES.

It has been demonstrated that SES has an impact on the development of EF (e.g., (Ardila, Rosselli, Matute, & Guajardo, 2005; Bradley & Corwyn, 2002; Sirin, 2005), such that lower SES is often associated with worse performance on measures of EF (Noble, Norman, & Farah, 2005). If bilingualism has a positive effect on EF and low SES has a negative impact, it is worthwhile considering SES in bilingualism EF research. Differences in SES could bias the results in two ways: (i) as a confound or (ii) as a limiting condition. Bialystok (2009) rejects the claim that not controlling for SES produces spurious results (Paap, 2014), explaining that, at least in her research, by sampling the bilingual and monolingual children from the same schools in economically

homogeneous middle-class neighbourhoods, she controlled for SES. Moreover, Bialystok and colleagues have specifically investigated the role of SES and bilingual status in children, suggesting that the effect of SES on EF exists and is important to consider in bilingual advantage research (e.g., Calvo & Bialystok, 2014). However, Paap and colleagues argue that SES does not have an impact on the bilingual advantage (e.g., Paap & Greenberg, 2013). These authors measured SES using parent's education level in a multiple task study. The authors used an antisaccade task, a Simon task, a flanker task, and a color-shape switching task, and they found that group differences did not change when SES was matched and when it was not.

Similar to the bilingual advantage reviewed thus far, the research related to SES and the bilingual advantage is mixed. Therefore, to ensure that differences in EF between monolingual and bilingual individuals is not the result of differences in SES, measurement and control for, either through matching or statistical analyses, the impact of SES in language groups is recommended.

Immigration Status. This variable describes whether or not an individual has immigrated to another country, typically where the research is being conducted. Engel de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok (2012) investigated the bilingual advantage in low-income Portuguese-Luxembourgish bilingual and Portuguese monolingual immigrant children who completed two visuospatial tests of working memory (i.e., an odd-one-out task and a dot matrix task), an abstract reasoning task (i.e., the Raven's progressive coloured matrices), a selective attention task (i.e., a sky search task from the test of everyday attention), and an interference suppression task (i.e., a flanker task). They then used principal component analysis and found two broad

cognitive factors of EF: (i) representation (abstract reasoning and working memory), and (ii) control (selective attention and interference suppression) and found a cognitive advantage for the control factor, but not the representation factor, despite having low Portuguese vocabulary scores for bilingual children. They reported that cognitive advantages are possible even with a seemingly low degree of proficiency in both languages and in spite of low SES. They further purported that these results clearly show that, in spite of facing many linguistic and environmental challenges, bilingual immigrant children present strengths in nonlinguistic cognitive domains, including interference suppression and selective attention.

Several studies with adult monolingual and bilingual participants have poorly accounted for immigrant status (e.g., Bialystok et al., 2008; Fernandez et al., 2013; Fernandez et al., 2014; Garbin et al., 2010; Kousaie & Phillips, 2012; Kousaie et al., 2014; Luk, Green, Abutalebi, & Grady, 2011; Mor, Yitzhaki-Amsalem, & Prior, 2014; Tao et al., 2011). For example, Bialystok and colleagues (2008) and Mor and colleagues (2014) simply commented on the prevalence of immigrants in the bilingual sample and did not address immigrant status differences in their discussion. Fernandez and colleagues (2013, 2014) simply reported information regarding the birthplace of the participant and the birthplace of their parents, but similar to Bialystok and colleagues (2008) and Mor and colleagues (2014), this information was not addressed in their discussion. Of these studies, however, two made a purposeful methodological decision to include only non-immigrant individuals (e.g., Garbin et al., 2010; Kousaie et al., 2014), while the remainder included a mix of immigrant and non-immigrant proportions.

To our knowledge, only one study has directly tested immigrant status as an independent variable of older adult monolinguals and bilinguals. In a study investigating the role of bilingualism in relation to the onset of dementia, Bialystok, Craik, and Freedman (2007) compared 184 monolingual and bilingual individuals diagnosed with dementia. In their analyses, they included immigrant status as an independent variable and found no statistical differences between groups, suggesting immigrant status did not impact the onset of dementia for monolinguals and bilinguals.

Taken together, there is limited, yet mixed, evidence for the relationship between bilingualism, immigrant status, and cognitive outcomes. The studies reviewed either do not account for immigrant status, or they simply document the prevalence without including immigrant status as a variable of interest in the study. Taken together, immigrant status should be included as demographic information in order to assess its impact on findings and should be incorporated into the discussion of findings. Additionally, further research is needed to determine the relationship between bilingualism, cognitive changes, and immigrant status.

Translator Status. This variable describes whether or not a bilingual functions as a translator for their two languages. Most research is conducted on professional translators; however, nonprofessional translators, such as children of immigrant families, also exist. Unlike noninterpreters, Garcia (2014) discusses how simultaneous interpreters face several unique communication scenarios in which: (i) ongoing source language input must be processed at the same time that previous instances of input are being translated and spoken; (ii) production of the output message must be in the target language, likely through inhibition of competing representations in the other language; and (iii) code-

mixing must be deliberately avoided. Compared to bilingual noninterpreters, simultaneous interpreters are subject to more stringent and difficult language scenarios, hypothetically resulting in further enhancement of EF (i.e., “a double advantage”). Garcia reviewed the available empirical literature and hypothesized that simultaneous interpreters have specific advantages in four areas: (i) working memory; (ii) attentional allocation (i.e., attention switching); and (iii) inhibition (Garcia, 2014). Based on his review, however, he found mixed results for working memory, while stating general enhancement of WM capacity for handling concurrent storage and processing. Further, he found no particular advantage for inhibition or attention switching.

Second language acquisition stage. It is unclear whether or not the bilingual advantage is related to the acquisition of a second language or whether it is related to the ongoing requirement for EF components to control language output. Linck and colleagues (2008) compared monolinguals to bilinguals at various stages of L2 learning and measured inhibitory control ability; however, the only solid finding was that bilinguals (when placed as one group) had superior inhibitory processes compared to monolinguals, and learning stage did not play a role in the effect. Contrary to this, a study by Sullivan, Janus, Moreno, Astheimer, and Bialystok (2014) investigated the effect of early stage second-language training on executive control using event-related potentials (ERP). Monolingual English-speaking students were tested on a go/nogo task, sentence judgment task, and verbal fluency, before and after six months of Spanish instruction; the control group consisted of students enrolled in introductory Psychology. After training, the Spanish group showed larger P3 amplitude on the go/nogo task and smaller P600 amplitude on the judgment task compared to pre-test while the controls showed no

differences between pre- and post-learning conditions, suggesting that initial L2 learning has an impact on P3 and P600 amplitudes. The authors argued that the larger P3 waveform reflected the requirement for greater attentional resources on those nogo trials compared to go trials, while the smaller P600 waveform reflected smaller P600 waveforms indicated less effortful processing on the judgment task. Together, this suggests that the learning stage may provide the most robust changes to the EF system, and further, subtle differences elicited in the neurophysiological markers may be elucidated when more sensitive measurement (i.e., ERP) is used to assess EF differences.

Music Training. Miyake and Shah (1999) report that music training involves working memory, selective attention, inhibition, switching, updating, and monitoring, all of which are considered to be components of EF. Learning to play music is similar to L2 acquisition in that a musician must learn to *read*, *speak* (i.e., play), and *understand* (i.e., differentiate tones and melodies) music, and thus, it has been hypothesized that musical training offers a similar EF fortification experience (Bialystok & Depape, 2009; Moreno et al., 2014). Like bilingualism, music training has also demonstrated a positive influence on EF (e.g., Bialystok & Depape, 2009; George & Coch, 2011; Hargreaves & Aksentijevic, 2011; Moreno et al., 2011, 2014). For example, Bialystok and colleagues (2009) compared bilinguals, trained musicians, and monolingual controls and generally found that bilinguals and trained musicians performed better than monolingual controls. In particular, Bialystok and colleagues (2009) used a trail making test (measuring alternating mental sets), Simon arrows task (measuring visual-motor inhibition), and an auditory Stroop task (measuring auditory interference control) and found the following results: (i) on the auditory task, musicians outperformed the other groups, including the

bilinguals, in that their RTs were faster and their ability to resolve conflict was better if that conflict was based on pitch; and (ii) musicians and bilinguals performed similarly to each other and were both superior to monolinguals on the Simon arrows task. However, to our knowledge, there have not been studies that have investigated the potential additive effect of a bilingual trained musician, though it could be hypothesized that the addition of music training is similar to the addition of a third language and thus no additional benefit is measured, suggesting a possible ceiling effect for the fortification of EF.

Summary of additional variables. Taken together, there are many variables that may have a differential impact on the effect of the bilingual advantage. In an attempt to provide descriptive information for the purpose of shedding light on the mixed nature of the bilingual advantage in young adults, it is recommended, at the very least, that future research on the bilingual advantage report characteristics on the following variables: AoA, language proficiency, SES, immigrant status, music training, and L2 acquisition stage. Of these variables, age of acquisition, language proficiency, and acquisition stage may have the most impact on the development of EF *neural networks*, while the other variables appear to have some – variable - influence on EF outcomes.

Opponents of the Bilingual Advantage

There are authors who do not support the idea that the bilingual advantage exists (e.g., Paap & Greenberg, 2013; Paap & Liu, 2014; Paap, 2014a, 2014b; Paap, Sawi, Dalibar, Darrow, & Johnson, 2014, 2015). These authors blame small sample sizes for the (few) positive findings in the young adult literature, misinterpretations of the data, a confirmation bias to report positive findings, as well as a reluctance to conduct and report exact replications. Further, they argue that most studies do not include multiple measures

of EF, and thus, cannot ascertain convergent or divergent validity of results related to EF components. That is, that indicators of inhibition, for example, do not correlate across measures, suggesting that they are not measuring similar constructs across participants.

In one study, Paap and Greenberg (2013) compared bilinguals to monolinguals on 15 indicators of EF from a few commonly used tasks in the bilingual literature (e.g., a Simon task, a flanker, a color-shape switching task, an antisaccade task, and the Raven's Advanced Matrices test). From each of these tasks, the authors used several indicators (e.g., RT and accuracy on the antisaccade task; RT from the Raven's task; accuracy and RT interference effect from the flanker task and the Simon task; mixing cost; switching cost from the color-shape switching task) that were supposed to represent similar EF components. Their goal was to investigate the bilingual advantage, as well as the convergent validity of similar indicators across tasks; for example, accuracy from the antisaccade task and accuracy from the flanker task both reflect inhibition. The authors did not find evidence for a bilingual advantage, and, reported several instances of a bilingual disadvantage. Further, they conducted subsample analyses matching SES and fluency factors, yielding similarly null results. Lastly, their indicators of EF did not correlate with each other, and thus, the authors argued that the absence of consistent cross-task correlations undermines the interpretation that these are valid indicators of domain-general abilities. In conclusion, the authors claimed "the research findings testing for bilingual advantages in EF do not provide coherent and compelling support for the hypothesis that the bilingual experience causes improved [EF]" (p. 256) and further, that most studies are problematic because they only measure one component of EF and cannot provide a measurement of convergent validity. A second study with 120 participants,

three tasks, and 13 indicators of EF purported to confirm these findings (Paap et al., 2014). However, it is important to note that these authors may not have had entirely separable language groups. That is, in both studies, if a participant rated their proficiency in two (or more) languages as a 4 or more they were classified as bilingual. If a participant rated their proficiency in English as a 4 or more and rated all other languages as 3 or less they were classified as monolingual. Participants who did not meet either classification criteria were excluded from further data analyses.

These authors neglect to account for the vast childhood and older adult literature, where the bilingual advantage is much more pronounced and the literature is less mixed. Furthermore, the mixed young adult literature lacks significant research that may help resolve the empirical issues. For example, all of the research has been conducted using computerized or neuropsychological measures of EF, which assess cognitive processing through behavioural output (i.e., RT and accuracy). Following the idea that young adults are at their peak EF, subtle differences may not be revealed through the use of information processing outcomes (e.g., RT, accuracy, switch cost, etc.), and as such, alternative forms or methods of measurement are needed (e.g., neurophysiological markers).

Looking forward: A multi-level approach to the measurement of the bilingual advantage

In an attempt to resolve some of the shortcomings within the field of bilingual advantage research, we propose a novel framework for assessing differences from a multi-level approach to measurement. This framework was inspired by the research literature on endophenotypes, where endophenotypes (i.e., measurable, nonclinical

makers of gene expression) and intermediate endophenotypes (e.g., neurophysiological markers) are used to indicate underlying genetic risk and gene expression (e.g., Crosbie, Perusse, Barr, & Schachar, 2008). The majority of the research on the bilingual advantage in young adults, and children for that matter, has been conducted using computerized neuropsychological tasks that are purported to measure particular cognitive functions, such as inhibition, switching, updating, and planning, to name a few. Accordingly, these studies rely on RT and accuracy, or variations of these measures (e.g., interference effect on the flanker task), as outcome measures for comparisons between groups. Although a review of the methodological advantages and disadvantages of these methods is beyond the scope of this review, it is necessary to highlight these measurements have several shortcomings, including: (i) assumptions about the underlying cognitive/information processing that results in such outcomes; (ii) no direct measurement of neural processing; and (iii) increased variability as a result of error of measurement. In combination with the other shortcomings of measuring the bilingual advantage in young adults (e.g., peak EF in young adulthood), assessing the bilingual advantage at a cognitive processing level may not be sensitive enough to measure subtle differences in processing differences between monolingual and bilingual young adults, resulting in variability across studies and various cognitive or executive functions. See figure 2 for a schematic representation of the relationship between genes, environment, and ultimately executive behaviours. That is, there is a hierarchical relationship with increasing complexity between genes and executive behaviours, such that genes create proteins, which cumulate to produce cells/networks/structures, which subsequently produce information processing/cognitive functions, and so on.

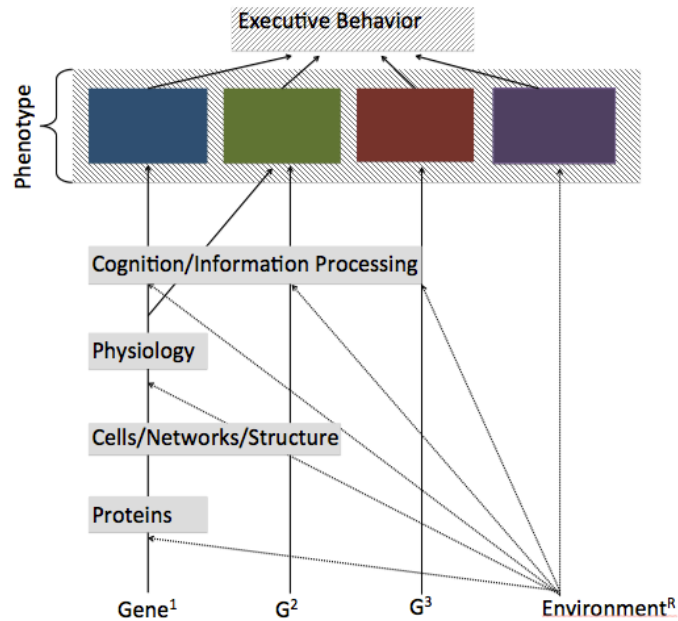


Figure 2. Adapted from Crosbie et al., 2008 and Garcia-Barrera, Frazer, & Areshenkoff, 2012. This figure represents the relationship between genes, environment, and executive behaviours.

Following the idea of increasing complexity from genes to executive behaviours, and the idea that most published studies have targeted the information processing *level*, it could be hypothesized that each *level* could represent a form of measurement (Garcia-Barrera et al., 2012; Trujillo-Orrego, Garcia-Barrera, Holroyd, & Pineda, 2011). For example, directly measuring *executive behaviours* could be conducted using behavioural rating scales, such as the Behavior Rating Inventory of Executive Functions (BRIEF) or the Behavioral Assessment System for Children (BASC). Furthermore, measuring *physiological* differences between monolingual and bilingual EF could be conducted with the use of event-related potential (ERP) paradigms. Following this logic, it is hypothesized that as one moves through each level of measurement, from the most to least complex, the difference between EF mechanisms in bilinguals and monolinguals would become greater; that is, by decreasing the complexity of what is being measured, the subtle differences between groups will be more clearly and more consistently

revealed. To our knowledge, this framework for measuring the bilingual advantage has never been proposed and, and more importantly, it may offer value in reducing the measurement variability and outcome inconsistencies already found in the bilingual advantage literature.

Also to our knowledge, no studies have yet been published to assess executive behaviour differences in children or young adults; however, research in our lab is underway. Behavioural rating scales are useful for assessing executive behaviours (Duggan, Garcia-Barrera, & Muller, in press) and further, some argue that behavioural rating scales measure different aspects of EF that are not picked up by performance-based measures (e.g., Toplak, West, & Stanovich, 2013). At this point, there is no empirical evidence to support the validity and efficacy of measuring the bilingual advantage using behavioural rating scales, as a result of a paucity of research using such measures. However, within the current proposed framework, measuring executive behaviours is the most complex, and thus, the most likely to neglect the subtle bilingual advantage differences. Moving down the hierarchy, a substantial amount of research has been conducted at the information processing level, which has yielded mixed results and is providing some preliminary evidence for the current framework. That is, as measurement approaches the lower levels (proteins and genes), subtle differences may be less diluted, but not necessarily easily captured. Following this logic, measuring physiology (e.g., neurophysiological markers in an ERP) is yielding differences between monolinguals and bilinguals (e.g., Martin et al., 2013; Moreno, Rodríguez-Fornells, & Laine, 2008; Moreno et al., 2014). Converging evidence has stemmed from several studies that have investigated brain structure differences in bilingual and monolingual participants,

yielding significant differences in structures related to EF (e.g., Luk et al., 2010). A recent review of the neuroplasticity associated with L2 acquisition was conducted by Li, Legault, and Litcofsky (2014), indicating that: (i) second language experience-induced brain changes, including increased gray matter density and white matter integrity, can be found in children, young adults, and the elderly; (ii) can occur rapidly with short-term language learning or training; and (iii) are sensitive to age, age of acquisition, proficiency or performance level, language-specific characteristics, and individual differences. Completing the framework, research into genetic expression and bilingual experience interactions has also found differences between monolinguals and bilinguals. For example, Hernandez, Greene, Vaughn, Francis, & Grigorenko (2015) provided preliminary results from genotype sampling of bilingual and monolingual individuals that revealed different distributions in allele frequencies of the DRD2/ANKK1 taq1A polymorphism. Bilinguals had twice as many DRD2/ANKK1 taq1A+ (A1+) alleles compared to controls, which, the authors suggest, has been linked to dopamine and EF. The authors plan to use the same genetic analysis and compare performance on EF tasks between subgroups (e.g., bilinguals with A1+, monolinguals with A1+, bilinguals with A1-, and monolinguals with A1-). However, further research in behavioural outcomes and empirical measurement is yet-to-be published.

Recommendations for future research

Include measurements of additional variables. It is becoming clear that EF is a multidimensional spectrum that is possibly impacted by a range of variables, such as bilingualism, AoA, language proficiency, SES, immigrant status, music training, and stage of L2 acquisition. Together, these variables will ultimately have an impact on

cognitive/information processing outcomes of EF. Further, the nature of the effect of the bilingual experience on EF is unclear, and as such, it is imperative for researchers to consider the multitude of variables that impact the bilingual experience in order to elucidate the relationship between bilingualism and EF. In fact, some authors have posited the question, how much bilingual experience is enough to effect change on EF? Additional cross-sectional and longitudinal studies are necessary to shed light on these questions. For example, between-subject cross-sectional studies aimed at different years of L2 experience will provide insights into potential trajectories of the fortification of EF, in that group differences may arise between those without L2 experience and those with 5 years, 10 years, etc., with no differences between those with 5 years, 10 years, etc. of L2 experience. If findings like this emerge, it would suggest the initial learning period has the most influence on EF compared to ongoing L2 experience. Longitudinal studies can corroborate such findings by demonstrating within-subject effects (i.e., bilingual individuals would show fortification of EF between time 1 and time 2 but no additional advantages after that, while monolinguals would show no differences over time). In fact, preliminary longitudinal research studies have been conducted, revealing positive changes to EF after learning a second language (e.g., Sullivan, Janus, Moreno, Astheimer, & Bialystok, 2014).

Use well-supported models of EF to build a strong foundation for measurement. The 3-factor model posited by Miyake and colleagues (e.g., Friedman et al., 2008; Miyake & Friedman, 2012; Miyake et al., 2000), which includes inhibition, switching, and updating, provides a framework upon which EF advantages can be measured and assessed with bilingual and monolingual samples. In terms of assessing the

diversity of EF, there are a number of individual research studies investigating the role of specific EF components. In young adults, the research-to-date is heavy within inhibitory control mechanisms, with a recent increase in measuring switching. As such, it will be important to develop the research literature using updating tasks as a component of EF, as this is largely lacking and could provide valuable information when assessing subtle differences in young adults.

Further, it appears necessary to continue to obtain measurements of the three components as an attempt to capture the diverse nature of EF, continuing to strengthen our understanding of the interplay between EF and language networks in bilinguals. Including all three components within the same study/sample will allow for a more comprehensive measurement of EF, facilitating an exploration of the unitary and diverse nature of EF across monolingual and bilingual samples. As such, comprehensive measurement of the components and adapting select statistical procedures, such as factor analysis or structural equation modelling, can facilitate more robust, and potentially impactful, comparisons between groups (e.g., invariance analysis), something that has yet to be documented within bilingual research.

Produce meta-analyses and/or systematic reviews. Empirical evidence for and against the bilingual advantage is cumulating, especially for inhibition as a component of EF. The field would benefit from a more systematic review of the inhibition literature, combining studies to further elucidate the mixed nature of the bilingual advantage in young adults. As the literature grows for other EF components, these too should be subject to systematic empirical analyses.

Include additional language groups. The majority of research reviewed in this paper includes bilinguals who have established proficiency in an L2 and use that language on an ongoing, and consistent basis. However, recent research on young adults who learned a second language in adulthood has suggested that differences in EF advantages appear within the learning phase of second language acquisition (Martin et al., 2013). Thus, this begs the question of whether or not the bilingual advantage is the result of simply learning a second language or if it is the result of the consistent use of two languages. Further research on individuals who are acquiring a second language and its impact on EF is needed to address the potential impact of initial learning on EF fortification. Furthermore, it would be interesting to assess the long-term effects of the bilingual advantage; that is, does the advantage persist when a second language is no longer used? To assess such a question, and provide further insight into the advantage itself, one may include a group of participants who had learned a second language to the point of daily use/fluency, but no longer use the language.

Conclusions

In conclusion, this chapter reviews the bilingual advantage in young adults, outlining the research literature using Miyake and colleagues (2000) three-factor model of EF. This review highlights the mixed literature for inhibitory control, as well as the lack of literature for attention switching and updating working memory. Additional variables that impact EF, especially in the context of bilingualism, were reviewed, and further, recommendations for future research were provided. The current review does not solve, nor did it attempt to solve, the ongoing debate about the presence of the variability of the bilingual advantage in young adults; however, this paper offers some advice to

assist in clarifying this issue, including suggestions for additional research groups, variables, longitudinal measurement in research designs, as well as conducting meta-analytic reviews to consolidate the mounting literature, especially for inhibition. Unfortunately, without the necessary research evidence, there still exist several explanations for the mixed literature in young adults, including: (i) the idea that, regardless the number of languages they manage, individuals in this age range have similar levels of EF that are not fortified by bilingual language processing because they are at their developmental peak (i.e., ceiling) for this set of processes; (ii) the tasks used to assess EF are not sensitive enough to measure subtle differences in EF processes; and (iii) bilingualism does not foster EF enhancement across the lifespan. It appears that the latter of these hypotheses is not likely, given the less mixed nature of the presence of the bilingual advantage in children and older adults (e.g., Bialystok et al., 2009; Bialystok, Martin, et al., 2005; Bialystok, 2011; Kroll & Bialystok, 2013). Figure 3 represents a proposed developmental trajectory of EF, and thus, the bilingual advantage across the lifespan, portraying its presence in childhood and older adulthood, as well as its subtlety or variable “absence” in young adulthood.

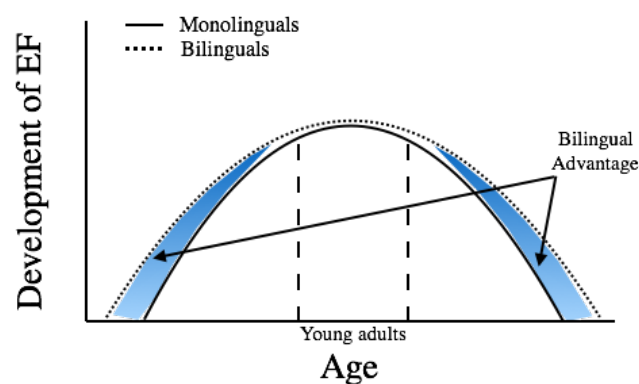


Figure 3. This figure represents a proposed developmental trajectory of the development of EF. The solid line represents the theoretical trajectory for monolinguals, while the dotted line represents that of bilinguals. The difference (i.e., shaded area) represents the bilingual advantage.

A novel multi-level approach for understanding and measuring the bilingual advantage was suggested here, which may help to elucidate the mixed nature of the young adult bilingual advantage literature. Following this logic, the following chapters of this dissertation aim to measure the bilingual advantage at three levels of measurement: (i) executive behaviors using behavioural rating scales (Chapter 2); (ii) information processing using computerized tasks of EF (Chapter 3); and (iii) physiology using an event-related potential paradigm (Chapter 4).

Chapter 2

**Examination of the bilingual advantage in young adults using two behavioral rating
scale measures of executive function**

Examination of the bilingual advantage in young adults using two behavioral rating scale measures of executive function

The Bilingual Advantage

The bilingual advantage, which suggests that individuals who fluently speak a second language have enhanced EF, has been widely studied and discussed (e.g., Bialystok, 2011; Kroll & Bialystok, 2013; Valian, 2014). Unfortunately for young adults, the research literature is mixed, casting doubt about whether or not the bilingual advantage phenomenon actually exists (Paap & Greenberg, 2013). Despite the mixed nature of the results, some authors (e.g., Bialystok, 2011; Valian, 2014) argue that the bilingual advantage does, in fact, exist. For example, Valian (2014) posits that the variability in the research findings, especially for young adults, suggest that the cognitive benefits experienced by being bilingual compete with other cognitively enhancing experiences to varying degrees. For example, several variables (e.g., physical exercise, music training, education, etc.) have been found to have an impact on EF, including, but not limited to, socioeconomic status (e.g., Sirin, 2005), immigrant status (e.g., Engel de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012), and music training (e.g., Moreno & Farzan, 2015; Moreno, Wodniecka, Tays, Alain, & Bialystok, 2014). Therefore, depending on the composition of a given research sample, the other cognitively enhancing experiences may be more plentiful in the monolingual than bilingual group (or sufficiently plentiful in both groups), so that the unique benefits of bilingualism are invisible. Further, EF development is at its peak in young adulthood; thus, various cognitively enhancing experiences would likely have been accumulated, producing similar peak performance in young adults. If second language acquisition leads

to EF enhancement in young adults, than it is either invisible in this age range, or the subtle differences are not captured by the current measurement strategies (i.e., computerized tasks of EF).

In the following sections, EF measurement is discussed, briefly reviewing the current approaches used in the literature, and suggesting a novel approach to assessing the bilingual advantage. Briefly, chapter one of this dissertation posited a novel framework for assessing the bilingual advantage in young adults by measuring EF from various levels of measurement (e.g., executive behaviors, cognitive processing, and neurophysiological markers). To our knowledge, there have been no studies published to date using behavioral rating scale measures of EF, despite their increasing utility in clinical populations for producing more ecologically valid measures of EF. Thus, the current study is the first to investigate executive behavior differences in bilingual and monolingual individuals.

Executive functions are a set of cognitive processes that aim to manifest goal-directed behaviours in novel and/or unpredictable situations (Jurado & Rosselli, 2007). A wide range of cognitive abilities and capacities is included and different researchers have proposed somewhat different repertoires, functions, or models of EF. For example, EF can refer to cognitive flexibility, inhibition, working memory, problem-solving, reasoning, and planning. Many research studies, especially in bilingual research, select individual components to compare performance of monolingual and bilingual individuals. Although assessing a unique aspect of EF has its benefits (e.g., more simple and concise), a major shortcoming is the lack of assessment of relationships between proposed executive components. Miyake and colleagues offer a way to conceptualize EF that

facilitates the investigation of both the unitary and diverse aspects of EF. These authors (e.g., Miyake & Friedman, 2012; Miyake et al., 2000) provide compelling factor analytic evidence that EF is both unitary (i.e., correlations between components or a “common” EF component) and diverse (i.e., made up of inhibition, switching, and updating mechanisms at the information processing level and problem solving, behavioural control, emotional control, working memory at the executive behaviour level). In their study, Miyake and colleagues (2000) used a factor analytic approach of many computerized tasks to propose that EF were made up of three components: inhibition, switching, and updating working memory. In a follow-up study, Miyake and Friedman (2012) suggested that only switching and updating working memory provided unique component contributions to EF, and inhibition was subsumed under the “common factor” because it did not provide any additional unique variance. Despite the finding that inhibition may be a key component in most EF tasks, it continues to be measured as a component on EF.

Measuring EF

There are several approaches to measuring EF, which converge onto two levels of measurement: (1) EF tests appropriate for measuring cognitive/information processing, such as laboratory performance tests, including computerized tasks and neuropsychological measures (i.e., paper and pencil measures); and (2) EF tests appropriate for measuring *executive behaviours*, such as rating scales. Laboratory-based tests, such as the Stroop task, the go/nogo task, and the Simon task, have served as the traditional methods for assessing EF, especially in the field of bilingualism, and there are several reviews of these studies (e.g., Bialystok, Craik, Green, & Gollan, 2009; Bialystok, 2011; Hilchey & Klein, 2011; Kroll & Bialystok, 2013). While these tests have numerous

advantages, including the ability to measure specific components of EF (e.g., inhibition, switching, and updating), they are not always sensitive to everyday manifestations of executive dysfunction outside the test setting (i.e., lack ecological validity) (Toplak, West, & Stanovich, 2013). Related to the lack of sensitivity to executive dysfunction, laboratory-based tests may be too narrow to capture more complex or larger behavioral differences between bilingual and monolingual young adults. Further, these laboratory tests sometimes indicate executive impairments in individuals not demonstrating any day-to-day functional difficulties (Pennington & Ozonoff, 1996). In response to the limitations of traditional approaches, researchers have developed several different behavioral rating scales, designed to capture everyday manifestations of EF at the level of executive behaviors in a more *ecologically valid* way (Isquith, Roth, & Gioia, 2013; Silver, 2014). Although rating scales were initially intended to complement traditional laboratory tests, recent research indicates rating scales have their own value, providing unique, but complementary information about different aspects of EF (Isquith et al., 2013; McAuley, Chen, Goos, Schachar, & Crosbie, 2010; Silver, 2014; Toplak et al., 2013).

Having only gained a strong body of research support and considered as an integral component of EF assessment within the past decade (Gioia, Kenworthy, & Isquith, 2010; Silver, 2014), self-rating scales of EF are not widely used, especially in bilingual research. Despite the growing support, the assessment of EF via rating scales is associated with a number of challenges, including individual bias, personal and cognitive characteristics of raters, environmental influences affecting ratings, inconsistent conceptualizations and definitions of the construct, and difficulty in parsing out specific

EF deficits (Grace & Malloy, 2001; Isquith et al., 2013). Further, most scales have been developed under the conceptualization of capturing executive *dys*function in behaviors. While some of these challenges are inherently associated with rating scales, this form of measurement offers a novel method for measuring the bilingual advantage in young adults that is yet to be explored. Specifically, using the multi-level approach to measuring the bilingual advantage, rating scales provide information at the most complex layer of the model.

Executive behaviors versus executive function. Behavior rating scales are thought to measure EF at the level of executive behaviors, which involve the culmination of several specific cognitive processes that work in concert to produce behavior. As such, the conceptualization of EF as measured by behavior rating scales is sometimes different than those that measure the more specific cognitive processes. For example, a latent variable approach to EF measurement, such as that used by Miyake and colleagues (2000), can be applied to the measurement of executive behaviors. This approach, which will be reviewed in more detail below for the two executive behavior measures used in the study, represents the integration of isolated EF components that represent overarching behaviors. For these reasons, executive behavior measurements, which are generally conducted using behavior-rating scales, often have somewhat different representations at the level of the latent variables. For this reason, the selection of relevant latent variables deviates slightly from Miyake and colleague's (2000) three-factor model.

The BRIEF-A. The Behavior Rating Inventory of Executive Function-Adult Version (BRIEF-A) is a questionnaire designed to assess executive functioning in everyday life. The BRIEF-A was developed as an extension of the original BRIEF (Gioia,

Isquith, Guy, & Kenworthy, 2010) to cover adults aged 18–90 and has both self- and informant-report forms (Roth, Isquith, & Gioia, 2005). There are nine subscales that constitute the BRIEF-A (Roth, Lance, Isquith, Fischer, & Giancola, 2013): (i) Inhibit, which measures behaviors associated with the ability to “control impulses and appropriately stop verbal, attentional, physical behavior at the proper time” (p. 426); (ii) Shift, which measures behaviors associated with the ability to “move freely from one situation, activity, or aspect of a problem to another as the situation demands, as well as think flexibly to aid problem-solving” (p. 426); (iii) Emotional control, which measures behaviors associated with the ability to “modulate one's emotional responses appropriately” (p. 426); (iv) Self-Monitor, which measures behaviors associated with the ability to “recognize the effect of one's own behavior on others” (p.426); (v) Initiate, which measures behaviors associated with the ability to “begin a task or activity without external prompting and independently generate ideas” (p. 426); (vi) Working memory, which measures behaviors associated with the ability to “hold information in mind in order to complete a task, as well as stay with, or stick to, an activity” (p. 426); (vii) Plan/organize, which measures the ability to “anticipate future events, set goals, develop steps ahead of time to carry out a task, organize information and behavior to achieve an objective, as well as carry out tasks in a systematic manner” (p. 426); (viii) Task monitor, which measures that ability to “assess performance during or after finishing a task for mistakes” (p. 426); and (ix) Organization of materials, which measures the ability to “keep workspace and living areas in an orderly manner, as well as keep track of materials needed for tasks” (p. 426). In a recent factor analytic assessment of the BRIEF-A (Roth et al., 2013), these scales form three factors: (i) the Metacognition factor, which is

composed of the Initiate, Working memory, Plan/organize, Task monitor, and Organization of materials scales; (ii) the Behavioral Regulation factor, which consists of the Inhibit and Self-monitor scales; and (iii) an Emotional Regulation factor, which is composed of the Emotional Control and Shift scales.

The BASC-2 Self-Report of Personality-College (BASC-2-SRP-COL). The BASC-2 (Reynolds & Kamphaus, 2004) is behavioral rating scale system that allows for a multimethod (i.e., multiple raters) and multidimensional (i.e., multiple aspects of behavior and personality) approach to measuring behaviors and self-perceptions of children and young adults. This rating scale was originally created to assess child behavior and emotion in order to facilitate differential diagnosis of pertinent childhood disorders and to assist with treatment plan design. Furthermore, when used as a comprehensive system, the BASC provides the clinician a more integrated and complete understanding of a child (Garcia-Barrera, Duggan, Karr, & Reynolds, 2014). Broadly, the BASC-2 consists of two categories of scales: clinical and adaptive. Overall, the clinical scales measure behaviors deemed as maladaptive or dysfunctional, with higher scores representing more negative characteristics that may impact functioning in one or more settings (e.g., home and/or school).

Recently, Garcia-Barrera and colleagues (2011) derived a four-factor screener of executive function from the original BASC teacher rating scale. The four factors include: (i) Problem solving, which measures planning, decision-making, conflict resolution, and the orientation of behavior towards goal achievement; (ii) Attentional Control, which measures behaviors related to focusing, sustaining, and shifting attention in response to current task demands; (iii) Behavioral control, which measures behaviors related to

inhibition and impulse control; and (iv) Emotional control, which measures behavioral inhibition but differs in the emotional saliency of the inhibited responses (i.e., controlling or delaying emotional responses). Given the support for Garcia-Barrera et al.'s (2011) four-factor model, a similar screener from the BASC-2-SRP-COL has been derived (Duggan et al., in press) and serves as an effective executive behavior screener for young adults. Further, this approach has been successful in the derivation of an EF screener with other versions of the BASC, suggesting its utility in measuring executive behaviors across version, raters, gender, age, and time (Garcia-Barrera, Kamphaus, & Bandalos, 2011; Garcia-Barrera et al., 2014; Sadeh, Burns, & Sullivan, 2012).

Comparison of BRIEF and BASC-2-SRP-COL scales in young adults. There are similarities and differences between the BRIEF-A and BASC-2-SRP-COL. First, both scales use a latent variable approach to executive behavior measurement, each with a slightly different theoretical and scale outcome. However, in a recent evaluation of the convergent validity between these two scales (Duggan et al., in press), the conceptualization and item content of the BRIEF-A and BASC-2-SRP-COL scales were reviewed and the scales determined to be most similar were isolated as target scales for comparison. The Plan/Organize scale (BRIEF-A) was selected as the scale best representing Problem Solving (BASC-2-SRP-COL). The Working Memory scale (BRIEF-A) was selected as the one most similar to Attentional Control (BASC-2-SRP-COL). The Inhibit scale (BRIEF-A) was selected as most closely representing Behavioral Control (BASC-2-SRP-COL). And finally, the Emotional Control scale (BRIEF-A) was selected as best representing Emotional Control (BASC-2-SRP-COL). As a result of the previously indicated convergence (Duggan et al., in press), representing overlap in

measures between the tests, it would be predicted that similar findings for factors that measure the same constructs. In addition to these target scales, the BRIEF-A also contains five additional scales, considered less similar to the four BASC factors: Organization of Materials, Shift, Initiate, Task Monitor, and Self-Monitor; these scales were not selected for comparison in the current study because of their lack of similarity to the BASC factors. Higher t-scores on both scales represent worse performance on executive behaviours (i.e., executive dysfunction).

The BASC was originally created as a broadband measure that captures dysfunctional behaviors, adaptive skills, clinical disorders, and personality traits, while the BRIEF was created specifically to measure executive behavior dysfunction. Despite these theoretical differences, both measures have demonstrated utility in measuring executive behaviors in clinical and nonclinical samples (e.g., Barkley & Murphy, 2010; Toplak et al., 2013). Also despite their theoretical differences, there is relevant overlap between executive behavior components, which renders both measures useful for measuring and contrasting self-rating executive behaviors in young adult bilinguals.

Objective

Using two behavioral rating scale measures of EF, we examined the bilingual advantage in young adults. In another study using the same sample (see Chapter 3), EF differences in inhibition and trends in switching were found, such that early bilinguals performed better than late bilinguals and monolinguals. The current study investigated whether or not these differences were detectable at the executive behavioral level.

Hypotheses

Hypothesis one. Early bilinguals will demonstrate an advantage (i.e., lower t-scores) on the BASC-2-SRP-COL Attentional control scale and on the BRIEF-A Working Memory scale over late bilinguals and monolingual controls.

Hypothesis two. Inhibition components on the BRIEF-A Inhibit Scale and the BASC-2-SRP-COL Behavioral control scale will reveal enhanced executive behaviors (i.e., lower t-scores) in both bilingual groups compared to monolingual controls.

Hypothesis three. Early bilinguals, late bilinguals, and monolinguals will not demonstrate differences on the remaining scales (i.e., BASC-2-SRP-COL Emotional control scale, BRIEF-A Emotional control scale, BASC-2-SRP-COL Problem solving scale, and the BRIEF-A Plan/Organize scale).

Method

Participants and Measures

One hundred and ninety-nine university students between the ages of 18 and 25 were recruited through a psychology research participant pool at the University of Victoria to take part in a larger study evaluating the convergent validity of several executive measures. The present study reports the results from a subsample of 97 participants who completed both the BASC-2-SRP-COL and the BRIEF-A. Participants were screened and excluded from enrolment if they were not between the ages of 18 and 25 (inclusive) or if they reported a significant history of neurologic or psychiatric disturbance (e.g., traumatic brain injury, seizures, mental illness), developmental disorder (e.g., Attention Deficit Hyperactivity Disorder, Fetal Alcohol Spectrum Disorder, Autism), learning disability, or substance abuse. In addition, participants completed a bilingual history questionnaire, which allowed for the creation of three distinct groups (i) monolinguals; (ii) early bilinguals; and (iii) late bilinguals. The study was approved by the Human Research Ethics Board at the University of Victoria. All participants were informed of the study's procedures, risks, and benefits and provided written consent before participating.

Participants completed a bilingual questionnaire (see Appendix B) to assess their relative level of proficiency in each additional language as compared to L1 (dominant language). L1 was assumed to be the strongest in reading, writing, and comprehension. Participants reported level of proficiency in L2 on a 10-point Likert-scale, with the anchors "not at all proficient" (1) and "Completely proficient" (10), and total number of languages, as well as when and how (e.g., school, home) they learned each language. Of

the 97 participants, there were 63 bilingual and 22 monolingual. Empirical evidence suggests differences in underlying neural development (e.g., Hull & Vaid, 2007) and in cognitive outcomes related to age of acquisition (AoA) (e.g., Luk, De Sa, & Bialystok, 2011). Therefore, bilinguals were separated into two groups: (i) early learners, defined as individuals who learned their second language at or before the age of six, and (ii) late learners, defined as those who learned their second language at or after the age of 12; this is the definition used in Hull and Vaid's (2007) meta-analytic review, which, from a developmental perspective, was used to delineate two bilingual groups. Individuals who learned their second language between the ages of six and 11 were excluded from analysis ($n = 12$). A summary of the demographic characteristics of the monolingual subsample, early bilingual, and late bilingual subsamples is provided in Table 1. No significant differences existed between groups in age and language proficiency (bilingual groups only).

The bilingual participants spoke a variety of second languages. See Table 2 for the proportion of L1 and L2 languages spoken by the bilingual groups. Monolingual participants were native English speakers and had not gained substantial proficiency in a second language. Some monolingual individuals had some second language education after the age of 12; however, their responses on a language proficiency questionnaire indicated poor self-rated proficiency (<2/10 on the Likert-scale) in a second language and thus were included in the monolingual group.

Table 1. *Demographic Information*

	Group		
	Early Bilinguals Mean(SD)	Late Bilinguals Mean(SD)	Monolinguals Mean(SD)
N	44	19	22
Age	20.09(1.74)	20.95(2.04)	21.23(1.90)
Sex	73% female	84% female	68% female
AoA (L2)	3.23(2.43)	14.16(2.54)	-
Proficiency L2 - speak	7.98(1.99)	7.37(2.27)	-
Proficiency L2 - comprehend	8.84(1.27)	8.13(2.10)	-
Proficiency L2 - write	6.36(3.13)	6.58(2.69)	-
Proficiency L2 - read	6.86(3.32)	7.45(2.50)	-

Note. The demographic information is displayed for the participants in the current study, including: number of participants, mean age, sex distribution, and age of acquisition for each of the three groups. AoA = age of acquisition; L2 = second language; - Represents data that was not measured in monolinguals.

Table 2. *Dominant and Nondominant Language Spoken by Participants*

Languages	Group			
	Early Bilinguals (%)		Late Bilinguals (%)	
	L1	L2	L1	L2
Afrikaans	2.33	4.65	-	-
Chinese	16.27	-	31.58	10.52
Danish	-	-	-	5.26
Dari	-	2.33	-	-
English	60.47	32.56	42.11	47.37
French	-	30.23	-	15.79
German	-	-	-	5.26
Hindi	-	2.33	-	-
Korean	2.33	-	11.11	-
Polish	-	4.65	-	-
Punjabi	6.97	11.62	-	-
Portuguese	-	-	5.26	-
Spanish	-	2.33	5.26	10.52
Swiss-German	-	4.65	-	-
Tagalog	2.33	-	-	-
Teochew	2.33	-	-	-
Turkish	-	-	5.26	-
Vietnamese	-	2.33	-	5.26

Note. This table displays the proportion of L1 and L2 languages spoken by each bilingual group. ^A represents mandarin and Cantonese.

The BASC-2-SRP-COL. This form contains 185 items, 68 of which consist of dichotomous items to be rated either true or false, and 117 items to be rated on a four-point Likert scale ranging from 1 (never) to 4 (almost always; Reynolds & Kamphaus, 2004). While most other ratings scales for individuals ages 18 to 25 are designed for use

in all adults (e.g., 18 to 90), the items on the BASC-2-SRP-COL are specifically designed to capture many behavioral, emotional, self-concept formation issues that are specific to the developmental period representing the transition from adolescence into early adulthood.

The BRIEF-A. The BRIEF-A is the most commonly used and empirically supported rating scale for the assessment of behavioral manifestations of executive dysfunction in adults ages 18 to 90 (Isquith et al., 2013). The BRIEF-A contains 75 items rated on a three-point Likert scale (Never, Sometimes, and Often), which contribute to nine scales: inhibit, shift, emotional control, self-monitor, initiate, working memory, plan/organize, task monitor, and organization of materials.

Statistical Analyses

There are four BASC executive function factors that have been reliably derived: Problem Solving, Attentional Control, Behavioral Control, and Emotional Control (Duggan, Mueller, & Garcia-Barrera, 2016; Garcia-Barrera et al., 2011, 2013; Sadeh et al., 2012). These four factors were compared to their corresponding scales from the BRIEF-A. Factor raw scores for each of the BRIEF and BASC scale were obtained for each participant, which were then converted to *t*-scores. The Statistical Package for the Social Sciences (SPSS) version 21.0 was used for all analyses. The subsample *t*-scores, for each factor, were compared between the three experimental groups using one-way ANOVAs. An alpha level of 0.05 was used for all analyses; a Bonferroni correction was used for all significant post-hoc analyses.

Results

Means and standard deviations (SD) are shown in Table 3 and t-tests are shown in Table 4.

Table 3. Mean and SD Table

Variables	Groups		
	Early Bilingual Mean (SD)	Late Bilingual Mean (SD)	Monolingual Mean (SD)
<i>BASC</i>			
Problem Solving	51.03 (15.20)	54.68 (11.26)	51.00 (9.93)
Attentional Control	52.88 (14.07)	50.51 (9.90)	48.89 (10.49)
Behavioral Control	51.29 (14.05)	47.31 (11.35)	50.21 (8.49)
Emotional Control	53.30 (14.42)	52.62 (9.57)	49.30 (7.15)
<i>BRIEF-A</i>			
Plan/Organize	52.45 (9.27)	55.84 (11.20)	52.83 (7.45)
Working Memory	54.38 (8.29)	53.42 (9.69)	54.52 (10.07)
Inhibit	54.91 (10.28)	56.58 (10.72)	54.91 (8.65)
Emotional Control	51.02 (10.17)	57.47 (9.03)	52.62 (7.63)

Note. This table displays the means and SD for each of the executive behavior variables for each group.

Table 4. ANOVA Table

Variables	One-way ANOVAs			
	<i>df</i>	<i>F</i>	<i>p</i>	η_p^2
<i>BASC-2</i>				
Problem Solving	2, 79	0.567	0.569	0.014
Attentional Control	2, 79	0.814	0.447	0.019
Behavioral Control	2, 79	0.703	0.498	0.017
Emotional Control	2, 79	0.846	0.433	0.020
<i>BRIEF-A</i>				
Plan/Organize	2, 79	0.976	0.381	0.024
Working Memory	2, 79	0.091	0.913	0.002
Inhibit	2, 79	0.205	0.815	0.005
Emotional Control	2, 79	3.14	0.049	0.074

Note. The individual F tests are presented above, including the degrees of freedom (*df*), *F*, the *p* value, and proportion of explained variance (η_p^2). Bolded items were significant at $p < 0.05$.

Post-hoc analyses of the parametric tests for the BRIEF-A Emotional control scale revealed significantly higher scores for late bilinguals than early bilinguals, $p < 0.05$. Monolinguals were not significantly different than either group.

Discussion

Despite differences on laboratory-based tests of EF at the cognitive/information processing level of measurement between bilinguals and monolinguals (e.g., Bialystok et al., 2004, 2008, 2014; Bialystok & Depape, 2009; Bialystok, 2006b; Blumenfeld & Marian, 2013; Colzato et al., 2008; Garbin et al., 2010; Kousaie et al., 2014; Marzecová et al., 2013), which helped to guide hypothesis formation for measurement at the *executive behavior* level, the first hypothesis, that early bilinguals would demonstrate an advantage (on the BASC-2-SRP-COL Attentional control scale and on the BRIEF-A Working Memory scale over late bilinguals and monolingual controls, was not supported. The second hypothesis, that inhibition components on the BRIEF-A Inhibit Scale and the BASC-2-SRP-COL Behavioral control scale would reveal enhanced executive behaviors both bilingual groups compared to monolingual controls, was also not supported. However, part of the third hypothesis, that no differences would be revealed for the BASC Problem solving scale and the BRIEF-A plan/organize scale was supported.

Despite a prediction of the contrary, a significant difference on the BRIEF-A Emotional control scale was found, but no significant difference on its analogous BASC scale, suggesting that the scale, at least when the sample is separated into language groups, may not measure the same construct. Post-hoc analysis of this finding revealed a difference between late bilinguals and early bilinguals, such that late bilinguals scored higher (i.e., more problematic emotional control) than early bilinguals; however, the other groups did not differ significantly. The former finding may lead one to suggest that late bilinguals have worse emotional control compared to early bilinguals. However, the mean T-score values for all groups are within the average or expected range (Roth et al.,

2005), suggesting that this difference may not represent a behaviorally significant difference; rather, it may simply reflect a meaningless statistical difference, such as a type I error as a result of multiple F tests.

Some authors (Isquith et al., 2013; Silver, 2014) suggest that behavioral rating scales are more ecologically valid ways of measuring EF. Toplak, West, and Stanovich (2013) compared 20 studies that used both laboratory performance-based (e.g., computerized tasks, neuropsychological paper and pencil measures) and behavioral rating scale measures of EF, and found that there is little correlation between them. From this, they suggest that these actually reflect different levels of measurement or *cognitive levels*. They propose that performance-based tasks, such as the Stroop and go/nogo tasks, measure algorithmic or information processing mechanisms, similar to the cognitive/information processing level of the multi-level approach to measuring the bilingual advantage. This level of measurement involves coding mechanisms, perceptual registration, working memory, long-term memory, and other specific cognitive mechanisms involved in processing information. They argue that behavioral rating scales, on the other hand, measure a different level or type of processing, referred to as the *reflective level*, which is concerned with the goals, belief structures, and choices of action related to the person. Together, these additional metacognitive processes assist with optimal decision-making and goal attainment. That is, in a hierarchical perspective, this metacognitive level pulls together the specific processes, ultimately resulting in more complex processing and, thus, outcomes. Moreover, these authors argue that the reflective level is more related to executive-related functions (i.e., executive behaviors), while the cognitive processing level is more related to *supervisory* processes. Relating

this to the bilingual advantage, the subtle enhancement effects that may be offered by learning a second language at the cognitive processing level may not be revealed using such a broadband measure of EF, despite being an ecologically valid measurement of EF.

A similar argument specifically related to understanding the bilingual advantage and multiple levels of measurement have been suggested. This approach contends that the bilingual advantage (related to EF) can be measured at multiple levels with varying degrees of complexity (e.g., executive behaviors, cognitive/information processing, physiology, etc.). The findings from the current research, namely, that no differences were found between groups for most of the scales, provide preliminary evidence for the multi-level approach. It is predicted in this multi-level approach that executive behaviors that the more complex the level of measurement (i.e., further from cells/networks/structures), the more *diluted* the bilingual advantage becomes, and thus, more difficult to reveal the subtle underlying differences between monolingual and bilingual young adults. Behavioral rating scale measurement of EF is the measurement of executive *behaviors*, which are the most complex form of EF. That is, as demonstrated in Figure 4 below, executive behaviors are the cumulative result of several underlying levels of measurement/processing. As is evidenced by the crossing black arrows, multiple cognitive processes culminate to produce executive behaviors.

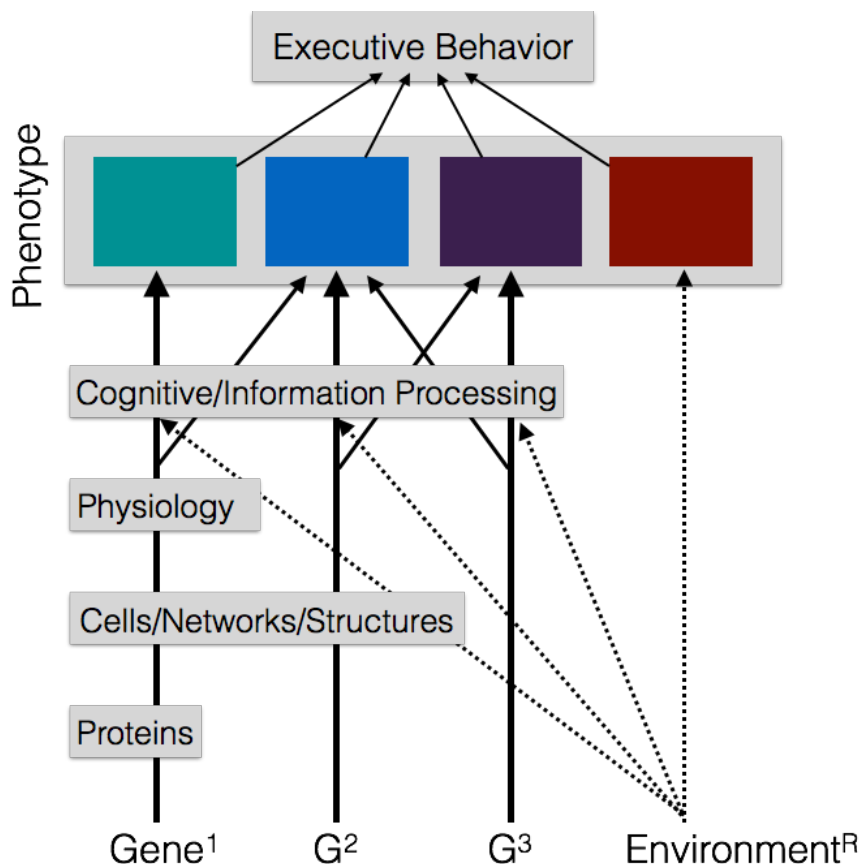


Figure 4 represents a multilevel perspective, where multiple interactions between genes, proteins, and structures interact with the environment during the production of cognitive processing, ultimately resulting in executive behaviors. This model was adapted from Crosbie et al. 2008 (model for endophenotypes) and Garcia-Barrera, Frazer, & Areshenkoff, 2012.

That is, to our knowledge, this is the first study to directly compare the behavioral outcomes of EF through rating scales (i.e., executive behaviors); we did not find significant or relevant differences between groups, suggesting that, despite purported underlying difference at the cognitive processing level (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Craik, & Ryan, 2006; Bialystok, Craik, & Luk, 2008; Bialystok & Depape, 2009), these differences do not provide additional advantages at a higher- or executive behavior- level of processing. The idea that behavioral rating scales are more ecologically valid measures of EF that relate to rational goal pursuit, and more behaviourally-relevant behaviors that are manifest in day-to-day activities, it begs the

question of whether or not the bilingual advantage is worth ongoing study. However, despite a lack of differences at the level of executive behaviors that may be relevant to an individual's daily function (i.e., ecological validity), understanding how bilingual brains differ from monolingual brains furthers our understanding of language processing and the complex interactions between underlying neural systems. This *basic* knowledge is important and relevant to advancing our understanding of bilingualism.

Some authors argue the bilingual advantage does not actually exist at the cognitive processing level (e.g., Paap & Liu, 2014; Paap, 2014; Paap, Sawi, Dalibar, Darrow, & Johnson, 2014). Not finding differences in executive behaviors may help to support that line of evidence. However, the multi-level approach contends that differences would not manifest at this level as a result of the complexity of executive behaviors and the possible subtlety of the bilingual advantage in young adults. The research literature for the bilingual advantage in young adults is mixed; therefore, further research that empirically examines the already available literature is needed. Moreover, this study opens a novel line of research not yet conducted, calling for more research on executive behaviors to expand on this novel idea.

Limitations. The current study cannot solely address the mixed nature of the bilingual advantage literature, and in fact, did not attempt to do so. This study is one piece of a line of literature that, as a whole, seeks to shed light on the debate of whether or not the bilingual advantage exists in young adults. Although the current study provides some preliminary evidence in line with the multi-level approach for assessing the bilingual advantage, it is only one piece of a multi-part puzzle. In terms of specific limitations, there were small and unequal sample sizes, which may have hindered finding

significant and relevant differences between groups. Another limitation of the current study is the lack of significant findings, which may hinder its dissemination in and valuable information to the bilingual advantage literature. Known as the *file drawer problem*, which delineates the major problem in scientific research that null results do not get published (Rosenthal, 1979). Taken out of the context of the multi-level approach, the null findings lose their potential value to understanding the mixed nature of the bilingual advantage literature in young adults.

Future directions. It is clear that more research that empirically synthesizes the available literature is needed to help settle the debate of whether or not the bilingual advantage can be documented for young adults at the cognitive processing level (i.e., using performance-based EF measures). It is also clear that more research on executive behavior rating scale comparisons between bilingual and monolingual individuals is needed in order to corroborate the current findings. Moreover, in order to support the argument made in the current study that behavioral rating scales and performance-based measures of EF do in fact measure different aspects of EF, a comparison of performance-based measures (i.e., cognitive/information processing level) using the same participant pool who completed the rating scales (i.e., executive behaviour level) would be necessary. A direct correlational comparison between performance on cognitive processing EF tasks and behavioral rating scales would strengthen this argument, especially given that some of the executive behaviors measured in the current study have analogous cognitive processing measures (e.g., inhibition, problem-solving, and attentional control). The multi-level approach predicts that there would be little correlation, as found by Toplak

and colleagues (2013), further confirming that they measure different, yet uniquely important, aspects of EF.

Chapter 3

The bilingual advantage in young adults:

Examining the importance of age of acquisition when measuring executive function

The bilingual advantage in young adults: Examining the importance of age of acquisition when measuring executive function

Models of the bilingual neural networks contended that second language comprehension and production was controlled through either a process of (i) inhibition (Green, 1998), (ii) relative levels of activation (Soares & Grosjean, 1984), or (iii) through other modes of network information processing, such as the Bilingual Interactive Activation+ (BIA+; e.g., Dijkstra & van Heuven, 2002). These models seek to describe the process through which one language network is activated during language perception and production in light of the contention that both language networks are simultaneously activated. These models differ in the underlying mechanisms through which this occurs, whether it is through reactive inhibition of the nontarget language (Green, 1998), a difference in the relative level of activation of the target versus nontarget language (Soares & Grosjean, 1984), or through neural network activation and language tagging (Dijkstra & van Heuven, 2002). However, rather than relying on internal systems and/or processes within the language networks, more recent research literature has suggested that the bilingual brain may effectively take advantage of another control system already set in place within prefrontal cortex and its networks: executive function (EF) (Bialystok, Craik, Green, & Gollan, 2009). Recent evidence also posits that the constant utilization of this control system in the bilingual brain results in earlier development of EF components (e.g., Bialystok & Viswanathan, 2009) and enhanced EF components through out the lifespan (Bialystok, 2011). These benefits are known as the ‘bilingual advantage.’

The ‘bilingual advantage,’ which posits that individuals who speak a second language have enhanced EF, has been well established in the literature, especially for

children (e.g., Bialystok, 2011). It is thought that this advantage arises as a result of the need to manage the perception and processing of two languages. Further, this advantage is strongly demonstrated in older adults in the form of delayed onset of cognitive impairment, dementia, and other signs of cognitive aging (e.g., Bialystok, 2008, 2011, 2012). However, the literature is mixed for young adults. Although not mutually exclusive, there are several interpretations for these findings: (i) individuals in this age range have similar levels of EF that are not fortified by bilingual language processing because they are at their developmental peak for this set of processes; (ii) the tasks used to assess EF are not sensitive enough to measure subtle differences in EF processes; and (iii) bilingualism does not foster an enhancement in young adults. This study examines the bilingual advantage of three specific EF components at the level of cognitive/information processing for early and late young adult bilinguals compared to monolingual controls: (i) inhibition; (ii) switching; (iii) updating; as well as (iv) slightly more complex outcomes of EF (i.e., problem solving).

Executive function

Executive function is an umbrella term and serves as a general referent to a set of cognitive abilities that are drawn upon when encountering novel problems or goal-oriented activities (Anderson, 2008). Recent research has focused on the developmental course, dimensionality, and malleability of EF; the latter of which is prominent in the bilingual advantage literature. Research findings have indicated that EF follows a protracted developmental course (e.g., Romine & Reynolds, 2005), showing evidence of both unity and diversity (e.g., Garon, Bryson, & Smith, 2008; Miyake & Friedman, 2012). Research has also demonstrated that EF is malleable, showing deficits in response to

stress (e.g., Arnsten, 2000) and enhancements in response to structured activities (e.g., Bryck & Fisher, 2012) and second language acquisition (e.g., Bialystok, Craik, Green, & Gollan, 2009). The research literature on second language acquisition suggests that bilingual language networks must recruit the EF system to manage attention to the target language while updating language demands and inhibiting the nontarget language within the context of linguistic and environmental selection constraints. While it is well accepted that EF is a unitary and diverse construct (Miyake & Friedman, 2012; Miyake et al., 2000), the individual components that reflect the diversity of EF are still debated. A theoretical interpretation that is well supported suggests that EF consists of, at least, three core components: inhibition, updating, and shifting (Miyake et al., 2000). Although this model of EF suggests diversity in multiple core components, performance across these components is generally correlated within individuals, suggesting interactions between the components in various tasks.

A model of EF. Miyake and colleagues (2000) have provided strong empirical evidence for the diversity of EF; here, each component is briefly defined.

Inhibition. Many authors distinguish between suppression of an over-practiced, automatic, or ‘prepotent’ behavioral response, and the ability to suppress or ignore information that is irrelevant but is currently interfering with or eliciting a conflicting response on the immediate task. The former of these two processes has generally been named response inhibition (e.g., Barkley, 1999; Nigg, 2003) or behavioral inhibition (Barkley, 1997); while the latter process is typically referred to as interference control (Friedman & Miyake, 2004), conflict resolution (Posner & DiGirolamo, 1998), or executive attention (Posner & Rothbart, 2007). It is important to differentiate types of

inhibition from reactive suppression and spreading activation models of inhibition because the latter types of inhibition do not involve deliberate control.

Switching. The construct of ‘attention’ is complex and is conceptualized as involving multiple functions and processes, and thus, cannot be described as a single system. As such, there are many models and theories of attention (see Neumann, 1996). Specifically, Posner and Rothbart (2007) argue for three attention systems, including a system for orienting attention, one for maintaining a state of system alertness, and finally a system of attention under the influence of executive control (i.e., executive attention). The latter attentional system is the most similar to the conceptualization of executive attention proposed by Miyake and colleagues (2000). Miyake’s conceptualization of EF includes *shifting*, also referred to as attention switching or task switching, which concerns shifting back and forth between multiple tasks, operations, or mental sets.

Updating. Working memory (WM) is a cognitive system dedicated to the *temporary* processing, maintenance, and integration of information during the performance of everyday cognitive tasks (Baddeley, 2003b). Given its purpose of maintaining task-relevant information in mind, WM is critical because it allows individuals to hold information received from the environment or retrieve information stored in long-term memory, and subsequently maintain that information if it is relevant to the task goals (Unsworth & Engle, 2007). If information becomes irrelevant, it is deleted and replaced with new, relevant information (i.e., updating). An individual can use and organize updated information to execute goal-directed behaviors, which is one of the main outcomes of the interactions of EF. It is necessary for any conceptualization of EF to include a short-term storage component that maintains current (updated), task-

relevant information until it is no longer needed, according to the goals of a given task at a given time. The essence of updating, as proposed by Miyake and colleagues (2000), is the requirement to actively manipulate relevant information in working memory, rather than passively store information.

Measuring the bilingual advantage

In bilinguals, researchers have proposed that simultaneous activation of both languages occurs when using only one of the languages, and thus, the language system is required to manage this nonselectivity in neural activation. To assist with this conflict, Bialystok and colleagues proposed that the language networks make use of the EF networks to facilitate this management (Bialystok et al., 2009). Within these integrated network connections, the constant use of the EF system to manage language inhibition, switching, and updating is purported to provide a fortifying effect for EF. This bilingual advantage has been prominently documented in childhood, where bilingual children outperform monolingual children on tasks of inhibition, working memory, and task switching (Bialystok, 2011; Kroll & Bialystok, 2013). Further, evidence suggests that the advantage extends to older adults, where they experience enhanced EF compared to monolingual older adults, delayed onset of dementia (up to five years) and less decline in other cognitive functions (Bialystok, 2011). The current study tests the fortification of the EF in young adults, where the EF system should be at its peak performance. The first question to be tested is whether or not the bilingual advantage is present across different EF components in this age range. The young adult bilingual advantage research literature is briefly reviewed for inhibition, switching, and updating, as well as for the influence of the age of acquisition (AoA) on EF enhancement. Reviews have been conducted in

chapter one; to reduce redundancy, the current review focuses on tasks that tap into the underlying EF components as well as those that are typically used in the bilingual advantage literature.

Inhibition. Most research in the literature have used tasks that measure *conflict resolution* or *interference control* aspects of inhibition, such as the Simon task (e.g., Bialystok et al., 2004; Bialystok & Depape, 2009; Bialystok, Martin, et al., 2005; Linck et al., 2008), the Stroop task (e.g., Bialystok et al., 2008; Bialystok, Poarch, Luo, & Craik, 2014; Blumenfeld & Marian, 2013; Costa, Hernández, & Sebastián-Gallés, 2008; Kousaie & Phillips, 2012; Kousaie et al., 2014), and a flanker task (e.g., Luk, De Sa, & Bialystok, 2011). However, expanding the inhibition literature, the ability to deliberately inhibit, dominant, automatic responses when necessary, otherwise known as *response inhibition*, is often measured using a version of the go/nogo task. To tap into inhibition and expand on the existing literature, we chose a traditional form of the go/nogo task where letters are presented on screen and a participant must inhibit a button press only to a specified letter. Tasks that similarly tap into response inhibition have found positive results; a bilingual advantage was found using a modified antisaccade task (Bialystok, 2006) and an inhibition of return task (Colzato et al., 2008).

Switching. The ability to shift back and forth between multiple tasks, operations, or mental sets can be measured using nonlinguistic switch tasks (e.g., Garbin et al., 2010), global-local tasks (e.g., Bialystok, 2010), trail making tasks (e.g., Bialystok, 2010), dichotic-listening tasks (e.g., Soveri, Laine, Hämäläinen, & Hugdahl, 2010), the lateralized attention network test (e.g., Tao et al., 2011) and a color-shape switch task (Prior & Macwhinney, 2009). Three studies using a task switch paradigm in bilingual

young adults have found a bilingual advantage (Garbin et al., 2010; Prior & Macwhinney, 2009; Soveri et al., 2010); offering hopeful outcomes for future investigation. Although other switch tasks, such as the color-shape switch task, were used in the bilingual advantage literature, we chose the local-global task, as it is commonly used to measure one's ability to *switch* between global and local features of a stimulus. As well, the selection of this task complements and expands the existing literature.

Updating. This component requires one to actively manipulate relevant information in working memory, rather than passively store information. Tasks typically used to measure this component include the n-back task (e.g., Chatham et al., 2011), the keep track task (Yntema, 1963), the letter memory task (Morris & Jones, 1990), and the tone monitoring task (Gardiner & Parkin, 1990). The bilingual advantage literature for updating advantages is lacking. In the bilingualism literature, there are various *working* memory tasks that authors have used, including – but not limited to – the n-back task (e.g., Soveri, Rodriguez-Fornells, & Laine, 2011), the alpha span task (e.g., Kramer, 2011), Corsi block test (e.g., Bialystok, Craik, & Luk, 2008; Wodniecka, Craik, Luo, & Bialystok, 2010), and self-ordered pointing (e.g., Bialystok et al., 2008). The majority of these tasks, with the exception of the n-back task, do not measure the *updating* component of working memory, which is the *executive* nature of this component, at least as posited by Miyake and colleagues (2000, 2012). The n-back task has been used to measure updating. For example, Chatham and colleagues (2011) used it to run simulations and Miyake and colleagues (2012) have used it to measure their updating component. In order to tap updating, we chose a letter-variant of the n-back task, where individuals have to keep two letters in WM while releasing old, irrelevant information.

Complex EF tasks. Dependence between EF constructs to produce a behavioral outcome is an issue related to EF measurement; this issue is known as the *task-impurity problem* (Miyake & Friedman, 2012; Miyake et al., 2000). Many traditional EF tasks involve more than one component to produce an accurate and/or efficient response. These tasks tend to measure more complex descriptions of EF such as problem solving, planning, and organization, and involve the use of inhibition, switching, and updating components. Complex measures of EF that are often employed to measure EF and frontal lobe function include the Wisconsin Card Sorting Task (Stuss et al., 2000) or other card sort tasks (Bialystok & Martin, 2004), the Iowa gambling task (IGT; Toplak, Sorge, Benoit, West, & Stanovich, 2010), and various versions of the tower task (Hull, Martin, Beier, Lane, & Hamilton, 2008), to name a few. To our knowledge, only one version of the card sorting task has been employed in the young adult bilingual advantage literature (e.g., Kousaie et al., 2014). To expand on this literature and add to our understanding of complex EF task performance in young adult bilinguals, we selected three complex tasks, including a card-sorting task, the IGT, and a tower task.

Age of acquisition. Age of acquisition (AoA) has been suggested to drive a different neural representation of those bilinguals who learned both languages at different points in their life. Soares & Grosjean (1984) were the first to methodologically define late bilinguals. Since then, Hull and Vaid (2007) conducted a meta-analysis using the following operational definition: early (or infant) bilinguals learned L2 by the age of 6 and late (or adult) bilinguals learned L2 after at or after the age of 13. Although this operational definition of AoA was selected for the current study, different operational definitions are used across studies; the latter issue has been discussed in detail in chapter

1. Despite this variability in operational definitions, several studies have found differences when AoA is considered (e.g., Luk, De Sa, & Bialystok, 2011; Pelham & Abrams, 2014; Tao, Marzecová, Taft, Asanowicz, & Wodniecka, 2011), while a recent study found no differences (Kalia et al., 2014). Despite some variability in empirical findings, significant differential effects have been found for early and late bilinguals. We hypothesize that learning an L2 early in life may produce more fortified enhancements of the EF system while it is developing, resulting in long lasting benefits throughout an individual's life. Evidence to support this contention would arise from early bilinguals performing better on EF tasks than late bilinguals and monolinguals. Alternatively, learning a second language later in life and in line with the development of EF, late bilinguals may, in fact, train their EF system to the same extent as early bilinguals and hence display a similar cognitive benefit related to the training. This may arise from a late bilingual's need for more adaptation to the demands of managing and producing a second language once L1 networks are already established. Evidence to support this contention would arise from early and late bilinguals performing better on EF tasks than monolinguals. A high level of proficiency of both languages is required when comparing groups, as a result of the influence of language proficiency on language networks (e.g., Mechelli et al., 2004; Newman et al., 2012; Nichols, 2013; Perani et al., 1998).

The present study

The idea that bilingual brains require the use of an EF system to manage the input and output of more than one language is becoming well-accepted. Evidence suggests that the adaptation of this system from repeated and consistent multi-language use has an enhancing effect on EF in children and older adults. What is less clear is whether or not

the advantage is present in young adults, where the functioning of the EF system is thought to be at its peak. The age at which an individual learns a second language may have an impact on the development of the brain (Hull & Vaid, 2007), especially for those who learn their second language early in development. Differential effects have been found on EF tasks for early and late bilinguals (e.g., Luk, De Sa, & Bialystok, 2011; Pelham & Abrams, 2014; Tao, Marzecová, Taft, Asanowicz, & Wodniecka, 2011). Thus, measuring AoA in bilingual individuals may provide comparison groups that better distinguish the bilingual advantage in young adults, a group in which this advantage is inconsistently found.

Although empirically confirming the structure of EF is not the primary goal of the current research, comparing monolinguals and bilinguals using a multicomponent perspective may provide some additional insight into the debate of the structure of EF, in addition to the primary goals of the current research. Bialystok (2011) indicated that the research strategy should be to demonstrate that monolingual and bilingual participants achieve different levels of success in performing EF tasks and that these differences cannot be attributed solely to the functioning of the individual core components. Instead, monolingual/bilingual performance differences on EF tasks may be traced to differences in the recruitment of the entire EF network (i.e., interactions), not just to the individual components of that network. In essence, Bialystok is reiterating the debate about the impurity of EF tasks, and reinforcing the idea of the *interactions* between EF components. That is, constant use of specific EF components, such as inhibition and switching, may lead to an indirect enhancement of other individual components (i.e., updating), as well as other more complex outcomes of EF (e.g., problem solving); thus resulting in

fortification of the entire EF system. Therefore, measuring the bilingual advantage from a multi-component perspective at the cognitive/information processing level may provide insight into function of the individual components and their interactions at this level.

The current study uses several computerized tasks: three specific computerized tasks to measure each of the individual EF components in Miyake's proposal: The go/nogo task measures inhibition, the local-global task measures switching, and the n-back task measures updating. Together, these tasks measure the cognitive/information processing level of measurement, as proposed by the multi-level approach to measuring the bilingual advantage. As well, a card sort task, a tower task, and a version of the Iowa Gambling Task were selected to assess the hypothesized transfer of EF fortification to more complex interactions of the EF system. In relation the multi-level approach, the cognitive/information processing level can be divided into two sublevels where the complex tasks are theoretically situated between the cognitive/information processing level and the executive behaviour level.

Hypotheses. Following the multi-level approach to measuring the bilingual advantage, performance measurement is being conducted at the *cognitive/information processing level* of measurement. A global hypothesis for the individual EF components is that the results will be mixed; specific hypotheses for each component are reviewed below. However, for the more complex EF tasks at this level of measurement, the results will reveal no significant differences, as a result of their theoretical placement between the cognitive/information processing level and executive behaviors level.

Hypothesis one. Research on the inhibitory control of bilinguals is extensive, demonstrating a strong effect of children, mid-aged adults, and older adults (Bialystok et

al., 2005). Although the research for young adults is mixed, the inclusion of AoA as an additional variable may reveal differences between groups. Early bilinguals will show an advantage over late bilinguals and monolinguals in inhibitory control mechanisms, as a result of the contention that the system is fortified early in development and maintaining the need for inhibitory mechanisms throughout life.

Hypothesis two. In our literature review, we identified three studies on young adults that found a bilingual advantage for attention switching (Garbin et al., 2010; Prior & Macwhinney, 2009; Soveri et al., 2010). Following the logic of fortification of the bilingual EF system for inhibition, we hypothesize a similar advantage for switching mechanisms in early bilinguals, compared to late bilinguals and monolinguals.

Hypothesis three. Updating has been suggested to play a role in verbal processes (e.g., Szmalec, Brysbaert, & Duyck, 2012). Updating may facilitate bilingual language processing by keeping *online* information relevant to the current conversation. Thus, we hypothesize an advantage for the updating component in bilinguals. In keeping with the idea that early learning fortifies the early development of EF, the bilingual advantage will be apparent in early bilinguals but not late bilinguals.

Hypothesis four. Although it is suggested that constant use of the individual EF components produces general fortification of the EF system in the bilingual brain, we hypothesize no differences between groups on all of the complex EF tasks. Following the multi-level approach, the subtle advantage in young adults would be washed out by the complexity of these tasks.

Hypothesis five. AoA appears to play an important role in bilingual language processing and neural network development. Following the logic that learning a second

language early in life fortifies the development of EF components as a result of the earlier need, use, and thus, development of this system, we hypothesize that inclusion of AoA in the analysis will facilitate the identification of differences of the effect of bilingualism on cognition in young adults. That is, when early and late bilinguals are collapsed together and compared to monolinguals, significant differences between monolingual and bilingual groups will dissolve.

Method

Participants

200 undergraduate students were included in this study¹⁶ ($M_{age} = 20.49$, $SD = 1.9$, 77% female). Of these participants, there were 96 bilingual and 79 monolingual participants who met the required language criteria. Further, the bilingual group was separated into two groups: early learners, defined as individuals who learned their second language before the age of six, and late learners, defined as those who learned their second language after the age of 12; individuals who learned their second language between the ages of six and 11 were excluded from analysis ($n = 25$). All participants had normal or corrected-to-normal vision and none had a history of head injury. Participants were recruited via an online undergraduate psychology recruitment source and were given course credit for participation in the study. The background variables of the three groups are presented in Table 5 below.

The bilingual participants spoke a variety of second languages, including: Afrikaans (1.61%), Cantonese (0.81%), Chinese (1.61%), Danish (0.81%), Dari (0.81%), English (28.23%), French (34.68%), German (1.61%), Igbo (1.61%), Hindi (2.42%), Japanese (0.81%), Mandarin (1.61%), Polish (1.61%), Punjabi (9.68%), Serbian (0.81%), Spanish (6.45%), Swiss German (1.61%), and Vietnamese (1.61%). Monolingual participants were native English speakers and had not gained substantial proficiency in a

¹⁶ The current study is an analysis of a larger study, which was collected to do a secondary validation of INTERACT (a model of EF) by measuring the components from various levels of measurement. As such, a total of 200 participants completed behavioral rating scales (e.g., the BASC-2-College, BRIEF, and DEX scale), five computerized cognitive tasks thought to measure each of the five EF components, and four computerized 'traditional' EF tasks that also measure complex EF components. The larger study was conducted as a follow up to the original validation study by Frazer (2012). A bilingual questionnaire was included in this study in order to conduct secondary analyses and answer questions about the bilingual advantage using a large sample and the INTERACT model.

second language. Some monolingual individuals had some second language education after the age of 12; however, their responses on a language proficiency questionnaire indicated poor self-rated proficiency in a second language and thus were included in the monolingual group. Importantly, the Early and Late bilingual groups were significantly different in respect to self-rated proficiency for speaking and comprehending L2; further, all included participants met a minimum cut-off score of self-rated proficiency of 6.

Table 5. *Demographic Information for Chapter 3*

	Group		
	Early Bilinguals Mean(SD)	Late Bilinguals Mean(SD)	Monolinguals Mean(SD)
N	64	32	79
Age	20.02(1.59)	20.84(2.13)	20.73(1.90)
Sex	76.6% female	84.4% female	81.0% female
AoA (L2)	3.38(2.29)	13.25(3.04)	-
Proficiency L2 - speak	7.78(2.04)*	6.22(2.54)*	-
Proficiency L2 - comprehend	8.70(1.44)*	6.92(2.76)*	-
Proficiency L2 - write	6.00(3.28)	5.79(2.69)	-
Proficiency L2 - read	6.66(3.32)	6.61(2.73)	-

Note. The demographic information is displayed for the participants in the current study, including: number of participants, mean age, sex distribution, and age of acquisition for each of the three groups; AoA = age of acquisition; L2 = second language; - represents data that was not measured in monolinguals; * represents significant difference $p < 0.05$

Materials

Participants completed a bilingual questionnaire (see Appendix C) to determine their relative level of proficiency in each additional language other than L1. L1 was assumed to be the participant's dominant language, and the strongest in reading, writing, and comprehension; L2 proficiency was compared to L1. Participants self-report recorded total number of languages, as well as when and how (e.g., school, home) they learned each language.

Stimuli & Procedure

Three computerized tasks, one that measures each of the three EF components posited by Miyake et al. (2000), were selected for the current study. Further, three additional tasks of more general executive function were included. All participants completed these six tasks in a single experimental session lasting approximately 60 minutes. All computerized tasks were presented on a Lenovo G560 laptop computer, with a 15.5-inch screen. Experimental scripts and data collection were managed by Psychtoolbox-3 (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) and Matlab[®] (2012b, The Mathworks, Natick, MA) to ensure accurate response time measurement. Participants were seated approximately 30 cm from the screen. Participants were also administered additional questionnaires and computerized tasks; however, the methodology and results of those tasks were part of a larger study and are not presented in the current study.

Specifically, the component-specific tasks include: (i) the go/nogo, which measures inhibition; (ii) the Local-Global, which measures attention switching; and (iii) the n-back, which measures updating working memory. The three complex tasks include: (a) the Iowa Gambling Task; (b) the tower task; and (c) the Berg's card sort task. These tasks are described in detail below.

Go/nogo Task. This task (Donders, 1868, 1969) required participants to respond to visual stimuli (i.e. a single letter appearing in the middle of the computer screen, presented at a rate of approximately one letter every 1,400 msec) by pressing one response key repeatedly to all letters ('go stimuli'), but refraining from key press when the letter "j" appears in the screen (rarer, 'nogo stimuli'). Participants were instructed to

respond as quickly as possible to stimuli. In an initial block of 35 trials, only go stimuli were presented to allow participants to develop a prepotent response tendency to press the key associated with most letters. In a second 100 trial block, nogo stimuli were randomly presented among go stimuli, although much less frequently than the go stimuli; the probability of nogo trials was 30%. As a result, top-down inhibitory control was required to prevent participants from responding according to the prepotent tendency to respond to go stimuli, in order to execute the less automatic response. The outcome measure is total number of correct commissions (i.e., accurately withhold response to 'j'; percentage of correct commissions); this value was calculated by dividing the number of correct commissions by the total number of nogo trials.

Local-Global Task. This task used 'Navon figures' (Navon, 1977) (see Figure 5) to allow participants to switch between features. Globally, these figures represent geometric shapes (e.g., a square); however, the lines comprising this overall shape are different 'local' geometric shapes (e.g., triangles). Participants were instructed to respond by pressing a number key on the keyboard, according to the number of lines contained in either the global or local geometric figure (circle = 1, triangle = 3, square = 4). Participants were instructed to attend to the global figure if the stimulus was blue, but attend to the local figures if the stimulus was red. Therefore, participants had to 'switch' between attending to the global features and the local features of the figure. The first two trial blocks (35 trials each) consisted of only blue stimuli (global cue) and only red stimuli (local cue), respectively. The third block consisted of 100 trials of both blue and red stimuli, presented at random. The outcome measure was the difference in accuracy

between block 3 (collapsed across local and global) and the average accuracy (collapsed across local and global) from blocks 1 and 2.

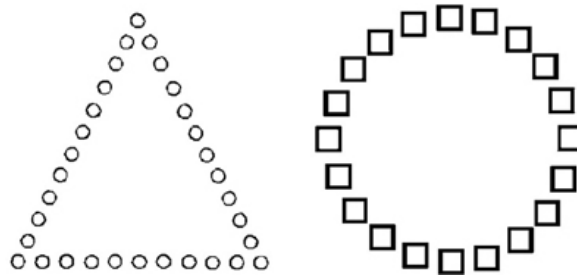


Figure 5. A schematic representation of Navon figures, where a “global” triangle is made up of “local” circles on the right and vice versa on the left.

N-back Task. In this task (Baddeley, 2003b), the subjects were instructed to press the key representing “yes” whenever the current letter they see on the screen is the same as the letter presented 2 trials back (e.g., L-T-L). As the letters are appearing in a constant stream, subjects were required to constantly update the letters they have held in working memory. There was one block with 70 trials; the outcome measure was the accuracy of total number correct two-back identifications (i.e., correct button presses for “hit” trials); this value was created by dividing the number of correct two-back identifications by the total number of hit trials.

Iowa Gambling Task (IGT). The IGT (see Figure 6) was used to assess participants’ reward learning in response to diverse levels of valence in outcomes. This task consisted of 100 trials with decks identical to those used in Bechara et al. (1994). After selecting a deck, participants received feedback in the form of a net gain or loss (in contrast to another common practice of presenting gains and losses simultaneously, and requiring the participant to calculate the net outcome). In the event that all of the cards in a deck were exhausted, the deck was reinitialized so that participants were free to select from all 4 decks at all times, without restriction. The outcome measure was the difference

between number of selections from the good deck minus the number of selections made from the bad deck in the last 25 trials. The difference over these last 25 trials was more useful because learning has not occurred early in the task.

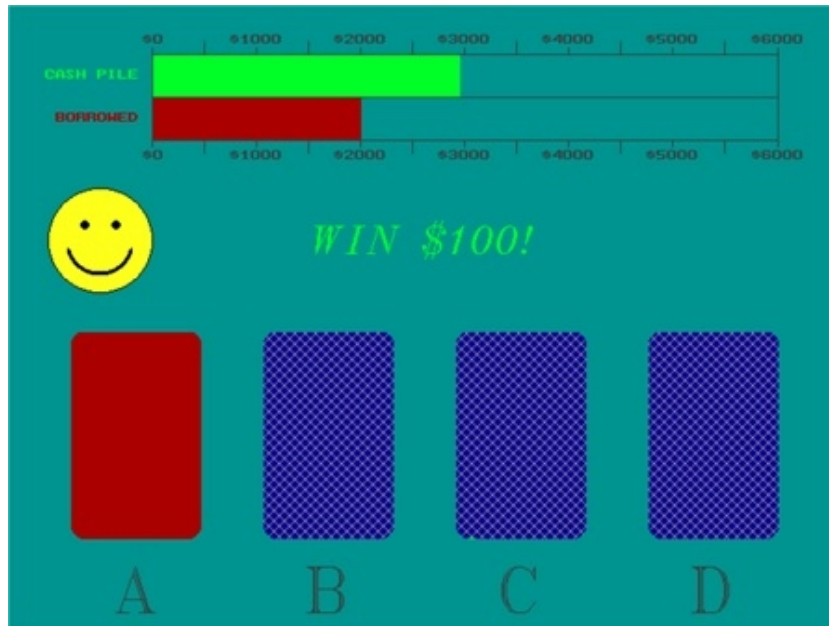


Figure 6. Illustration of the IGT, demonstrating feedback after selecting deck 'A.'

Tower of London (TOL) Task. A computerized version of the TOL from the Psychology Experiment Building Language (PEBL) platform version 0.13, (see Figure 7)(freely available at <http://pebl.sourceforge.net/>) was used to assess higher-order problem solving abilities. It was originally developed by Shallice (1982) and is regarded as a measure both sensitive and specific to planning abilities. In this task, a set of colored disks was presented at the top of the computer screen (i.e., the task goal) and a set of colored disks, in a different arrangement, was presented at the bottom of the screen. Participants were required to move the disks by using the mouse to click on the disk and move it to the desired column and to match the bottom set to the top set (Figure 3), using the smallest number of moves possible. The starting pattern of disks required 1 to 13 moves to complete the pattern and there were 12 patterns to solve, for a minimum total of

104 moves. A discontinue rule was not set during the task; however, participants were allowed to discontinue on their own volition, or the experimenter would discontinue the task after it was clear they were far off from completing the item. Once an item was discontinued, the task was discontinued overall. There was no maximum number of moves. The outcome measure was the total number of moves across all items; discontinued performance was coded as missing.

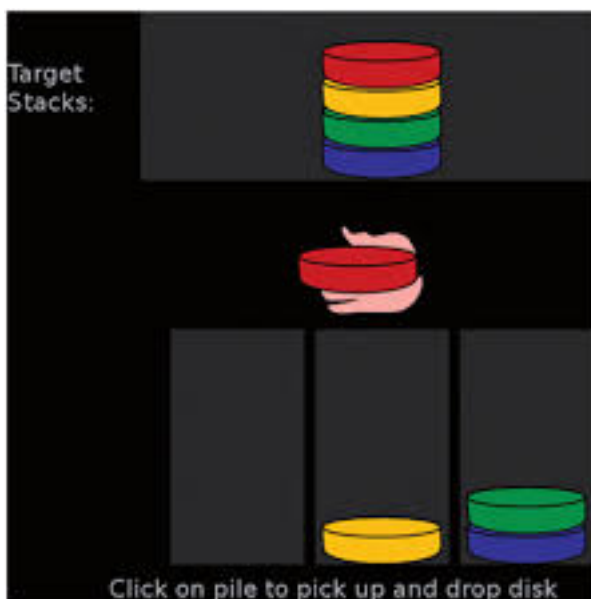


Figure 7. Illustration of the Tower of London (PEBL version, 0.12) with four disks.

Berg's Card Sorting Task (BCST). The BCST is a computerized version of the Wisconsin card sorting task (WCST; see Table 8) and is open-source, available at <http://pebl.sourceforge.net/>. It was used to assess the general executive abilities of abstract reasoning and shifting cognitive sets by utilizing environmental feedback. The standard BCST consists of a 128-card deck displayed on a computer screen. Each card contains a different combination of one of four shapes, colors, and quantities (Figure 4). Four key cards are displayed at the top of the screen as a guide to help determine which of the four stacks the deck's up-card is sorted to. The deck is revealed one card at a time,

and the visible card is matched to key cards depending on the particular rule (unknown to the participant) for a given set. After ten cards have been successfully matched, the set is completed and the sorting rule changes (also unknown to the participant). The new rule must be discovered using trial and error via on-screen feedback received after each card is sorted. After each card is sorted, the participant is provided with feedback regarding whether it was sorted correctly, according to the current rule. This process continues until the participant either sorts all 128 cards, or until the participant successfully completes nine categories, whichever comes first. The outcome measure was the number of categories completed.

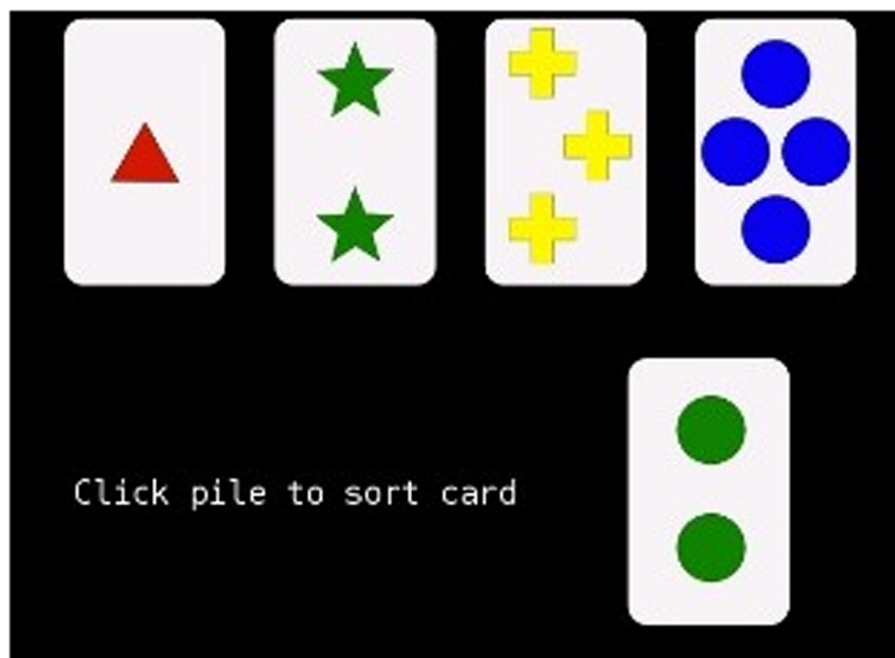


Figure 8. An illustration showing stimuli from the Berg Card Sorting Test (PEBL, version 0.12).

Statistical analyses

Analyses were conducted using SPSS 21.0 statistical analysis package. A series of one-way between subject analysis of variance (ANOVA) tests were used to analyze group differences for the converted scores for each of the computerized tasks. Pearson r

correlations were also used to assess the relationship between the individual EF components and, separately, the complex EF tasks; correlation analyses were conducted using the sample (i.e., without group separation). Bonferroni corrections were used for significant post-hoc group comparisons.

Results

Table 5 displays the means, standard deviations, and one-way ANOVA used to compare the effect of language status on the outcome measure for each of the individual and complex EF tasks in early bilinguals, late bilinguals, and monolinguals.

Table 6. *Table of Means, SDs, and ANOVAs Results for Three Groups*

Tasks	Early Bilingual		Late Bilingual		Monolingual		One-way between subjects ANOVA			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	Sig.	η_p^2
Go/No-go	0.49	0.35	0.34	0.28	0.37	0.30	3.240	2, 168	0.035	0.039
LG	0.41	0.87	0.45	0.39	0.46	0.30	0.119	2, 155	0.888	0.002
n-Back	0.68	0.22	0.63	0.28	0.74	0.20	2.792	2, 168	0.064	0.032
BCST	3.86	1.25	3.83	1.13	4.04	0.95	0.581	2, 162	0.561	0.007
IGT	5.26	13.91	3.87	14.76	6.94	13.42	0.610	2, 168	0.544	0.007
TOL	279.51	102.85	274.27	121.63	265.99	74.31	0.314	2, 151	0.731	0.004

Note. Means and standard deviations for the three language groups, including the one-way ANOVAs for each of the tasks used in the current study. Bolded values were significant at $p < 0.05$; LG = Local/Global task.

Inhibition

An ANOVA was significant. Post-hoc analyses using a Bonferroni test revealed a significant difference between early and late bilinguals, such that early bilinguals ($M = 0.49$, $SE = 0.35$) performed better than late bilinguals ($M = 0.34$, $SE = 0.28$). Post-hoc analyses revealed a significant difference between early bilinguals and monolinguals, such that early bilinguals ($M = 0.49$, $SE = 0.35$) performed better than monolinguals ($M = 0.37$, $SE = 0.30$). No difference was found between late bilinguals and monolinguals; together, this supported our hypothesis that early bilinguals would demonstrate better performance on the go/nogo task measuring inhibition compared to late bilinguals and

monolinguals. When speaking and comprehension proficiency values were included as covariates in a separate one-way between-subjects ANOVA to compare the effect of language status (Early versus Late bilingual) on inhibition, only a significant trend was found, $F(1, 89) = 3.128, p = 0.080, \eta_p^2 = 0.034$, suggesting the possible fortifying influence of proficiency on inhibition advantages.

Switching

An ANOVA did not reveal a significant effect of language status on switching; this did not support our hypothesis that early bilinguals would perform better on the local-global task compared to late bilinguals and monolinguals.

Updating

Although an ANOVA did not reveal a significant effect of language status on updating, a trend was found. Despite this trend, these results did not support our hypothesis that early bilinguals would perform better on the n-back task compared to late bilinguals and monolinguals.

Complex EF Tasks

A multivariate analysis of variance (MANOVA) was conducted to examine the effect of language status (three groups) on the complex EF tasks. A significant multivariate effect was not found, Wilk's Lambda = 0.980 $F(6, 294) = 0.486, p = 0.819, \eta_p^2 = 0.010$. These results supported our hypothesis that no differences would be found for more complex tasks of EF.

Analyses with two language groups

Individual EF Tasks; inhibition, switching, and updating. The bilingual language groups were collapsed into one larger group and comparisons between this

group and monolinguals were assessed using one-way between subjects ANOVAs. As is evidenced in Table 7 below, no significant differences were found between the bilinguals and monolinguals, except for on the n-back task, where the monolingual group ($M = 0.738$, $SE = 0.208$) performed better than the bilingual group ($M = 0.661$, $SE = 0.238$), suggesting a bilingual disadvantage for updating working memory processes, which did not support the hypothesis that bilinguals would perform better than monolinguals. Of particular note, significant differences on the go/nogo task were erased when the late and early bilinguals were combined, partially supporting our hypothesis that AoA has a positive impact on EF outcomes at the cognitive/information processing level of measurement.

Table 7. *Table of Means, SDs, and ANOVAs for Two Groups*

Tasks	Bilingual		Monolingual		One-way between subjects ANOVA			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	Sig.	η_p^2
Go/Nogo	0.434	0.328	0.373	0.305	1.510	1, 168	0.221	0.009
Local/Global	0.429	0.721	0.454	0.302	0.070	1, 155	0.791	0.000
n-Back	0.661	0.238	0.738	0.208	4.891	1, 168	0.028	0.028
BCST	4.099	0.897	3.828	1.221	2.468	1, 162	0.118	0.015
IGT	7.257	13.324	4.781	14.000	1.346	1, 168	0.248	0.008
TOL	278.299	108.616	264.349	74.232	0.803	1, 151	0.372	0.005

Note. Raw score means and standard deviations for the combined-bilinguals and monolinguals, including the one-way ANOVAs for each of the tasks used in the current study. Bolded values were significant at $p < 0.05$.

Complex EF Tasks. Another separate MANOVA was conducted to examine the effect of language status (2 groups: bilingual vs. monolingual) on the complex EF tasks.

A significant multivariate effect was not found, Wilk's Lambda = 0.970, $F(3, 172) =$

1.753, $p = 0.158$, $\eta_p^2 = 0.030$. This finding supported the hypothesis that no differences would be found between bilinguals and monolinguals for more complex EF tasks.

Correlations

Experimental groups were collapsed together to conduct Pearson r correlations between the individual EF component tasks (Table 8) to assess the interactions between the individual components; this revealed no significant correlations between tasks. The same correlation analysis was conducted for the complex EF tasks (Table 9), revealing significant correlations between tasks.

Table 8. *Correlation Table Between Individual EF Components*

		Go/nogo	n-Back	Local/Global
Go/nogo	Pearson r	1		
	Sig.	-		
	N	171		
n-Back	Pearson r	-0.112	1	
	Sig.	0.144	-	
	N	171	171	
Local/Global	Pearson r	0.084	-0.090	1
	Sig.	0.295	0.259	-
	N	158	158	171

Note. This table displays the Pearson r correlations between the individual EF component tasks, including the go/nogo (inhibition), n-back (updating), and local/global (switching) tasks. No significant correlations were found.

Table 9. *Correlation Table Between Complex EF Tasks*

		IGT	BCST	TOL
IGT	Pearson r	1		
	Sig.	-		
	N	171		
BCST	Pearson r	0.341	1	
	Sig.	0.000	-	
	N	164	165	
TOL	Pearson r	-0.204	-0.386	1
	Sig.	0.003	0.000	-
	N	153	154	154

Note. This table displays the Pearson r correlations for the complex EF tasks, revealing significant, moderate correlations between all of the tasks. Bolded values represent significant at $p < 0.01$.

Discussion

The hypothesis that early bilinguals would display better inhibition than late bilinguals and monolinguals was supported. The hypotheses that early bilinguals would display better switching and updating mechanisms than late bilinguals and monolinguals were not supported. The hypothesis that no differences would be revealed on the complex tasks of EF was also supported. The hypothesis that previously significant task differences would dissolve when the late and early bilingual groups were collapsed into one bilingual group was supported for inhibition. However, this hypothesis was not confirmed for updating, and in fact, the opposite finding was demonstrated. That is, a statistical trend was found when three language groups were used and a significant difference, suggesting a bilingual *disadvantage*, was found when only two language groups were used. Correlations between tasks found that the individual EF tasks are not correlated with each other, suggesting that they measure separable components of EF (i.e., *diversity* of EF), while the complex EF tasks were all correlated with each other, suggesting that these tasks involve a combination of EF components and are measuring the *unity* of EF. A more detailed discussion of the findings is outlined below.

Inhibition

Bilingual advantages for inhibition appear to be the most consistent for bilingual young adults (Bialystok et al., 2009; Bialystok et al., 2005; Bialystok, 2007, 2012). Most of this literature has typically used inhibition tasks that reflect the *conflict resolution* aspect of this component (e.g., the flanker, Stroop, and Simon tasks). The Simon and Stroop tasks both involve *prepotent* or overlearned responses (i.e., response inhibition), in addition to conflict resolution, thus, rendering them somewhat similar to the go/nogo

task used in the current study. Tasks that have investigated response inhibition have found a bilingual advantage using a modified antisaccade task (Bialystok, 2006) and an inhibition of return task (Colzato et al., 2008). Both of these tasks involve the suppression of a prepotent response, measuring response inhibition, similar to the go/nogo task.

Together with the current findings, it appears that response inhibition mechanisms are advantaged for bilinguals, and particularly, early bilinguals. The current study sheds some light on the nature of the mixed literature, such that the inclusion and exclusion of particular variables play a role in whether or not differences are found. For example, AoA, when included as a language group in analyses of the go/no go task, reveals significant differences for early bilinguals, but not late bilinguals, suggesting that learning a second language well before EF has fully developed may provide long-lasting advantages to inhibition mechanisms. There were two potential confounds to this finding: First, self-rated L2 speaking and understanding proficiency were significantly different between groups where early bilinguals were more proficient than late bilinguals. This finding may lead one to suggest that L2 proficiency is the underlying reason for the effect, especially given that significant differences between the early and late bilinguals were reduced to statistical *trends* when L2 proficiency was statistically controlled. However, the current study used a minimum self-rating of proficiency, such that individuals below five were excluded. This minimum cut-off provides some protection against proficiency differences because relatively high level of proficiency was assumed at baseline. Second, the idea that early bilinguals have at least seven additional years of bilingual experience may provide the advantage, rather than the process of early L2 acquisition, per se. Empirical evidence of this latter variable has yet to be investigated.

In a research study investigating the relationship of AoA on the bilingual advantage findings for inhibition, Pelham and Abrams (2014) found that, despite statistically different levels of proficiency, early and late bilinguals outperformed monolinguals on the lateralized attention network test (LANT), which measures the conflict resolution and response inhibition components of inhibition; the latter of these components is similar to the response inhibition require in the no/nogo task. Tao and colleagues (2011) found that early and late bilinguals also outperformed monolinguals, but that the difference between early bilinguals and monolinguals seemed to be qualitatively different from the difference between late bilinguals and monolinguals. On the one hand, early bilinguals showed a reduced conflict cost in response time measures (RT) but not in error rates, as well as an advantage in overall RT. The authors argue that conflict resolution and *monitoring* processes (reflected by overall RT advantages) are fortified in early bilinguals as a result of the early consolidation of two languages. On the other hand, there was an advantage for late bilinguals in conflict resolution both in terms of RT and error rates, without significant differences in overall RT, suggesting that late bilinguals do not experiencing the same fortification of monitoring processes as early bilinguals. Taken together with the current study's findings, early bilinguals more consistently demonstrate differences in inhibition compared to monolinguals; however, when it comes to late bilingual performance, findings related to inhibition continues to be mixed.

Switching

To our knowledge, only one study has been conducted to compare early and late bilingual to monolingual performance on a switching task. Kalia *et al.* (2014) used the

auditorily cued number-numeral task to measure several components of EF, including inhibition and switching ability. Consistent with the current study's results, they did not find significant switching advantages for either bilingual group, suggesting similar switching ability between bilinguals and monolinguals. Extending these findings with only two language groups, Moradzadeh, Blumenthal, and Wiseheart (2014) did not find any differences in switching ability. This is contrary to several other studies who have investigated switching advantages comparing bilinguals (without controlling for AoA) and monolinguals (i.e., Marzecová et al., 2013; Prior & Macwhinney, 2009; Soveri et al., 2010) and did find a bilingual advantage.

The current results suggest that bilingualism does not confer cognitive advantages under certain conditions, such as switching. Alternatively, and consistent with predictions of the multi-level approach for assessing the bilingual advantage, the current negative findings support the idea that empirical findings at the cognitive processing level may be mixed as a result of potentially confounding variables (e.g., music training, SES, etc.) that influence EF outcomes in addition to tasks at this level not being sensitive enough to reveal the bilingual advantage in young adults. In the current study, the sensitivity in measuring the bilingual advantage of the switch task was further complicated with the use of block-by-block switch cost, rather than trial-by-trial switch costs, the former of which tends to be a less sensitive measure. More research testing effects of potential confounding variables is needed to assess the impact these variables on switching advantages in bilingual young adults.

Updating

Using the n-back task, the current study is the first, to our knowledge, to compare updating processes, as proposed by Miyake and colleagues (2000; 2012), in bilingual and monolingual participants. In the current study, the findings did not support our hypothesis that early bilinguals would perform better than late bilinguals and monolinguals. Instead, we found a statistical trend between the three language groups (early, late, and monolinguals), with the trend favouring a monolingual advantage. The trend became significant when the bilingual groups were collapsed into one group and compared to monolinguals. This reason for this finding is not entirely clear; however, there are several possible explanations.

It is tempting to suggest that the apparent bilingual disadvantage found in the current study is due to the use of letters (i.e., verbal information) in the n-back task, because previous researchers have reported a bilingual disadvantage in verbal tasks (e.g., Bialystok, Craik, Green, & Gollan, 2009; Fernandes, Craik, Bialystok, & Kreuger, 2007). This disadvantage is proposed to stem from competing co-activated language representations and the processing cost associated with conflict resolution (e.g., Dijkstra & Van Heuven, 2002; Dijkstra et al., 1998; Green, 1998; Libben & Titone, 2009; Spivey & Marian, 1999). However, in the current study, the go/nogo task also used letters and results still revealed a bilingual advantage for inhibition; as well, other studies using executive tasks with verbal stimuli found advantages (e.g., Bialystok et al., 2008). Thus, this does not appear to be a likely explanation.

Another explanation is that the current results are an accurate representation for updating processes in bilinguals. Chatham *et al.* (2011) identified that at the cognitive level, this task is thought to involve numerous executive processes: (i) active

maintenance of the last n items; (ii) updating of new items so that they can be actively maintained; (iii) rapid binding of items to their serial order so that responses are based on the match between the current item and the n -back item and not between items matching at a non- n lag; and (iv) resolution of any proactive interference arising from non- n lag items. Consistent with the contention of numerous executive processes and assuming a ceiling effect for capacity with the updating system, the addition of monitoring language options may increase the demand on the EF system. This increased demand overloads its capacity, thus resulting in constant, diminished capacity, even for seemingly non language-related tasks (e.g., n -back task). This interpretation is supported by the current findings. Alternatively, there is no extra demand placed on the EF system because inhibitory processes have already suppressed the non-target language, and thus, updating mechanisms process and manipulate a single language, as they would for monolinguals. The current results do not support this contention. More research investigating updating processes in bilinguals and monolinguals is necessary before making any assertions, especially since the literature for more prevalent processes (e.g., inhibition) continues to be mixed, yielding an array of positive, neutral, and negative findings.

Complex EF tasks

The results of the complex EF tasks supported our hypothesis that no differences would exist between language groups. The finding supports the multi-level approach's assumption that, as tasks become more complex, the subtle bilingual advantage in young adults is more difficult to measure. The complex EF tasks used in the current study, despite being measurements at the *cognitive/information processing level*, are theoretically situated between this level and executive behaviors as a result of their

complexity. Any differences in EF provided by the bilingual experience are further diluted by the required complexity of these tasks. Similar to the EF literature for inhibition, switching, and updating, the research is mixed for complex EF tasks. For example, Kalia *et al.* (2014) used the auditorily cued number-numeral task, which is a deductive rule-use sorting task that requires participants to hold stimulus–response rules in working memory and suppress a prepotent response. This task is complex because it broadly assesses response inhibition, updating, and deductive reasoning mechanisms. In their study, all three language groups (early bilinguals, late bilinguals, and monolinguals) performed similarly. In another study, Kousaie *et al.* (2014) found a partial bilingual disadvantage using the Wisconsin Card Sorting Task (WCST), where francophone monolinguals outperformed English monolinguals and bilinguals. The WCST was similar to the BCST used in the current study and the same outcome measure was used in both tasks (i.e., total number of categories completed). The authors voice difficulty in interpreting this finding, but suggest it is related to the language environments of the language groups. This argument is not convincing, suggesting rather, that the finding may have been spurious, especially in light of the fact that differences between groups were within 0.5 points of a category. Francophone monolinguals were not included in the current study; thus, if that group is removed from the interpretation, Kousaie and colleagues results are similar to those in the current study. Despite these similarities and differences, more research is needed for complex EF tasks in the young adult bilingual advantage literature.

Limitations

One limitation is that early and late bilingual groups differed in level of self-rated proficiency. Although Tao, Marzecova, Taft, Asanowicz, and Wodniecka (2011) and Linck, Hoshino, and Kroll, (2008) studied samples that differed in L1 and L2 proficiency levels and differences in proficiency did not influence their results, other studies have found relevant differences. For example, in a study by Pelham and Abrams (2014), early and late bilinguals differed in their self-rated speaking proficiency. These authors argue that, although late bilinguals, as a sample, may have difficulty developing L2 proficiency compared to early bilinguals, late bilinguals may also be poor self-reporters of their L2 proficiency. In fact, these authors found that, of their initial 40 late bilingual participants who self-reported high proficiency in L2, interviews in that language revealed inadequate proficiency for 10 of those participants. Thus, differences in self-rated proficiency between early and late bilinguals may be an under-representation of the true differences in proficiency. Together with the current results, proficiency may account for the group differences, such that late bilinguals had similar results to monolinguals. Considering the paucity of studies that have compared early and late bilinguals, it is challenging to draw firm conclusions, but an inconsistent pattern in the findings underscores the need for additional research in the area.

There are many confounding variables that have been demonstrated to have an impact on EF, including SES, music training, translator status, and immigrant status. Characterizing a sample, especially young adult bilinguals, helps to explain the multiple factors that contribute to EF advantages. For example, it has been posited that a potential confound in bilingualism research is SES (e.g., Calvo & Bialystok, 2014), and that

controlling for differences on this variable can attenuate the bilingual advantage.

However, studies that have controlled for SES using child (e.g., Barac & Bialystok, 2012) and adult (e.g., Emmorey, Luk, Pyers, & Bialystok, 2008) populations show a bilingual advantage. Despite the evidence that suggests bilingual samples demonstrate EF advantages when SES is controlled, SES was not measured in the current study, which renders it as a possible confound. Similarly, other additional variables such as music training, translator status, and immigrant status, were also not considered in the current sample.

Another limitation of the current study was the use of global switch cost as the switch task outcome measure, rather than local switch cost. This measure subtracts the average RT for mixed blocks (i.e., blocks that contain both switch and nonswitch trials) from the average RT for nonmixed blocks (i.e., blocks that contain only trials for one of the tasks). Although this measurement has been used in other studies, it is a less sensitive measure of attention switching than local switch cost (i.e., the average RT for nonswitch trials across mixed blocks subtracted from the average RT for switch trials across mixed blocks) (Monsell, 2003).

Conclusions and Future research

Taken together, the current results of the specific and complex EF tasks raise questions about the reliability, robustness, and specificity of the purported bilingual advantage. There are several possible reasons why the data reported here do not support previous findings, all of which imply that the bilingual advantage may be less robust, or more task-specific than previously suggested. The bilingual advantage was only observed in one of the specific EF tasks (i.e., the go/nogo task) and not the other specific or

complex EF tasks, suggesting that previously observed bilingual EF advantages may be driven by a specific, rather than a global, aspect of EF. For example, a bilingual advantage may only arise when the task requires a certain degree of inhibition (e.g., response inhibition). Relating this supposition to Miyake and colleagues' (2012) conceptualization of EF where differences in inhibition were entirely explained by what is common to all EF tasks, the variability of the bilingual advantage may be further explained, in part, by an advantage typically seen in inhibition-specific tasks. That is, inhibition is suggested to be used in all EF, and thus, inhibition-specific advantages would benefit performance on other tasks. When a task is measuring another EF component (e.g., "updating-specific"), complexity is increased (i.e., involving inhibition and updating mechanisms), facilitating variability in reported bilingual advantage findings. This argument supports the complexity notion argued in the multi-level approach to measuring the bilingual advantage.

The findings of the current study also lend support to the multi-level approach. As well, the current findings support the idea that the bilingual advantage in young adults is mixed, perhaps as a result of tasks at the cognitive/information processing level of measurement not being sensitive enough to pick up in subtle EF advantages offered by bilingual status when young adults are at their peak EF performance. Further to this idea, the complex EF tasks used in the current study are theoretically situated between the measurement levels such that they are closer to executive behavior measurement and further from cognitive/information processing (i.e., more impure; task impurity problem) than the specific information processing tasks. In order to shed light on the potential explanations provided here, more research is needed, including empirical measurement at

a level of processing *lower* than cognitive processing (e.g., physiology) to test predictions that (i) if the bilingual advantage is circumscribed to a specific EF, then a bilingual advantage will be found only for inhibition; or (ii) if the bilingual advantage is diluted by cognitive processing task measurement, then a bilingual advantage will be demonstrated across electrophysiological measurements of the specific components.

Chapter 4

**Measuring the neurophysiological markers of executive function advantages in
bilingual young adults: Inhibition, shifting, and updating working memory**

Measuring the neurophysiological markers of executive function advantages in bilingual young adults: Inhibition, shifting, and updating working memory

Executive function and the bilingual advantage

Executive function (EF) has come to be an umbrella term used to describe various underlying cognitive processes, including working memory, attention, inhibition, planning, problem solving, self-monitoring, self-regulation, to name a few. Research suggests that a combination of interconnected networks within the prefrontal cortex, parietal cortex and subcortical structures facilitate the process of executive control over behavior (Fuster, 2000). Although there have been many definitions for the term EF (see Goldstein, Naglieri, Princiotta, & Otero, 2014), there appears to be a common thread: EFs are the combination of cognitive abilities that direct or *control* purposeful, goal-directed behavior. Many models of EF have been proposed; however, there is a growing consensus from factor analytic studies that three abilities constitute the core of EF, which include: (i) inhibition, which is resisting a strong inclination to do one thing in order to do what is most appropriate or needed (i.e., achieve a goal), such as completing a task despite wanting to go out and play, withholding a salient response, or saying what is polite rather than something socially inappropriate or hurtful; (ii) updating working memory, which is often referred to as holding information in mind, updating, and working with it, as in mentally manipulating objects or ideas, doing mental math calculations, or relating what you know to what you learned or read earlier; and (iii) switching, which is the ability to nimbly adjust to changed demands or priorities, or being able to disengage from one stimulus and reengage with another. Miyake and colleagues

(2000) reported the results of an influential confirmatory factor analysis of several EF measures in adults that pointed to these three core components: inhibition, updating (i.e., ‘working memory’), and shifting (i.e., attention shifting).

Researchers in bilingualism have hypothesized that bilingual language processing relies on components of the EF system for its management, especially during the initial stages of second language (L2) acquisition (e.g., Bialystok et al., 2009; Bialystok, 2011; Craik & Bialystok, 2006; Kroll & Bialystok, 2013; Stocco, Yamasaki, Natalenko, & Prat, 2012; van Heuven, Schriefers, Dijkstra, & Hagoort, 2008). The language networks continue to use EF components to manage the perception and production of two languages (Bialystok et al., 2009). During language processing involving a bilingual individual, it is well accepted that two responses are simultaneously activated in the bilingual brain (i.e., simultaneous activation) (e.g., Green, 1998). Simultaneous activation would cause a problem – or conflict - for the language production system. To produce the correct output, the language networks would be required to reduce this conflict (i.e., conflict resolution) and *inhibit* the nontarget language. Also during this conversation, the interlocutors must keep the current topic of speech in mind and formulate responses, which may be facilitated by working memory processes (Juffs & Harrington, 2011); as such, this process would be holding the recent speech in mind and updating response options - potentially *in both languages*. If one of the bilingual speakers was interrupted by a speaker of her other language, she would be required to *update* the language context (from L2 to L1), while inhibitory processes release inhibition of L2 in order to inhibit L1, all while attentional resources are *switched* from L1 to L2; this latter process would likely enhance the *updating* and *inhibitory* processes through the interactions of the EF system.

Given evidence to suggest that language networks recruit the EF networks to manage bilingual language perception and production, it has been posited that bilinguals would experience an advantage in these functions compared to their monolingual counterparts (e.g., Bialystok, 2011; Kroll & Bialystok, 2013). In fact, this advantage has been demonstrated for children (e.g., Bialystok et al., 2004; Bialystok, 2010; Martin-Rhee & Bialystok, 2008), young adults (e.g., Linck et al., 2008; Tao et al., 2011), and older adults (e.g., Bialystok et al., 2008). However, there is limited neurophysiological evidence to support this advantage.

Electroencephalography/Event-related potentials

Electrical brain activity can be recorded by placing electrodes on an individual's scalp. Event-related potentials (ERPs) are obtained by presenting the participant with: (i) stimuli; and/or (ii) a particular task to complete. During presentation or task performance, electrical potentials (brain waves) are recorded from the start of the stimulus (i.e., stimulus-locked ERPs) or a response onset (i.e., a button press; response-locked ERP) for X amount of time (e.g., 600 ms) after the stimulus or response onset. These potentials are then averaged over a large number of trials of the same type, yielding an ERP to that particular stimulus or response. The ERP is a sequence of positive- and negative-going deflections, and include several waveforms, such as the N100, P200, N200, and N400. These waveforms have been distinguished on the basis of their polarity (i.e., positive or negative), timing (i.e., latency in milliseconds) of the onset or the peak, their duration, and/or distribution across the scalp (i.e., at which positions on the scalp a waveform is smallest or largest). Waveforms are commonly called components, which are thought to

be a reflection of the neural mechanisms involved in certain functional (i.e., cognitive or perceptual) processes.

The involvement of EF in bilingual language processing has been established in theory by Bialystok and colleagues (2009). This theory predicts that learning a second language requires the adaptation of EF to assist in the processing of the two languages. However, the majority of studies have measured EF at a *cognitive/information processing* level. There are several major limitations to measuring performance at this level, which include: (i) a potentially interfering secondary task, needed in order to obtain data (i.e., button press); (ii) compensation strategies from other cognitive systems, such as using verbal mediation during a visual task, can be used to enhance outcomes (i.e., accuracy and response time); and (iii) response measures may not be sensitive enough to pick up on subtle differences between groups. Therefore, comparing groups on measures of EF, especially groups that are likely at their peak EF performance (such as young adults), may produce similar response outcomes (i.e., response times and accuracy). In contrast, a measurement level theoretically *closer* to the underlying neural processing/neural structures of the groups may reveal differences not present from measures of executive behavior or cognitive processing. The neurophysiological evidence for the requirement and use of three EF components in bilinguals is reviewed: inhibition, switching, and updating, before turning to a review of the differences in these control mechanisms between monolingual and bilingual individuals.

Several reviews of the neurophysiological evidence for EF components in bilingual language processing have been conducted (Hervais-Adelman, Moser-Mercer, & Golestani, 2011; Jansma, Rodriguez-Fornells, Moller, & Munte, 2004; Kaan, 2007;

Moreno, Rodríguez-Fornells, & Laine, 2008; Rodríguez-Fornells et al., 2006). All reviews have consistently corroborated the use of inhibition processes and switching mechanisms during bilingual language production and comprehension. Kaan (2007) also provides a summary review of the major ERP components relevant to research on speech perception (mismatch negativity), word and sentence comprehension (N400, left anterior negativity, P600), and word production (lateralized readiness potential, N200) (Kaan, 2007). Further, Kaan highlighted several advantages of using ERPs to study language processing in bilinguals. First, ERPs allow researchers to collect a continuous stream of data with a temporal accuracy of a few milliseconds (i.e., the sampling rate is typically between 250 and 512 Hz - samples per second - in language-related experiments), which matches the fast rate of language comprehension and production. Second, ERP responses obtained are multi-dimensional, which allows researcher to make qualitative inferences concerning the nature of the underlying neural processes being recorded. Third, when using ERPs, there may not be an interfering secondary task required to obtain data because neural responses to stimulus presentation are measured (i.e., stimulus-locked ERP). In typical behavioral studies, participants are asked to press a button, which induces additional processing requirements and may also induce particular processing strategies. Lastly, Kaan suggests that being able to record neural processing to visually- and aurally-produced stimuli, as well as language production, is an advantage of ERPs in the study of bilingual language processing.

Although a detailed review of the neurophysiological literature for *bilingual language processing* is beyond the scope of the current study and would be redundant given several published and comprehensive reviews, a brief review of the relevant

neurophysiological evidence for three EF components (i.e., updating, inhibition, and switching) is outlined below, in an attempt to consolidate some of the tasks and waveforms investigated during bilingual language production.

Inhibition. In a go/nogo task, participants are asked to make a response when a given stimulus-related condition is fulfilled (i.e., various letters are displayed on screen; go trials) and otherwise withhold their response (i.e., to a specific letter, such as ‘x’; nogo trials). Using ERPs, researchers can analyze the brain responses related to the inhibitory commands (i.e., nogo trials) and compare it directly with the non-inhibited or go trials. The difference waveform can be computed by subtracting the go condition from the nogo condition. There are only a few overt measures of behavioral response inhibition (e.g., percentage of commission errors; number of correctly inhibited responses); however, these do not directly measure inhibitory processes, unlike the neurophysiological response to nogo trials. No-go trials elicit an N200 component, which is a negative-going deflection in the ERP waveform observed approximately 200ms after stimulus onset. The exact role of the N200 is debated, such that Nieuwenhuis, Yeung, van den Wildenberg, and Ridderinkhof (2003) argue that the presence of the N200 reflects response inhibition, while Donkers and van Boxtel (2004) argue that it reflects conflict-monitoring; both thought to reflect aspects of inhibition. Further, it has been suggested that enhancement of N200 component reflects interference effects in bilingual tasks, and has been commonly observed (see below for brief review). However, more recent evidence (Huster, Westerhausen, Pantev, & Konrad, 2010; Smith, Johnstone, & Barry, 2008) suggests that the N200 may, in fact, reflect response selection, and that a later component, the P300, may reflect inhibitory cognitive components. Despite the debate, evidence from

studies on bilinguals is demonstrating a greater elicitation of the N200 during the need to inhibit information across languages (Rodriguez-Fornells, Rotte, & Heinze, 2002; Rodriguez-Fornells et al., 2005; Rodriguez-Fornells et al., 2006).

Kaan (2007) summarizes two studies that have investigated the role of the N200 in bilingual language processing. In an ERP experiment in which highly fluent, early Spanish-German bilinguals were asked to respond when the German name of a picture started with a vowel, but to withhold their response when it started with a consonant, or vice versa (Rodriguez-Fornells et al., 2005).

Some of the items created a conflict such that the name would start with a vowel in German, but with a consonant in Spanish. Compared to monolinguals, who would not have any familiarity with Spanish, bilinguals showed an enhanced N200 for both go and nogo items in the conflict condition, suggesting that they suffered from interference from their other language (Spanish) at the stage of phonological encoding, and that language selection has not been completed at this stage (Kaan, 2007, p. 585).

A similar experiment was carried out in which the go/nogo decision was based on gender information (Rodriguez-Fornells et al., 2006). For this study, the conflict items were feminine in one language and masculine in the other language. Again, compared to monolinguals, the bilinguals showed a larger N200 for both go and nogo conflict items, suggesting that the gender information was accessed in both languages.

Expanding on the review by Kaan (2007), Moreno and colleagues (2008) reviewed the same studies as Kaan (2007) (e.g., Rodriguez-fornells et al., 2005; Rodriguez-Fornells et al., 2006) that used the go/nogo paradigm to investigate the role of

inhibitory or conflict resolution processes during language perception¹⁷. Moreno and colleagues concluded that the reviewed studies showed an increase in a frontal negativity (i.e., difference wave between go and nogo trials) on conflict conditions, which might be related to the need for inhibition during simultaneous activation of the nontarget and target lexical candidates. Further, the reviewed studies revealed a long-lasting negative frontocentral component appearing 400ms after stimulus onset, which the authors suggested might reflect the amount of cognitive control (i.e., inhibitory control) required to process a specific task in one language. According to Green's inhibitory control model (Green, 1998), access to the non-dominant language representation involves greater suppression or inhibition of the dominant language, and therefore, enhanced control might be required to overcome the applied inhibition when naming in the dominant language.

Switching. Bilinguals frequently switch between languages in daily conversation depending on conversation context, target language goals, and occasional code switches¹⁸. This switch, whether in production or perception, often comes with a cost in accuracy or response time (e.g., Bialystok et al., 2009). Similar to the neurophysiological literature for inhibitory control in bilingual language processing, several reviews of the neurophysiological evidence for language switching have been conducted on bilingual individuals (e.g., Hervais-Adelman et al., 2011; Jansma et al., 2004; Moreno et al., 2008; Rodriguez-Fornells et al., 2006). Some studies have investigated the N400 component

¹⁷ It should be noted that the studies reviewed used a difference wave approach to measuring inhibitory control processes in this task.

¹⁸ Code switches are a particular form of language switching, whereby multilingual speakers electively employ words from alternative languages within utterances, while respecting the syntactic structure of the carrier language (Moreno et al., 2008)

(e.g., Alvarez, Holcomb, & Grainger, 2003), which is thought to be sensitive to semantic aspects of word processing, particularly to the cloze probability¹⁹ of a word as the N400 is greater in the case of unexpected words (e.g., Kaan, 2007); however, there is an ongoing controversy about the exact functional interpretation of the N400 in bilingual language switching (e.g., Moreno et al., 2008). Other studies investigated the N250 component (Van Der Meij, Cuetos, Carreiras, & Barber, 2011), which is thought to reflect the mapping of sublexical information (e.g., ordered letter combinations) onto whole-word orthographic representations (Holcomb & Grainger, 2006). In their study, Van Der Meij and colleagues (2011) recorded ERPs in low and high proficiency bilinguals while they read English sentences, half of which contained Spanish adjectives. Both groups of participants demonstrated similar N250 waveforms to the language code switch, suggesting similar language-specific orthographic processing regardless of proficiency. The N600 component, which is thought to reflect stimulus re-evaluation (e.g., Kuipers & Thierry, 2010), has also been investigated in language processing. Kuipers and Thierry (2010) suggest that bilinguals engaged in a process of reinterpreting the stimuli after a language switch, while monolinguals did not. In their study, Kuipers and Thierry used a variant of the oddball paradigm involving picture-word pairs with mixing of English (75% probability) and Welsh words (25% probability) and found the early detection of a language change in bilinguals triggered a significant P600 modulation by Welsh words, while a P600 was not elicited in monolinguals because the words were meaningless.

¹⁹ Cloze probability is defined as the probability of the target word completing that particular sentence frame

Moreno and colleagues (2008) reviewed ERP studies related to switches of language either in language perception (Alvarez et al., 2003; Jackson, Swainson, Mullin, Cunnington, & Jackson, 2004; Moreno, Federmeier, & Kutas, 2002; Proverbio, Leoni, & Zani, 2004; Rodriguez-Fornells et al., 2002) or production (Jackson et al., 2001; Rodriguez-fornells et al., 2005; Rodriguez-Fornells et al., 2006). Similar to other reviews (e.g., Hervais-Adelman et al., 2011), the authors concluded that the literature is mixed in terms of paradigms used, neurophysiological markers measured, and outcomes. Despite this, there is a consensus among bilingual researchers that the data suggest that language switching in production requires active inhibition of a nontarget language, and that the ERPs related to language switching and to withholding responses during non-linguistic go/nogo tasks are quite similar. For example, Jackson *et al.* (2001) used a visually cued naming task where bilinguals switched between naming digits in L1 or L2. These authors found an N200 elicited on switch trials similar to an N200 elicited on a go/nogo task, which the authors suggest is related to the inhibition required for a language switch. Further, the research evidence suggests there a bias in switch tasks where bilingual individuals often show enhancement in neurophysiological markers for L1 – L2 switches, but not L2 – L1 switches.

Recent studies have furthered the understanding of the neurophysiological functioning during language switches (e.g., Van Der Meij, Cuetos, Carreiras, & Barber, 2011; Verhoef, Roelofs, & Chwilla, 2010). Verhoefs and colleagues (2010) used a picture naming task where the switch cue preceded the picture, rather than changing the color of the picture, in order to assess endogenous language control in Dutch-English bilingual speakers. The authors argued that if evidence could be provided for an endogenous

language switch, bilinguals would hypothetically be able to bias the target language in advance. In their study, a cue–stimulus paradigm with an interval of 750 ms was used to measure cue-locked ERPs as an index of endogenous control. The authors investigated endogenous attention switches. Cue-locked ERPs showed an early posterior negativity (i.e., a negative-going deflection between 200–350 ms) for switch compared to repeat trials for L2 but not for L1, and a late anterior negativity (i.e., a positive-going deflection between 350–500 ms) for switch compared to repeat trials for both languages. The authors proposed that early switch–repeat effect (i.e., early posterior negativity between 200–320 ms) might reflect disengagement from the nontarget native language, whereas the late switch–repeat effect (i.e., late anterior negativity between 350–500 ms) reflects engaging in the target language. A second study investigated language code switches while late bilinguals with high (mean age of acquisition = 8.2 years) and low proficiency (mean age of acquisition = 8.7 years) read English sentences with a Spanish adjective in the middle of the sentence (Van Der Meij et al., 2011). These authors investigated the N250, N400, and LPC to the language code switch and found that highly proficient bilinguals demonstrated a N400 effect that extended to left anterior electrodes, which was followed by larger LPC amplitudes at posterior sites: P7 and P8. The authors concluded that proficiency modulates the different processes triggered by language switches. Taken together, the reviews and these recent findings suggest that language switching involves some differential mechanisms compared to monolinguals, and further, that some variables, such as proficiency and perhaps age of acquisition, may play a role in the underlying neural processing of language switches.

Updating. During a conversation, the interlocutors must keep the current topic of speech in mind and formulate responses, which may be facilitated by updating mechanisms (Juffs & Harrington, 2011); these mechanisms would be holding the recent speech in mind and updating response options. The need for updating mechanisms in language processing becomes increasingly more complex with the addition of a second language. Juffs and Harrington (2011) argue for the involvement of updating in bilingual language processing, extensively reviewing all behavioral evidence for the use of various memory processes in bilinguals. Using positron emission tomography (PET), Kim and colleagues (2002) investigated updating mechanisms, using different two-back tasks with visually presented stimuli (i.e., simple pictures, English words, and Korean words) in English-Korean bilinguals who were highly proficiency in L1 and somewhat proficient in L2. The experimental tasks required the subjects to keep in mind both the identity and the order of the presented objects and to continuously update the mental record with each subsequently presented stimulus. In light of no behavioural outcome differences, the PET results revealed differential activation for L1 (i.e., the anterior portion of the right dorsolateral prefrontal cortex and the left superior temporal gyrus) and L2 (i.e., the posterior portion of the right dorsolateral prefrontal cortex and the left inferior temporal gyrus); these activated brain regions were also different from the brain pattern activation to simple pictures. These results suggest that the right dorsolateral prefrontal cortex and left temporal lobe may be organized into two discrete, language-related functional systems. Further, this evidence supports the contentions by Juffs and Harrington (2011) that updating processes are integrated into language processing.

To our knowledge, there are no neurophysiological studies that have investigated the role of updating in bilingual language processing. The current study will be the first to investigate whether or not the neurophysiological response to updating processes are advantaged in bilinguals compared to their monolingual counterparts.

Summary. There is enough neurophysiological evidence of the neural markers of language control to suggest several similarities between bilingual language control and control of other EFs (Hervais-Adelman et al., 2011). Evidence suggests the inhibitory mechanisms in language switching are similar to those observed during inhibition of overt responses in nogo tasks (Moreno et al., 2008); the evidence also suggests there are other switching-specific mechanisms captured in the ERP waveforms (e.g., N400, N250, and P600). The former finding favors bilingual psycholinguistic models that incorporate some form of inhibition of nontarget language in bilingual word production (de Bot, 2004; Green, 1998). Neurophysiological evidence for inhibitory processes during bilingual language processing is becoming clearer, as indicated by N200 waveform findings. However, neurophysiological evidence for switching is less clear and the same line of neurophysiological evidence for updating is completely lacking.

Neurophysiological differences of EF in bilinguals and monolinguals

Although important in understanding the role of language processing in the bilingual brain, the neurophysiological studies of bilingual language processing do not address whether or not executive abilities are enhanced in bilinguals; rather, they suggest the need to use inhibition and switching mechanisms (and likely updating mechanisms, though the research evidence is lacking) to manage bilingual language processing. In a review of the bilingual brain, Bialystok and colleagues (2009) provide a detailed review

of bilingual executive control, including interference control and task switching; however, these authors do not include a review of studies investigating the neurophysiological responses of EF components in bilinguals. To our knowledge, only three published studies have been conducted that investigated the neurophysiological response of *inhibition* in bilingual and monolingual participants (Fernandez, Tartar, Padron, & Acosta, 2013; Fernandez, Acosta, Douglass, Doshi, & Tartar, 2014; Moreno, Wodniecka, Tays, Alain, & Bialystok, 2014). There is a paucity of studies investigating the neurophysiological responses of other EF component differences between monolinguals and bilinguals, such as task switching and working memory. This paucity of neurophysiological research extends to the child bilingual literature, where the majority of the research on the bilingual advantage *at the cognitive/information level* has been conducted (e.g., Bialystok, 2011).

Inhibition. Fernandez, Tartar, and Acosta (2013) conducted the first study, to our knowledge, to directly link this bilingual advantage to a neural correlate of inhibition. Fifteen monolingual and 13 bilingual young adult participants completed a simple non-verbal, auditory go/nogo task to elicit the N200 and P300 ERP components as neural markers of cognitive inhibition; the authors argued that both components have been linked to non-motor, cognitive inhibition in monolinguals. Results revealed significantly greater N200 amplitude at central electrode sites (i.e., greater inhibition) in bilinguals compared to monolinguals during nogo trials even though both groups performed the task equally well (i.e., accuracy of withholding a button press). As further evidence that inhibitory mechanisms were at play, neither of the go trials ERP components distinguished the groups. The results revealed that language proficiency, which was

measured using a standardized measures of proficiency in each language in bilinguals and just English in monolinguals (i.e., the Oral Vocabulary subtest of the Bilingual Verbal Ability Test), were positively correlated with N200 amplitude, suggesting that inhibition in bilinguals is moderated by second language proficiency. A second study by these authors (Fernandez et al., 2014), which included a visual go/nogo task, supported their auditory go/nogo finding, while finding no effect for the visual inhibition task.

Moreno and colleagues (2014) compared early adult bilinguals and musicians to monolinguals in order to demonstrate that the bilinguals and musicians exhibit behavioral advantages on specific EF tasks (e.g., inhibition-specific; e.g., go/nogo task) (Moreno et al., 2014). These authors used a visual go/nogo task to measure the behavioral and event-related potential (ERP) responses of these three groups. Results showed similar levels of behavioral responses (i.e., accuracy and response time) across groups, but differences in the N200 were found between the three groups. In particular, results revealed larger N200 amplitudes coupled with an increased late positivity wave relative to monolinguals, while musicians demonstrated smaller N200 amplitudes coupled with larger P200 amplitudes compared than monolinguals. Although they found differential ERP responses between musicians and bilinguals, the authors demonstrated that both groups were different than controls in the way in which their executive control system processes the information. Consistent with research suggesting that the N200 reflects a conflict detection signal (or inhibition of the prepotent response), the authors concluded that “the larger amplitude produced by bilinguals would suggest that they are more sensitive in detecting existing response competition or allocating resources to resolve conflict than controls” (p. 6).

These authors did not take into consideration age of acquisition for the bilingual or musician group.

Taken together, there is evidence to suggest that the go/nogo task and related N200 and P300 are task relevant components to measure inhibition in bilinguals. Modified versions of the go/nogo task have been used to capture the need for the inhibitory processes during bilingual language comprehension, and further, this task has been used to measure advantages (i.e., larger waveforms) in bilinguals compared to monolinguals.

Updating & Switching. To our knowledge there have been no studies that have investigated the neurophysiological correlates of the bilingual advantage for these EF components; the current study is the first to measure such components. The n-back task is commonly used to measure *updating* processes (e.g., Chen, Mitra, & Schlaghecken, 2008; Fan, Hsu, & Cheng, 2013; Nakao, Kodabashi, Yarita, Fujimoto, & Tamura, 2012; Ozen, Itier, Preston, & Fernandes, 2013) and it has been used to capture ERP components, including the P300, which is a positive-going deflection in the ERP component that peaks about around 300 ms after stimulus onset and is thought to measure the maintenance and updating of information in working memory.

To measure the neurophysiological components of switching, several types of tasks have been used, including a task-switch version of the Stroop task (Eppinger, Kray, Mecklinger, & John, 2007), a cued match/mismatch discrimination task (Kieffaber & Hetrick, 2005), and an odd/even number switch task (Friedman, Nessler, Johnson, Ritter, & Bersick, 2008); the latter of these tasks is the most similar to the Navon-figures task used in Chapter 3. Throughout all these tasks, versions of the P300, including the P3a and

P3b, were conceptualized to measure switching mechanisms, such that larger P300 waveforms reflect the mechanisms required to switch from one task to another. In particular, larger P3a amplitudes are believed to reflect activity associated with attention engagement, which is required for the neural systems to attend to the task and initiate the switch, while the P3b is thought to reflect context-updating operations and subsequent memory storage (Polich, 2007).

Statement of the problem

Behavioural evidence for the bilingual advantage for young adults is mixed, where some studies find advantages for bilinguals (e.g., Fernandez et al., 2013; Hernández et al., 2010; Kousaie & Phillips, 2012a; Tao et al., 2011), while others find no differences between groups (Bialystok, 2006a; Blumenfeld & Marian, 2013; Costa et al., 2009; Paap, Johnson, & Sawi, 2015)). In an attempt to help solve the problem of mixed evidence, authors have turned to the neurophysiological markers of EF to see if differences exist where overt behavioural evidence may not. This solution has yielded positive results, where the only three studies, to our knowledge, have found greater inhibitory control for young adult bilinguals, despite not finding differences in the traditional, overt behavioral outcomes. Following the multi-level approach to measuring the bilingual advantage, the bilingual advantage evidence should become clearer as levels of measurement closer to neural structures, such as neurophysiology, are used.

In order to further assess the hypothesis that the mixed nature of behavioural evidence of the bilingual advantage may be, in part, due to a lack of sensitivity from performance-based outcomes measures of EF, another outcome measure, d-prime (d'), was included. This measure is derived from signal detection theory and is thought to be a

more sensitive measure of accuracy as it accounts for a signal-to-noise ratio within the accuracy equation (McNicol, 2005). For example, in a go/nogo task, two outcome measures are possible: (i) typical accuracy measures, which are created by dividing the number correctly withheld responses to nogo trials (i.e., true negative) by total number of nogo trials; and (ii) d' measure, which is created by subtracting the z -score normed proportion of properly withheld nogo trials (i.e., true negative) from the z -score normed proportion of responding to nogo trials (i.e., miss). Therefore, as error rates increase, more 'noise', and thus less 'signal detection' is included within the equation, yielding a smaller d' value. To our knowledge, this measure has not been used in the bilingual advantage literature, but has been used extensively in the cognitive psychology literature as a measure with greater sensitivity to dysfunction (e.g., Aston-Jones & Cohen, 2005; Haatveit et al., 2010; Schmidt & Vorberg, 2006). Given its sensitivity, it was included in addition to typical EF performance-based outcome measures, such as accuracy and response time. Following the multi-level approach posited in Chapter 1, this measure may provide a better proxy for executive advantages in bilinguals.

Further, several important variables have been identified as confounding variables when investigating the bilingual advantage and thus, were included in the present study. These include age of acquisition (AoA), socioeconomic status (SES), and music training. AoA is the age at which a bilingual learned his/her L2 (e.g., Hull & Vaid, 2007; Luk et al., 2011; Tao et al., 2011), and is operationally defined as early bilinguals who learned L2 at or before the age of 6 and late bilinguals who learned L2 at or after the after of 13., the findings for AoA are mixed: at times, early bilinguals demonstrate an advantage over late bilinguals and monolinguals (e.g., e.g., Luk et al., 2011), whereas some studies have

found an advantage for early and late bilinguals (e.g., Pelham & Abrams, 2014; Tao et al., 2011), while other studies have found no advantages for early and late bilinguals (e.g., Kalia, Wilbourn, & Ghio, 2014).

The present study uses the three-factor model of EF posited by Miyake et al. (2000), monolingual, early bilingual and late bilingual participants engaged in three computerized tasks with the goal of measuring neurophysiological markers of inhibition, switching, and updating (working memory). To measure inhibition, we used a traditional go/nogo task where participants responded with a button press to letter stimuli, except for the letter 'j.' To measure switching, we used a task switch paradigm where participants switched from responding to numbers using two different rules: (i) greater or less than 5; and (ii) odd or even. And finally, to measure updating (working memory), we used a traditional n-back task where participants had to indicate when a letter matches a letter two or three letters back in a series.

Hypotheses

Given the multidimensional approach we are taking to measure neurophysiological markers of EF to assess the bilingual advantage, we have several hypotheses, as follows.

Inhibition hypotheses. The bilingual advantage appears to be the most well-established for inhibition (e.g., Bialystok et al., 2005; Fernandez et al., 2013; Moreno et al., 2014) , and, may represent that bilingual language networks must constantly inhibit the nontarget language during language processing. The research evidence supports the idea that both languages are simultaneously activated during language production. Therefore, for inhibition, we hypothesized that early bilinguals would demonstrate an

advantage in the neurophysiological marker (i.e., larger N200 amplitude on nogo trials) compared to their late bilingual and monolingual counterparts. No difference would be found for latency. Further, we predicted that early bilinguals would demonstrate an advantage over late bilingual and monolinguals on the more sensitive behavioral outcomes measures of inhibition (i.e., d-prime) for the same task; no differences would be found between groups for percentage of commission errors on nogo trials (i.e., accuracy), as it is a less sensitive measure of inhibition. Given these behavioral differences, larger N200 peak amplitudes would reflect better inhibitory control.

Switching hypotheses. For task switching, the research literature using behavioral outcome measures (e.g., accuracy, response time, and switch cost) is mixed, which may represent that language switching is more variable, compared to inhibitory control. Despite these findings, we hypothesized that early bilinguals would demonstrate greater switching ability reflected by a larger P3a amplitude compared to late bilingual and monolingual participants, which would suggest that early bilinguals demonstrate more activity associated with engagement of attention (Polich, 2007). Further, early bilinguals would also show an advantage in later switching processing, reflected by larger P3b amplitudes compared to the other two language groups. The P3b is thought to reflect context-updating operations and subsequent memory storage (Polich, 2007), which would be in line with what bilinguals would need to do during a language switch: update the current language context and store that information to avoid unwanted slips. We predicted no differences in the behavioral outcome measures for attention switching (i.e., accuracy and response time switch costs) in this sample of young adults.

Updating hypotheses. Similar to the process of updating following a switch, updating (working memory) is likely an executive ability that is used during bilingual language production such that the interlocutors must keep the current topic of speech in mind and formulate responses; as such, this process would be holding the recent speech in mind and updating response options. In addition, a bilingual speaker would also be holding and updating contextual information to ensure the correct language is produced. Following this logic, we hypothesized that the neural markers of updating working memory (i.e., peak P300 amplitude) would be larger for early bilinguals than late bilinguals and monolinguals in the high demand situations (i.e., three-back). However, in the low demand situation (i.e., two-back), no differences between groups would be found. Similarly to the other two tasks, we predict no differences in the low sensitivity behavioral outcome measures (i.e., accuracy of two- and three-back) between groups, whereas we predicted early bilinguals would demonstrate better performance in the high demand condition when d' is used; no differences would be found for d' in the low demand condition.

Method

Questionnaires²⁰

Demographics. Demographics details were collected to allow any differences between groups to be identified: age, biological sex, socioeconomic status (SES), musical training history, physical activity, and whether or not they immigrated to Canada. Participants also recorded their history of brain injury, history of neurological conditions, prescription medications being taken, and if any previous diagnoses had been given to the participant (i.e., ADHD, FASD, Autism, Developmental Disorder, Learning Disability, Speech/language difficulties, and any other diagnosed psychological conditions); participants were excluded if they endorsed any of these demographic characteristics, with the exception of prescription birth control, because of their potential influence on neurophysiological outcomes.

Language questionnaire. A language questionnaire was constructed based on an empirical review of bilingual language questionnaires from Li, Sepanski, and Zhao (2006), and was used to obtain participant information in order to classify the bilinguals as either early (i.e., learned second language at or before age of 6) or late (i.e., learned second language at or after age 13). Bilingual participants provided information relating to language experience, so that effects of individual differences in factors such as proficiency and usage could be examined. Participants were asked to record their perceived level of fluency in reading, writing, and understanding in each of the languages that they speak²¹ using a 5-point scale (1= poor, 3 = functional, and 5 = excellent/fluent).

²⁰ See appendix C for the questionnaire package that was given to participants to complete prior to the computerized experiment.

²¹ A cutoff proficiency score of 4 in speaking and understanding for both languages was used to ensure highly proficiency bilinguals in both groups.

Self-ratings were also provided for the amount of daily usage of each language (expressed in percentages) and whether L2 was used daily, weekly, monthly, or rarely; the frequency of daily L2 use was expressed in percentages for group comparison purposes. Information on translator status was collected for group comparison purposes (i.e., whether or not the individual worked as a professional translator). Bilingual translators have demonstrated differential outcomes on EF measures (Dong & Xie, 2014). For each language, participants were also asked to record other characteristics, such as whether or not learning occurred formally (i.e., lessons) or informally (i.e., at home or at work), and duration of learning circumstance.

Participants

For the current study, a total of 76 participants were included in three groups: (i) monolinguals ($n = 32$); (ii) early bilinguals ($n = 23$); (iii) late bilinguals ($n = 21$). However, two participants from the late bilingual group did not complete testing, one as a result of stating they had an allergy to the conductive gel used during administration, and one for stating they learned their language between the ages of 7-11. All participants had normal or corrected-to-normal vision and no a history of head injury, Attention Deficit/Hyperactivity Disorder (ADHD), Fetal Alcohol Spectrum Disorder (FASD), Autism Spectrum Disorder (ASD), developmental disorder, or psychiatric disorder. Participants were recruited through an online undergraduate psychology recruitment source (SONA) at the University of Victoria, and through advertisements. For those recruited through SONA, participants were given course credit for participation in the study. For those recruited through advertisements, participants were given \$20 CAN for their participation. The Human Research Ethics Board of the University of Victoria

approved the study and participants provided written informed consent prior to participation. Table 10 presents the sociodemographic characteristics for the three groups and the language characteristics of the two bilingual groups. Table 11 presents the percentages of the L1 and L2 languages spoken by participants in the early and late bilingual groups.

Table 10. *Participant Demographics*

Demographics	Group		
	Early Bilinguals Mean(SD)	Late Bilinguals Mean(SD)	Monolinguals Mean(SD)
N	23	19	32
Age	20.35 (1.75)	20.29 (1.55)	20.28 (1.85)
Sex (% female)	78.3%	61.9%	68.8%
AoA of L2	3.13(2.87)	13.26(2.28)	-
SES	3.23(0.69)	3.05(0.87)	3.34(0.79)
% Music training ²²	39.1%	47.6%	31.3%
% Physical activity	69.6%	61.9%	84.4%
% Translator	4.5%	31.6%	0%
% Immigrated	69.6%	89.5%	18.8%
L1 proportion of use	66.74(19.11)	69.87(20.59)	-
L2 proportion of use	28.39(15.78)	31.24(19.79)	-
L2 proportion of SS who use daily	59.1%	81%	-
Proficiency L2 - speak	4.22(0.74)	3.91(0.70)	-
Proficiency L2 - comprehend	4.65(0.57)	4.05(0.78)	-
Proficiency L2 - write	3.26(1.29)	3.19(0.93)	-
Proficiency L2 - read	3.70(1.36)	3.71(0.78)	-

Note. The demographic information is displayed for the total sample of participants in the current study, including: number of participants, mean age, sex distribution, socioeconomic status, age of acquisition of L2, and self-rated proficiency for each of the three groups. AoA = age of acquisition; L2 = second language; SES = Socioeconomic status; - represents data that was not measured in monolinguals.

²² Independent samples t-tests were conducted between those with and without music training within each of the three empirical groups using all of the outcome measures. All t-tests were not significant, with the exception of t-test comparing late bilinguals with and without music training with *d'* as the outcome measure (on 3-back trials), $t(40) = -2.202, p = 0.042$.

Table 11. *Dominant and Nondominant Language Spoken by Participants*

Languages	Group			
	Early Bilinguals (%)		Late Bilinguals (%)	
	L1	L2	L1	L2
English	60.87	39.13	26.32	78.95
Chinese ^A	30.43	17.39	43.48	10.53
Tagalog	4.35	-	-	-
Farsi	-	4.35	5.26	-
Spanish	-	-	5.26	-
Korean	-	-	5.26	-
Russian	-	-	5.26	-
French	-	4.35	-	10.53
Arabic	4.35	4.35	-	-
German	-	4.35	-	-
Polish	-	4.35	-	-
Punjabi	-	4.35	-	-
Fillipino	-	4.35	-	-

Note. This table displays the proportion of L1 and L2 languages spoken by each bilingual group. ^A represents mandarin and Cantonese

Stimuli & Procedure

Participants were provided with both written and verbal instructions that explain the procedure and emphasize that they should maintain correct posture and minimize head movement and eye blinks during the EEG/ERP tasks. Following this, participants were asked to complete the demographics/language history questionnaire, the Raven's progressive matrices task, and the letter fluency task while they are seated comfortably in a room under dim lighting, following which the electroencephalography (EEG) set up took place. Once EEG set up was complete, participants completed the three individual computerized tasks while the EEG was being measured; each of these tasks measured a specific component of EF. The three computerized EF ERP tasks were counterbalanced across participants. Participants viewed all stimuli from a distance of about 70 cm from the computer monitor; a 70 cm string was used to measure the distance between their forehead and the computer monitor. Stimuli were displayed on a 17-inch LCD computer

monitor using MatLab presentation software (MATLAB 11.0, The MathWorks, Inc., Natick, Massachusetts, US).

Control task: Verbal fluency. Following a specific rule, participants were asked to overtly generate as many English words as possible within 60 s (e.g., Heaton, Miller, Taylor, & Grant, 2004; Portocarrero, Burright, & Donovanick, 2007; Tombaugh, Kozak, & Rees, 1999). Letter fluency was assessed on the first three trials: participants were first asked to produce words that start with the letter F, the A, and then the letter S. On the last trial, category fluency was assessed where participants were asked to produce animal names. Participants were asked to avoid producing names of people or places as well as the same word with different endings. Four outcome measures were obtained: (i) a total score for letter fluency, which consisted of the total correct items for all three letters; (ii) a total score for category fluency, which consisted of the total correct items for this trial; (iii) a total error score for letter fluency, which consisted of the total number of errors made across the three letter trials; and (iv) a total error score for the category trial.

Control task: Raven's advanced progressive matrices. This task was used to estimate participants' nonverbal intelligence and ensure differences in EF neural markers are not related to differences in nonverbal intelligence, we used a version of the modified Raven's progressive matrices, which has been frequently used in the bilingual advantage literature (e.g., Bialystok et al., 2004; Bialystok & Martin, 2004; Colzato et al., 2008; Costa et al., 2008; Engel de Abreu et al., 2012; Hernández et al., 2010; Tao et al., 2011). This computerized task is a modified version of Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998). In order to solve novel problems, this task requires participants to organize and sequence novel information, by deducing rules and applying

them (see Carpenter, Just, & Shell, 1990). Participants were presented with a series of visual puzzles; each displayed in the form of a 3X3 matrix of geometric figures and patterns. One segment of a larger pattern was always missing and participants were required to deduce the pattern of the missing piece, given the properties of the available matrix. Participants selected the piece that best completed the matrix from a set of multiple-choice options. Participants were asked to solve eight different matrices according to standard procedures; one block of eight trials. To select a response, participants pressed numbers 1-8, which corresponded to one of eight options. The matrix stayed on screen until the participant made a response. Immediately following their response, the next matrix appeared on screen. The outcome score was the total number (maximum score = 8) of correct items.

Experimental task one: Measuring inhibition. Participants completed a traditional go/nogo task (Donders, 1969) to measure inhibition (see Figure 9). The task consisted of two blocks; in the first block, which had 50 trials, participants were asked to respond as quickly as possible to any letter appearing in the center of the computer screen by pressing the 'L' key with their dominant hand. All trials in this block were "go" trials, which will be used to establish baseline reaction times. In the second block, which consisted of 150 trials, participants were asked to respond by pressing the 'L' key as quickly as possible when they see a letter, except when the letter 'J' (the target stimulus) appears (i.e., the "nogo" trials in the mixed condition). Thirty percent of the trials in this condition were "nogo" trials (i.e., 45 trials). Letters appeared on screen for 1000 ms; Inter-trial interval time was variable at approximately 1400 ms post-stimulus offset. Task completion time was approximately six minutes. Participants were not provided with

feedback on their performance. Three scores were derived as outcome measures: (i) serial reaction time (SRT), which was the average response time (RT) for all the ‘go’ trials in the first block (i.e., 50 trials); (ii) accuracy for ‘nogo’ trials (i.e., number of correctly withheld responses divided by 45 trials); and (iii) d' , which was calculated by subtracting the z-score normed proportion of true negative trials from the z-score normed proportion of miss trials²³.

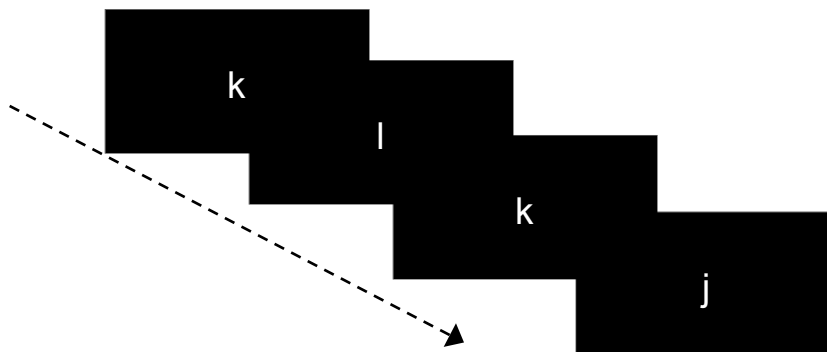


Figure 9. Schematic example of the computerized go/nogo inhibitory task. The arrow indicates the progression of time. The first three trials are examples of “go” trials, and the last trial is an example of a “nogo” trial.

Experimental task two: Measuring attention switching. To examine shifting ability, participants completed a More-Than/Less-Than Task (i.e., deciding whether a digit is more or less than 5) and an Odd/Even Task (i.e., deciding whether a digit is odd or even). This task was modeled after a similar task by Friedman, Nessler, Johnson, Ritter, and Bersick (2008). Both tasks, which occur within the same block, consisted of multiple trials involving a single white digit (1 – 9, excluding 0 and 5) displayed centrally on a black background. A new digit appeared for each successive trial, with the digit displayed

²³ This formula was rendered from Signal Detection Theory and is thought to be a more sensitive measure of accuracy as it accounts for error trials. Correctly withheld responses to ‘No-go’ trials (i.e., target absent) were coded as ‘True Negative’ (TN) trials, while withheld responses to ‘go’ trials (i.e., target present) were coded as Miss trials. As such, the formula was: $d' = z(\text{accuracy of TN}) - z(\text{accuracy of Miss})$. For probability values that equalled 0 or 1, values were corrected so that: (a) if the number is 1, it was replaced with $1 - (1/[2*45])$ and (b) if number is 0, it was replaced with $1/[2*105]$, where 45 = total possible number of TN trials and 105 = total possible number of miss trials.

on the screen until the participant provided a response; inter-stimulus duration was 500 ms between stimulus offset and stimulus onset. Depending on the specific task during that trial, the words “more less” or “odd even” were displayed below the digit to prompt the participants regarding task instructions. For a switch trial, two changes from the previous trial occurred: (i) the digit was surrounded by a rectangle (or the rectangle was removed), which indicated the type of task to be completed; and (ii) the reminder prompts changed to those associated with the new task set. Together, these two tasks comprised the shifting condition, which consisted of a total of four mixed blocks. All together, there were six blocks. The first two blocks consisted of the same task for 60 trials (i.e., either all *More/Less* or all *Odd/Even*). The third to sixth blocks included 120 trials with 24 task-switch trials in each block for a total of 480 trials (384 nonswitch and 96 switch), requiring participants to switch from one task to another when a white rectangle surrounds or is removed from the displayed number; participants completed the same task while the rectangle was present (e.g., more/less than) or not present (e.g., odd/even). Switch trials occurred every five trials, four nonswitch trials between each switch. This task took approximately 15 minutes to complete. Two outcome measures were derived: Switch cost scores, which was the difference between the pre-switch trial and the switch trial, were calculated for RT and accuracy; these scores were calculated using the average RT and accuracy from blocks four through six. Following task completion, participants were asked if they recognized a pattern in switch trials.

Experimental task three: Measuring updating working memory. Participants completed a letter variant of a computerized n-back to measure how well participants can update, hold, and discard information. This task was modeled after a similar n-back task

(Ozen et al., 2013). Participants were shown a series of centered white-colored letters in 100-point font, one at a time, on a black background computer screen. At the beginning of the experiment, participants were instructed to indicate when the current letter matched the letter shown either two (easy; two-back) or three (difficult; three-back) trials prior by pressing the ‘A’ key, thus keeping up to three trials (i.e., letters) in working memory; participants pressed the ‘L’ key for all other trials. Participants were shown block-specific instructions on screen before each block; the 2-back and 3-back tasks were presented separately in blocks. Two of the blocks were the 2-back task while the other two blocks were the 3-back task. Response options were counterbalanced. Trials without a button press were rejected from further analysis, as they would not have been coded into one of the four categories for response (i.e., Hit, miss, false positive, and true negative). There were four blocks of 75 trials each, for a total of 300 trials; there was a 30% probability of match for target trials, resulting in approximately 25 trials for each block. Each trial started with the onset of a letter, which stayed on screen for 750 ms. Between each trial, an inter-stimulus period (black screen) was shown for 1000 ms. The total task length was approximately 15 minutes, including time for self-paced breaks between blocks.

Five outcome measures were derived: (i) the accuracy of hits for the combined 2-back blocks; (ii) the accuracy of hits for the combined 3-back blocks; (iii) a d' measure²⁴ was calculated for the combined 2-back blocks; (iv) a d' measure was calculated for the

²⁴ This formula is similar to that used in the go/nogo task, except that correct identification of target trials were coded as ‘Hits’ and incorrect target trials (i.e., incorrectly indicating that the current letter was the same as 2 or 3 letters ago) were coded as ‘False positives (FP).’ As such, the formula was: $d' = z(\text{accuracy of Hits}) - z(\text{accuracy of FP})$. For probability values that equalled 0 or 1, values were corrected so that: (a) if the number is 1, it was replaced with $1 - (1/[2 * 100])$ and (b) if number is 0, it was replaced with $1/[2 * 50]$, 100 = total number of FP and 50 = total number of hits.

combined 3-back blocks; and finally, (v) the load effect was calculated on accuracy scores by subtracting 3-back accuracy from 2-back accuracy. Since participants were instructed to respond to all trials with a button press, trials without responses were not coded, and thus, not included in the accuracy analysis (allowing for the differentiation between true negatives and unintentional misses/nonresponses and reducing bias by the ambiguity of non-response trials; Wilks, 2011).

Data acquisition & ERP analyses

Table 12. *ERP Rejection and Correction Table*

	Groups								
	Early Bilinguals			Late Bilinguals			Monolinguals		
	% Trials removed	No. Hjorth	No. Excluded	% Trials removed	No. Hjorth	No. Excluded	% Trials removed	No. Hjorth	No. Excluded
GNG	4.51 (9.27)	3	2	1.76 (3.24)	1	1	2.61 (5.76)	2	2
Switch	6.50 (11.19)	7	3	6.82 (11.76)	2	3	2.98 (3.95)	2	1
n-Back ²⁵	7.82 (12.52)	7	5 / 13*	8.00 (18.29)	3	6 / 6*	4.84 (13.54)	2	3 / 11*

Note. Displays the mean and SD (in brackets) for the percent of trials removed using the rejection criteria for each group in each of the three tasks. The percent of participants excluded as a result of missing more than 10% of their data, as well as the number of participants where the Hjorth's Nearest Neighbor algorithm was used, are also included. GNG = go/nogo task. *Represents the number of participants excluded in the 3-back; this number is larger because the excluded participants had less than 10 trials per averaged waveform.

The EEG was recorded using a montage of 37 electrode sites in accordance to the extended international 10-20 system (Jasper, 1958). Signals were acquired using Ag/AgCl ring electrodes mounted in a nylon electrode cap with an abrasive, conductive gel (EASYCAP GmbH, Herrsching-Breitbrunn, Germany), which was used to decrease the impedance from the scalp to the electrode. Signals were amplified by low-noise electrode differential amplifiers with a frequency response of DC 0.017-67.5 Hz (90 dB-

²⁵ For the n-back task, participants' accuracy scores for the 3-back blocks were low (see table 4), resulting in a small number of useable ERP trials for the averaging procedure. As such, participants with less than 10 trials per average waveform were excluded.

octave roll off) and digitized at a rate of 250 samples per second. Digitized signals were recorded to disk using BrainVision Recorder software (Brain Products GmbH, Munich, Germany). Inter-electrode impedances were maintained below 20 k Ω . Two electrodes were placed on the left and right mastoids, which were used to create (offline) the average reference of the two mastoids for all the EEG recordings. Postprocessing and data visualization was performed using BrainVision Analyzer software (Brain Products GmbH, Munich, Germany). The electro-oculogram (EOG; eye movements) were recorded for the purpose of artifact correction; horizontal EOG was recorded from the outer canthi of both eyes, and vertical EOG was recorded from the suborbit of the right eye and electrode channel Fp2. Artifact rejection criteria for all tasks were as follows: The gradient criterion was set such that voltage steps of maximally 35 μ V were allowed per sampling point; the absolute difference voltage per 200 ms segment should not exceed 150 μ V; the lowest allowed activity was 0.1 μ V(max–min) per 200 msec; and amplitudes had to be between –150 and 150 μ V. See Table 3 for the mean percent of rejected trials and the associated SD, as well as the percent of participants that were excluded as a result of missing more than 10% of the trials in a given task. The digitized signals were filtered using a fourth-order digital Butterworth infinite impulse response (IIR) filter with a bandpass of different frequencies for each of the target waveforms (see below). Hjorth's nearest neighbor, which is an algorithm that averages surrounding electrodes to estimate the electric potential at a given electrode, was also used to correct malfunctioning electrodes; the number of participants in which this procedure was used is shown in Table 12.

Inhibition. To investigate inhibitory processes, the N200, which is a negative-going waveform that occurs around 200 ms post stimulus presentation, and the P300, which is a positive-going waveform that occurs approximately 300 ms post stimulus onset, were measured. The peak amplitude of this waveform is typically more negative for nogo trials than go trials. In order to reduce high and low frequency noise, the time-averaged ERPs were filtered using a fourth-order digital Butterworth IIR filter with a passband of 0.1 to 20 Hz, 24 dB-octave roll off, and a 60 Hz notch filter. Baseline waveforms were established 200 ms prior to stimulus onset and ERP epochs extended 600 ms post stimulus onset. Two separate ERPs from the go/nogo task were sorted and averaged together for each participant: (i) averaged ERPs for correct ‘go’ stimuli from the second block, and (ii) averaged ERPs for correct ‘nogo’ stimuli (which were only present during the second block). Thus, peak amplitudes were calculated for each participant by finding the most negative “peak” in the 200-350 ms time window for the N200. The P300 peak amplitudes were calculated by finding the most positive peak between 350 ms – 500 ms. Only correct trials were included in the averaged waveforms. The latencies for these peak measures were also calculated for each participant. These measures were calculated at central midline sites (i.e., Fz, FCz, Cz, CPz), where these components are typically maximal (e.g., Moreno et al., 2014).

Attention switching. To examine shifting ability, the P300 at frontal-central sites (i.e., P3a) and the P300 at posterior central sites (i.e., P3b) were measured. The P3a and P3b nomenclature has been used to describe the two separable and distinct P300 ERP waveforms (Polich, 2007), and thus, is used to assist in the differentiation between the P300 in other tasks (inhibition and updating). The underlying interpretation of these

waveforms is task dependent, and thus, represent different cognitive mechanisms (Luck, 2014). The P300 is a positive-going deflection in the ERP waveform that occurs approximately 300-600 ms following stimulus onset; thus, the P3a and the P3b are both positive-going deflections in the ERP waveform, but have different time courses. The P3a peak amplitude was measured in a time window of 300-450 ms at frontal central sites (i.e., Fz and FCz), and the P3b peak amplitude was measured in a time window of 450-600 ms at posterior central sites (i.e., Cz, CPz, and Pz). The time-averaged ERPs were filtered using a fourth-order digital Butterworth IIR filter with a passband of 0.1 to 30 Hz, 24 dB-octave roll off, and a 60 Hz notch filter. The stimulus-locked data were epoched offline with 200 ms pre- and 800 ms post-stimulus periods. To simplify the analyses and capture the major effects, the behavioral and ERP data from two critical trial types were subjected to statistical analyses: (i) the pre-switch trial and (ii) the switch trial; only the correct trials from blocks four – six were used for waveform averages.

Updating. To investigate updating working memory processes, the P300 waveform, which is a positive-going deflection in the ERP waveform that peaks approximately 300 ms following stimulus onset (i.e., the onset of the letter on the computer screen), was measured. The digitized signals were filtered using a fourth-order digital Butterworth IIR filter with a passband of 0.1-30 Hz, 24 dB-octave roll off, and a 60 Hz notch filter. A 1000ms epoch of data extending from 200 ms prior to 800 ms following the onset of each stimulus were extracted from the continuous data file for analysis. Only correct trials were analyzed. Stimulus-onset trials were segmented and averaged by stimulus type (target and nontarget) and level of difficulty (two- and three-back), thus resulting in four averaged waveforms for each participant in the three groups.

Peak amplitudes were calculated for each P300 average waveform for each participant by finding the most positive peak amplitude in the 300-500 ms time window for the P300 component. The latencies for these peak measures were also calculated for each participant. These measures were calculated at posterior-central midline sites (i.e., Cz, CPz and Pz).

Results

Analyses were conducted using SPSS 21.0 statistical analysis package.

Bonferroni corrections were used for significant post-hoc analyses. Table 12 shows the means, SD, and one-way analysis of variance (ANOVA) results for each group for the behavioral outcome measures.

Demographic and language between-group comparisons

Values associated with the distribution of demographic variables, including percentages and means, are displayed in Table 9. The comparisons between the three groups on demographic characteristics, including age, biological sex, SES, physical activity, and musical training, did not reveal significant differences²⁶. Immigration status, however, did reveal a significant difference between groups, where the late bilingual group had a higher prevalence of people who immigrated, $\chi^2(2, N = 74) = 27.73, p < 0.001$. This is not surprising given the prevalence of international students attending the University of Victoria. The comparisons between early and late bilinguals on language characteristics did not reveal significant differences for percentage of L1 use, percentage of L2 use, and for three of the group language subskills (speak, read, and write). However, comparisons did reveal that early bilinguals had higher proficiency in L2 on comprehension than late bilinguals, $t(40) = 2.870, p < 0.05$. As well, comparisons revealed a significant difference between early and late bilinguals on translator status, where late bilinguals had a higher prevalence of translators than early bilinguals, $\chi^2(1, N = 42) = 5.56, p = 0.018$. Comparisons also revealed a significant difference between the

²⁶ One-way ANOVAs were used to compare values for age, SES, percentage of L1 use, percentage of L2 use, and for three of the group language subskills, whereas chi-square (χ^2) tests were used to compare categorical values for biological sex, physical activity, immigrant status, translator status, and musical training.

age of acquisition, $t(40) = -12.481, p < 0.001$, which was expected by the design of the study.

Behavioral data

Table 13. Behavioral Data: means, SDs, and one-way ANOVA results

Behavioral Data	Group			One-way ANOVAs			
	Early Biling. Mean (SD)	Late Biling. Mean (SD)	Monoling. Mean (SD)	<i>F</i>	<i>df</i>	Sig.	η_p^2
<i>Control Vars: Fluency Tasks</i>							
Total Correct Letter	37.87 (12.29)	35.11 (11.69)	41.22 (11.69)	1.643	2, 71	0.201	0.044
Total Correct Animal	18.13 (5.51)	17.42 (5.92)	21.13 (3.29)	4.449	2, 71	0.015	0.111
Total Errors Letter	0.87 (1.63)	0.95 (1.47)	0.67 (0.97)	0.332	2, 71	0.719	0.009
Total Errors Animal	0.17 (0.49)	0.26 (0.65)	0.22 (0.55)	0.132	2, 71	0.877	0.004
<i>Control Vars: Matrices</i>							
Total Correct	4.30 (1.82)	4.42 (1.74)	3.77 (1.28)	1.227	2, 70	0.299	0.034
<i>Inhibition</i>							
% Correct Hits	0.82 (0.15)	0.86 (0.11)	0.82 (0.10)	1.033	2, 71	0.361	0.028
<i>d'</i>	3.30 (0.64)	3.52 (0.47)	3.13 (0.54)	3.056	2, 71	0.053 ²⁷	0.079
<i>Switching</i>							
Switch cost Ac	0.03 (0.06)	0.03 (0.06)	0.04 (0.05)	0.136	2, 71	0.873	0.004
Switch cost RT	-0.37 (0.24)	0.43 (0.56)	-0.33 (0.16)	0.553	2, 71	0.578	0.015
<i>Updating</i>							
2-back Ac	0.37 (0.16)	0.42 (0.16)	0.42 (0.13)	0.982	2, 71	0.400	0.025
2-back <i>d'</i>	1.36 (0.53)	1.60 (0.57)	1.41 (0.46)	1.268	2, 71	0.288	0.034
3-back Ac	0.25 (0.12)	0.27 (0.12)	0.28 (0.12)	0.359	2, 71	0.699	0.010
3-back <i>d'</i>	0.79 (0.45)	0.95 (0.28)	0.83 (0.50)	0.699	2, 71	0.501	0.019
Load effect	0.12 (0.13)	0.16 (0.17)	0.14 (0.10)	0.366	2, 71	0.695	0.010

Note. This table shows the mean, SDs, and one-way ANOVAs for the behavioral data (i.e., RT, accuracy, total correct, *d*-prime) for all tasks for each of the three groups. Bolded values denote significant differences, $p < 0.05$, with Bonferroni correction. Ac = Accuracy; Biling. = Bilinguals; Monoling. = Monolinguals; RT = response time (ms); Vars = variables.

Given a significant one-way ANOVA for effect of language status on total correct for category fluency (animals), pairwise group comparisons were conducted. These tests

²⁷ As a result of this trend, a series of pairwise comparisons between groups revealed a significant difference between monolinguals and late bilinguals, $p < 0.05$, only for the inhibition component (using *d'*) such that late bilinguals outperformed monolinguals; no other significant differences were found between groups.

revealed a significant difference between monolinguals and late bilinguals, $p < 0.05$, where monolinguals outperformed late bilinguals. Post-hoc analyses approached significance, $p = 0.076$, between monolinguals and early bilinguals; post-hoc analyses were not significant between early and late bilinguals.

ERP data

Table 14. *Midline Electrode Mean Voltages*

	Groups														
	Early Bilinguals (μV)					Late Bilinguals (μV)					Monolinguals (μV)				
	Fz	FCz	Cz	CPz	Pz	Fz	FCz	Cz	CPz	Pz	Fz	FCz	Cz	CPz	Pz
a) Go/nogo Task															
<i>N200</i>															
Nogo	-	-	-	-	1.11	-	-	-	0.66	2.03	0.28	0.08	1.02	1.61	2.39
	2.68	2.89	2.21	0.33		0.84	0.96	0.06							
Go	-	-	-	0.96	2.33	-	0.04	1.31	2.52	3.06	-	-	0.95	2.12	3.15
	1.67	1.09	0.39			1.03					1.06	0.32			
<i>P300</i>															
Nogo	6.82	10.65	11.71	12.88	12.74	6.58	11.11	12.97	12.71	12.76	7.37	12.78	14.92	14.52	14.28
Go	3.12	5.65	7.04	7.49	7.70	2.32	5.17	7.16	7.81	7.27	3.03	5.75	7.86	8.60	8.76
b) Switch Task															
<i>P3a</i>															
Pre	2.50	3.01	3.35	-	-	2.63	3.95	4.16	-	-	4.11	4.21	3.72	-	-
Sw	3.53	4.81	4.23	-	-	4.03	5.11	5.68	-	-	5.80	7.03	7.09	-	-
<i>P3b</i>															
Pre	-	-	3.07	4.19	4.69	-	-	4.78	4.84	5.20	-	-	3.72	4.32	4.78
Sw	-	-	3.37	4.85	5.75	-	-	3.67	5.33	6.06	-	-	4.63	5.92	6.48
c) n-Back task															
<i>2back – P300</i>															
Tar	-	-	7.81	9.18	10.84	-	-	12.47	13.94	14.00	-	-	11.23	12.36	13.06
Non	-	-	3.55	4.74	6.11	-	-	4.13	5.46	5.96	-	-	4.74	6.24	7.58
<i>3back – P300</i>															
Tar	-	-	11.00	16.33	13.17	-	-	15.05	13.00	16.74	-	-	11.19	12.09	12.78
Non	-	-	1.79	3.04	4.77	-	-	3.66	5.43	6.88	-	-	4.75	6.55	7.60

Note. Bolded values represent peak amplitude along midline. This table shows the peak amplitudes (μV) for the central midline electrode sites for the waveforms of interest for each of the three tasks: (a) go/nogo task; (b) switch task; and (c) n-back task. – represent electrode site values not measured for that waveform; Pre = preswitch; Sw = switch; Tar = target; Non = nontarget.

Table 15. *Peak Amplitude and Latency Means and SDs*

	Group					
	Early Bilinguals		Late Bilinguals		Monolinguals	
	Mean (SD)		Mean (SD)		Mean (SD)	
	Amp. (μ V)	Lat. (ms)	Amp. (μ V)	Lat. (ms)	Amp. (μ V)	Lat. (ms)
a) Go/nogo task						
<i>N200 at site FCz</i>						
Nogo trials	-2.89 (5.46)	289.57 (30.02)	-0.96 (2.56)	281.47 (35.08)	0.08 (3.67)	276.65 (32.81)
Go trials	-1.09 (4.76)	268.52 (33.56)	0.04 (3.45)	265.90 (39.15)	-0.32 (3.60)	258.70 (27.02)
<i>P300 at site Cz</i>						
Nogo trials	11.71 (5.69)	414.26 (28.92)	12.97 (4.11)	422.74 (35.33)	14.92 (5.98)	413.55 (34.00)
Go trials	7.04 (4.61)	405.74 (30.96)	7.16 (2.69)	413.26 (45.30)	7.86 (5.18)	408.77 (33.62)
b) Switch task						
<i>P3a at site Cz</i>						
Preswitch trials	3.35 (3.60)	361.40 (37.15)	4.16 (3.84)	382.25 (42.16)	3.72 (3.79)	357.47 (28.35)
Switch trials	4.23 (3.60)	364.20 (34.10)	5.68 (3.97)	370.25 (35.45)	7.09 (4.01)	366.13 (40.70)
<i>P3b at site Pz</i>						
Preswitch trials	4.69 (1.86)	513.40 (27.39)	5.20 (2.54)	516.25 (24.85)	4.78 (2.58)	524.23 (29.74)
Switch trials	5.75 (3.13)	516.40 (34.90)	6.06 (3.98)	517.75 (33.85)	6.48 (3.98)	531.20 (38.33)
c) n-Back task						
<i>P300 2back at site Pz</i>						
Target trials	10.84 (4.92)	417.11 (53.24)	14.00 (4.81)	408.30 (47.43)	13.06 (5.41)	387.57 (53.14)
Nontarget trials	6.11 (3.48)	403.56 (55.22)	5.96 (1.84)	406.16 (50.98)	7.59 (4.13)	397.43 (64.20)
<i>P300 3back at site Pz</i>						
Target trials	13.17 (4.93)	398.80 (52.16)	16.74 (9.76)	391.39 (41.28)	12.78 (5.99)	402.80 (58.04)
Nontarget trials	4.77 (3.43)	382.00 (53.21)	6.88 (3.63)	393.23 (60.51)	7.60 (2.93)	388.80 (54.16)

Note. This table shows the mean (μ V) and SDs of the ERP data (i.e., peak amplitude and latency) for each task and for each of the three groups. Peak amplitude values are in μ V and latencies are in ms; Amp. = amplitude; Lat. = latency.

Inhibition. Table 14a shows the mean peak amplitudes for the central midline electrode sites for each of the N200 and P300 waveforms by trial type (go vs. nogo).

Figures 10-16 show the grand average N200 and P300 waveforms for go and nogo trials

for each of the three groups; scalp distributions for peak amplitude activity for either the N200 or P300. Several ANOVAs were conducted on various ERP components and latencies; means and SDs are shown in Table 15a.

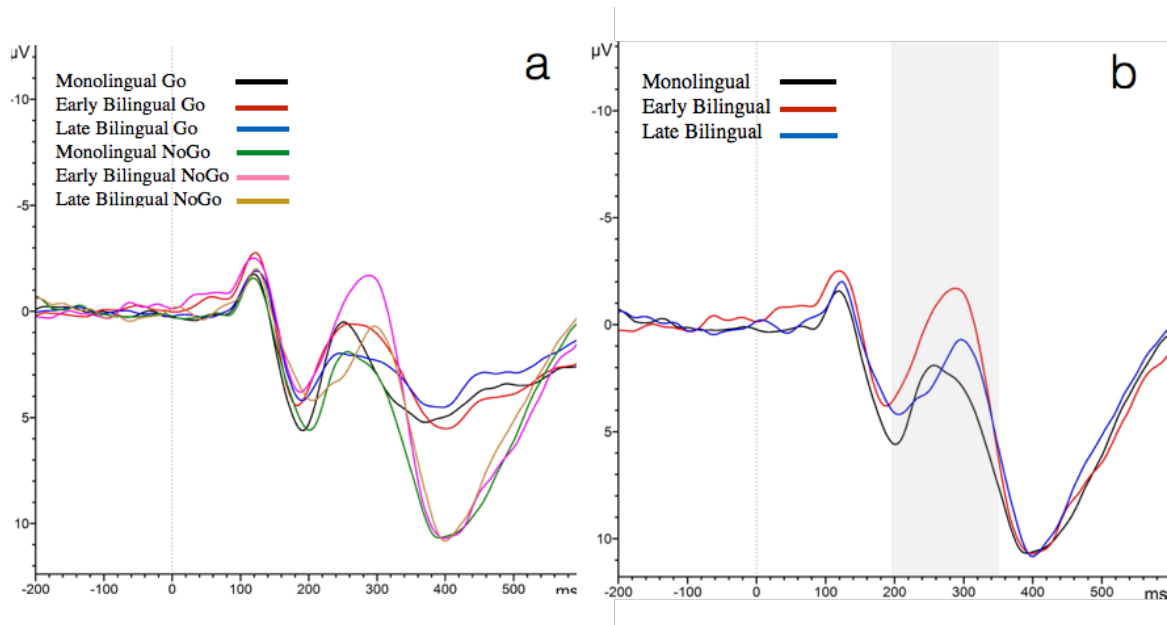


Figure 10. Represents (a) the stimulus-locked go and nogo trial grand average waveforms for all three groups; and (b) the grand average waveforms for only nogo trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.

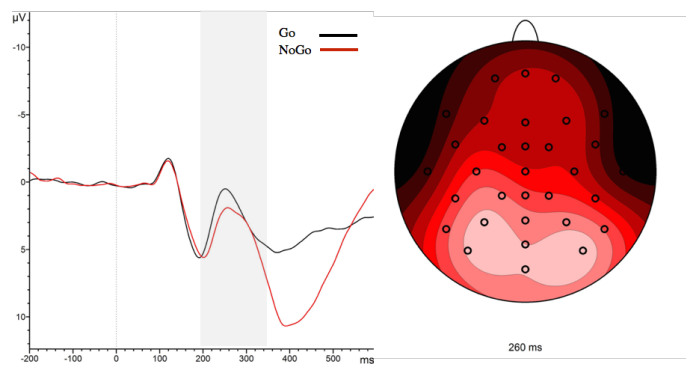


Figure 11. Represents stimulus-locked go and nogo trial grand average waveforms for monolinguals at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.

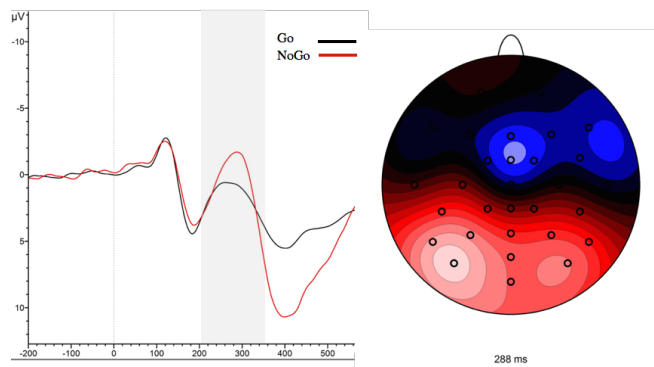


Figure 12. Represents the stimulus-locked go and nogo trial grand average waveforms for early bilinguals at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.

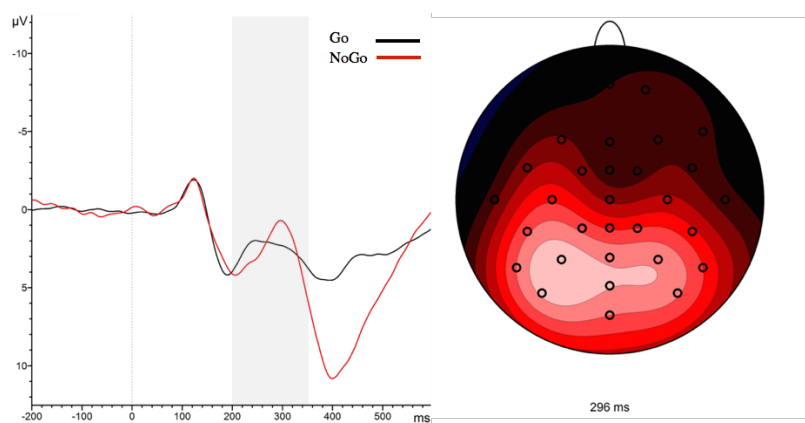


Figure 13. Represents the stimulus-locked go and nogo trial grand average waveforms for late bilinguals at electrode site FCz. Zero represents the onset of the stimulus; the shaded area represents the time window for the N200.

N200 peak amplitude. A repeated measures ANOVA was completed with Trial Type (2: go vs. nogo) as the within-subject factor and Group Membership (3 groups) as the between-subject factor to examine the effect of language status on N200 peak amplitude at electrode site FCz, where it was maximal (see Table 5a). The ANOVA revealed a significant main effect, $F(1, 70) = 5.885, p < 0.05, \eta_p^2 = 0.084$, such that the peak amplitude for nogo trials ($M = -1.05, SE = 4.36$) was larger than go trials ($M = -0.25, SE = 3.96$); this was an expected main effect, suggesting the task properly elicited

differential processing for go and nogo trials. The ANOVA also revealed a significant interaction between group membership and trial type, $F(2, 70) = 3.389, p < 0.05, \eta_p^2 = 0.108$. See Table 6 for means and SD; pairwise comparisons of the group differences for the N200 peak amplitudes for nogo trials revealed a significant difference between monolinguals and early bilinguals, where early bilinguals produced larger N200 amplitudes, $p < 0.05$; no differences were found between early and late bilinguals, no differences were found between late bilinguals and monolinguals.

N200 latencies. A repeated measures ANOVA was completed with Trial Type (2: go vs. nogo) as the within-subject factor and Group Membership (3 groups) as the between-subject factor to examine the effect of language status on N200 peak latency at electrode site FCz. The ANOVA revealed a significant main effect, $F(1, 64) = 15.815, p < 0.001, \eta_p^2 = 0.198$, such that the latency of nogo trials ($M = 281.97$ ms, $SE = 31.69$) peaked later than go trials ($M = 264.42$ ms, $SE = 33.12$). An interaction between trial type and group was not found, $F(2, 64) = 0.007, p = 0.993, \eta_p^2 = 0.000$.

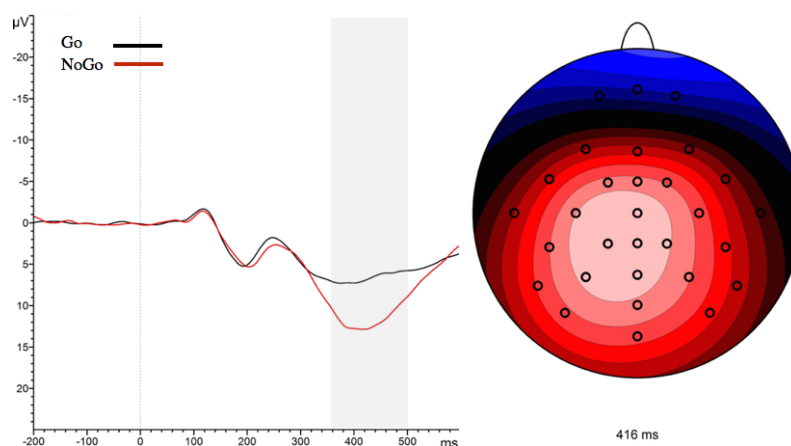


Figure 14. Represents the go and nogo trial grand average waveforms for monolinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300.

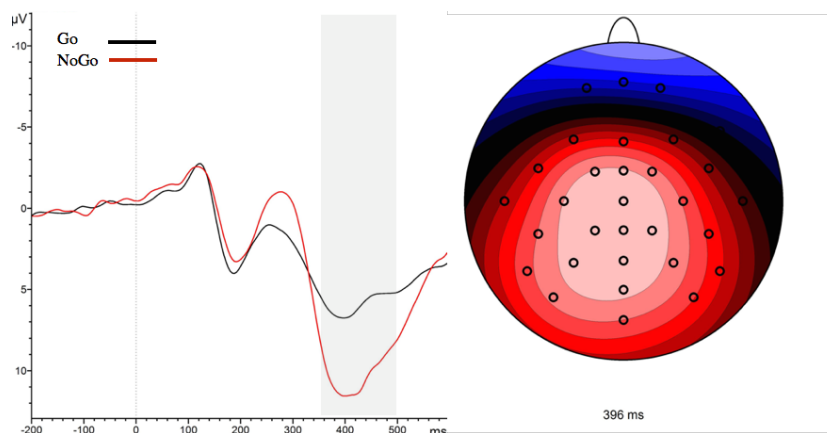


Figure 15. Represents the go and nogo trial grand average waveforms for early bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300.

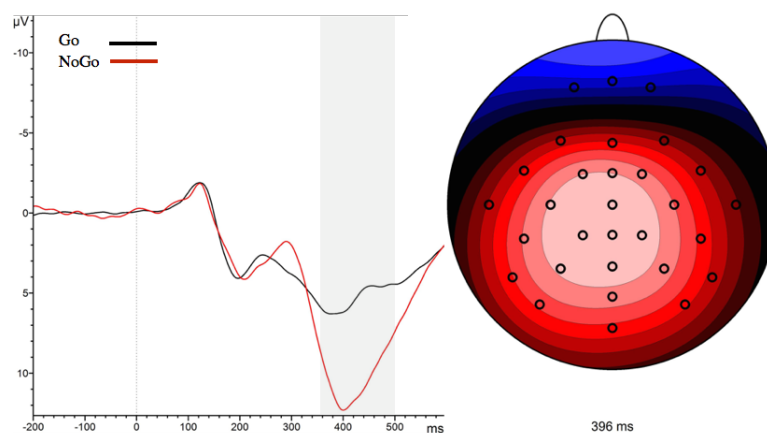


Figure 16. Represents the go and nogo trial grand average waveforms for late bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300.

P300 peak amplitude. A repeated measures ANOVA was completed with Trial Type (2 go vs. nogo) as the within-subject factor and Group Membership (3 groups) as the between-subject factor to examine the effect of language status on P300 peak amplitude at electrode site Cz, where it was generally maximal (see Table 5a). It revealed a significant main effect of trial type, $F(1, 65) = 68.422, p < 0.001, \eta_p^2 = 0.513$, where the peak amplitude for nogo trials was larger ($M = 13.56 \mu\text{V}, SE = 5.53$) than that of go trials ($M = 7.68 \mu\text{V}, SE = 4.29$); this was an expected difference. A significant interaction

between trial type and group membership was not found, $F(2, 65) = 1.385, p = 0.258, \eta_p^2 = 0.041$.

P300 latencies. A repeated measures ANOVA was completed with Trial Type (2: go vs. nogo) as the within-subject factor and Group (3) as the between-subject factor to examine the effect of language status on P300 peak latency at electrode site Cz. The ANOVA did not reveal a significant main effect of trial type, $F(1, 65) = 1.603, p = 0.210, \eta_p^2 = 0.024$, which was an expected finding. The ANOVA also did not reveal a significant interaction between trial type and group membership, $F(2, 65) = 0.092, p = 0.913, \eta_p^2 = 0.003$.

Switching. Table 14b shows the mean peak amplitudes for the central midline electrodes for each of the P3a and P3b waveforms by trial type (preswitch and switch). Figures 17-24 show the grand average P3a and P3b waveforms for target and nontarget trials for each of the three groups. Scalp distributions are present for the three groups and represent the peak amplitude activity on switch trials for either the P3a or P3b. Several ANOVAs were conducted on various ERP components and latencies; means and SDs are shown in Table 15b.

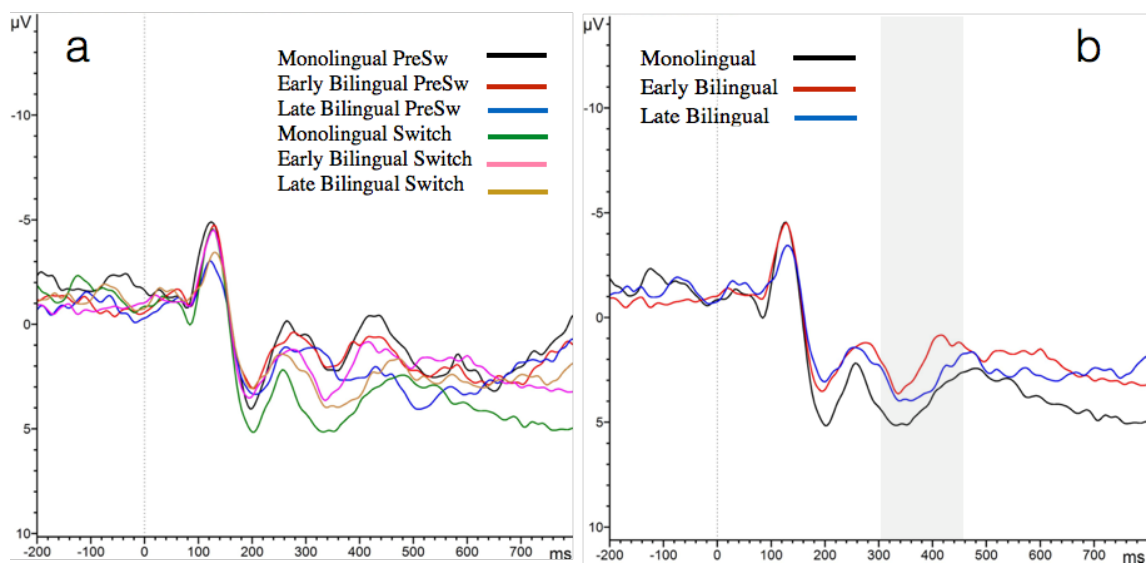


Figure 17. Represents (a) the stimulus-locked pre-switch and switch trial grand average waveforms for all three groups; and (b) the stimulus-locked grand average waveforms for Switch trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window. PreSw = pre-switch trials.

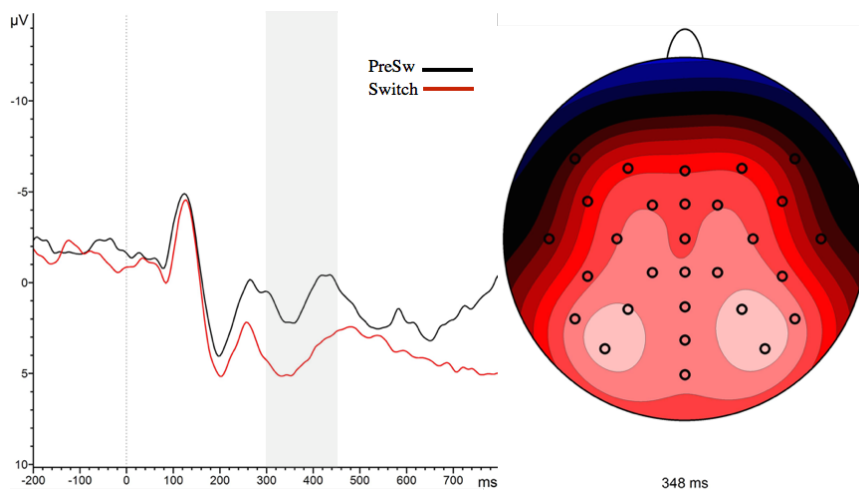


Figure 18. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for monolinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window.

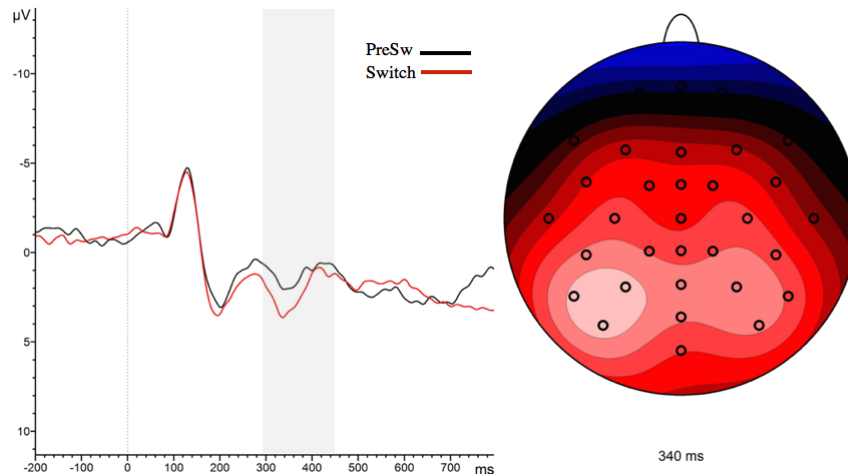


Figure 19. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for early bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window.

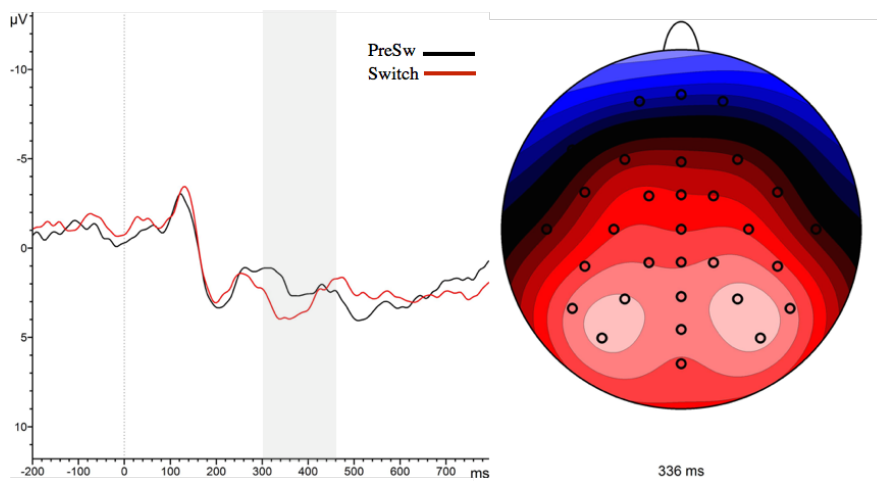


Figure 20. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for late bilinguals at electrode site Cz. Zero represents the onset of the stimulus; the shaded area represents the P3a time window.

P3a peak amplitude. A repeated measures ANOVA was completed with electrode (3: electrodes Fz, FCz, and Cz) and trial type (2: preswitch vs. switch) as the within-subject factors and group membership (3 groups) as the between-subjects factor to examine the effect of language status on P3a peak amplitude. Results revealed a significant three-way interaction, $F(4, 126) = 3.846, p < 0.05, \eta_p^2 = 0.109$, and a two-way interactions between trial type and electrode, $F(2, 62) = 3.799, p < 0.05, \eta_p^2 = 0.109$. As well, two significant main effects were found, including trial type, $F(1, 63) = 20.799, p$

<0.001, $\eta_p^2 = 0.248$, and electrode, $F(1, 63) = 18.174$, $p < 0.001$, $\eta_p^2 = 0.370$. No other main effects or interactions were significant.

As a result of the three-way interaction, three separate repeated measures ANOVAs with trial type (2: preswitch vs. switch) as the within-subject factors and group membership (3 groups) as the between-subjects factor were conducted to examine the effect of language status on P3a peak amplitude at each of the three electrode sites (Fz, FCz, and Cz). At site Fz, results revealed a significant main effect of trial type, $F(1, 63) = 12.107$, $p < 0.001$, $\eta_p^2 = 0.161$, but no interaction effect was found, $F(2, 63) = 0.275$, $p = 0.761$, $\eta_p^2 = 0.009$. The main effect revealed a larger amplitude to switch trials ($M = 4.69 \mu\text{V}$, $SE = 3.96$) compared to preswitch trials ($M = 3.26 \mu\text{V}$, $SE = 4.06$), as expected. At site FCz, results revealed a significant main effect of trial type, $F(1, 63) = 20.357$, $p < 0.001$, $\eta_p^2 = 0.244$, but no interaction effect was found, $F(2, 63) = 1.386$, $p = 0.761$, $\eta_p^2 = 0.009$. The main effect revealed a larger amplitude to switch trials ($M = 5.89 \mu\text{V}$, $SD = 4.11$) compared to preswitch trials ($M = 3.78 \mu\text{V}$, $SE = 4.20$), as expected. At site Cz, results revealed a significant main effect of trial type, $F(1, 63) = 22.771$, $p < 0.001$, $\eta_p^2 = 0.265$, as well as a significant interaction effect, $F(2, 63) = 4.164$, $p < 0.05$, $\eta_p^2 = 0.117$. The main effect revealed a larger amplitude to switch trials ($M = 5.88 \mu\text{V}$, $SE = 4.07$) compared to preswitch trials ($M = 3.76 \mu\text{V}$, $SE = 3.70$), as expected. To elucidate this latter interaction, a one-way ANOVA with group (3) as the between subjects factor was conducted to assess the impact of language status on the peak P3a at site Cz. It revealed a significant difference, $F(2, 63) = 3.277$, $p < 0.05$, $\eta_p^2 = 0.094$. Pairwise comparisons revealed a significant difference between monolinguals ($M = 7.09 \mu\text{V}$, $SD = 4.02$) and

early bilinguals ($M = 4.23 \mu\text{V}$, $SE = 3.60$); no other pairwise comparisons were significant.

P3a latency. A repeated measures ANOVA was completed with electrode (3: electrodes Fz, FCz, and Cz) and trial type (2: preswitch vs. switch) as the within-subject factors and Group (3) as the between-subjects factor to examine the effect of language status on P3a latencies. Results revealed a significant two-way interaction between trial type and group, $F(2, 63) = 3.175$, $p < 0.05$, $\eta_p^2 = 0.092$, showing that mean P3a peak latencies for preswitch trials were earliest for monolinguals ($M = 356.58 \text{ ms}$, $SE = 5.83$), slightly later for early bilinguals ($M = 369.40 \text{ ms}$, $SE = 7.15$), and delayed for late bilinguals ($M = 381.58 \text{ ms}$, $SE = 7.99$); mean P3a peak latencies for switch trials were similar. These simple main effects were averaged across electrode sites. No other main effects or interactions were significant.

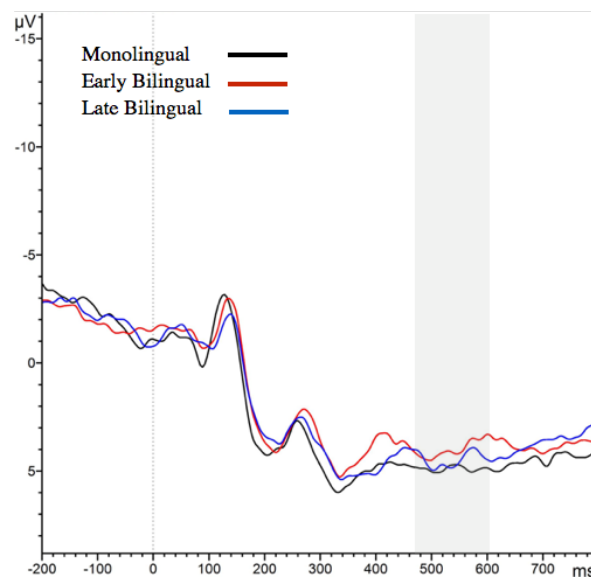


Figure 21. Represents the grand average waveforms for switch trials for monolinguals, early bilinguals, and late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b.

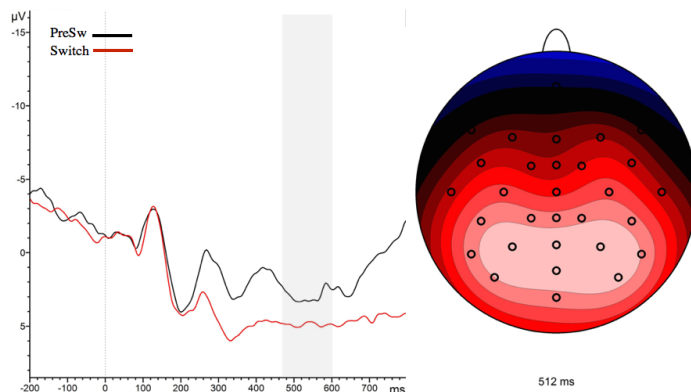


Figure 22. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for monolinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b.

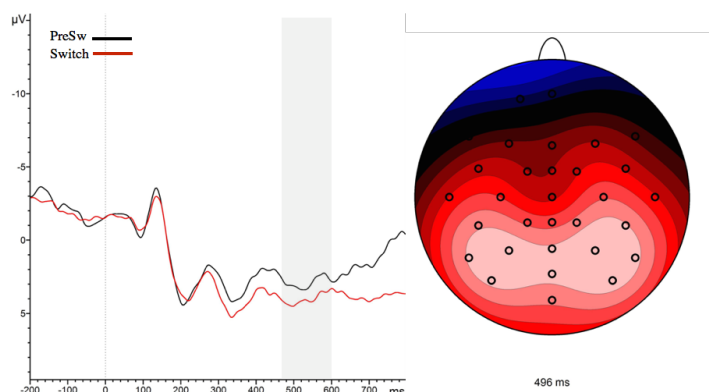


Figure 23. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for early bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b.

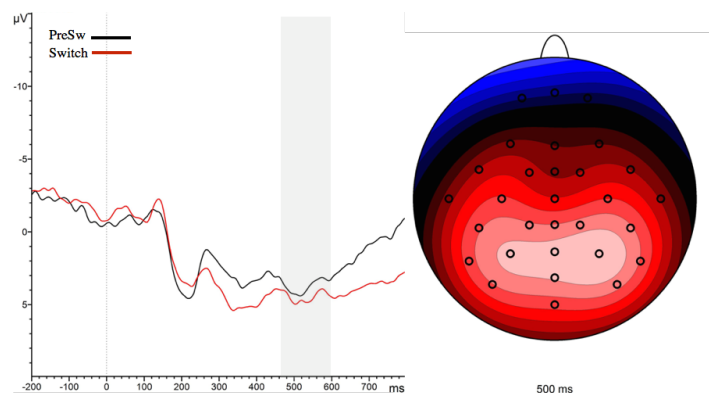


Figure 24. Represents the stimulus-locked pre-switch and switch trial grand average waveforms for late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P3b.

P3b peak amplitude. A repeated measures ANOVAs was completed with electrode (3: electrodes Cz, CPz, and Pz) and trial type (2: preswitch vs. switch) as the

within-subject factors and Group (3) as the between-subjects factor to examine the effect of language status on P3b peak amplitude. Results revealed the following significant main effects and interactions: (i) trial type, $F(1, 63) = 4.403, p < 0.05, \eta_p^2 = 0.065$; (ii) electrode, $F(1, 63) = 34.586, p < 0.001, \eta_p^2 = 0.527$; and (iii) an electrode by trial type interaction, $F(2, 62) = 10.131, p < 0.001, \eta_p^2 = 0.246$. The main effect for trial type showed P3b peak amplitude for switch trials was larger ($M = 5.12 \mu\text{V}, SE = 0.46$) than preswitch trials ($M = 4.40 \mu\text{V}, SE = 0.30$); this was an expected difference. The main effect for electrode showed that the P3b peak amplitude for site Pz was largest ($M = 5.49 \mu\text{V}, SE = 0.36$), compared site CPz ($M = 4.90 \mu\text{V}, SE = 0.36$), which was larger than site Cz ($M = 3.90 \mu\text{V}, SE = 0.36$). The two-way interaction between electrode and trial type showed that P3b peak amplitudes were smallest at site Cz for switch trials ($M = 3.90 \mu\text{V}, SE = 0.47$) compared to switch trials at site CPz ($M = 5.37 \mu\text{V}, SE = 0.73$) and Pz ($M = 6.09 \mu\text{V}, SE = 0.48$); P3b peak amplitude means were similar for preswitch trials. No other main effects or interactions were found.

P3b latency. A repeated measures ANOVA was completed with electrode (3: electrodes Cz, CPz, and Pz) and trial type (2: preswitch vs. switch) as the within-subject factors and Group (3) as the between-subjects factor to examine the effect of language status on P3b latencies. No main effects or interactions were found.

Updating. Table 14c shows the peak amplitudes (μV) for the central midline electrode sites for each of the P300 waveforms for the 2- and 3-back tasks by trial types for each group. Several ANOVAs were conducted on the P300 peak amplitudes and latencies at sites Cz, CPz, and Pz; Table 15c shows the mean and SDs. Figures 25-28 show the 2-back grand average waveforms for the P300; Figures 29-32 show the 3-back

grand average waveforms for the P300. Main effects of trial type were predicted and represent the underlying differences in updating processes required for each trial type. Significant interaction effects involving group membership and trial type were predicted and represent the group differences in working memory processes.

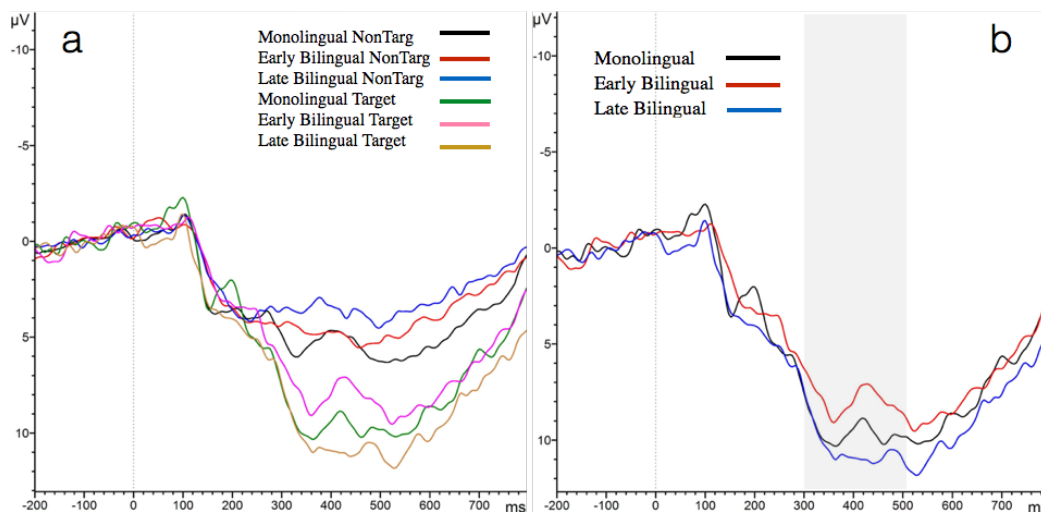


Figure 25. Represents (a) the 2-back target and nontarget trial grand average waveforms for all three groups at electrode site Pz; and (b) the grand average waveforms for 2-back target trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300. NonTarg = Nontarget trials.

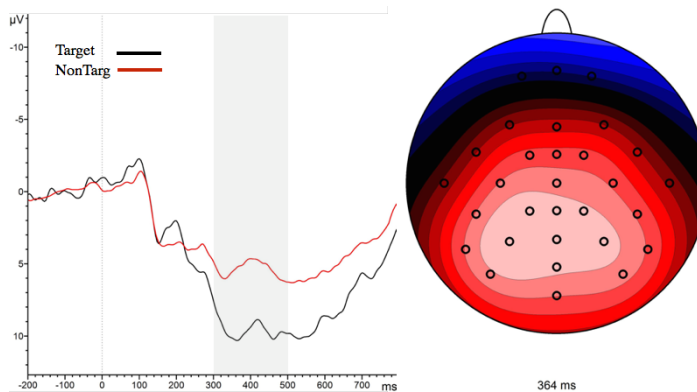


Figure 26. Represents the 2-back target and nontarget trial grand average waveforms for monolinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials.

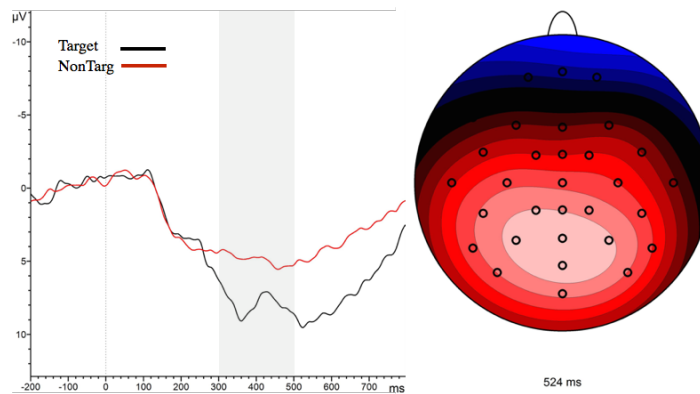


Figure 27. Represents the 2-back target and nontarget trial grand average waveforms for early bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials.

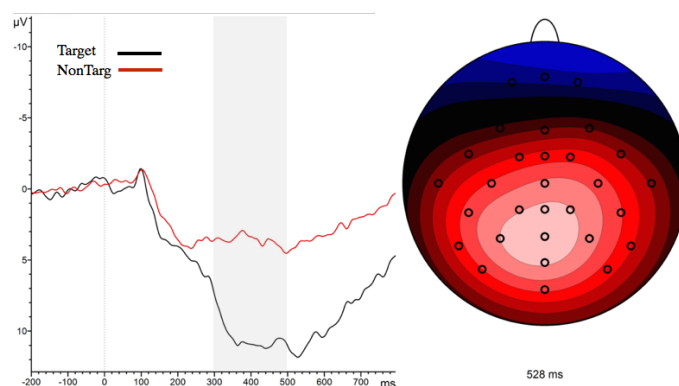


Figure 28. Represents the 2-back target and nontarget trial grand average waveforms for late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials.

2-back P300 peak amplitude. A repeated measures ANOVA was completed with electrode (3: electrodes Cz, CPz, and Pz) and trial type (2: target vs. nontarget) as the within-subject factors and Group (3) as the between-subjects factor to examine the effect of language status on P300 peak amplitude in the 2-back task. Results revealed two significant main effects, including trial type, $F(1, 56) = 121.77, p < 0.001, \eta_p^2 = 0.685$, and electrode, $F(2, 55) = 21.98, p < 0.001, \eta_p^2 = 0.444$. The main effect for trial type revealed larger P300 amplitudes for target trials ($M = 11.65 \mu\text{V}, SE = 0.70$) compared to nontarget trials ($M = 5.39 \mu\text{V}, SE = 0.51$). The main effect for electrode revealed P300 peak amplitudes that were largest for site Pz ($M = 9.60 \mu\text{V}, SE = 0.61$), slightly smaller at

site CPz ($M = 8.66 \mu\text{V}$, $SE = 0.56$), and smallest at site Cz ($M = 7.32 \mu\text{V}$, $SE = 0.61$). No other main effects or interactions were significant.

Exploratory Analyses: Although a three-way interaction was not found, three separate repeated measures ANOVAs with trial type (2: target vs. nontarget) as the within-subject factors and Group (3) as the between-subjects factor were conducted to examine the effect of language status on P300 peak amplitude at each of the three electrode sites (Fz, FCz, and Cz). At site Cz, results revealed a significant main effect of trial type, $F(1, 56) = 124.99$, $p < 0.001$, $\eta_p^2 = 0.691$, and a significant interaction effect between trial type and group membership, $F(2, 56) = 3.75$, $p < 0.05$, $\eta_p^2 = 0.118$. The main effect revealed a larger amplitude to target trials ($M = 10.46 \mu\text{V}$, $SE = 5.80$) compared to nontarget trials ($M = 4.24 \mu\text{V}$, $SE = 4.80$). Pairwise comparisons of the group membership by trial type interaction revealed no significant differences. At site CPz, results revealed a significant main effect of trial type, $F(1, 56) = 120.44$, $p < 0.001$, $\eta_p^2 = 0.683$, and an interaction between between trial type and group, $F(2, 55) = 3.46$, $p < 0.05$, $\eta_p^2 = 0.110$. The main effect revealed a larger amplitude to target trials ($M = 11.74 \mu\text{V}$, $SE = 5.30$) compared to nontarget trials ($M = 5.61 \mu\text{V}$, $SE = 3.90$). Pairwise comparisons of the group membership by trial type interaction revealed no significant differences. At site Pz, results revealed a significant main effect of trial type, $F(1, 56) = 86.19$, $p < 0.001$, $\eta_p^2 = 0.606$, but no significant interaction effect was found, $F(2, 55) = 1.942$, $p = 0.153$, $\eta_p^2 = 0.065$. The main effect revealed a larger amplitude to target trials ($M = 12.59 \mu\text{V}$, $SE = 5.20$) compared to nontarget trials ($M = 6.78 \mu\text{V}$, $SE = 3.58$). Also for exploratory purposes and to enhance more specific statistical power to reveal differences should they exist, a one-way ANOVA was conducted to evaluate the effect of

language status on peak P300 amplitude at site Pz (where the peak was maximal for all groups; see table 8). Results did not reveal a significant difference between groups.

2-back P300 latency. A repeated measures ANOVA was completed with electrode (3: electrodes Cz, CPz, and Pz) and trial type (2: target vs. nontarget) as the within-subject factors and Group (3) as the between-subjects factor to examine the effect of language status on P300 latency in the 2-back task. Results revealed no significant main effects or interactions.

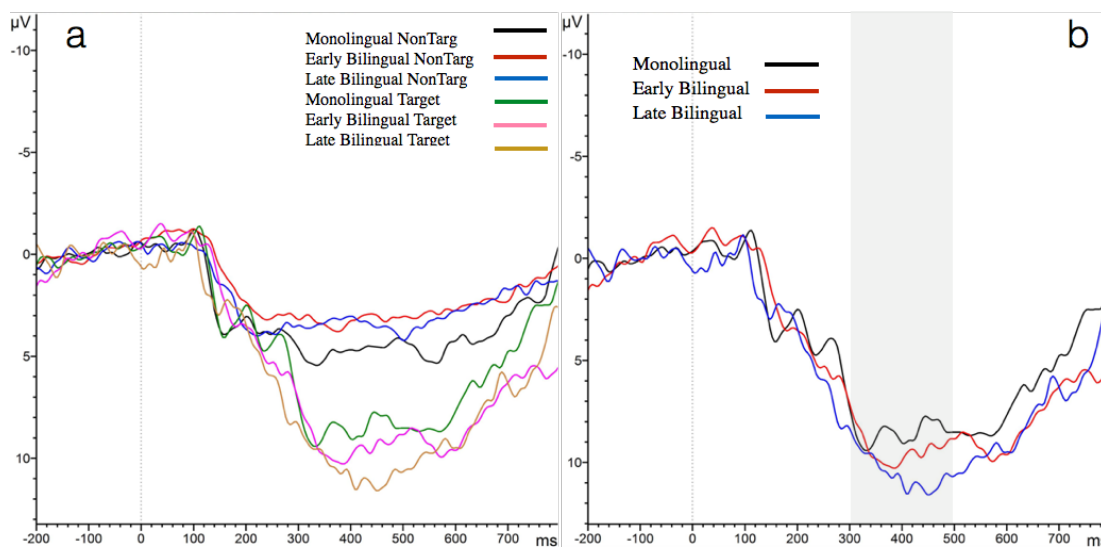


Figure 29. Represents (a) the 3-back target and nontarget trial grand average waveforms for all three groups; and (b) the grand average waveforms for 3-back target trials for monolinguals, early bilinguals, and late bilinguals. Both figures represent activity at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300. NonTarg = Nontarget trials.

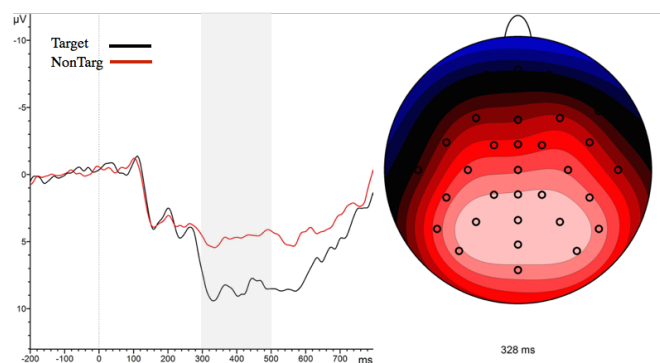


Figure 30. Represents the 3-back target and nontarget trial grand average waveforms for monolinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials.

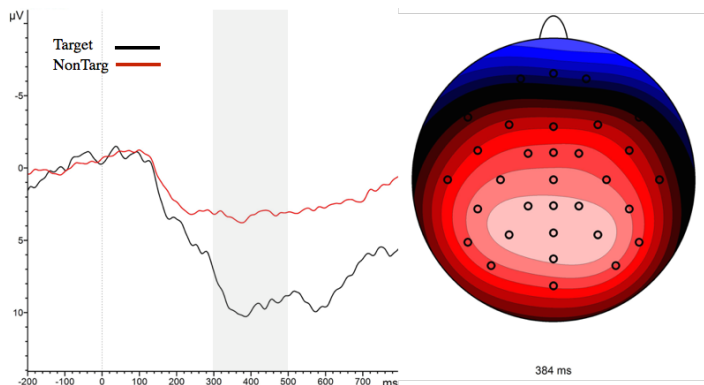


Figure 31. Represents the 3-back target and nontarget trial grand average waveforms for early bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials.

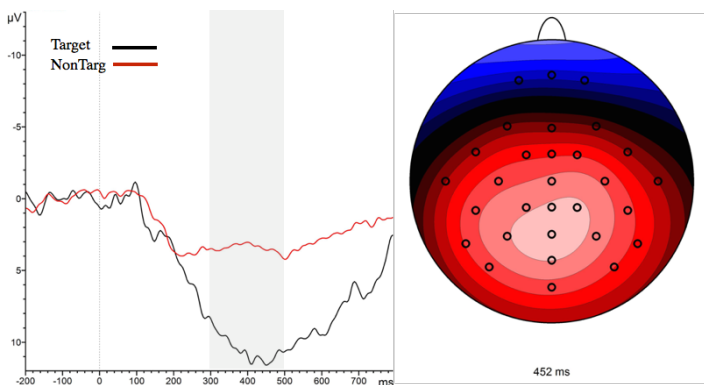


Figure 32. Represents the 3-back target and nontarget trial grand average waveforms for late bilinguals at electrode site Pz. Zero represents the onset of the stimulus; the shaded area represents the time window for the P300; NonTarg = nontarget trials.

3-back P300 peak amplitude. A repeated measures ANOVAs was completed with electrode (3: electrodes Cz, CPz, and Pz) and trial type (2: target vs. nontarget) as the within-subject factors and group membership (3) as the between-subjects factor to examine the effect of language status on P300 peak amplitude in the 3-back task. Results revealed two significant main effects, including trial type, $F(1, 40) = 70.539, p < 0.001$, $\eta_p^2 = 0.638$, and electrode, $F(2, 39) = 21.815, p < 0.001$, $\eta_p^2 = 0.528$. The main effect for trial type revealed larger P300 amplitudes for target trials ($M = 13.45 \mu\text{V}, SE = 0.98$) compared to nontarget trials ($M = 4.93 \mu\text{V}, SE = 0.48$). The main effect for electrode revealed P300 peak amplitudes that were largest for site Pz ($M = 10.38 \mu\text{V}, SE = 0.60$),

slightly smaller at site CPz ($M = 9.27 \mu\text{V}$, $SE = 0.66$), and smallest at site Cz ($M = 7.93 \mu\text{V}$, $SE = 0.73$). A significant two-way interaction was found for trial type by group membership, $F(2, 40) = 3.341$, $p < 0.05$, $\eta_p^2 = 0.143$, which showed that late bilinguals had larger P300 peak amplitudes on target trials ($M = 16.04 \mu\text{V}$, $SE = 1.99$) than early bilinguals ($M = 12.05 \mu\text{V}$, $SE = 1.69$) and monolinguals ($M = 12.25 \mu\text{V}$, $SE = 1.38$); mean values were similar between groups for nontarget trials. Another significant two-way trial by electrode interaction was found, $F(2, 39) = 4.366$, $p < 0.05$, $\eta_p^2 = 0.183$, showing generally larger P300 peak amplitudes for target trials at sites Pz ($M = 14.34 \mu\text{V}$, $SE = 0.93$), CPz ($M = 13.54 \mu\text{V}$, $SE = 0.98$), and Cz ($M = 12.47 \mu\text{V}$, $SE = 1.12$) compared to nontarget trials at sites Pz ($M = 6.41 \mu\text{V}$, $SE = 0.47$), CPz ($M = 5.00 \mu\text{V}$, $SE = 0.53$), and Cz ($M = 3.38 \mu\text{V}$, $SE = 0.51$). No other main effects or interactions were significant.

Exploratory analyses. Although a three-way interaction between electrode, trial type, and group membership was not found, three separate repeated measures ANOVAs with trial type (2: target vs. nontarget) as the within-subject factors and group membership (3 groups) as the between-subjects factor were conducted to examine the effect of language status on P300 peak amplitude at each of the three electrode sites (Cz, CPz, and Pz). At site Cz, results revealed a significant main effect of trial type, $F(1, 40) = 75.280$, $p < 0.001$, $\eta_p^2 = 0.653$, and an interaction between trial type and group membership that approached significance, $F(2, 39) = 3.151$, $p = 0.054$, $\eta_p^2 = 0.136$. The main effect revealed a larger amplitude to target trials ($M = 11.60 \mu\text{V}$, $SE = 7.57$) compared to nontarget trials ($M = 3.94 \mu\text{V}$, $SE = 3.25$). A one-way ANOVA that analyzed the effect of language status on mean differences of target trials was not significant, $F(1, 40) = 1.341$, $p = 0.273$, $\eta_p^2 = 0.063$. At site CPz, results revealed a

significant main effect of trial type, $F(1, 40) = 70.048, p < 0.001, \eta_p^2 = 0.637$, and an interaction effect between trial type and group membership, $F(2, 39) = 3.640, p < 0.05, \eta_p^2 = 0.154$. The main effect revealed a larger amplitude to target trials ($M = 12.88 \mu\text{V}, SE = 7.34$) compared to nontarget trials ($M = 5.80 \mu\text{V}, SE = 3.29$). A one-way ANOVA that analyzed the effect of language status on mean differences of target trials was not significant, $F(1, 40) = 1.387, p = 0.262, \eta_p^2 = 0.065$. At site Pz, results revealed a significant main effect of trial type, $F(1, 40) = 58.944, p < 0.001, \eta_p^2 = 0.596$, but despite approaching significance, no significant interaction between trial type and group membership effect was found, $F(2, 39) = 2.975, p = 0.062, \eta_p^2 = 0.129$. The main effect revealed a larger amplitude to target trials ($M = 13.56 \mu\text{V}, SE = 7.18$) compared to nontarget trials ($M = 7.07 \mu\text{V}, SE = 3.26$).

3-back P300 latency. A repeated measures ANOVAs was completed with electrode (3: electrodes Cz, CPz, and Pz) and trial type (2: target vs. nontarget) as the within-subject factors and Group (3) as the between-subjects factor to examine the effect of language status on P300 latency in the 3-back task. Results revealed no significant main effects or interactions.

Discussion

The present study measured the neurophysiological differences between three language groups: (i) early bilinguals, (ii) late bilinguals, and (iii) monolinguals. This study used three tasks to measure three different EF components: inhibition (go/nogo task), updating (n-back task) and attention switching (switch task). The groups were well matched; no differences were found on control variables, with the exception of semantic fluency and self-rated proficiency for L2 comprehension. For semantic verbal fluency, monolinguals outperformed late bilinguals, while no differences were found between early and late bilinguals or early bilinguals and monolinguals. This finding supports previous literature indicating that bilinguals, regardless of age of acquisition, perform worse on semantic fluency tasks (e.g., Kousaie, Sheppard, Lemieux, Monetta, & Taler, 2014). For proficiency, early bilinguals had higher proficiency in L2 on comprehension than late bilinguals. Mixed empirical results have been found for the contention that high language proficiency in both languages results in better performance on EF measures when compared to individuals with lower proficiency in one language (e.g., Friesen, Luo, Luk, & Bialystok, 2014; Luk & Bialystok, 2013; Luo, Luk, & Bialystok, 2010; Perani et al., 1998); however, it is generally accepted that language groups should be matched on proficiency levels if it is not a main independent variable. Most studies include language proficiency at least as a control variable (e.g., Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Costa, Santesteban, & Ivanova, 2006; Emmorey, Luk, Pyers, & Bialystok, 2008; Martin-Rhee & Bialystok, 2008; Prior & Macwhinney, 2009), which was the case in the current study. Unforeseen statistical differences in self-rated proficiency have been accounted for through statistical control procedures to reduce the

statistical effect (e.g., Moreno et al., 2014; Tao et al., 2011). The addition of L2 proficiency for comprehension as a statistical control did not change the results of tests between early and late bilinguals.

Our findings partially support the first set of hypotheses regarding inhibition, in that larger N200 peak amplitudes on nogo trials (requiring inhibition) were found for early bilinguals, but not late bilinguals or monolinguals. As well, better performance for late bilinguals, when compared to monolinguals, was found when d' was used as the outcome measure. Similarly, our findings also provide partial support our hypotheses regarding attention switching, in that no differences were found between groups using behavioral measures of performance, while smaller P3a peak amplitudes were found for early bilinguals compared to monolinguals. Finally, none of the hypotheses regarding updating processes were supported by our findings. These results are discussed in more detail below.

Inhibition

First, the current study successfully replicated the frequently observed “go/nogo effect” in which the amplitudes of the N200 and P300 components on nogo trials are larger relative to the waveform amplitudes elicited on the go trials (e.g., Bokura, Yamaguchi, & Kobayashi, 2001; Donkers & van Boxtel, 2004; Smith et al., 2008). This finding confirms that our task successfully elicited inhibition processes to demonstrate that the processing of nogo stimuli is qualitatively and/or quantitatively different to that of go stimuli. Second, the current study found larger N200 amplitudes for early bilinguals on nogo trials compared to late bilinguals and monolinguals. This finding supports and extends earlier studies between bilinguals and monolinguals that found similar results,

including larger N200 amplitudes (e.g., Fernandez et al., 2013; Fernandez, Acosta, Douglass, Doshi, & Tartar, 2014; Kousaie & Phillips, 2012; Moreno et al., 2014).

Only a handful studies investigating N200 amplitudes in bilingual and monolinguals have been conducted. In a study by Kousaie & Phillips (2012), multiple ERP tasks were used to measure inhibition, including the Stroop task, the Simon task, and the Erikson flanker task; however, only the Stroop task, which measures *conflict resolution*, demonstrated smaller N200 responses in bilinguals compared to monolinguals, while the other tasks did not reveal differences in N200 amplitude. The authors argued that the smaller N200 amplitudes indicated that bilinguals required less conflict monitoring than controls and that bilinguals were better able to selectively attend to the relevant aspects of the stimulus. Another study found larger N200 amplitudes for bilinguals using a traditional go/nogo task, similar to the one in the current study (Moreno et al., 2014). Other authors (Fernandez et al., 2013; Fernandez et al., 2014) found larger N200 amplitudes for bilinguals compared to monolinguals using an auditory go/nogo task, but no differences were found using a visual version of the task. Fernandez and colleagues (2013; 2014) argued that their larger N200 findings supported the hypothesis that the N200 reflects non-motor cognitive inhibition (Falkenstein, Hoormann, & Hohnsbein, 1999) rather than conflict monitoring (Nieuwenhuis et al., 2003). Falkenstein and colleagues (1999) found that those participants with high false-alarm rates (i.e., those who often pressed the button to non-targets; categorized as ‘poor performers’) exhibited smaller nogo N200 amplitudes compared to those who had fewer false alarms (‘good performers’). Fernandez and colleagues (2013) hypothesized that

bilinguals would function similarly to “good performers” and produce larger N200 amplitudes.

In these studies, the potentially confounding variables of interest, such as AoA, language proficiency, L1 and L2 frequency of use, music training, and physical activity were inconsistently reported and/or addressed. For example, only one study mentioned AoA (i.e., Fernandez et al., 2013); these authors proposed that this variable may account for their incongruent results across two studies. In their first study, participants were *all* early bilinguals who learned both languages before age 6, while there was a mix of early and late bilingual participants in the second study (Fernandez et al., 2014). In terms of language proficiency, Fernandez et al. ensured equivalent English proficiency across both language groups. Other variables of interest were not accounted for at all, including immigration and translator status. Translator status has been demonstrated to promote enhanced EF (Engel de Abreu et al., 2012; Garcia, 2014). Thus, it would be tempting to argue that a higher prevalence of translator status (i.e., someone who conducted in-the-moment translation in their two languages) and immigrant status (i.e., someone who immigrated to Canada) in the late bilingual group had an impact on the current results; however, late bilingual group had a higher prevalence of these variables, yet these participants produced behavioral and neurophysiological responses similar to monolinguals (i.e., no advantage).

In the current study, percentage of time spent engaging in physical activity was not significantly different between groups, suggesting that it did not affect group outcomes. In contrast, Stroth and colleagues (2009) found smaller N200 peak amplitudes, interpreted as reflecting more efficient neural processing, for physically active

individuals compared to sedentary ones. Further, the authors stated that the increased efficiency of the EF system was achieved because physical activity reduces the amount of effort required to response monitor. Following the logic that the language neural systems of bilinguals are under constant conflict (Green & Abutalebi, 2013), one hypothesis may be that bilinguals develop a more efficient system to deal with such conflict, similar to physically fit individuals, and thus, produce smaller N200 amplitudes. However, the bilingual brain encounters a problem that is not present in physical activity: conflict between two simultaneously activated language options. Therefore, a different neural response for inhibition may be required. Theory suggests there is a difference in the neural networks of bilingual and monolingual individuals, in that bilingual networks are constantly required to produce *reactive suppression* of task irrelevant information (i.e., the equally activated nodes/response options in the nontarget language) (Green & Abutalebi, 2013). In relation to underlying neurophysiology and the interpretation of the waveforms, the N200 is thought to reflect the detection and resolution of interference or response selection at early stages, whereas the P300 is thought to reflect the execution of a response or its suppression (Huster et al., 2010). As such, it could be argued that the EF system is trained to produce more inhibition to suppress activation of the nontarget language, a process that is similar to the *detection and resolution of interference* (as expressed by the N200 amplitude). Together, this may explain the differences observed between bilinguals versus monolinguals, including the larger amplitude of the N200 in early bilinguals and the absence of differences in the P300 amplitude.

The current study also found distinct patterns of neurophysiological response between early and late bilinguals, suggesting differences in neurophysiological

processing. Previous research has demonstrated functional brain organization differences in early bilinguals from late bilinguals to process and manage two languages (Hull & Vaid, 2007). These authors argue that relatively greater right hemisphere involvement in early bilinguals is related to developing and applying metalinguistic knowledge from a very early age in order to monitor the language environment and use the appropriate language. Further, similar to the differences in lateralization between early and late bilinguals, differences in the neurophysiological markers of inhibition may be related to the sheer amount of experience early bilinguals have with handling two language systems, which may intensify or accelerate the automatization of language processes and create different functional and/or structural pathways than those of less experienced bilinguals (Sebastian-Galles et al., 2005). Other authors have proposed similar arguments in the face of empirical evidence that revealed different neural activation patterns in early bilinguals from late bilinguals (Garbin et al., 2010).

The current study yielded an unexpected trend in the d' measure for inhibition. This measure was derived from signal detection theory and is thought to be a more sensitive measure of accuracy as it accounts for more of the possible types of error or *noise* in the signal-to-noise ratio (McNicol, 2005) than traditional accuracy measures (i.e., proportion correct responses, which only accounts for miss trials). This study was the first, to our knowledge, to use d' as an outcome measure to investigate the bilingual advantage for inhibition. Contrary to our prediction that early bilinguals would demonstrate better inhibition (i.e., larger d' value) compared to late bilinguals and monolinguals, on our behavioral computerized task, late bilinguals exhibited superior inhibitory control over monolinguals and no differences emerged between early and late

bilinguals. One argument is that late bilinguals demonstrated better performance relative to monolinguals when a more *sensitive* measure of inhibition is used. However, given the contrary finding, the relative novelty of this measure in the bilingual advantage literature, the fact that initial statistical tests did not fully reach significance, and the higher prevalence of translators in the late bilingual group, interpretation is deferred at the present time, as erroneous conclusions could be made over a potentially spurious or confounded finding.

Switching

Firstly, the current study successfully replicated the frequently observed switch effect in which the amplitudes of the P3a and P3b components for switch trials are larger relative to the waveform elicited by the pre-switch trials (e.g., Barcelo, Escera, Corral, & Periáñez, 2006; Eppinger et al., 2007; Kieffaber & Hetrick, 2005). This finding suggests that the present task successfully elicited switching processes. Further, this finding indicates that the underlying neural processing of switch stimuli is qualitatively and/or quantitatively different to that of pre-switch stimuli, presumably as a result of the requirement to *switch* attention. The current study is the first, to our knowledge, to investigate the neurophysiological markers of attention switching in relation to the bilingual advantage. We hypothesized that early bilinguals would demonstrate (a) greater switching ability as reflected by a larger P3a amplitude, and (b) an advantage in later switching processing as reflected by larger P3b amplitudes, compared to late bilingual and monolingual participants. We predicted that there would be no differences in the behavioral data as the literature is mixed, and these measures may not be sensitive enough to identify subtle differences between groups.

In contrast to our predictions, early bilinguals demonstrated smaller P3a peaks compared to monolinguals and there were no differences in the amplitude of the P3b waveforms on switch trials or in behavioural data. Larger P3a amplitudes are believed to reflect activity associated with attention engagement, which is required for the neural systems to attend to the task and initiate the switch, while the P3b is thought to reflect context-updating operations and subsequent memory storage (Polich, 2007). There are two possible interpretations of the smaller P3a amplitude in early bilinguals: (i) smaller amplitudes reflect an early bilingual *disadvantage* in early attention mechanisms; or (ii) in light of similar behavioral outcomes, smaller amplitudes reflect *more efficient* attention engagement mechanisms.

In considering these interpretations, the neurophysiological literature for attention switching mechanisms in bilinguals is large and growing, with many studies supporting the notion that languages switches requires attention switching processes (Abutalebi et al., 2007; Alvarez et al., 2003; Blackburn, 2013; Christoffels, Firk, & Schiller, 2007; Jackson et al., 2001; Kuipers & Thierry, 2010; Moreno et al., 2002; Proverbio et al., 2004; Rodriguez-Fornells et al., 2002; Van Der Meij et al., 2011; Verhoef et al., 2010; Wang, Xue, Chen, Xue, & Dong, 2007). These studies also propose that similar to other switching tasks, language switches produce a *cost*, such as a response time cost. However, this also means that fortification of switching mechanisms could arise through the constant use of switching mechanisms in language switches. In fact, a bilingual advantage for behavioral outcome measures of attention switching has been demonstrated in several studies (e.g., Kousaie et al., 2014; Prior & Macwhinney, 2009; Tao et al., 2011; Wang, Kuhl, Chen, & Dong, 2009). As such, it seems counterintuitive to argue that the

current findings support a bilingual disadvantage for attention switching mechanisms, even in light of null results for task switching behavioral outcome measures. Given the current findings, however, it is possible that early bilinguals do experience a taxing of their neural system as a result the excitation of two equally appropriate options for language selection. It may be that a bilingual's executive network is monitoring the situation for two things above and beyond that of a monolingual: (i) cues for the target language, and (ii) cues for a language switch. These monitoring processes, which would include the engagement of attention and/or other attention mechanisms, would reduce the neurophysiological response of the system (i.e., tax the system and produce smaller P3a amplitude). In support of this argument, Harmony et al. (2000) found reduced P3a amplitudes as the attention load increased, meaning that the P3a amplitude was inversely correlated to the demands on the system. In light of similar behavioural performance across groups, there does not appear to be an overall capacity reduction (i.e., performance cost) for bilinguals compared to monolinguals. This suggests: (i) the system has developed adequate compensatory strategies for dealing with the additional system tax, and/or (ii) behavioural outcome measures are not sensitive enough to pick up on subtle differences in processing of a taxed system. Taken together, smaller P3a amplitudes in early bilinguals may reflect a neural processing difference based on the unique qualities of early bilingual neural circuitry.

On the other hand, smaller P3a amplitudes may actually reflect a bilingual advantage, meaning a more efficient system, for attention switching. The P3a reflects the early engagement of the attentional system to initiate the reallocation of attention (i.e., an attentional switch; Polich, 2007). Early bilinguals may develop a more efficient system as

a result of the constant requirement to switch between languages, which promotes the earlier development and integration of EF and language networks. Further, in line with this argument and following the contention and empirical finding that inhibition may actually represent a core EF (Miyake & Friedman, 2012), the smaller P3a peak amplitudes for early bilinguals may be related to the neurophysiological finding for inhibition (i.e., larger N200 peak amplitudes on nogo trials). That is, inhibition is required in a switch task, such that previously relevant task information needs to be inhibited in order to appropriately switch and respond to the correct task. If larger inhibitory signals are sent in the early bilingual neural networks, then there would be less effort required by the system for the allocation of attention mechanism (i.e., more efficient switch mechanisms and thus, smaller peak P3a peak amplitudes). The initial hypotheses for the P3a amplitude may have been misinformed without the consideration of the theoretical overlap between inhibition and switching mechanisms in the switch task.

Updating

The current study successfully replicated the frequently observed effect in which the amplitude of the P300 waveform for correct n-back target trials is larger relative to the waveform elicited by the nontarget trials (e.g., Ozen et al., 2013); this finding was expected and suggests the task successfully elicited updating working memory processes. Further, this finding indicates that the underlying neural processing of correct n-back target stimuli is qualitatively and/or quantitatively different to that of nontarget stimuli, likely as a result of the requirement to *update* the task information in order to identify the target and produce a behavioral response. In relation to the bilingual advantage literature, the current study is the first, to our knowledge, to investigate the neurophysiological

markers of updating working memory. We hypothesized that the peak P300 amplitude would be larger for early bilinguals than late bilinguals and monolinguals in the high demand situations (i.e., three-back), but not in the low demand situation (i.e., two-back). As well, we predicted that early bilinguals would demonstrate better performance in the high demand condition when d' is used, but no differences would be found for d' in the low demand condition. Finally, we predicted that there would be no between group differences in the low sensitivity behavioral outcome measures (i.e., accuracy of two- and three-back).

Aside from the expected trial-type differences, none of the group difference predictions were supported, which likely reflects poor task sensitivity and performance across groups, rather than truly similar group performance. The first indication that a problem arose with the task itself was low mean percentage correct for the two-back (0.40) and three-back (0.25) tasks across groups. Compared to previous studies in our lab using this task, these mean scores were extremely low. However, a methodological change was made to the task in the current study - in addition to response options for target trials, a button press was added for nontarget trials. This additional response was added in order to capture important behavioral information about the processing of nontarget trials. That is, without a response to nontarget trials, it is difficult to ascertain whether or not a participant successfully rejected the trial or if the individual missed the trial to due some extraneous factor (e.g., lapse of attention). Further, a lack of a button press makes it difficult, if not impossible, to average ERP waveforms for truly correctly rejected nontarget trials. Despite these valid reasons for including a nontarget button press, it appears that change may have placed an additional tax on the working memory

system (e.g., which key to press on which trial). This additional system tax likely not only lowered task performance across groups, but also resulted in the exclusion of some participants from the two-back task (14 out of 76 participants) and many participants from the three-back task (30 out of 76 participants) who had less than 10 trials per waveform. This increased the interindividual variability, as well as the intraindividual variability, reduced the power, and likely rendered potentially significant ANOVAs effects insignificant. Indeed, several interaction effects in the current study, although exploratory in nature, approached and reached significance.

Although an interpretation of the results is not warranted given the methodological concerns, the exploratory findings are briefly reviewed. In the 2-back task, significant group by trial type interaction effects were found at two electrode sites (Cz and CPz); however, follow-up results did not reveal significant group differences in peak P300 amplitude for target trials. Similar results were found for the 3-back task, such that significant group by trial type interaction effects were found at the same two electrode sites. Also similar to the 2-back task, follow-up results did not reveal significant group differences. Group by trial type interaction effects were not found at the electrode site where the P300 was maximal (site Pz) for all groups for both conditions.

Limitations

There are several limitations to this study. First, these findings must be interpreted with caution because of small and unequal sample sizes. Unfortunately, sample sizes across studies investigating the bilingual advantage tend to be small, especially in ERP studies (e.g., Fernandez et al., 2013; Fernandez et al., 2014; Moreno et al., 2014). Further, the group sample sizes tend to be unequal, especially in studies investigating AoA. It is

often difficult to find highly proficient late bilinguals, and thus, this group is often smaller in research samples (e.g., Kalia et al., 2014; Tao et al., 2011). Secondly, outliers may have influenced the behavioural and neurophysiological data across the three tasks, as outliers were not removed prior to statistical analysis. Outlying data points, if more prevalent or extreme in one group, can distort the mean values and distort the differences between groups. In order to account for this, participant data can be plotted to visually represent the possibility of outliers; data points that extend beyond two SD beyond the sample mean should be removed (Dixon, 1953).

Although this study is one of the first in the bilingual advantage literature, to our knowledge, to include three EF tasks as posited by Miyake and colleagues (2000), it has been argued that bilingual advantage researchers should include multiple measures and/or tasks that target the same construct to provide convergent validity (e.g., Paap & Greenberg, 2013; Paap, 2014a, 2014b). Although this is ideal, the convergent validity of similar EF tasks is often poor (Chan, Shum, Touloupoulou, & Chen, 2008; Toplak et al., 2013). One possible way to circumvent this issue is to use a latent variable approach, in which a set of manifest variables (i.e., related measures) are combined into latent variables to compare performance across groups. Looking forward, studies that include multiple ERP tasks that require inhibition, switching, and updating, as well as larger sample sizes, are necessary to further our understanding of the bilingual advantage and neurophysiology in bilinguals.

Another potential limitation of the current study is the use of a self-rating scale for language proficiency, as it may be too insensitive to assess proficiency and subject to individual bias. However, this method of measuring language proficiency is relatively

common and is used across bilingual advantage studies (e.g., Bialystok & Depape, 2009; Bialystok, 2006; Blumenfeld & Marian, 2013; Colzato et al., 2008; Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Kousaie & Phillips, 2012; Kousaie et al., 2014; Salvatierra & Rosselli, 2010; Linck et al., 2008; Luk et al., 2011). Further, this measurement technique likely provides a *good enough* estimate of proficiency, especially when attempting to ensure a minimum proficiency cutoff for matching purposes. In terms of specific proficiency concerns in the current study, early and late bilinguals differed significantly on L2 comprehension proficiency where early bilinguals had slightly higher scores than late bilinguals. It is possible that this had an effect on the outcomes of this study; however, it is unlikely that the effect was substantial given that both groups met a minimum proficiency cutoff for inclusion in the study. This was further confirmed by post-hoc analysis using L2 comprehension proficiency as a covariate, which did not change any of the findings.

Interestingly in the current study, late bilinguals had a higher prevalence of language translation use (i.e., translator status) and daily use of languages, which may lead one to argue that they would have a switching network advantage as a result of exercising this function more frequently. However, this group did not differ from the monolingual or early bilingual groups on the behavioural outcome measure and P3a peak amplitude. Perhaps, then, the higher prevalence of these variables in this group produced a potential confound in the current study whereby these variables produced P3a activity that was more similar to early bilinguals and less similar to monolinguals. It is clear that the novelty of this area of research signals the need for further research to elucidate the

nature of the P3a, as well as the impact of such variables, in early bilingual, late bilingual, and monolingual individuals.

The study of bilingualism is hampered by a number of methodological problems, which is related to the complex interplay and variability of social, biological, psychological, and chance factors. Together, this complex interplay can make generalizations difficult. For example, another demographic variable that may have had a confounding impact on the current results and highlights this complex interplay of variables was immigration status. Moving beyond linguistics itself, immigration involves adaptations to the cultural environment and learning of new customs, a context that varies greatly from one speech community to another. Thus, in addition to learning a second language, the immigrant bilingual must also learn and adapt to these additional factors, which may have an additional enhancing effect on EF.

In the current study, the late bilingual group had a larger proportion of immigrants than the other two groups, with the early bilingual group having a larger proportion than the monolingual group. It was expected that the monolingual group would have the smallest proportion. To exemplify the complex interplay of factors, late bilinguals in the current study were mainly from Chinese culture, moving to Canada to complete post-secondary education. Thus, in addition to learning English, they had to adjust and adapt to west coast Canadian culture, as well as learning and adjusting to the new city, neighbourhood, and university lifestyle. These adjustments would be greater for late bilinguals than early bilinguals because the late bilingual would have had more time to learn the customs and culture of their original country, and thus, greater adaptation would be required to adjust to the new context. That is, the late bilinguals in the current study

would have spent approximately 18 years learning Chinese culture, whereas early bilingual immigrants would have moved to Canada, and thus, started adapting to Canadian culture, prior to the age of six.

Given the additional adaptation and learning required when one immigrates, it could be argued that these adaptations would enhance EF; for example, Kharkhurin (2008) found divergent thinking, which was enhanced in bilinguals compared to monolinguals, was moderated and further enhanced by language proficiency, AoA, and the length of exposure to the new cultural settings. However, research has found that immigration and cultural differences did not have an additive effect on the bilingual advantage (e.g., Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Bialystok & Viswanathan, 2009). In the current study, the late bilingual group had the largest proportion of immigrants, which would lead one to predict that this group may experience greater enhancement; however, the results did not align with this hypothesis, and in fact, the results are more in line with Bialystok and colleagues (2010, 2009). Nonetheless, as a result of immigration status not being an independent variable, it is difficult to ascertain the exact influence of immigration status on the current data, and thus, it is a limitation. Future studies should endeavour to incorporate and empirically manipulate the influence of immigration status as it relates to the bilingual advantage.

There were several other task-specific limitations. Related to the go/nogo task, behavioural (d') and neurophysiological (N200 amplitude) were contradictory. Using d' as an indicator of inhibition, late bilinguals exhibited superior inhibitory control over monolinguals and no differences emerged between early and late bilinguals, whereas using N200 peak amplitude as an indicator of inhibition, early bilinguals had larger N200

amplitudes than late bilinguals and monolinguals. Given that the d' measure of inhibition has rarely been used in bilingual advantage research, more research using this outcome measure is necessary. In regards to the n-back task and the methodological problems identified in the current study, interpretation of the findings were reserved at the present time. Although this is a limitation and is disappointing given the novel nature of this task in the bilingual advantage, it is hoped that other researchers will build on the idea of assessing updating differences, both at an information processing level and neurophysiological level.

Conclusions and future directions

To our knowledge, this is the first study to incorporate AoA as a primary variable in assessing the neurophysiological differences of inhibition, switching, and updating in bilingual individuals. This is also the first study to use neurophysiological measures of switching and updating processes to evaluate the bilingual advantage effect. In terms of inhibition, this study supports and extends the mounting evidence for larger N200 amplitudes in bilinguals. However, at the present time, these positive neurophysiological findings for bilinguals are restricted to versions of the go/nogo task and the Stroop task. In terms of switching mechanisms, tentative differences between early bilinguals and monolinguals were found for the P3a component. Finally, in terms of updating mechanisms, no group differences were found, which was potentially a result of overall poor task performance and methodological problems.

The behavioral outcomes measures in the current study generally aligned with our predictions; no differences were found between groups, except for a trend using d' where late bilinguals outperformed monolinguals. The current study's findings did not align

with previous research that has found bilingual advantages for inhibitory tasks (e.g., Bialystok, 2006; Colzato et al., 2008; Tao et al., 2011) and switching tasks (e.g., Kousaie et al., 2014; Prior & Macwhinney, 2009; Tao et al., 2011; Wang, Kuhl, Chen, & Dong, 2009). Unfortunately, no prior studies have been conducted using updating working memory tasks, except for the study presented in chapter 3, which actually demonstrated a partial bilingual disadvantage. The current results did not replicate the finding using a modified n-back task, but performance using the current methodology was problematic, which may have rendered the task too difficult.

A bilingual advantage effect for early bilinguals was found in the go/nogo and switch task that demonstrated neurophysiological differences between groups. It was consistently found in these results that early bilinguals had significantly different peak amplitudes than monolinguals in the go/nogo and switch tasks. As well, peak amplitude differences were found between early bilinguals and late bilinguals in the go/nogo task, but no differences were found between early and late bilinguals as well as late bilinguals and monolinguals in the switch task. Given the relative novelty of this neurophysiological evidence using AoA, an analysis of the mean values and grand average waveforms was done: late bilinguals mean values for peak amplitudes of the N200 and P3a fell in between the other two groups, rendering it difficult for post-hoc analyses to reveal the relatively smaller group differences. Further, these data trends may suggest a *layered or graded* effect, such that early bilinguals demonstrate the largest effect, while late bilinguals demonstrate a slightly smaller effect, and monolinguals demonstrate the smallest effect. However, stronger support for this contention would come from

statistically significant difference between all groups in the layered effect; future research investigating this contention is required.

Overall, results from two of the three tasks demonstrated between group differences at the *neurophysiological* level. In contrast, the behavioural outcome measures yielded null or mixed results at the *cognitive/information processing level* of measurement. The lack of group difference using behavioural outcome measures in the go/nogo and n-back tasks did not align with those reported in chapter 3, contributing to the mixed nature of the findings reported at this level of measurement. As the level of measurement was conducted at less complex levels of measurement (i.e., neurophysiology), group differences were elucidated where group differences were not found at higher levels of measurement (i.e., cognitive/information processing). The alignment of these results suggests that the tasks used to measure the bilingual advantage in young adult bilinguals may not be sensitive enough to pick up on subtle differences between groups. Together, these findings provide additional preliminary support for adopting a multi-level approach to assess the bilingual advantage (see Figure 33 below), and support the contention that the bilingual advantage may be too subtle in young adults to be consistently revealed at the information processing level of EF measurement.

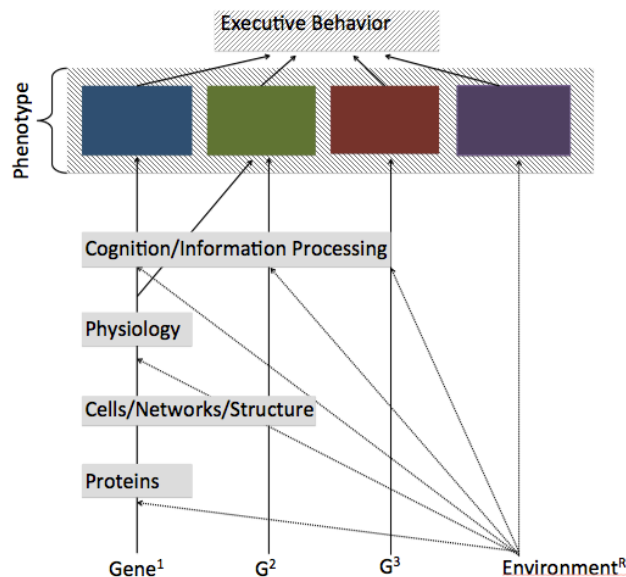


Figure 33. This figure was adapted from Crosbie et al. (2008) and from Garcia-Barrera, Frazer, & Areshenkoff (2012). It represents the relationship between genes, environment, and executive behaviours, and provides a theoretical framework from which one can assess the bilingual advantage effect at varying levels of measurements, as outlined in Chapter 1 of this dissertation.

The neurophysiological research literature for the bilingual advantage is relatively novel, and as such, further research elucidating the characteristics of neurophysiological responses for each the various EF components in bilinguals is needed. For example, conducting additional studies that measure the different neurophysiological responses of inhibition underlying different tasks, such as the Simon and Erikson flanker tasks. As well, novel studies should aim to investigate other inhibition tasks, such as antisaccade task and attention network test (ANT) to assess the *spread* of the bilingual advantage across types of inhibition. As neurophysiological research mounts for each of the EF components, future research should aim to consolidate such knowledge through the use of meta-analyses and systematic reviews. Moreover, research that accounts for other variables (e.g., immigrant status, musical training, physical exercise, translator status, and SES) that may impact and/or modify the characteristics of the bilingual advantage is needed.

Epilogue

This dissertation sought to shed light on the mixed nature of the bilingual advantage in young adults by reviewing the extant literature, positing a novel multi-level approach, and empirically investigating the bilingual advantage across the various levels of measurement. To achieve this, a compilation of four complete manuscripts was prepared. This series of research was successful in assessing and generally supporting the hypotheses developed from the multi-level approach, offering a novel way to conceptualize measurement of the bilingual advantage in young adults and organize the findings in a coherent manner. Further, this multi-level approach shed light on the mixed nature of the literature and proposing that more complex forms of measurement (e.g., executive behaviors, information processing) are likely not sensitive enough to reveal subtle differences between monolingual and bilingual individuals. Thus, the multi-level approach directs researchers towards an exciting line of future research to explore other levels of measurement, such as neurophysiology as well as structural and functional brain imaging.

Although the veracity of the bilingual advantage has been criticized (Paap & Greenberg, 2013; Paap & Liu, 2014; Paap, 2015; Paap, Sawi, Dalibar, Darrow, & Johnson, 2014, 2015), there is mounting evidence to suggest that it does exist, but it is subtle in young adults, and this multi-level approach assists in explaining the mixed nature of the young adult bilingual advantage literature. Further, early bilinguals may demonstrate specific and unique underlying neural changes as a result of executive and language neural system integration, as well as the development of executive systems earlier than typical. Further research investigating these differences is needed, and

following the multi-level approach may help to guide hypotheses and interpretation of subsequent research findings. To expand the current findings and to provide a highlight of the empirical research that is already available for additional levels of the approach, research investigating the functional differences, a level below neurophysiology, that used functional magnetic resonance imaging (fMRI) in young adult bilinguals has revealed differences in the underlying activation of separate brain regions in early bilinguals from monolinguals (Garbin et al., 2010). Further, structural differences have been found in the frontal regions of bilingual young adults, but not in monolinguals or multimodal bilinguals (sign language), suggesting enhanced cortical changes and a bilingual advantage (Olulade et al., 2015).

In conclusion, it is clear that more research in these areas is needed, as the depths of this multi-level approach posited here are novel and not well explored. However, this dissertation has assisted in the organization of the extant literature and revitalized the hope that the young adult bilingual advantage, in its elusive nature, exists.

References

- 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 : The Official Journal of the Society for Neuroscience*, 27(50), 13762–9. doi:10.1523/JNEUROSCI.3294-07.2007
- Altarriba, J., & Heredia, R. R. (2008). *An introduction to bilingualism: Principles and processes*. New York, NY: Taylor & Francis.
- Alvarez, R. P., Holcomb, P. J., & Grainger, J. (2003). Accessing word meaning in two languages: An event-related brain potential study of beginning bilinguals. *Brain and Language*, 87(2), 290–304. doi:10.1016/S0093-934X(03)00108-1
- Anderson, P. (2008). Towards a developmental model of executive function. In V. Anderson, R. Jacobs, & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A ...* New York, NY: Taylor & Francis. Retrieved from http://books.google.com/books?hl=en&lr=&id=f8p5AgAAQBAJ&oi=fnd&pg=PA3&dq=Towards+a+developmental+model+of+executive+function&ots=RhaIBZtmVU&sig=I-GliYbutMZ7mU3xquGag_ca7S8
- Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology*, 28(1), 539–560. doi:10.1207/s15326942dn2801_5
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8(4), 170–7. doi:10.1016/j.tics.2004.02.010

- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of the locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28(1), 403–450. doi:10.1146/annurev.neuro.28.061604.135709
- Baddeley, A. (2003a). Working memory and language: an overview. *Journal of Communication Disorders*, 36(3), 189–208. doi:10.1016/S0021-9924(03)00019-4
- Baddeley, A. (2003b). Working memory: looking back and looking forward. *Nature Reviews. Neuroscience*, 4(10), 829–39. doi:10.1038/nrn1201
- Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29(4), 699–714. doi:10.1016/j.ecresq.2014.02.003
- Barcelo, F., Escera, C., Corral, M. J., & Periáñez, J. A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of Cognitive Neuroscience*, 18(10), 1734–1748. doi:10.1162/jocn.2006.18.10.1734
- Barkley, R. A. (1997). Behavioral Inhibition, Sustained Attention, and Executive Functions: Constructing a Unifying Theory of ADHD, 121(1).
- Barkley, R. A. (1999). Response inhibition in attention-deficit hyperactivity disorder. *Mental Retardation and Developmental Disabilities*, 184, 177–184.
- Barkley, R. A., & Murphy, K. R. (2010). Impairment in Occupational Functioning and Adult ADHD: The Predictive Utility of Executive Function (EF) Ratings Versus EF Tests. *Archives of Clinical Neuropsychology*, 25(3), 157–173. doi:10.1093/arclin/acq014
- Bialystok, E. (2006a). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology/Revue Canadienne de*

Psychologie Expérimentale, 60(1), 68–79. doi:10.1037/cjep2006008

- Bialystok, E. (2006b) Second-language acquisition and bilingualism at an early age and the impact on early cognitive development. Rev ed. In: Tremblay RE, Barr RG, Peters RDeV, eds. *Encyclopedia on Early Childhood Development* [online]. Montreal, Quebec: Centre of Excellence for Early Childhood Development; 2008:1-4. Available at: http://www.child-encyclopedia.com/documents/BialystokANGxp_rev.pdf.
- Bialystok, E. (2007). Cognitive Effects of Bilingualism: How Linguistic Experience Leads to Cognitive Change. *International Journal of Bilingual Education and Bilingualism*, 10(3), 210–223. doi:10.2167/beb441.0
- Bialystok, E. (2008). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3. doi:10.1017/S1366728908003477
- Bialystok, E. (2009). Claiming evidence from non-evidence: A reply to Morton and Harper. *Developmental Science*, 12(4), 499–501. doi:10.1111/j.1467-7687.2009.00868.x
- Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: beyond inhibition. *Developmental Psychology*, 46(1), 93–105. doi:10.1037/a0015466
- Bialystok, E. (2011). Reshaping the mind: the benefits of bilingualism. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale*, 65(4), 229–35. doi:10.1037/a0025406
- Bialystok, E. (2012). Consequences of bilinguals on cognitive development. In *Aspects and Implications on Bilingualism* (pp. 417–433).

- Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development : Official Journal of the Cognitive Development Society*, *11*(4), 485–508. doi:10.1080/15248372.2010.516420
- Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*(2), 459–64. doi:10.1016/j.neuropsychologia.2006.10.009
- Bialystok, E., Craik, F. I. M., 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), 40–9. doi:10.1016/j.neuroimage.2004.09.044
- Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual Minds. *Psychological Science in the Public Interest*, *10*(3), 89–129. doi:10.1177/1529100610387084
- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. doi:10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., & 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), 1341–54. doi:10.1037/0278-7393.32.6.1341
- 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), 859–73. doi:10.1037/0278-7393.34.4.859

- Bialystok, E., & Depape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology. Human Perception and Performance*, 35(2), 565–574. doi:10.1037/a0012735
- Bialystok, E., & Martin, M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, 3, 325–339. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-7687.2004.00351.x/full>
- 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. doi:10.1177/13670069050090010701
- Bialystok, E., Poarch, G. J., Luo, L., & Craik, F. I. M. (2014). Effects of bilingualism and aging on executive function and working memory, 29(3), 696–705. doi:10.1037/a0037254
- Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, 112(3), 494–500. doi:10.1016/j.cognition.2009.06.014
- Blackburn, A. (2013). *A study of the relationship between code switching and the bilingual advantage: Evidence that language use modulates neural indices of language processing and cognitive control.*
- Blumenfeld, H. K., & Marian, V. (2013). Cognitive control in bilinguals: Advantages in Stimulus–Stimulus inhibition. *Bilingualism: Language and Cognition*, 80(1), 1–20. doi:10.1017/S1366728913000564
- Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for

- response inhibition in a Go/NoGo task. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *112*(12), 2224–2232.
doi:10.1016/S1388-2457(01)00691-5
- Bradley, R., & Corwyn, R. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371–99. doi:10.1146/annurev.psych.53.100901.135233
- Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288. doi:10.1016/j.cognition.2013.11.015
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*(2), 282–98. doi:10.1111/j.1467-7687.2008.00675.x
- Carpenter, P., Just, M., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*. doi:10.1037/e586602011-001
- Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201–216.
doi:10.1016/j.acn.2007.08.010
- Chatham, C. H., Herd, S. A., Brant, A. M., Hazy, T. E., Miyake, A., O'Reilly, R., & Friedman, N. P. (2011). From an executive network to executive control: a computational model of the n-back task. *Journal of Cognitive Neuroscience*, *23*(11), 3598–619. doi:10.1162/jocn_a_00047
- Chen, Y., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the

- N-back task: An investigation using ERPs, *Clinical Neurophysiology*, *119*, 1546–1559. doi:10.1016/j.clinph.2008.03.003
- Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: an event-related brain potential study. *Brain Research*, *1147*, 192–208. doi:10.1016/j.brainres.2007.01.137
- 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), 302–12. doi:10.1037/0278-7393.34.2.302
- Cook, V. (2002). *Portraits of the L2 user* (Vol. 1). Multilingual Matters.
- 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), 135–49. doi:10.1016/j.cognition.2009.08.001
- Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition*, *106*(1), 59–86. doi:10.1016/j.cognition.2006.12.013
- Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(5), 1057–74. doi:10.1037/0278-7393.32.5.1057
- Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews. Neuroscience*, *15*(5), 336–45. doi:10.1038/nrn3709

- Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in Cognitive Sciences*, *10*(3), 131–8. doi:10.1016/j.tics.2006.01.007
- Crosbie, J., Perusse, D., Barr, C. L., & Schachar, R. J. (2008). Validating psychiatric endophenotypes: Inhibitory control and attention deficit hyperactivity disorder. *Neuroscience and Biobehavioral Reviews*, *32*(1), 40–55.
doi:10.1016/j.neubiorev.2007.05.002
- Cushen, P. J., & Wiley, J. (2011). Aha! Voila! Eureka! Bilingualism and insightful problem solving. *Learning and Individual Differences*, *21*(4), 458–462.
doi:10.1016/j.lindif.2011.02.007
- Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). doi:10.1017/S1366728902003012
- Dixon, W. J. (1953). Processing Data for Outliers. *Biometrics*, *9*(1), 74–89.
- Donders, F. (1969). On the speed of mental processes. *Acta Psychologica*. Retrieved from <http://www.sciencedirect.com/science/article/pii/0001691869900651>
- Donders, F. C. (1868). Over de snelheid van psychische processen. *Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool 1868*; 30: 412–431. 1868–1869, Tweede reeks, II: 92–120. Reprinted as Donders, Franciscus C.(1969). On the speed of mental proces. *Acta Psychologica*.
- Dong, Y., & Xie, Z. (2014). Contributions of second language proficiency and interpreting experience to cognitive control differences among young adult bilinguals. *Journal of Cognitive Psychology*, *26*(5), 506–519.
doi:10.1080/20445911.2014.924951

- Donkers, F. C. L., & van Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition*, *56*(2), 165–76.
doi:10.1016/j.bandc.2004.04.005
- Duggan, E. (2014). *Assessing the Behavioral Aspects of Executive Functioning across the Lifespan: Review of Rating Scales and Psychometric Derivation of a Screener for Young Adults*. Retrieved from <http://dspace.library.uvic.ca:8080/handle/1828/5656>
- Duggan, E. C., Garcia-Barrera, M. A., & Muller, U. (in press). Derivation, replication, and validity analyses of a screener for the behavioral assessment of executive functions in young adults. *Assessment*.
- 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), 1201–6. doi:10.1111/j.1467-9280.2008.02224.x
- Engel de Abreu, P. M. J., 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), 1364–71.
doi:10.1177/0956797612443836
- Eppinger, B., Kray, J., Mecklinger, A., & John, O. (2007). Age differences in task switching and response monitoring: Evidence from ERPs. *Biological Psychology*, *75*(1), 52–67. doi:10.1016/j.biopsycho.2006.12.001
- Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychologica*, *101*(2-3), 267–291.
doi:10.1016/S0001-6918(99)00008-6
- Fan, Y., Hsu, Y., & Cheng, Y. (2013). Sex matters: N-back modulates emotional

mismatch negativity. *Cognitive Neuroscience and Neuropsychology*, 24(9), pp. 457-663. doi:10.1097/WNR.0b013e32836169b9

Feng, X., Bialystok, E., & Diamond, A. (2009). Do bilingual children show an advantage in working memory. *Retrieved January*, 1–41. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Do+Bilingual+Children+show+an+Advantage+in+Working+Memory#0>

Fernandez, M., Acosta, J., Douglass, K., Doshi, N., & L. Tartar, J. (2014). Speaking Two Languages Enhances an Auditory but Not a Visual Neural Marker of Cognitive Inhibition. *AIMS Environmental Science*, 1(2), 145–157. doi:10.3934/Neuroscience.2014.2.145

Fernandez, M., Tartar, J. L., Padron, D., & Acosta, J. (2013). Neurophysiological marker of inhibition distinguishes language groups on a non-linguistic executive function test. *Brain and Cognition*, 83(3), 330–6. doi:10.1016/j.bandc.2013.09.010

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. *Behavioral and Brain Functions*, 6, 5. doi:10.1186/1744-9081-6-5

Fox, E. (1996). Cross-Language Priming from Ignored Words : Evidence for a Common Representational System in Bilinguals. *Journal of Memory and Language*, 35, 353–370. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0749596X96900200>

Friedman, D., Nessler, D., Johnson, R., Ritter, W., & Bersick, M. (2008). Age-related changes in executive function: an event-related potential (ERP) investigation of

- task-switching. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 15(1), 95–128. doi:10.1080/13825580701533769
- 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), 101–35. doi:10.1037/0096-3445.133.1.101
- Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology. General*, 137(2), 201–225. doi:10.1037/0096-3445.137.2.201
- Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2014). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience*, (July 2014), 1–13. doi:10.1080/23273798.2014.918630
- Fuster, J. M. (2000). Executive frontal functions. *Experimental Brain Research*, 133(1), 66–70. doi:10.1007/s002210000401
- Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., *et al.* (2010). Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *NeuroImage*, 53(4), 1272–8. doi:10.1016/j.neuroimage.2010.05.078
- Garcia-Barrera, M. A., Kamphaus, R. W., & Bandalos, D. (2011). Theoretical and statistical derivation of a screener for the behavioral assessment of executive functions in children. *Psychological Assessment*, 23(1), 64–79. doi:10.1037/a0021097

- Garcia-Barrera, M. A., Duggan, E. C., Karr, J. E., & Reynolds, C. R. (2014). Examining executive functioning using the Behavior Assessment System for Children (BASC). In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of Executive Functioning* (1st ed., pp. 283–299). New York, NY. doi:10.1007/978-1-4614-8106-5
- Garcia-Barrera, M., Frazer, J., & Areshenkoff, C. (2012). Theoretical derivation and empirical validation of an integrative neuropsychological theory of executive-related abilities and component transactions (INTERACT) [Abstract]. *Journal of the International Neuropsychological Society*, 18(S2), 61.
doi:10.1017/S1355617712001063
- Garcia, A. (2014). The interpreter advantage hypothesis: Preliminary data patterns and empirically motivate. *Translation and Interpreting Studies*, 9(2), 1932–2798.
- Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, 18(6), 579–583.
doi:10.3758/BF03197100
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, 5(FEB), 1–14.
doi:10.3389/fpsyg.2014.00011
- George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, 49(5), 1083–1094. doi:10.1016/j.neuropsychologia.2011.02.001
- 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), 245–62. Retrieved from
<http://www.ncbi.nlm.nih.gov/pubmed/1827830>
- Gioia, G. A., Kenworthy, L., & Isquith, P. K. (2010). Executive Function in the Real World: BRIEF lessons from Mark Ylvisaker. *The Journal of Head Trauma Rehabilitation*, 25(6), 433–439. doi:10.1097/HTR.0b013e3181fbc272
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2010). TEST REVIEW Behavior Rating Inventory of Executive Function Behavior Rating Inventory of Executive Function. *Child Neuropsychology*, 6(3), 235–238.
doi:10.1076/chin.6.3.235.3152
- Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A History of Executive Functioning as a Theoretical and Clinical Construct. In *Handbook of Executive Functioning* (pp. 3–12). Springer.
- Grace, J., & Malloy, P. F. (2001). *Frontal Systems Behavior Scale Professional Manual*. Lutz, FL: Psychological Assessment Resources.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81. Retrieved from
http://journals.cambridge.org/abstract_S1366728998000133
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(March 2015), 1–16.
doi:10.1080/20445911.2013.796377
- Guttentag, R. E., Haith, M. M., Goodman, G. S., & Hauch, J. (1984). Semantic processing of unattended words by bilinguals: A test of the input switch mechanism. *Journal of Verbal Learning and Verbal Behavior*, 23(2), 178–188.

doi:10.1016/S0022-5371(84)90126-9

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A.

(2010). The validity of d prime as a working memory index: results from the “Bergen n-back task”. *Journal of Clinical and Experimental Neuropsychology*, 32(8), 871–880. doi:10.1080/13803391003596421

Hargreaves, D. J., & Aksentijevic, A. (2011). Music, IQ, and the executive function.

British Journal of Psychology, 102(3), 306–8; discussion 309–12.

doi:10.1111/j.2044-8295.2011.02029.x

Harmony, T., Bernal, J., Fernández, T., Silva-Pereyra, J., Fernández-Bouzas, A., Marosi,

E., *et al.* (2000). Primary task demands modulate P3a amplitude. *Cognitive Brain Research*, 9(1), 53–60. doi:10.1016/S0926-6410(99)00044-0

Heaton, R. K., Miller, S. W., Taylor, M. J., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead-Reitan battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults*. Lutz, FL: PAR.

Hernandez, A. E., Greene, M. R., Vaughn, K. a., Francis, D. J., & Grigorenko, E. L.

(2015). Beyond the bilingual advantage: The potential role of genes and environment on the development of cognitive control. *Journal of Neurolinguistics*, 35, 109–119. doi:10.1016/j.jneuroling.2015.04.002

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2010).

The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition*, 13(03), 315–325.

doi:10.1017/S1366728909990010

- Hervais-Adelman, A. G., Moser-Mercer, B., & Golestani, N. (2011). Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. *Frontiers in Psychology*, 2(September), 234. doi:10.3389/fpsyg.2011.00234
- 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. doi:10.3758/s13423-011-0116-7
- Hilchey, M. D., Saint-Aubin, J., & Klein, R. M. (in press). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing*. Cambridge: Cambridge University Press.
- Holcomb, P., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Cognitive Neuroscience, Journal of*, 18(10), 1631–1643. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6794990
- Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: a structural equation modeling approach. *Neuropsychology*, 22(4), 508–22. doi:10.1037/0894-4105.22.4.508
- Hull, R., & Vaid, J. (2007). Bilingual language lateralization: A meta-analytic tale of two hemispheres. *Neuropsychologia*, 45(9), 1987–2008. doi:10.1016/j.neuropsychologia.2007.03.002
- Huster, R. J., Westerhausen, R., Pantev, C., & Konrad, C. (2010). The role of the cingulate cortex as neural generator of the N200 and P300 in a tactile response

- inhibition task. *Human Brain Mapping*, 31(8), 1260–71. doi:10.1002/hbm.20933
- Isel, F., Baumgaertner, A., Thrän, J., Meisel, J. M., & Büchel, C. (2010). Neural circuitry of the bilingual mental lexicon: effect of age of second language acquisition. *Brain and Cognition*, 72(2), 169–80. doi:10.1016/j.bandc.2009.07.008
- Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of Rating Scales to the Assessment of Executive Functions. *Applied Neuropsychology: Child*, 2(2), 125–132. doi:10.1080/21622965.2013.870014
- Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, 4(02), 169–178. doi:10.1017/S1366728901000268
- Jackson, G. M., Swainson, R., Mullin, A., Cunnington, R., & Jackson, S. R. (2004). ERP correlates of a receptive language-switching task. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, 57(2), 223–40. doi:10.1080/02724980343000198
- Jansma, B. M., Rodriguez-Fornells, A., Moller, J., & Munte, T. F. (2004). Electrophysiological studies of speech production. *Trends in Linguistics and Monographs*, 157, 361–396.
- Juffs, A., & Harrington, M. (2011). *Aspects of working memory in L2 learning. Language Teaching* (Vol. 44). doi:10.1017/S0261444810000509
- Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology Review*, 17(3), 213–33. doi:10.1007/s11065-007-9040-z
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual

- differences in working memory. *Psychological Review*, 99(1), 122–149.
- Kaan, E. (2007). Event-Related Potentials and Language Processing: A Brief Overview. *Language and Linguistics Compass*, 1(6), 571–591. doi:10.1111/j.1749-818X.2007.00037.x
- Kalia, V., Wilbourn, M. P., & Ghio, K. (2014). Better early or late? Examining the influence of age of exposure and language proficiency on executive function in early and late bilinguals. *Journal of Cognitive Psychology*, 26(7), 699–713. doi:10.1080/20445911.2014.956748
- Kaushanskaya, M., & Marian, V. (2007). *Age-of-Acquisition Effects in the Development of a Bilingual Advantage for Word Learning. Proceedings of the 32nd Annual Boston University Conference on Language Development.*
- Kharkhurin, A. V. (2008). The effect of linguistic proficiency, age of second language acquisition, and length of exposure to a new cultural environment on bilinguals' divergent thinking. *Bilingualism: Language and Cognition*, 11(2), 225–243. doi:10.1017/S1366728908003398
- Kieffaber, P. D., & Hetrick, W. P. (2005). Event-related potential correlates of task switching and switch costs. *Psychophysiology*, 42(1), 56–71. doi:10.1111/j.1469-8986.2005.00262.x
- Kim, J.-J., Kim, M. S., Lee, J. S., Lee, D. S., Lee, M. C., & Kwon, J. S. (2002). Dissociation of working memory processing associated with native and second languages: PET investigation. *NeuroImage*, 15(4), 879–91. doi:10.1006/nimg.2001.1025
- Klein, D., Mok, K., Chen, J.-K., & Watkins, K. E. (2014). Age of language learning

- shapes brain structure: a cortical thickness study of bilingual and monolingual individuals. *Brain and Language*, *131*, 20–4. doi:10.1016/j.bandl.2013.05.014
- Kousaie, S., & Phillips, N. A. (2012a). Ageing and bilingualism: Absence of a “bilingual advantage” in Stroop interference in a nonimmigrant sample. *The Quarterly Journal of Experimental Psychology*, *65*(2), 356–369. doi:10.1080/17470218.2011.604788
- Kousaie, S., & Phillips, N. A. (2012b). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research*, *1446*, 71–90. doi:10.1016/j.brainres.2012.01.052
- Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience*, *8*(July), 250. doi:10.3389/fnbeh.2014.00250
- Kramer, R. (2011). *Effects of bilingualism on inhibitory control and working memory: a study with early and late bilinguals*. Florianópolis: Universidade Federal de Santa Catarina. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:EFFECTS+OF+BILINGUALISM+ON+INHIBITORY+CONTROL+AND+WORKING+MEMORY:+A+STUDY+WITH+EARLY+AND+LATE+BILINGUALS#5>
- Kroll, J., & de Groot, A. (2005). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford, New York: Oxford University Press. Retrieved from <http://psycnet.apa.org/psycinfo/2005-08338-000>
- Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon*. John Benjamins Publishing Company Amsterdam.

- Kroll, J. F., & Bialystok, E. (2013). Understanding the Consequences of Bilingualism for Language Processing and Cognition. *Journal of Cognitive Psychology*, 25(5), 37–41. doi:10.1080/20445911.2013.799170
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174. doi:10.1017/CBO9781107415324.004
- Kroll, J. F., Sumutka, B. M., & Schwartz, A. I. (2005). A cognitive view of the bilingual lexicon: Reading and speaking words in two languages. *International Journal of Bilingualism*, 9(1), 27–48. doi:10.1177/13670069050090010301
- Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism (Cambridge, England)*, 13(3), 373–381. doi:10.1017/S136672891000009X
- Kuipers, J. R., & Thierry, G. (2010). Event-related brain potentials reveal the time-course of language change detection in early bilinguals. *NeuroImage*, 50(4), 1633–8. doi:10.1016/j.neuroimage.2010.01.076
- Kumar Mishra, R. (2014). Let's not forget about language proficiency and cultural variations while linking bilingualism to executive control. *Bilingualism: Language and Cognition*, 18(October 2014), 39–40. doi:10.1017/S1366728914000455
- Salvatierra, J. L., & Rosselli, M. (2010). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism*, 15(1), 26–37. doi:10.1177/1367006910371021
- Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second

- language learning: anatomical changes in the human brain. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 58, 301–24.
doi:10.1016/j.cortex.2014.05.001
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–10.
Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16956095>
- Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Mental Lexicon*, 3(3), 349–374. doi:10.1075/ml.3.3.06lin.
- Linck, J., Osthus, P., Koeth, J., & Bunting, M. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861–83. doi:10.3758/s13423-013-0565-2
- Luck, S. (2014). *An introduction to the event-related potential technique* (Second Edi.). MIT Press.
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5), 605–621. doi:10.1080/20445911.2013.795574
- Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, 14(04), 588–595. doi:10.1017/S1366728911000010
- Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2011). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, 27(10), 1479–1488.
doi:10.1080/01690965.2011.613209

- Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29–41. doi:10.1016/j.cognition.2009.08.014
- 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*(01), 81–93. doi:10.1017/S1366728907003227
- Martin, C. D., Strijkers, K., Santesteban, M., Escera, C., Hartsuiker, R. J., & Costa, A. (2013). The impact of early bilingualism on controlling a language learned late: an ERP study. *Frontiers in Psychology*, *4*(November), 815. doi:10.3389/fpsyg.2013.00815
- Marzecová, A., Bukowski, M., Correa, Á., Boros, M., Lupiáñez, J., & Wodniecka, Z. (2013). Tracing the bilingual advantage in cognitive control: The role of flexibility in temporal preparation and category switching. *Journal of Cognitive Psychology*, *25*(5), 586–604. doi:10.1080/20445911.2013.809348
- McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society : JINS*, *16*(3), 495–505. doi:10.1017/S1355617710000093
- McNicol, D. (2005). *A Primer of Signal Detection Theory*. Psychology Press. Retrieved from <https://books.google.com/books?id=HPN5AgAAQBAJ&pgis=1>
- Mechelli, A., Crinion, J. T., Noppeney, U., O’Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain; Proficiency in a second language and age at acquisition affect grey-matter density. *Nature*,

431(7010), 757. doi:10.1038/431757a

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition*, 17(01), 89–117. doi:10.1017/S1366728913000084

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. doi:10.1177/0963721411429458

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

Miyake, A., & Shah, P. (1999). *Mechanisms of Active Maintenance and Executive Control*. (A. Miyake & P. Shah, Eds.) (Vol. 21). Cambridge University Press.
Retrieved from
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.6736&rep=rep1&type=pdf>

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
doi:10.1016/S1364-6613(03)00028-7

Mor, B., Yitzhaki-Amsalem, S., & Prior, A. (2014). The Joint Effect of Bilingualism and ADHD on Executive Functions. *Journal of Attention Disorders*.
doi:10.1177/1087054714527790

Moradzadeh, L., Blumenthal, G., & Wiseheart, M. (2014). Musical Training, Bilingualism, and Executive Function: A Closer Look at Task Switching and Dual-

- Task Performance. *Cognitive Science*, n/a–n/a. doi:10.1111/cogs.12183
- Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): an electrophysiological study of code switching. *Brain and Language*, 80(2), 188–207. doi:10.1006/brln.2001.2588
- Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, 21(6), 477–508. doi:10.1016/j.jneuroling.2008.01.003
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, 22(11), 1425–33. doi:10.1177/0956797611416999
- Moreno, S., & Farzan, F. (2015). Music training and inhibitory control: a multidimensional model. *Annals of the New York Academy of Sciences*, 1337(1), 147–152. doi:10.1111/nyas.12674
- Moreno, S., Wodniecka, Z., Tays, W., Alain, C., & Bialystok, E. (2014). Inhibitory control in bilinguals and musicians: event related potential (ERP) evidence for experience-specific effects. *PloS One*, 9(4), e94169. doi:10.1371/journal.pone.0094169
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *Journal of Psychology*, 81(2), 111–121. doi:10.1017/CBO9781107415324.004
- Nakao, Y., Kodabashi, A., Yarita, M., Fujimoto, T., & Tamura, T. (2012). Temporal activities during P3 components on the working memory-related brain regions: n-back ERP study. *Proceedings of the IEEE-EMBS International Conference on*

Biomedical and Health Informatics, 25(C), 424–427.

- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 383, 353–383. Retrieved from <http://www.sciencedirect.com/science/article/pii/0010028577900123>
- Neumann, O. (1996). Theories of Attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of Perception and Action* (Vol. 3, pp. 389–446). New York, NY: Elsevier. doi:10.1016/S1874-5822(96)80027-2
- Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience*, 24(5), 1205–23. doi:10.1162/jocn_a_00143
- Nichols, E. S. (2013). The influence of proficiency and age of acquisition on second language processing: An fMRI study of Mandarin-English bilinguals.
- Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, Affective & Behavioral Neuroscience*, 3(1), 17–26. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12822595>
- Nigg, J. T. (2003). Response inhibition and disruptive behaviors. *Annals of New York Academy of Sciences*, 2, 170–182.
- Noble, K., Norman, F., & Farah, M. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children, 1, 74–87.
- Olulade, O. a, Jamal, N. I., Koo, D. S., Perfetti, C. A., LaSasso, C., & Eden, G. F. (2015).

- Neuroanatomical Evidence in Support of the Bilingual Advantage Theory. *Cerebral Cortex*, 1–9. doi:10.1093/cercor/bhv152
- Ozen, L. J., Itier, R. J., Preston, F. F., & Fernandes, M. A. (2013). Long-term working memory deficits after concussion: electrophysiological evidence. *Brain Injury : [BI]*, 27(11), 1244–55. doi:10.3109/02699052.2013.804207
- Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology*, 26(3), 1–14. doi:10.1080/20445911.2014.891597
- Paap, K. R. (2015). The neuroanatomy of bilingualism: Will winds of change lift the fog? *Language, Cognition and Neuroscience*, 3798(October), 1–4. doi:10.1080/23273798.2015.1082607
- 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. doi:10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69, 265–278. doi:10.1016/j.cortex.2015.04.014
- Paap, K. R., & Liu, Y. (2014). Conflict resolution in sentence processing is the same for bilinguals and monolinguals: The role of confirmation bias in testing for bilingual advantages. *Journal of Neurolinguistics*, 27(1), 50–74. doi:10.1016/j.jneuroling.2013.09.002
- Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: problems

in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, 5(September), 1–15.

doi:10.3389/fpsyg.2014.00962

Paap, K., Sawi, O., Dalibar, C., Darrow, J., & Johnson, H. (2015). Beyond Panglossian Optimism: Larger N2 Amplitudes Probably Signal a Bilingual Disadvantage in Conflict Monitoring. *AIMS Neuroscience*, 2(1), 1–6.

doi:10.3934/Neuroscience.2015.1.1

Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 40(2), 313–25. doi:10.1037/a0035224

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 37(1), 51–87. Retrieved from

<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1996094975>

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., *et al.* (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain: A Journal of Neurology*, 121, 1841–52. Retrieved from

<http://www.ncbi.nlm.nih.gov/pubmed/9798741>

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–48. doi:10.1016/j.clinph.2007.04.019

Portocarrero, J. S., Burright, R. G., & Donovan, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical*

Neuropsychology, 22(3), 415–422. doi:10.1016/j.acn.2007.01.015

Posner, M. I., & DiGirolamo, G. J. (1998). Conflict, target detection and cognitive control. *The Attentive Brain*, 401–423.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology*, 58, 1–23. doi:10.1146/annurev.psych.58.110405.085516

Poullisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production* (Vol. 20). John Benjamins Publishing.

Prior, A., & Macwhinney, B. (2009). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13(02), 253. doi:10.1017/S1366728909990526

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: an ERP study. *Neuropsychologia*, 42(12), 1636–56. doi:10.1016/j.neuropsychologia.2004.04.013

R. Paap, K., M. Sawi, O., Dalibar, C., Darrow, J., & A. Johnson, H. (2014). The Brain Mechanisms Underlying the Cognitive Benefits of Bilingualism may be Extraordinarily Difficult to Discover. *AIMS Neuroscience*, 1(3), 245–256. doi:10.3934/Neuroscience.2014.3.245

Raven, J., Raven, J. C., & Court, J. H. (1998). Manual for Raven's progressive matrices and vocabulary scales. *Manual for Raven's Progressive Matrices and Vocabulary Scales*.

Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavior assessment system for children*. Bloomington, MN: Pearson Assessment.

- Rodriguez-fornells, A., Lugt, A. Van Der, Rotte, M., Britti, B., Heinze, H., & Mu, T. F. (2005). Second Language Interferes with Word Production in Fluent Bilinguals : Brain Potential and Functional Imaging Evidence, 422–433.
- Rodriguez-Fornells, A., Rotte, M., & Heinze, H. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, 415(February). Retrieved from <http://www.nature.com/nature/journal/v415/n6875/abs/4151026a.html>
- Rodriguez-Fornells, A., De Diego Balaguer, R., & Munte, T. F. (2006). Executive control in bilingual language processing. *Bilingual Language Processing*, 133–190. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9922.2006.00359.x/full>
- Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology. General*, 126(3), 211–27. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9281831>
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. doi:10.1037/0033-2909.86.3.638
- Roth, R. M., Isquith, P. K., & Gioia, G. A. (2005). *Behavior Rating Inventory of Executive Function-Adult Version™(BRIEF-A™)*. Lutz, FL: Psychological Assessment Resources.
- Roth, R. M., Lance, C. E., Isquith, P. K., Fischer, A. S., & Giancola, P. R. (2013). Confirmatory factor analysis of the behavior rating inventory of executive function-adult version in healthy adults and application to attention-deficit/ hyperactivity disorder. *Archives of Clinical Neuropsychology*, 28(5), 425–434.

doi:10.1093/arclin/act031

Sabourin, L., Brien, C., & Burkholder, M. (2013). The effect of age of L2 acquisition on the organization of the bilingual lexicon: Evidence from masked priming.

Bilingualism: Language and Cognition, 17(03), 542–555.

doi:10.1017/S1366728913000643

Sadeh, S. S., Burns, M. K., & Sullivan, A. L. (2012). Examining an executive function rating scale as a predictor of achievement in children at risk for behavior problems.

School Psychology Quarterly, 27(4), 236–246. doi:10.1037/spq0000012

Schmidt, T., & Vorberg, D. (2006). Criteria for unconscious cognition: three types of dissociation. *Perception & Psychophysics*, 68(3), 489–504.

doi:10.3758/BF03193692

Schwanenflugel, P. J. (1986). Interlingual Semantic Representational Facilitation : Evidence for a Common System in the Bilingual Lexicon, *618*, 605–618.

Sebastian-Galles, N., Echeverria, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals.

Journal of Memory and Language, 52(2), 240–255. doi:10.1016/j.jml.2004.11.001

Silver, C. H. (2014). Sources of data about children's executive functioning: review and commentary. *Child Neuropsychology: A Journal on Normal and Abnormal*

Development in Childhood and Adolescence, 20(1), 1–13.

doi:10.1080/09297049.2012.727793

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453.

doi:10.3102/00346543075003417

- Smith, J. L., Johnstone, S. J., & Barry, R. J. (2008). Movement-related potentials in the Go/NoGo task: the P3 reflects both cognitive and motor inhibition. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *119*(3), 704–14. doi:10.1016/j.clinph.2007.11.042
- Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: the effect on lexical access. *Memory & Cognition*, *12*(4), 380–6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6503701>
- Soliman, A. M. (2014). Bilingual advantages of working memory revisited: A latent variable examination. *Learning and Individual Differences*, *32*, 168–177. doi:10.1016/j.lindif.2014.02.005
- Soveri, A., Laine, M., Hämäläinen, H., & Hugdahl, K. (2010). Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm. *Bilingualism: Language and Cognition*, *14*(03), 371–378. doi:10.1017/S1366728910000118
- Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is There a Relationship between Language Switching and Executive Functions in Bilingualism? Introducing a within group Analysis Approach. *Frontiers in Psychology*, *2*(August), 183. doi:10.3389/fpsyg.2011.00183
- Stocco, A., Yamasaki, B., Natalenko, R., & Prat, C. S. (2012). Bilingual brain training: A neurobiological framework of how bilingual experience improves executive function. *International Journal of Bilingualism*, *18*(1), 67–92. doi:10.1177/1367006912456617
- Stroth, S., Kubesch, S., Dieterle, K., Ruchow, M., Heim, R., & Kiefer, M. (2009).

Physical fitness, but not acute exercise modulates event-related potential indices for executive control in healthy adolescents. *Brain Research*, *1269*, 114–124.

doi:10.1016/j.brainres.2009.02.073

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., & Hamer, L. (2000).

Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*, 388–402.

Sullivan, M. D., Janus, M., Moreno, S., Astheimer, L., & Bialystok, E. (2014). Early stage second-language learning improves executive control: Evidence from ERP.

Brain and Language, *139*, 84–98. doi:10.1016/j.bandl.2014.10.004

Szmalc, A., Brysbaert, M., & Duyck, W. (2012). Working memory and (second)

language processing. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism: Theoretical and applied approaches*. (pp. 74–94). Cambridge:

Cambridge University Press. doi:10.1080/13670050.2014.941727

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*(June), 123. doi:10.3389/fpsyg.2011.00123

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and

education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*(2), 167–177. doi:10.1016/S0887-6177(97)00095-4

Toplak, M. E., Sorge, G. B., Benoit, A., West, R. F., & Stanovich, K. E. (2010).

Decision-making and cognitive abilities: A review of associations between Iowa Gambling Task performance, executive functions, and intelligence. *Clinical*

Psychological Review, 30(5), 562–81. doi:10.1016/j.cpr.2010.04.002

- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 54(2), 131–143. doi:10.1111/jcpp.12001
- Trujillo-Orrego, N., Garcia-Barrera, M. A., Holroyd, C. B., & Pineda, D. A. (2011). Examining the feasibility of N200 as an endophenotype candidate in a genetic isolate of familiar type ADHD. [Abstract]. *Journal of the International Neuropsychological Society*, 3, 184.
- Tse, C., & Altarriba, J. (2014). The relationship between language proficiency and attentional control in Cantonese-English bilingual children: Evidence from Simon, Simon switching, and working memory tasks. *Frontiers in Psychology*, 5(September), 1–18. doi:10.3389/fpsyg.2014.00954
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104–32. doi:10.1037/0033-295X.114.1.104
- Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, 18(November 2014), 3–24. doi:10.1017/S1366728914000522
- Van Der Meij, M., Cuetos, F., Carreiras, M., & Barber, H. a. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology*, 48(1), 44–54. doi:10.1111/j.1469-8986.2010.01039.x
- van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language

- conflict in the bilingual brain. *Cerebral Cortex (New York, N.Y. : 1991)*, 18(11), 2706–16. doi:10.1093/cercor/bhn030
- Verhoef, K., Roelofs, A., & Chwilla, D. (2010). Electrophysiological evidence for endogenous control of attention in switching between languages in overt picture naming. *Journal of Cognitive Neuroscience*, 22(8), 1832–1843. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2009.21291>
- Wang, Y., Kuhl, P. K., Chen, C., & Dong, Q. (2009). Sustained and transient language control in the bilingual brain. *NeuroImage*, 47(1), 414–22. doi:10.1016/j.neuroimage.2008.12.055
- 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), 862–70. doi:10.1016/j.neuroimage.2006.09.054
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. Academic Press.
- Wodniecka, Z., Craik, F. I. M., Luo, L., & Bialystok, E. (2010). Does bilingualism help memory? Competing effects of verbal ability and executive control. *International Journal of Bilingual Education and Bilingualism*, 13(5), 575–595. doi:10.1080/13670050.2010.488287
- Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 5(1), 7–17.

Appendix A

**(Chapter 1: Table of bilingual advantage studies in adults for inhibition, switching,
and updating)**

Study	Participants	Age of Acquisition	Language Proficiency Measure	SES	Immigrant Status	Bilingual Language Usage	Relative language balance	Task	Index of EF	Difference in EF index
Inhibition										
Bialystok (2006)	(a) 19 video-game players BL ²⁸ (22.2 years) (b) 17 video-game players ML (21.6 years) (c) 30 non-video-game players BL (22.0 years) (d) 31 non-video-game players ML (22.0 years)	Early Before age 5	Self-rating (1–10): at least 6 (spoken L1)	Not reported; Undergrad students	Not reported	L1 at home, L2 school or work	Balanced	(i) Simon squares task (ii) Simon arrows task	(i) Incongruent vs. congruent trials (ii) Incongruent vs. congruent trials	(i) No (ii) No
Bialystok et al., (2005)	(a) 10 English monolinguals (22-36 years) (b) 10 French-English bilinguals (22-36 years) (c) 10 Cantonese-English bilinguals (22-36 years)	Early Before age 6	Not reported; “English fluency was equivalent between groups”	Not reported	Not reported	L1 at home, L2 (English) in the community/school	Balanced	Simon task	(i) Incongruent vs. congruent trials (ii) fMRI	(i) Mixed results (ii) Brain Regions Yes

Bialystok et al. (2004)	<p>Study 1 (a) 10 mid-age Tamil–English BL (43.0 years) (b) 10 mid-age English ML (43.0 years) (c) 10 older Tamil–English BL (72.3 years) (d) 10 older English ML (71.6 years)</p> <p>Study 2 (a) 32 mid-age Tamil–English (20) or Cantonese–English (12) BL (42.6 years) (b) 32 mid-age English ML (42.6 years) (c) 15 older English–Tamil (9) or English–French (6) BL (70.2 years) (a) 15 older English ML (70.4 years)</p> <p>Study 3</p>	<p>Study 1 Later in childhood from age 6</p> <p>Study 2 Later in childhood from age 6</p> <p>Study 3 Early from childhood</p>	<p>Study 1 PPVT-R score: (English) (a) 91.8 (b) 91.0 (c) 91.9 (d) 85.8</p> <p>Study 2 PPVT-III score: (a) 86.0 (b) 85.4 (c) 81.4 (d) 79.7</p> <p>Study 3 PPVT-III score: (a) 91.0 (b) 89.1</p>	<p>Study 1, 2, 3 “Shared similar middle-class SES backgrounds”</p>	<p>Not reported</p>	<p>Study 1 44.0% (L1) 56.0% (L2)</p> <p>Study 2 51.7% (L1) 48.3% (L2)</p> <p>Study 3 50.0% (L1) 50.0% (L3)</p>	<p>Study 1, 2, 3 Balanced</p>	<p>Study 1, 2, 3 Simon Task</p>	<p>Study 1, 2, 3 Incongruent vs. congruent trials</p>	<p>Study 1 Yes</p> <p>Study 2 Yes</p> <p>Study 3 Yes</p>
-------------------------	---	---	--	--	---------------------	--	-----------------------------------	-------------------------------------	---	--

	(a) 10 French-English BL (40.6 years) (b) 10 ML English (38.8 years)									
Bialystok et al., (2006)	Study 1 & 2 (a) 24 ML (20.7 years) (b) 24 BL (20.7 years) (c) 24 ML (70.4 years) (d) 24 BL (70.4 years)	Young adults Before age 6 Older adults Before age 12	Not reported	Not reported; Undergrad students	Not reported	L1 at home, L2 in the community	Balanced	Study 1 Antisaccade task Study 2 Modified Antisaccade task	Study 1 Eye gaze Study 2 Eye gaze	Study 1 Bilinguals No Age Yes Study 2 Bilinguals Yes Age Yes
Bialystok et al., (2008)	(a) 24 young BL (19.7 years) (b) 24 young ML (20.7 years) (c) 24 older BL (68.3 years) (d) 24 older ML (67.2 years)	Early Before age 6 Late Before age 20	Self-rating (0-4): (a) 3.15 (L1) 3.83 (L2) (c) 3.65 (L1) 3.79 (L2)	Not reported	Young BL: 14 were immigrants before age 6 Older BL: 20 were immigrants before age 12	Use both L1 and L2 daily	Balanced	(i) Simon arrows ii) Stroop color-naming task	(i) Incongruent vs. congruent trials (ii) Incongruent vs. congruent trials	(i) No (ii) Yes
Bialystok & DePape,	95 SS (23.8 years)	Not reported	Self-rating (1-	Not reported	Not reported	Use both L1 and L2	Balanced	(i) TMT B	(i) RT	(i) No

2009	(a) 24 ML (b) 24 BL (c) 22 Instrumentalists (d) 25 Vocalists		5): (b) English 4.25 (b) Other language 3.42			daily English: 56% Other Language: 44%		(ii) Simon arrows (iii) Auditory Stroop	(ii) Incongruent vs. congruent trials (iii) Incongruent vs. congruent trials	(ii) Yes (iii) Mixed
Bialystok et al., 2014; Study 1	Study 1 (a) 27 Young ML (20.3 years) (b) 44 Young BL (20.4 years) (c) 25 Older ML (71.3 years) (d) 34 Older BL (67.6 years)	Study 1 (b) 5.02 years (d) 8.04 years	Study 1 Self-rating of relative balance (1-5): (b) 4.2 (d) 4.6	Not reported	Not reported	Use L1 and L2 daily	Balanced	(i) Paper Stroop task	(i) Interference cost (Interference time – color naming time/color naming time)	(i) Yes
Blumenfeld & Marian, 2014	Study 1 (a) 30 BL (22 years) (b) 30 ML (21.4 years)	Study 1 (a) 2.9 years	Study 1 Self-rating (0-10) (a) L1 9.4; L2 7.7 (b) L1 9.6	Study 1 Not reported; University participants	Study 1 Not reported	Study 1 L1 (English) 79.7% L2 (Spanish) 20.5%	Study 1 L1 slightly more dominant	Study 1 (i) Nonlinguistic Stroop task (ii) Nonlinguistic Simon task	Study 1 (i) Efficiency Score (RT/accuracy) (ii) Efficiency Score (RT/accuracy)	Study 1 (i) Yes (ii) No
	Study 2 (a) 60 BL (21.7 years) (b) 60 ML (22.2 years)	Study 2 (a) 4.4 years	Study 2 Self-rating (0-10)	Study 2 Not reported; University participants	Study 2 Not reported	Study 2 L1 (Spanish) 35%	Study 2 L2 slightly more dominant	Study 2 (i) Nonlinguistic Stroop task	Study 2 (i) Efficiency Score	Study 2 (i) No (ii) No

			(a) L1 9.4; L2 8.5 (b) L1 9.8			L2 (English) 65%		(ii) Nonlinguistic Simon task	(RT/accuracy) (ii) Efficiency Score (RT/accuracy)	
Colzato et al. (2008)	Study 1 (a) 16 Dutch-English BL (22 years) (b) 16 Spanish ML (22 years) Study 2 (a) 18 monolinguals (22 years) (b) 18 bilinguals (22 years)	Study 1 & 2 Early From birth	Study 1 & 2 Self-rating (1 – 10); 8.9	Study 1 & 2 “Shared similar middle-class SES backgrounds”	Study 1 & 2 Not reported	Study 1 & 2 Use L1 and L2 daily	Study 1 & 2 Balanced	Study 1 Stop signal task Study 2 Inhibition of return	Study 1 Inhibition of response (SSRT) Study 2 Cuing cost	Study 1 No Study 2 Yes
Costa et al., (2009)	Study 1 (a) 60 Catalan-Spanish BL (20.1 years) (b) 60 Spanish ML (20.0 years) Study 2 (a) 62 Catalan-Spanish BL (20.1 years) (b) 62 Spanish monolinguals	All SS: early From early childhood	Seven-point scale (1 = only L2, 7 = only L1) Study 1 (5.3) Study 2 (5.0)	Not reported	Not reported	Study 1 & 2 Use L1 and L2 daily	Study 1 & 2 Balanced	Study 1 Attention Network Test (i) 8% congruent (ii) 92% congruent Study 2 Attention Network Test (i) 50%	Study 1. Incongruent vs. congruent trials Study 2 Incongruent vs. congruent trials	Study 1 (i) No (ii) No Study 2 (i) No (ii) Yes

	(20.7 years)							congruent (ii) 75% congruent		
Fernandez et al., 2013	(a) 13 Spanish/English BL (20.46 years) (b) 15 English ML (22.67 years)	All BL Early: Before age 6	Oral vocabulary subtest of BVAT ²⁹	Not reported	11 BL were immigrants	English at School/work Spanish at home	Not reported	Auditory Go/No Go ERP ³⁰ task	N2 ³¹ on NoGo trials	Yes
Fernandez et al., 2014	(a) 17 English ML (20.41 years) (b) 18 Spanish/English BL (22.06 years)	(b) Mean AoA = 6.22 years; Four BL learned L2 after age 9	Oral vocabulary subtest of BVAT	Self-reported Household income	11 BL were immigrants	English at School/work Spanish at home	Not reported	(i) Auditory Go/No Go ERP task (ii) Visual Go/No Go ERP task	(i) N2 on NoGo trials (ii) N2 on NoGo trials	(i) Yes (ii) No
Gathercole et al., 2014 ³² ; Study 2	(a) 20 English ML (b) 42 sequential English/Welsh BL	All BL – early learners (from birth)	Not reported	Not reported	Not reported	Measured; not reported	Not reported	(i) Simon task	(i) Accuracy (i) RTs	(i) No (i) No

²⁹ BVAT: Bilingual Verbal Ability Test - The Oral Vocabulary subtest is one of three subtests that make up this standardized instrument, which is used to assess English language proficiency and overall verbal ability in non-native English speakers. This test is administered in English to quantify English language proficiency.

³⁰ ERP = Event-related potential.

³¹ The N2 component is a negative-going deflection in the ERP waveform that peaks between 150-350 ms post-stimulus; greater N2 amplitude in NoGo trials is thought to reflect enhanced inhibition.

³² Gathercole et al., 2014 presented a large study with several more age groups; only the young adult age group is presented here

	(c) 23 simultaneous English/Welsh BL									
Hernandez et al. (2010); Study 1	(a) 41 Catalan–Spanish BL (20.9 years) (b) 41 Spanish ML (21.4 years)	Early From early childhood	Not reported	Not reported	Not reported	Seven-point scale (1 = only L1; 7 = only L2) 5.1	Balanced	Numerical Stroop Task	Incongruent vs. neutral (interference effect)	Yes
Kousaie & Phillips, 2012	(a) 38 young ML (22.5 years) (b) 35 Young English/French BL (23.7 years) (c) 25 older ML (68.9 years) (d) 20 older English/French BL (71.9 years)	All BL learned both languages before age 8	(i) Self-Report scale (1-5): (b) L1 = 4.9; L2 = 4.2 (d) L1 = 4.9; L2 = 4.6 (ii) Animacy Judgment Task ³³	Not reported	Non-immigrant sample	L1 (English) and L2 (French) used daily	Balanced	Stroop Task	RT for each condition: (i) Word reading (ii) Color naming (iii) Incongruent color naming	(i) Yes (ii) Yes (iii) Yes
Kousaie et al., 2014	(a) 40 English ML (21.48 years) (b) 30 French	All BL learned both languages	(i) Self-reported scale (1-5)	Not reported	Non-immigrant sample	Not reported	Balanced	(i) Stroop task (ii) Simon	(i) Stroop interference (Color naming –	(i) Yes (ii) No

³³ The animacy judgment task is used as an objective measure of relative second language (L2) proficiency. Bilingual participants are presented with nouns on a computer monitor and were required to decide as quickly and accurately as possible whether each noun referred to something living or nonliving.

	ML (21.8 years) (c) 51 French/English BL (21.49 years)	before age 13	(ii) Animacy Judgment Task					task (iii) The SART ³⁴	incongruent color naming) (ii) Simon interference (RT congruent – RT incongruent)	(iii) No
Salvatierra & Rosselli, (2010)	(a) 66 Young ML (25.88 years) (b) 67 Young English-Spanish BL (26.67 years) (c) 42 Older ML (63.40 years) (d) 58 Older English-Spanish BL (64.84 years)	Early Before age of 6	Boston Naming Test (English & Spanish) Five-point rating scale (1 = poor; 5 = fluent)	Not reported	Not reported	Daily use of L1 & L2	Balanced and unbalanced	Simon Task (i) Simple condition (ii) Complex condition	(iii) RT Incongruent vs. congruent trials	Young (i) No (ii) No Older (i) Yes (ii) No
Linck et al., (2008) ³⁵	(a) 20 Pre-immersion BL (20 years)	Late After age of 12	10-point self-rating scale (1 =	Not reported; University	Not reported	Depends on group and stage of	Depends on group and stage	Simon Task	Incongruent vs. congruent trials	ML vs. BL Yes ³⁶

³⁴ The SART was used as a measure of response inhibition. For this task, participants were presented with the digits 1 through 9 on the computer screen and were required to press the space bar in response to every number except the number 3, for which no response was required.

³⁵ The study by Linck et al. (2008) is more complicated than a simple bilingual advantage; please refer to their study for a more detailed description of the their hypotheses and results.

³⁶ Linck et al. (2008) stated that they were the first to demonstrate a bilingual advantage of inhibitory control for late bilingual.

	(b) 28 Immersion BL (20.5 years) (c) 50 Post- immersion BL (21.3 years) (d) 23 Classroom learning BL (21.3 years) (e) 41 monolinguals (19.5 years)		poor; 10 = excellent)	students		learning	of learning			
Luk et al., (2010)	(a) 10 BL (20 years) (b) 10 ML (22 years)	Later in childhood From age 6	Self- rating (1– 10): 7.1 (L1) 7.8 (L2) PPVT-III score: (English) (a) 94.8 (b) 105.8			Use L1 and L2 regularly	Balanced	Flanker task (i) Baseline (ii) neutral (iii) congruent (iv) incongruent	Neutral vs. congruent trials	No
Luk et al., 2011	(a) 43 Early BL (21.1 years) (b) 42 Late BL (21.3 years) (c) 38 ML (21.0 years)	Early = before age 10; Active bilingualis m = age 5 Late = after age 10; active bilingualis m = age 15	Self- rating (1– 10) PPVT-III score: (English)	Not reported	(a) 47% immigrant s (b) 79% immigrant s (c) 21% immigrant s	Variable	Early BL = balanced Late BL = L1 dominant	Flanker task (i) RT congruent – RT control (ii) RT incongruent – RT control (iii) RT congruent – RT	(i) No (ii) Yes (Early BL > Late BL + ML) (iii) Yes (Early BL >	

									incongruent	Late BL + ML)
Mor et al., 2014	(a) 20 Control ML (24.25 years) (b) 20 ADHD ML (24.35 years) (c) 20 Control BL (24.8 years) (d) 20 ADHD BL (25.15 years)	BL group = before age 9	LEAP-Q ³⁷ Self-rating scale (1-10)	Not reported; College students	Unreported portion of BL group immigrated from the Former Soviet Union before age 9	“Continuous use of L1 and L2”	Not reported	(i) Numeric Stroop task (ii) Simon Arrows/Spatial Stroop task	(i) RT congruent vs. incongruent (i) Accuracy congruent vs. incongruent (ii) RT congruent vs. incongruent (ii) Accuracy congruent vs. incongruent	(i) No (i) No (ii) No (ii) No
Tao et al., 2011	(a) 36 early BL (18.9 years) (b) 30 late BL (20.8 years) (c) 34 ML (20.4 years)	Early BL: before age 6 Late BL: After age 12	Self-rating scale (1-7)	Percentile Scores (48 th – 77 th %ile)	All BL were immigrants	(a) 25% L1; 75% L2 (b) 59% L1; 40% L2	(a) L2 dominant (b) Balanced	(i) The LANT ³⁸ ; flanker condition	(i) congruent vs. incongruent	(i) Yes; (c)<(a)<(b)

Attention Switching

Bialystok et	(a) 24 young	Early	Self-	Not reported	Young	Use both	Balanced	(iii) SART	(iii) RT	(iii) No
--------------	--------------	-------	-------	--------------	-------	----------	----------	------------	----------	----------

³⁷ Participants completed a Hebrew translation of the Language Experience and Proficiency Questionnaire (LEAP-Q), which includes questions regarding language exposure and ratings of spoken language proficiency (1-10).

³⁸ The Lateralized Attention Network Task (LANT) provides indices for the efficiency of alerting, orienting, and executive networks.

al., (2008); third task	BL (19.7 years) (b) 24 young ML (20.7 years) (c) 24 older BL (68.3 years) (d) 24 older ML (67.2 years)	Before age 6 Late Before age 20	rating (0–4): (a) 3.15 (L1) 3.83 (L2) (c) 3.65 (L1) 3.79 (L2)		BL: 14 were immigrants before age 6 Older BL: 20 were immigrants before age 12	L1 and L2 daily			(iii) Errors on target trials	(iii) No
Garbin et al., 2010	(a) 19 Catalan-Spanish BL (22.55 years) (b) 21 Spanish ML (20.9 years)	All participants were early learners (before age 4)	Self-report; otherwise not reported	Not reported	Non-immigrants	“Used both languages continuously throughout their lives” (p.1273)	Balanced	(i) Color-shape switch task (ii) MRI scans	(i) Switch cost RT (Switch vs. nonswitch) (i) Switch Cost Accuracy (ii) Switch Cost activation (Left & Right IFG ³⁹) (ii) Switch Cost Activation (Putamen)	(i) Yes (i) Yes (ii) No (ii) Yes
Kousaie et al., 2014;	(a) 40 English ML (21.48)	All BL learned	(i) Self-reported	Not reported	Non-immigrant	Not reported	Balanced	(iv) WCST ⁴⁰	(iv) Categories	(iv) No

³⁹ IFG = Inferior Frontal Gyrus

⁴⁰ WCST = Wisconsin Card Sorting Task

Study 4	years) (b) 30 French ML (21.8 years) (c) 51 French/English BL (21.49 years)	both languages before age 13	scale (1- 5) (ii) Animacy Judgment Task		sample				completed	
Marzecova et al., 2013	(a) Hungarian- Polish BL (27 years) (b) Hungarian ML (27.3 years)	BL group = L1 and L2 from birth	Self- report scale (1- 10)	Not reported	Not reported	L1 dominant (69% daily use)	Balanced	(i) Social category switching task (SCST) ⁴¹	(i) Switch vs. no switch RT (i) Swtich vs. no swtich Accuracy	(i) Yes (i) Yes
Mor et al., 2014	(a) 20 Control ML (24.25 years) (b) 20 ADHD ML (24.35 years) (c) 20 Control BL (24.8 years) (d) 20 ADHD BL (25.15 years)	BL group = before age 9	LEAP-Q Self- rating scale (1- 10)	Not reported; College students	Unreporte d portion of BL group immigrate d from the Former Solviet Union before age 9	“Continuou s use of L1 and L2”	Not reported	(i) TMT B (ii) Task- switching paradigm	(i) RT (i) Accuracy (ii) Switch cost: switch vs. nonswitch in mixed block (ii) Mixing cost ⁴²	(i) No (ii) No
Moradzadeh et al., 2014;	(a) 45 ML musicians	Not reported	Self- rating	Matched based on	Not reported	Variable	Variable	(i) Quantity/ Identity task	(i) Local switch cost:	(i) No

⁴¹ The SCST measures the flexibility of switching between social categories depending on the type of repetition of stimuli features from trial to trial: complete repetition (both features are repeated), partial repetition (1 of the features is repeated), or complete alternation (both features are different than in the previous trial).

⁴² The mixing cost is the difference in performance between trials in the single-task blocks and nonswitch trials in the mixed blocks.

first task	(b) 36 ML nonmusicians (c) 36 BL musicians (d) 36 BL nonmusicians		scale (1-5) K-BIT-2 ⁴³ Vocab subtest	mother's education level					RT switch vs. nonswitch	
Prior & MacWhinney, 2009	(Mage = 22.01) (a) 44 ML (18.7 years) (b) 44 BL (19.5 years)	BL group = before age 6	PPVT Self-rating scale (1-10)	Not reported; university students	Not reported	BL group: 73% English use daily	Balanced	(i) Color-shape switch task	(i) Switch Cost RT (Switch vs. no switch) (ii) Switch Cost Accuracy (ii) Mixing cost ⁴⁴ RT (ii) Mixing cost Accuracy	(i) Yes (i) No (ii) No (ii) No

Updating Working Memory

Moradzadeh et al., 2014; third task	(a) 45 ML musicians (b) 36 ML nonmusicians (c) 36 BL musicians	Not reported	Self-rating scale (1-5) K-BIT-2 ⁴⁵	Matched based on mother's education level	Not reported	Variable	Variable	(iii) Dual N-back task	(iii) Accuracy z-score ⁴⁶	(iii) no
-------------------------------------	--	--------------	--	---	--------------	----------	----------	------------------------	--------------------------------------	----------

⁴³ K-BIT-2 = Kaufman Brief Intelligence Test 2.

⁴⁴ Mixing costs were defined as the difference between the performance in the single-task blocks and the performance on non-switch trials of each task in the mixed-task blocks.

(d) 36 BL
nonmusicians

Vocab
subtest

(Mage =
22.01)

⁴⁵ K-BIT-2 = Kaufman Brief Intelligence Test 2.

⁴⁶ See study description for an explanation of how they derived this value

Appendix B

(Bilingual Questionnaire used in Chapters 2 & 3)

Name: _____ Participant Number (*provided by researcher*): _____

Age: _____ Gender: M F Height (in feet and inches): _____ Weight (in pounds): _____

Handedness: Right Left Ethnicity: _____

Do you have normal/corrected vision? Yes No

Do you have normal hearing? Yes No

Has a professional ever diagnosed you with a Traumatic Brain Injury (not including a concussion)? Yes No

Has a physician ever diagnosed you with any neurological conditions (e.g., epilepsy)?
Yes No

Previous Diagnoses? (Please circle all that apply)

ADHD	Yes	No
FASD	Yes	No
Autism	Yes	No
Any Other Developmental Disorder Please specify:	Yes	No
Learning Disorder/Disability Please specify:	Yes	No
Have you ever received learning assistance or extra accommodations at school? In what areas (e.g. Math, Reading, Writing)?	Yes	No
Have you ever had any speech/language difficulties?	Yes	No
Any Psychological Conditions? Please Specify:	Yes	No

Are you currently taking any prescription medication? Yes No

- If YES, which type of medication:
- How often do you take it:
- How much do you take each time:

Have you consumed any alcohol in the past 48 hours? Yes No

- Please Specify approximately how much:

What is your native language? _____

Are you Bilingual/Multi lingual? Yes No

If yes, please respond to the questions below:

What is(are) your second language(s)? _____

For each of the second languages that you know please rate how well you can speak, understand, write and read the language in the table below. Please indicate proficiency using a 1-10 rating scale.

Not at all proficient 1 2 3 4 5 6 7 8 9 10 Completely proficient

Second Languages	Speaking	Understanding	Writing	Reading

For the languages you have listed, please indicate the age at which they were learned, and if applicable, whether you learned them by formal lessons (e.g., at school or a course), or by informal learning (e.g., at home, at work, from friends, living in another country/province).

Languages	Age	Lessons (Y/N)	Duration of Formal Learning	Informal (Y/N)	Duration of Informal Learning
L2:			<1 yrs 1-5 yrs 6-12 yrs >12 yrs		<1 yrs 1-5 yrs 6-12 yrs >12 yrs
L3:			<1 yrs 1-5 yrs 6-12 yrs >12 yrs		<1 yrs 1-5 yrs 6-12 yrs >12 yrs

Athletic History

Please list up to 5 sports in which you compete on the table below.

Sport	Competitive Level (e.g., Club, Intercollegiate, etc.)	Position?	Years playing?	Currently in a season of competition?
1.				Yes No
2.				Yes No
3.				Yes No
4.				Yes No
5.				Yes No

Concussion History

Have you ever had a concussion: Yes No

If yes, starting with your **most recent** concussion, please fill out as much information below as possible:

How many concussions have you suffered: _____

<i>Approximate Date of Concussion (mm/dd/yy)</i>	<i>Did you lose consciousness?</i>	<i>Were you playing sport?</i>	<i>Did you see a Doctor?</i>	<i>How long did it take to return to play?</i>
1) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
2) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
3) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
4) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
5) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
6) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
7) _____	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____

Physical Activity Levels

We are interested in finding out about the kinds of physical activities that people do as part of their everyday lives. The questions will ask you about the time you spent being physically active in the last 7 days. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the vigorous activities that you did in the last 7 days. Vigorous physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

1. During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling?

___ days per week **If you performed no vigorous physical activities, skip to question 3.*

2. How much time did you usually spend doing vigorous physical activities on one of those days?

___ hours per day ___ minutes per day Don't know/Not sure

Think about all the moderate activities that you did in the last 7 days. Moderate activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

3. During the last 7 days, on how many days did you do moderate physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.

___ days per week **If you performed no moderate physical activities, skip to question 5.*

4. How much time did you usually spend doing moderate physical activities on one of those days?

___ hours per day ___ minutes per day Don't know/Not sure

Think about the time you spent walking in the last 7 days. This includes at work and at home, walking to travel from place to place, and any other walking that you have done solely for recreation, sport, exercise, or leisure.

5. During the last 7 days, on how many days did you walk for at least 10 minutes at a time?

___ days per week **If you performed no walking, skip to question 7.*

6. How much time did you usually spend walking on one of those days?

___ hours per day ___ minutes per day Don't know/Not sure

The last question is about the time you spent sitting on weekdays during the last 7 days. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

7. During the last 7 days, how much time did you spend sitting on a weekday?

___ hours per day ___ minutes per day Don't know/Not sure

Appendix C

(History and Bilingual questionnaire used for Chapter 4)

Name: _____ Participant # (provided by researcher): _____

Age: _____ Gender: MF Height (in feet and inches): _____ Weight (in pounds): _____

Handedness: Right Left Ethnicity: _____

Do you have normal/corrected vision? Yes No

Do you have normal hearing? Yes No

Has a professional ever diagnosed you with a Traumatic Brain Injury (not including a concussion)? Yes No

Has a physician ever diagnosed you with any neurological conditions (e.g., epilepsy)?
Yes No

Previous Diagnoses? (Please circle all that apply)

Are you currently taking any prescription medication?	ADHD	Yes	No
	FASD	Yes	No
	Autism	Yes	No
	Any Other Developmental Disorder Please specify:	Yes	No
	Learning Disorder/Disability Please specify:	Yes	No
	Have you ever received learning assistance or extra accommodations at school? In what areas (e.g. Math, Reading, Writing)?	Yes	No
	Have you ever had any speech/language difficulties?	Yes	No
	Any Psychological Conditions? Please Specify:	Yes	No

Yes

No

- If YES, which type of medication:
- How often do you take it:
- How much do you take each time:

Have you consumed any alcohol in the past 48 hours? Yes No

- Please Specify approximately how much:

