



**University of
Nottingham**

UK | CHINA | MALAYSIA

**MODULATORY EFFECTS OF LANGUAGE
CONTEXT ON EXECUTIVE FUNCTIONING IN MULTILINGUALS**

CHOO KAM KITT

Presented to the Faculty of Science and Engineering of
University of Nottingham Malaysia
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY IN PSYCHOLOGY

October 2024

I. Abstract

The bilingual advantage hypothesis argues that bi/multilinguals have improvements in executive functioning compared to monolinguals. It is speculated that the constant daily management of simultaneous language activations results in a far transfer from language control to domain-general executive functions. Decades of research on this hypothesis remain inconclusive. A potential source of variation is bi/multilinguals' language use habits. Depending on the sociolinguistic landscape, some bi/multilinguals may experience single-, dual- or dense code-switching contexts. According to the Adaptive Control Hypothesis (ACH), being primarily immersed in one of these language contexts may lead to different cognitive outcomes. It is unclear how the predictions of the ACH manifest in a highly integrated language culture like Malaysia, whose population is more likely to be multilingual than bilingual. Multilinguals from highly integrated language cultures rarely stay in one type of language context. This thesis explores the bilingual advantage hypothesis and the ACH within such context.

First, it is important to understand people's language background and attitudes to language learning before investigating the cognitive implications of knowing more than one language. In Chapter 3, the thesis explores the Malaysian public's language attitudes. Following this exploration, the thesis presents a series of studies employing increasingly sensitive measures of cognitive abilities and paradigms, beginning with a simple Flanker task measuring basic reaction time (Chapter 4), to an Attention Network Task (ANT) using eye-tracking (Chapter 5), and, finally, a 2-back task using Event-related Potentials (ERPs, Chapter 6). The main findings from the survey study are that the Malaysian public, but more specifically Chinese, highly educated and higher socioeconomic status, generally has a positive language attitude and that higher education levels and greater language proficiency are associated with more positive language attitudes towards multilingualism. In the experimental studies, there was no clear-cut evidence to support the bilingual advantage hypothesis nor a cognitive adaptation effect of rapidly changing language contexts. Consequently, the studies also did not support the use of language entropy as a better predictor of task performance compared to traditional measures of bi/multilingualism. These findings are discussed in relation to current theories and arguments surrounding language control in multilinguals.

For my younger self: We.

Did.

It.

II. Acknowledgements

I want to express my deepest gratitude to my supervisors, Dr Jessica Price and Dr Wong Hoo Keat, for their continuous support, guidance, and encouragement throughout my doctoral research. Our bantering in the EEG lab, and on the MRT ride after work made the whole PhD journey more grounded and less lonely. Your expertise, insightful feedback, and unwavering commitment to my academic development have been invaluable. I am sincerely thankful for the opportunities you have provided me to grow as a researcher and for your patience and dedication in helping me navigate the challenges of this journey. I am also grateful for all those times where I didn't want to do the presentations, the submissions, the conferences, but you made me do them anyways. Looking back, these helped me in finding my ground as a researcher.

I also owe a debt of gratitude to all the participants who were so willing and eager to contribute to this research. Thank you for spending time completing the extensive language background questionnaires and cognitively taxing experimental tasks.

To my friends, Xiu Ling, Navila, Kim, Jonas, Markus, and Moon, who shared many hours laughing, crying, and ranting over the seemingly endless amounts of work, thank you for your presence. Especially to Xiu Ling, our many writing sessions in KL have made the thesis writing year so much less lonely and surprisingly enjoyable. Our future careers may lead us down different paths in life, but I will always remember the six of you who have stuck it out with me.

Lots of love goes to my parents, Ben and Rebecca, your steadfast love and support throughout the years have single-handedly pushed me to accomplish this PhD. Thank you for dealing with my outbursts and my emotional imbalances with kindness and patience. Thank you for listening even if half of the things I say don't make sense to you. You both inspire me to be better, this thesis is dedicated to both of you.

Finally, I thank God for helping me not just survive but thrive through the four years. You've prompted me to begin the PhD back in HELP in 2021, and so it is only fitting that I conclude this journey with You.

Declaration

I declare that this thesis is my work carried out under the normal terms of supervision.


.....
Choe Kam Kitt

III. Table of Contents

I.	Abstract.....	2
II.	Acknowledgements	4
III.	Table of Contents	6
IV.	Abbreviations	11
V.	List of Tables.....	14
VI.	List of Figures.....	15
	Chapter 1: Introduction	17
1.1	Defining and Measuring Multilingualism	21
1.2	What Makes a Multilingual “Multilingual”?.....	25
1.2.1	<i>Emergence And Duration of a Bilingual Effect on Cognition</i>	27
1.3	Defining and Measuring a Cognitive Effect in Executive Function	28
1.3.1	<i>The Issue of Task Validity</i>	31
1.3.2	<i>Which Came First? The Multilingual Chicken or The Executive Function Egg?</i>	35
1.4	Theoretical Framework	36
1.4.1	<i>Behavioural Theories</i>	37
1.4.2	<i>Neurophysiological Theories</i>	44
1.4.3	<i>Rapid Modulatory Effects of Language Contexts on The Language-Cognition Relationship</i>	48
1.5	Research Context.....	52
1.6	Research Questions and Hypotheses	54
	Chapter 2: General Methodology.....	56
2.1	Covariates	56
2.1.1	<i>Socioeconomic Status</i>	56
2.1.2	<i>General Cognitive Ability.....</i>	57
2.2	Language Proficiency Measures.....	58
2.3	Language History Questionnaire 3.0	61
2.4	Absence of Control Groups	63
2.5	General Statistical Approach	64
2.5.1	<i>Within-subjects ANOVA</i>	66
2.5.2	<i>Bayesian Within-subjects ANOVA</i>	67
2.5.3	<i>Forward-stepwise Multiple Regression Analysis</i>	68
2.6	Eye-tracking Methodology	70
2.7	Electroencephalography (EEG) Methodology	72

2.8	Summary.....	74
Chapter 3: Understanding Attitudes towards Bi/Multilingualism and Language Learning Experiences in Malaysia 75		
3.1	Introduction	75
3.1.1	<i>The Influence of Age.....</i>	78
3.1.2	<i>The Influence of Socioeconomic Status (SES), Language Proficiency, and Education Level</i>	79
3.2	Language Attitudes in Malaysia.....	81
3.3	Measuring Attitudes Towards Multilingualism and Language Learning Experiences	83
3.4	Current Study.....	84
3.5	Method.....	85
3.5.1	<i>Design.....</i>	85
3.5.2	<i>Materials.....</i>	85
3.5.3	<i>Participants</i>	86
3.5.4	<i>Procedure.....</i>	87
3.6	Results	88
3.6.1	<i>Exploratory Factor Analysis</i>	89
3.6.2	<i>Regression Analyses</i>	92
3.7	Discussion.....	93
Chapter 4: Relationship Between Language Context and Linguistic Distance on Executive Functioning in Multilinguals.....99		
4.1	Introduction	99
4.2	The Current Study	101
4.3	Experiment 2A.....	101
4.4	Method.....	102
4.4.1	<i>Design.....</i>	102
4.4.2	<i>Materials</i>	102
4.4.3	<i>Participants</i>	105
4.4.4	<i>Procedure</i>	107
4.5	Results	108
4.5.1	<i>Forward-stepwise Multiple Regression</i>	110
4.6	Discussion.....	112
4.7	Experiment 2B.....	113
4.8	Method.....	114
4.8.1	<i>Design.....</i>	114

4.8.2	<i>Materials</i>	114
4.8.3	<i>Participants</i>	115
4.9	<i>Results</i>	116
4.9.1	<i>Forward-stepwise Multiple Regression</i>	118
4.10	<i>Discussion</i>	119
4.11	<i>Experiment 2C</i>	120
4.12	<i>Method</i>	121
4.12.1	<i>Design</i>	121
4.12.2	<i>Materials</i>	121
4.12.3	<i>Participants</i>	121
4.13	<i>Results</i>	123
4.13.1	<i>Forward-stepwise Multiple Regression</i>	124
4.14	<i>Discussion</i>	125
4.15	<i>Comparison Between Experiment 2A, 2B and 2C</i>	127
4.16	<i>General Discussion</i>	129
4.16.1	<i>Language Context Rapidly Modulates Executive Functions</i>	129
4.16.2	<i>Linguistic Distance and the Multilingual Effect</i>	130
4.16.3	<i>Predicting Flanker Task Performance with MLD Scores</i>	134

Chapter 5: Does Linguistic Contexts Affect Multilinguals' Performance on the Attention Network Task? An Eye-Tracking Study.....136

5.1	<i>Introduction</i>	136
5.2	<i>Current Study</i>	139
5.3	<i>Method</i>	142
5.3.1	<i>Design</i>	142
5.3.2	<i>Materials</i>	142
5.3.3	<i>Participants</i>	146
5.3.4	<i>Procedure</i>	148
5.4	<i>Results</i>	149
5.4.1	<i>Behavioural Results – RT & Accuracy</i>	149
5.4.2	<i>Behavioural Results - the Attentional Networks</i>	151
5.4.3	<i>Eye-Tracking Results – the Attentional Networks</i>	152
5.4.4	<i>Regression analysis</i>	153
5.5	<i>Discussion</i>	154
5.5.1	<i>Language Contexts and Cognitive Control</i>	156

Chapter 6: Can Language Entropy Predict ERP Differences in Multilinguals During a Working Memory Task? 162

6.1	Introduction	162
6.2	Current Study.....	165
6.3	Method.....	167
6.3.1	<i>Design</i>	167
6.3.2	<i>Materials</i>	167
6.3.3	<i>Participants</i>	170
6.3.4	<i>Procedure</i>	171
6.3.5	<i>EEG Preprocessing</i>	172
6.4	Results	172
6.4.1	<i>2-Back Behavioural Performance</i>	172
6.4.2	<i>ERP Analyses</i>	173
6.4.3	<i>Forward-stepwise Multiple Regression</i>	180
6.5	Discussion.....	180
6.5.1	<i>Role Of Multilingual Experience on ERP Components</i>	184

Chapter 7: General Discussion 186

7.1	Language Contexts and Cognitive Control	188
7.1.1	<i>Does Bilingual Language Control Necessarily Require Executive Control?</i> 190	
7.1.2	<i>Reconciling The Opposing Findings: Is There Cognitive Control in Language Control?</i> 195	
7.2	The Complexities of Measuring Multilingualism	196
7.3	Future Research	198
7.4	Implications and Significance	201
7.5	Conclusion.....	204

References 205

Appendix 1: Scales and Questionnaires 266

MacArthur's SSS Scale	266
Raven's Standard Progressive Matrices (Adapted).....	267
LAQ-LL	275
LexTALE English Word List.....	275
LexTALE Chinese Characters List.....	278
Malay Cloze Test.....	282

LHQ	285
Multilingual Switching Questionnaire (MSWQ)	292
Appendix 2: Priming Stimuli List	295
English Word Priming List (Experiment 2A, 2B, 2C and 3)	295
Pinyin Word Priming List (Experiment 2A & 3)	309
Malay Words Priming List (Experiment 2B)	315
Welsh Words Priming List (Experiment 2C)	321
NoLanguage Stimuli (Experiment 4)	326
English Dialogues (Experiment 4)	327
Mandarin Dialogues (Experiment 4)	329
Mixed Dialogues (Experiment 4)	332
Appendix 3: Error Rate Analysis Output for Experiment 2A, 2B, and 2C	335
Experiment 2A - ANOVA (Err)	335
Descriptives	335
Experiment 2B - ANOVA (Err)	335
Descriptives	336
Experiment 2C - ANOVA (Err)	336
Descriptives	337
Appendix 4: IES Calculation	338
Appendix 5: K-Means Cluster Analysis	339

IV. Abbreviations

Abbreviation	Definition
ACC	Anterior Cingulate Cortex
ACH	Adaptive Control Hypothesis
ANEW	Affective Norms for English Words
ANOVA	Analysis of Variance
ANT	Attention Network Task
AoA	Age-of-Acquisition
AX-CPT	Ax-Continuous Performance Task
AS	Antisaccade Task
BAPSS	Bilingual Anterior to Posterior And Subcortical Shift
BEPA	Bilingual Executive-Processing Advantage
BF	Bayes Factor
BICA	Bilingual Inhibitory Control Advantage
BLC	Bilingual Language Control
BM	Bahasa Melayu
BSWQ	Bilingual Switching Questionnaire
CorCenCC	<i>Corpws Cenedlaethol Cymraeg Cyfoes (National Corpus of Contemporary Welsh)</i>
CS	Contextual Switch Tendency
CSST	Color-Shape Switching Task
CTI	Cue-Target Interval
CVAW	Chinese Valence-Arousal Words
DLP	Dual Language Program
DLPFC	Dorsolateral Prefrontal Cortex
DRM	Dynamic Restructuring Model
EEG	Electroencephalography
EF	Executive Function
EFA	Exploratory Factor Analysis

ERP	Event-Related Potential
ES	Executive Speed
fMRI	Functional Magnetic Resonance Imaging
HEI	Higher Education Institution
IC	Inhibitory Control
ICA	Independent Components Analysis
IES	Inverse Efficiency Score
IOR	Inhibition of Response
ITI	Intertrial Interval
L1	First Language
L1S	L1-Switch Tendency
L2	Second Language
L2S	L2-Switch Tendency
L3	Third Language
L4	Fourth Language
LANT	Lateralized Attention Network Task
LAQ-LL	Language Attitudes Questionnaire for Language Learners
LD	Linguistic Distance
LEAP-Q	Language Experience and Proficiency Questionnaire
LexCHI	Lexical Test for Advanced Learners Of Chinese
LexMAL	Lexical Test for Advanced Learners Of Malay
LexTALE	Lexical Test for Advanced Learners Of English
LexTALE_CH	Lexical Test for Advanced Learners Of English (Chinese Version)
LHQ	Language History Questionnaire
LLE	Language Learning Experience
LRP	Lateralized Readiness Potential
LSBQ	Language and Social Background Questionnaire
MacArthur SSS Scale	Macarthur Scale of Subjective Social Status Scale
MLD	Multilingual Language Diversity
MOT	Multiple Object Tracking
MSWQ	Multilingual Switching Questionnaire

NHST	Null Hypothesis Significance Testing
OCD	Obsessive-Compulsive Disorder
PASA	Posterior-Anterior Shift in Aging
PD	Proportion Dominance
PFC	Prefrontal Cortex
PPSMI	Pengajaran Dan Pembelajaran Sains Dan Matematik Dalam Bahasa Inggeris
PPVT	Peabody Picture Vocabulary Test
PWA	People With Aphasia
ROWPVT	Receptive One-Word Picture Vocabulary Test
RPM	Raven's Progressive Matrices
RT	Reaction Time
SAS	Supervisory Attentional System
SAT	Speed-Accuracy Tradeoff
SD	Standard Deviation
SES	Socioeconomic Status
SLA	Second Language Acquisition
SMA	Supplementary Motor Area
SSRT	Stop-Signal Reaction Time
STEM	Science, Technology, Engineering, and Mathematics
SUBTLEX-CH	Subtitle Frequency: Chinese
SUBTLEX-UK	Subtitle Frequency: English
US	Unintended Switch Tendency
VA	Valence-Arousal
VIF	Variance Inflation Factor
VoM	Value of Multilingualism
WM	Working Memory
YoU	Years of Use

V. List of Tables

Table 1. Cognitive Demands across Language Contexts	43
Table 2: Participant Demographic Measures	87
Table 3: Individual Item Reliability Statistics	89
Table 4: EFA Factor Pattern Matrix	91
Table 5: Experiment 2A Language Background Statistics	106
Table 6: Experiment 2A – Estimated Marginal Means by Congruency, by Language, and by Congruency * Language (with Standard Error).....	108
Table 7: Experiment 2B - Language Background Statistics	115
Table 8: Experiment 2B – Estimated Marginal Means by Congruency, by Language, and by Congruency * Language (with Standard Error).....	117
Table 9: Experiment 2C Participant Demographics.....	122
Table 10: Experiment 2C – Estimated Marginal Means by Congruency, by Language, and by Congruency * Language (with Standard Error).....	124
Table 11: Summary of Past Studies and Their LD Classification	131
Table 12: Participant Demographic Measures (Range)	147
Table 13 Descriptives of RT and Accuracy (with Standard Deviations) separated by Language, Congruency, and Cue Type.....	151
Table 14: Experiment 3 Participant Demographics	170
Table 15: Descriptives of RT and Accuracy (with Standard Deviations).....	173

VI. List of Figures

Figure 1: Overview of ANOVA and Regression Approach	66
Figure 2: Lee and Wagenmakers (2014) BF Scale Classification Scheme	68
Figure 3: Distribution of Responses on the Language Learning Experience and Value of Multilingualism Subscale.....	88
Figure 4: EFA Scree Plot Comparison	91
Figure 5: Sequence of NoLanguage and Experimental Blocks	104
Figure 6: <i>Experiment 2A - Scatterplot of MLD Scores Against Flanker Effect Across Language Blocks</i>	111
Figure 7: <i>Experiment 2B - Scatterplot of MLD Scores Against Flanker Effect Across Language Blocks</i>	118
Figure 8: <i>Experiment 2C - Scatterplot of MLD Scores Against Flanker Effect Across Language Blocks</i>	125
Figure 9: Comparison of Congruent and Incongruent Condition Mean RT (ms) Against Language Contexts Across Experiments 2A and 2B and 2C, ...	128
Figure 10: Overview of The Proposed Attentional Control Hierarchy. Adapted from Bialystok and Craik (2022)	137
Figure 11: Overview of the Modified ANT	144
Figure 12: <i>Boxplot of No-Cue and Double-Cue RT (ms) and First Cue-Saccade Latency (ms) across Language.</i>	152
Figure 13: <i>Boxplot of Congruent and Incongruent trials in (A) RT (ms) and (B) Flanker-Fixation Duration (ms) across Language</i>	153
Figure 14: Overview of the 2-Back Task.....	168
Figure 15: Visual Representation of Anterior (Green) and Posterior (Orange) Regions of Interest and Their Electrode Makeup.....	173
Figure 16: ERP Distribution of Target Stimuli at Anterior Sites.	175
Figure 17: ERP Distribution of Target Stimuli at Posterior Sites.....	176
Figure 18: ERP Distribution of Non-Target Stimuli at Anterior Sites	177
Figure 19: ERP Distribution of Non-Target Stimuli At Posterior Site.....	178
Figure 20: Plot of d' Values Across Language Contexts.....	179

Figure 21. <i>Overview of the Word Recognition and Processing System as Discussed in The BIA+ Model.</i>	192

Chapter 1: Introduction

...to understand the behavior of a complex system we must understand not only the behavior of the parts but how they act together to form the behavior of the whole. It is because we cannot describe the whole without describing each part, and because each part must be described in relation to other parts, that complex systems are difficult to understand (Bar-Yam, 1997, p. 1)

This thesis investigates two key research questions. First, can different language contexts (i.e., no language, single language, dual language) adaptively modulate executive function in multilinguals? Throughout this thesis, language contexts are either experimentally induced (e.g., as seen in Chapter 4, 5 and 6 experimental methodology) or arise naturally through social interactions, further details will be discussed in Section 1.4.1. Second, does language entropy predict cognitive task performance beyond traditional measures of multilingualism? These questions guide the current research into how multilingual experience shapes cognitive processing. This research is also motivated by the underrepresentation of highly multilingual cultures in bilingual-cognitive literature. The literature heavily focuses on WEIRD (Western, Educated, Industrialized, Rich, and Democratic) samples, making it unclear how the bilingual advantage hypothesis¹ might manifest in non-WEIRD populations.

Importantly, while it has traditionally been referred to as the bilingual “advantage” hypothesis, the current scientific discourse favors the terms “effect” or “adaptation.” This shift reflects a growing understanding that bilingualism is not a simple monolingual versus bilingual dichotomy. Instead, language influences cognition for everyone, regardless of language background. Labeling bilinguals as “advantaged” implies a monolingual disadvantage. Different language profiles may lead to different cognitive outcomes. It could be misleading to frame one as inherently superior when there is no clear benchmark for a bilingual control group, an issue that will be discussed further in Section 2.4. In this thesis, the term “bilingual advantage hypothesis” will be used when referring to past research. However, where appropriate, the term “bilingual

¹ Although the term “multilingual advantage hypothesis” might be more suited for this thesis, the original term “bilingual advantage hypothesis” will be used to prevent confusion and ensure consistency with previous literature. It is important to note that both are interchangeable and similarly refer to the debate surrounding cognitive enhancements in language users who speak more than two languages.

effect" will be used to more accurately describe the relationship between bilingualism and cognition.

This chapter will provide an overview of the standard challenges in defining and measuring bilingualism/multilingualism, followed by a review of the behavioural and neurophysiological models that attempt to explain whether the bilingual advantage hypothesis exists. In this thesis, the term “bilingualism” specifically refers to those who know two languages. On the other hand, “multilingualism” broadly encompasses language users who have acquired one or more non-native languages in addition to their native (L1) language. To maintain continuity with past research, the original label of “bilingualism” will be used where appropriate. However, the focus of this thesis is on multilingualism.

In the past, the ability to speak two languages used to carry heavy negative connotations for a child’s cognitive development. For example, early research on bilingual children demonstrated deficiency in several areas of academic measures compared to their monolingual counterparts, including writing composition and grammatical proficiency (Saer, 1923), overall vocabulary (Barke & Williams, 1938), and articulation (Carrow, 1957). Bilingual children used to be labelled as “mentally confused” and “language handicapped” (Saer, 1923). Peal and Lambert (1962) challenged these claims in a landmark paper, arguing that the evidence against bilingualism was based on widespread methodological problems leading to inaccurate conclusions. For example, evaluating past papers on the bilingual advantage hypothesis before 1962, some studies compared monolinguals and bilinguals without accounting for the differences in linguistic and cultural backgrounds. In Saer (1923), the Welsh translation of the Stanford-Binet test was used to test Welsh monolingual children and Welsh-English bilingual children. However, Saer (1923) did not account for the variations in language ability of the bilingual children, with some having better proficiency in Welsh and others in English. In Pintner (1932), bilingualism was determined purely based on Anglo-Saxon names, that is, children with Anglo-Saxon names were categorized as English (i.e., Monolinguals), whereas those with foreign names were classified as non-English (i.e., Bilinguals). This categorization meant that the monolingual group was relatively homogenous, whereas the bilingual group was made up of children from varying cultural backgrounds.

Fortunately, following Peal and Lambert’s (1962) study, developments of new experimental procedures accompanying stronger methodological rigour led to different

conclusions about the effects of bilingualism on development (Grosjean, 1984). Many studies have found that bilinguals have greater working memory (Cockcroft et al., 2017; Fyndanis et al., 2023), inhibitory control (Blumenfeld & Marian, 2011; Kang & Lust, 2018), and conflict monitoring (Abutalebi & Green, 2007; Abutalebi et al., 2011; Grundy et al., 2017; Zhou & Krott, 2018) compared to monolinguals.

Perceptions of bilingualism are less negative today than they were in the 1900s, scepticism persists about whether bilingualism is cognitively a “burden”. This scepticism is often based on the scientifically unfounded, but pervasive, idea of the brain as a vessel with finite storage capacity (Bak & Mehmedbegovic, 2017). Further exacerbating this fallacy is the belief that monolingualism is the desirable default linguistic state, or what Blommaert et al. (2006) termed as “normative monoglot ideologies” (see also Rothman et al., 2022). For example, in Colorado and Massachusetts, the debate on whether to continue funding bilingual education in public schools was contested by public sentiment which heavily favoured a monolingual English education (Fitzgerald, 2011).

The sentiments differ from country to country. Much like the United States, Germany (Gogolin, 2021) opposes multilingual education, whereas others, like Norway (Haukås et al., 2022), embrace it. According to Titone and Tiv (2022), different historical, cultural, and social contexts from the overarching macrosystems (see Bronfenbrenner’s Ecological Systems Theory; Bronfenbrenner, 1986) trickle down through the various systems and into the belief system of an individual. This process affects how language is used and perceived as either a resource or a burden in different countries (see also Nguyen & Hamid, 2017). Bourhis (1997) described 17th-century France and its aggressive linguistic policies legislated against non-French languages, resulting in the emergence of a unilingual France from what used to be a multilingual France. Such policies were successful because French policymakers ensured that the French language emerged as the prestigious standard form of the language. At that time, francophones had little reason to learn a second language, as the economy, education, and governance were all conducted in standard French. This example illustrates how macro-level language policies can shape public perceptions of multilingualism over time. However, although understanding public sentiments towards multilingualism is important, it is equally vital to explore the factors that drive these changes. This knowledge provides critical insights into actionable steps policymakers can take to create a more

positive and receptive environment for multilingualism and to address potential barriers that may hinder its implementation in education.

These cautious sentiments towards bi/multilingualism persist despite meta-analyses that support the existence of a “bilingual cognitive advantage” (for review, see Grundy, 2020; van den Noort et al., 2019). Importantly, this effect has been found throughout the stages of life from infants (D’Souza et al., 2020) to children (Blom et al., 2014), young adults (Jiao et al., 2019), and into older adults (Baumgart & Billick, 2018; Salvatierra & Rosselli, 2010), indicating that the effects of bilingualism may not be an artefact of normal cognitive maturation. However, there have also been studies that found no bilingual effects (Dick et al., 2019; Morton & Harper, 2007; Paap & Greenberg, 2013) as well as negative bilingual effects (Paap et al., 2015; Ratiu et al., 2017).

How can this contradiction be explained? A closer inspection of the literature revealed a surprising amount of convergence. If there were truly no effects of bilingualism on cognition, we would see spurious effects in other cognitive domains in a seemingly unpredictable pattern (Bak, 2016). However, if a bilingual effect were to be discovered, it is considerably more likely to be in the executive domain than in any other cognitive domain. A systematic review by van den Noort et al. (2019) found that a majority of bilingualism-cognition studies published between 2004 and 2018 reported a bilingualism effect in executive functions. Poarch and Krott (2019) noted that it is rare to find studies reporting a monolingual effect on cognition relative to bilinguals. One would assume that if all studies reporting the bilingual effect are false positives, we should expect an equal number of false positives for a monolingual effect. However, publication biases and questionable research practices may influence these findings (de Bruin et al., 2015; Paap, 2023; van den Noort et al., 2019). Therefore, although there is consistent evidence supporting the bilingual effect, more scientific rigour is needed to address biases and questionable research practices.

The bilingual effect argues that the day-to-day management of multiple languages enhances these processes. Notably, as highlighted by Bialystok and Craik (2022), it is unlikely that bilingualism leads to an increased cognitive resource capacity, whatever an “increased” capacity may look like, as it is something yet unquantifiable (Hommel & Dreisbach, 2022). Instead, it is more likely that bilingualism affects the efficiency of cognitive resource allocation, as demonstrated by studies using functional magnetic resonance imaging (fMRI). For instance, in cognitive tasks, such as the flanker task (Abutalebi et al., 2011), the Simon task (Berroir et al.,

2017), and a task-switching paradigm (Gold et al., 2013), bilinguals across the lifespan exhibited lesser brain activation to attain similar or superior performance levels compared to monolinguals. This reduced activation has been interpreted as a sign of increased cognitive efficiency in bilingual individuals. However, Grundy et al. (2017) reported that when first learning a new language, multilinguals may counterintuitively demonstrate greater frontal lobe activation to devote greater resources to managing language competition. With increasing proficiency and experience, cortical structures should be optimally remodelled (see Section 1.4.2), ultimately leading to reduced frontal lobe activation.

1.1 Defining and Measuring Multilingualism

With the notion of what makes a multilingual a multilingual shifting slightly from culture to culture, how can one reliably measure a moving target? Commonly, researchers have taken to categorizing multilingualism into a binary yes/no category, either you are a monolingual or a multilingual (Kremin & Byers-Heinlein, 2021). Such a reductionistic conceptualization of multilingualism may work if language competency were purely based on comprehension of its linguistic code. Then, standardized language tests, with their arbitrary cut-off scores, may be used to determine monolinguals and multilinguals, because knowledge of the linguistic system, such as language proficiency, can be objectively measured. However, the functional system, made up of abstract, and sometimes subconscious, sociocultural norms is much harder, or even impossible, to measure. Therefore, to conceptualize this abstract system, researchers typically use language usage (i.e., how often the language is used daily) as a proxy measure.

In Surrain and Luk's (2017) systematic review of the labels and descriptions used to describe bilinguals, they found that most of the literature between 2005 to 2015 reported language usage (79%) and proficiency (77%) followed by language history (67%) and language of schooling (60%). The challenges of defining who qualifies as multilingual are even greater. Currently, there is no clear, agreed-upon definition of when a bilingual person should be classified as multilingual. Importantly, less than half of the observed studies objectively evaluated proficiency, and even fewer (30%) defined the sociolinguistic context of the sample's population. Theories that seek to explain whether a bilingual effect exists typically focus on the interplay between managing certain aspects of the multilingual experience (e.g., language proficiency and AoA) and the amount of experience managing those demands (Macnamara & Conway, 2014).

Unlike the linguistic system, there are no reliable standardized tests to measure these areas. Researchers typically rely on researcher-generated self-report questionnaires to obtain measures in such domains. Take, for example, the measure of language use. Traditionally, this domain has been used to categorize bilinguals as either “balanced” or “unbalanced/dominant.” A balanced bilingual is characterized by equal use and proficiency in both languages (Yow & Li, 2015). In contrast, if a bilingual uses one language more frequently or is more proficient in one language, they are labelled as an “unbalanced” or “dominant” bilingual. However, this dichotomy presents a conceptual problem, as achieving a complete balance in language use and proficiency is impossible (Meisel, 2007). Language use and proficiency are both closely tied to sociolinguistic factors and are not static variables; they evolve in response to the surrounding environment and its pressures.

The above arguments make one thing clear: there is no simple way to measure multilingualism that eliminates subjectivity. Ideally, all aspects of a multilingual experience should be measured, but this level of inclusion is unrealistic due to the interdisciplinary nature of multilingualism and the need to account for both individual and societal perspectives (Cenoz, 2013a). A promising approach is to focus on the most relevant factors, such as comprehensive language assessments, to provide deeper insights into how these variables interact and influence EF. However, this approach carries an unavoidable risk: depending on the researcher and their chosen measure, an individual labelled as multilingual in one study might be considered monolingual in another. This risk may be mitigated by using psychometrically validated language background questionnaires, which offer a standardized approach to capture the degree of bilingualism.

Although these questionnaires rely on self-ratings, they are generally more sensitive than a simple 7-point self-rating scale on language proficiency and fluency (Paap & Greenberg, 2013). Most language background questionnaires require participants to state all languages that they know, the age at which they acquired these languages, and their self-rated language proficiency (across the four key domains: reading, writing, listening, speaking). The questionnaires, based on these measures, can differentiate between passive (receptive) and active (productive) bilinguals, something a simple “Are you bilingual?” question cannot do. Passive bilinguals are defined by their bilingual comprehension competency but monolingual production ability. In contrast, active bilinguals more closely resemble the prototypical bilingual, with competency in both comprehension and production (Hoff, 2021). A more in-depth discussion of the effects of active

vs passive bilinguals on the bilingual effect will be discussed later in Section 1.3. However, quantifying this distinction is not straightforward. Much like the concept of “language contexts,” determining whether a bilingual is active or passive is challenging. The lack of research in this area makes it difficult to provide a clear answer, but these preliminary measures offer a richer language history background.

Other questions included in such questionnaires are socio-linguistic environments, such as language immersion (i.e., the context of when they use the languages they know) and dominance (i.e., which language they prefer to use in each situation). These questions also provide important details as to whether a bilingual is dominant or balanced in their languages. Dominant bilinguals generally have a stronger, preferred language compared to their other languages. Balanced bilinguals, on the other hand, are typically fluent in and use both languages equally. However, Silva-Corvalán and Treffers-Daller (2016) argue that language dominance does not exist because dominance 1) is tied to the domains and functions of language and 2) is not a permanent trait and will change over time depending on a range of factors. A discussion on language dominance extends beyond the scope of this thesis. However, it is important to recognize here that although language dominance is a measurable trait and can differentiate between types of bilinguals, it is not a perfect measure and should be interpreted with caution.

A major drawback of self-rating questionnaires is that there is a possibility of participants under- or overestimating their proficiency levels (Grosjean, 2006). This possibility may be remedied by using other measures to triangulate and verify participants’ self-rating, such as using pre-existing examination scores. Language background questionnaires offer several advantages over standardized language proficiency tests. First, unlike proficiency tests that may not be accessible in every language, questionnaires only require participants to rate their proficiency in each language they speak. Second, administering language background questionnaires is quick and cost-effective compared to standardized tests (Baker, 2011). Third, language background questionnaires yield a comprehensive measure of language use across different contexts. They capture various aspects of bilingual experiences, including daily language use, Age of Acquisition (AoA), and code-switching tendencies.

Three of the most used questionnaires are the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007), the Language and Social Background Questionnaire (LSBQ; Anderson, Mak, Chahi, & Bialystok, 2017) and the Language History Questionnaire

(LHQ; Li et al., 2006, 2020). First, the LEAP-Q includes a wide range of questions relevant to language use, the context of language acquisition, such as immigration experience, and the degree of language exposure (Marian et al., 2007). To validate the suitability of LEAP-Q, the authors compared self-reported proficiency scores with objective test scores and then identified three orthogonal factors (L1 competence, late L2 learning, and L2 competence) using factor analyses. L1/L2 competence incorporates proficiency and preference for using either language, whereas late L2 learning encompasses years of learning and time spent in an L2 community. Although comprehensive, the external validity of LEAP-Q may be compromised due to small sample sizes and the use of orthogonal rotation in the factor analysis because the factors need to be uncorrelated with one another. These limitations were recognized by the authors, and suggestions to update the questionnaire to bring it up to current standards have been proposed (Kaushanskaya et al., 2019).

Second, the LSBQ (Anderson et al., 2017; Luk & Bialystok, 2013) uses a similar factor-analytic approach. It asks additional questions regarding language use in specific contexts (i.e., at home or work, internet use, when accessing healthcare services) and if language use varies depending on who the individual is interacting with (i.e., friends, family, or neighbours). A bilingualism composite score is calculated for every participant based on language use in different contexts. The higher the composite score, the more “bilingual” a participant is. Based on this score, participants are then labelled as either monolingual (composite score below -3.13), bilingual (composite score above 1.23), or ambiguous (e.g., receptive bilinguals, composite score between -3.12 and 1.22). Importantly, the psychometric validation of this questionnaire was performed on a sample of monolinguals and bilinguals, but not multilinguals (i.e., those who spoke more than three languages). Hence, the LSBQ is suited for monolingual and bilingual, but not multilingual, populations. The bilingualism composite score provided also enables researchers to measure bilingualism on a continuum, thereby providing superior data flexibility in contrast to the LEAP-Q. Recently, Gullifer and Titone (2019) advanced the bilingualism composite score by quantifying the propensity to use two or more languages in specific contexts, which they refer to as “language entropy” (see *languageEntropy* R package, Gullifer, 2020). Higher values index more balanced language use and greater language diversity.

Lastly, similar to LEAP-Q, Li et al. (2006) developed the original LHQ by condensing commonly asked questions across 41 published language history and bilingualism studies that measure important theoretical constructs in bilingualism research. The most recent version of this

questionnaire (LHQ 3.0, Li et al., 2020) includes an automatic web-based scoring system, which calculates aggregated scores for language proficiency, language dominance, and language immersion levels; using different weights when calculating the scores of interests. Based on participants' AoA of languages, language proficiency, immersion and dominance, and dominance ratio, a Multilingual Language Diversity (MLD) Score can be calculated using a formula based on Shannon's Entropy, similar to Gullifer and Titone's (2019) language entropy. The MLD scores range from zero to two, with two indicating a higher degree of multilingualism. However, the calculations for MLD differ importantly from the *languageEntropy* package by additionally accounting for language proficiency. The accuracy between iterations in determining a person's degree of multilingualism remains unclear. However, as proficiency was considered in the MLD score but not in *languageEntropy*, MLD scores should provide a more accurate reflection of a person's degree of multilingualism compared to *languageEntropy*. Both versions of language entropy factors in multiple languages beyond the traditional L2-only approach. This means that multilingualism can be accounted for through the composite score generated from Shannon's Entropy, which accounts for AoA, proficiency, immersion, dominance across all the languages of an individual. This makes the *languageEntropy* or the MLD score highly appropriate for investigating how the bilingual effect is exhibited in multilinguals.

Nevertheless, being able to capture a single value to determine participants' degree of multilingualism will allow for a more objective measure of their linguistic experience and how this linguistic experience could influence executive functioning (EF). Notably, one important limitation of the LHQ is that it is not suited for children and older adults (Anderson et al., 2018).

Furthermore, as the criteria used to group participants as either high or low proficient may be arbitrarily determined, it compounds the complexity of cross-study comparisons (Lehtonen et al., 2018). One way to resolve this issue is to include proficiency test results as a predictive variable in the model. This inclusion would provide us with a richer set of data as it avoids categorizing participants into two dichotomous groups. Getting a deeper understanding of the bilingual experience could provide a spotlight as to why there are mixed findings within the literature.

1.2 What Makes a Multilingual “Multilingual”?

It is difficult to pin down one exact definition of what it takes to be a multilingual because multilingualism can be characterized by several different factors, such as language proficiency,

immersion, dominance, AoA, or code-switching habits (DeLuca et al., 2019). However, the way these factors are labelled and measured is not always consistent (Surrain & Luk, 2017). For the sake of brevity, current definitions of multilingualism typically resemble that of the European Parliament (2019) as “the ability of societies, institutions, groups and individuals to engage regularly, with more than one language in their day-to-day lives” (p. 2).

In the quest to explore the cognitive effects of multilingualism, the definition of multilingualism must be explicit and measurable, and the European Parliament’s vague and overarching definition of multilingualism simply does not suffice. How then should multilingualism be defined? Furthermore, suppose there was a unanimous agreement on a definition of multilingualism, it would be crucially as important to address how such a definition may be quantitatively measured. The truth is, that multilingualism is difficult to consistently define, akin to water taking the shape of the container it is placed in, so it is that the definition of multilingualism would be relative to the culture from where the question is posed (Freeman et al., 2022). However, according to Cenoz (2013b), there are generally two overarching categories defining multilingualism: maximalism and minimalism. Maximalists would consider multilinguals as those with native-like competence in two or more languages. Conversely, minimalists embrace a more inclusive stance, acknowledging individuals as multilingual despite potential limitations in their language skills, provided they possess the ability to engage in functional communication with speakers of those languages. Both ideas carry with them their problems, but notice that both camps of thought agree on the requirement of some level of multi-language competence to be labelled as multilingual.

According to Canale and Swain (1980; see also Canale, 1983), language competence comprises four sub-competencies: 1) grammatical, 2) discourse, 3) sociolinguistic and 4) strategic competence. The linguistic system is made up of both grammatical (i.e., knowledge of the language’s lexis, syntax, phonology, etc.) and discourse competence (i.e., ability to create larger bodies of text or speech concerning the cohesion and coherence of the conversation). On the other hand, the functional system of a language is made up of both sociolinguistic (i.e., social norms and conventions that dictate discourse acceptability) and strategic competence (i.e., verbal and non-verbal communication techniques for dealing with breakdowns in interactions). These two systems and their competencies must be accounted for when defining multilingualism, because a language is not only the totality of its linguistic codes but also the rules and social norms that bind and

dictate its use in communication. For this thesis, a minimalist approach to the European Parliament's definition of multilingualism will be assumed. Multilinguals in this thesis are defined as individuals who actively use and are functionally competent in multiple languages in their daily lives.

1.2.1 Emergence And Duration of a Bilingual Effect on Cognition

Much of past research categorized participants into discrete categories of “monolinguals” and “bilinguals” (Hofweber et al., 2020a). Given that bilingualism is a multi-dimensional construct consisting of a variety of experience-based factors (EBF), such as language proficiency, dominance, and immersion (DeLuca et al., 2019; Luk & Bialystok, 2013), binary classification of bilingualism would obscure important theoretical findings because idiosyncratic variations in language experience are discarded in favour of a simple, reductionistic definition of bilingualism (Bialystok, 2021). For example, variability in an individual's L2 AoA, whether they learned both languages simultaneously or sequentially, may carry meaningful implications on their cognitive development trajectory. Against this backdrop, it is crucial that future research respects and integrates the inherent variability of bilingualism into their study designs. Conducting research that specifically exploits systematic variations in bilingual language experience could reveal new insights that might have been missed in traditional group comparison designs.

Recently, researchers have begun to shift to this new paradigm (DeLuca et al., 2019; Zirnstein et al., 2019; Sulpizio et al., 2020, Gullifer et al., 2020; van Heuven & Coderre, 2015; Grundy, 2020). This shift in perspective reflects a growing recognition that a binary approach is neither practical nor representative of the complex and nuanced nature of multilingualism (Rothman et al., 2022, but see Kremin & Byers-Heinlein, 2021, for an alternative account). For example, Hofweber et al. (2020a, 2020b) identified code-switching habits as a moderator of the bilingual effect, whereby individuals who frequently engaged in code-switching displayed better EF than those who infrequently code-switch (see Section 1.4.1 for more information on code-switching). Similarly, López-Penadés et al. (2020) also found that code-switching habits predicted EF abilities, above and beyond typical measures of bilingualism (i.e., L2 AoA and proficiency).

Gullifer and Titone (2019) and Li et al. (2021) describe bilingualism in terms of Shannon's Entropy (also called Information Entropy). Entropy is a concept from physics used to describe the randomness or uncertainty of a system; on the other hand, Shannon's Entropy (Ben-Gal & Kagan,

2021) describes the randomness or uncertainty of a system to produce a piece of specific information. If the information is completely random and unpredictable, it is said to have high Shannon's entropy; if the information is perfectly predictable, it is said to have low Shannon's entropy. Applying this concept to bilingualism, language entropy (Gullifer & Titone, 2019; Li et al., 2021) is a measure of "randomness" - or, in this case, "diversity" - in language use for an individual or setting. The likelihood of bilinguals using either one of their languages at any given time is much more random than a monolingual who is likely to use one language only always, making their language use perfectly predictable. Therefore, it can be said that bilinguals have higher language entropy than monolinguals. This "randomness" in language use raises important questions regarding bilingualism's effect on cognition, specifically about executive function (EF). To explore this variability, it is essential to first define cognition and understand why certain cognitive processes may be advantageous or disadvantageous.

1.3 Defining and Measuring a Cognitive Effect in Executive Function

“Cognition is a process (broadly speaking, a chain of events) in which bits of information are internalized (e.g., some specific type of input derived from sensory stimulation) and then get manipulated (processed, computed) in certain ways that are (hypothesized to be) necessary for supporting the agent's behaviour now or later.” (de Oliveira & Oliveira, 2022, p. 9)

The broader concept of EF is typically made up of inhibitory control, task-switching, and working memory cognitive processes (Miyake et al., 2000). If a cognitive task taps into these components of EF, then, in an experiment where two groups engage in the task, the group with more efficient components would deplete their cognitive resources slower than the other. This discrepancy, in most cases, would lead to observable differences in task performance (Ross et al., 2014). Such experimental designs leverage a widely accepted idea that cognition is a limited but renewable resource (Hommel & Dreisbach, 2022).

In the literature, there are a variety of tasks that tap into different aspects of EF and the design of these tasks tends to differ from one study to another. Traditionally, these tasks have been used to assess the bilingual effect. For instance, in addition to the original Attention Network Task (ANT; Fan et al., 2002; see Section 4.6 for an overview of this task), designed to evaluate orienting,

alerting, and executive functioning abilities, there are also modified versions like the revised ANT (ANT-R; Fan et al., 2009), which presents the flanker stimuli to the left and right of the fixation cross. There is also the Lateralized ANT (LANT; Greene et al., 2008), which presents the flanker stimuli not only at the sides but also vertically. It is therefore important to select the appropriate task and its psychophysical properties because even nuanced differences may significantly affect participants' performance. Examples of such subtleties are the number of practice trials (Bialystok et al., 2004), the ratio of congruent to incongruent trials (Lehle & Hübner, 2008), and the magnitude of cue-target intervals (CTIs; Stasenko et al., 2017). It is thus not surprising that the literature on the bilingual effect is filled with mixed findings, which has led researchers to argue that this effect is inconsistent and task-specific (Paap & Greenberg, 2013; Paap et al., 2015; Ware et al., 2020).

Tasks measuring inhibition can be grouped largely into two; one that requires participants to inhibit a distractor stimulus (interference suppression) and one that requires an inhibition of habitual or prepotent responses (response inhibition). Interference suppression is induced by presenting two conflicting stimuli, such as in the flanker task and ANT, whereas response inhibition arises from a stimulus-response conflict, such as in Stroop, Simon, and go/no-go tasks (also known as the stop-signal reaction time (SSRT) task; Brunetti et al., 2019). Although both types of tasks involve top-down inhibitory control, the neurological pathways recruited during task performance are different, as demonstrated in fMRI studies (e.g., Bunge et al., 2002). When presented with two conflicting stimuli, one needs to process both stimuli before deciding on which stimulus to focus on and respond to, resulting in a heavier reliance on attentional control. An example of such a task is the incongruent condition of the flanker task, where a target arrow is flanked by two arrows pointing in the opposite direction.

In contrast, response inhibition tasks, such as Stroop, present only one stimulus in every trial (i.e., a colour word, such as red, which may be printed in blue ink in the incongruent condition). The habitual response in Stroop is to read the word itself, but the task requires the participants to name the ink colour instead. The conflict is thus resolved by suppressing the automatic behaviour. Popular variations of the Stroop task include day-night Stroop (Escobar et al., 2018; also adapted to present animals and their sounds, called Silly Sound Stroop, Verhagen et al., 2017) and number Stroop (Antón et al., 2019; Hernández et al., 2010; Nielsen et al., 2019). These variations may be used when developmental factors need to be controlled, or when

researchers are interested in a specific cognitive process that is not readily measured in the original task design. For example, a young child may not be developmentally ready to produce the colour names, hence, a day-night Stroop version may be more appropriate. Whereas the original Stroop task measures language processing to some extent, the number Stroop task measures non-linguistic numerical cognition.

The basic idea behind Stroop variations is that participants need to inhibit habitual responses and produce a task-related response instead. In the Simon task, participants need to discriminate one stimulus (e.g., press the left key if a square appears) from another (e.g., press the right key if a circle appears). A stimulus-response conflict is created by presenting the stimulus on the opposite side of the screen as the response key, that is, showing a left-response stimulus on the right (and vice versa). Response inhibition may also be induced in a detection task, such as SSRT, where participants must press a key in response to a particular stimulus and not respond when a secondary stimulus is presented (e.g., Luk et al., 2010; Sullivan et al., 2014).

These two types of inhibition mirror the daily cognitive load faced by bilinguals. Bilinguals engage in interference suppression (i.e., suppressing cross-linguistic interference during language selection) and response inhibition (i.e., suppressing the production of the irrelevant language during speech). Previous research suggests that active bilinguals may reap greater long-term cognitive benefits because they often engage in both types of inhibition in linguistically diverse environments, which requires significant cognitive effort (Abutalebi & Green, 2007; Elliot et al., 2024; Rentería et al., 2019). In comparison, passive bilinguals may only need to engage in interference suppression during semantic processing and may not benefit to the same extent as active bilinguals. The majority of the

To test for participants' ability in switching, researchers typically use the dimensional change card sort (DCCS; Hillyard, 2009) task and switch task, in which participants need to flexibly switch between several rules of the task (e.g., either sort cards based on shape, number or colour; Carlson & Meltzoff, 2008; Javan & Ghonsooly, 2017; Park et al., 2018; Ross & Melinger, 2017). In switch tasks, participants need to keep two rules in mind and flexibly use either of these rules to respond accurately. Switch tasks can be divided into linguistic and non-linguistic versions, both of which are administered by Prior and Gollan (2011). In the non-linguistic switch task, participants were requested to discriminate two stimuli in terms of shape or colour, whereas, in the linguistic switch task, participants were asked to name digits in either their first language (L1) or

second language (L2). A variation of the switch task that incorporates spatial information is the heart-flower task, where participants must click on the key on the same side as the stimulus presentation when a heart is present but click on the key on the opposite side when a flower is present (Jaekel et al., 2019). As this task also required participants to inhibit habitual response (i.e., press on the opposite key), it is akin to a mixed presentation of congruent and incongruent conditions of the day-night Stroop task.

Finally, working memory is frequently measured using tasks, such as the Corsi block and digit span tasks. In a Corsi block task (or a similar version called Frog matrix; Thomas-Sunesson et al., 2018), participants are presented with several squares, some of which will light up sequentially in a random order. The task is to remember the sequence of the lighted squares (Antoniou et al., 2016; Calvo & Bialystok, 2014; Purić et al., 2017). The difficulty of the task can be modified accordingly, such as only requesting participants to remember which squares lighted up (i.e., the order is not important) or presenting numbers or pictures on the squares instead of blank squares (the former presentation style will likely aid remembering of squares). Due to such differences in task designs, a bilingual effect may or may not be observed. Digit span tasks are used to measure working memory without tapping into spatial memory, where participants are presented (visually or auditorily) with a series of digits, after which participants need to repeat either in the same order (forward span) or the reverse order (backward span). Therefore, it is crucial to consider the specific cognitive processes being assessed by the experimental tasks employed and their relevance to bilingual language control when investigating the presence or absence of a bilingual effect.

1.3.1 The Issue of Task Validity

According to Martin-Rhee and Bialystok (2008), it is important to distinguish between stimulus-inhibition and response-inhibition tasks because dealing with two competing languages is akin to resolving two conflicting stimuli but not the same as inhibiting habitual responses. Therefore, the bilingual effect should prevail only in stimulus-inhibition tasks and not response-inhibition tasks. Such a distinction between inhibitory tasks is widely accepted (e.g., Kałamała et al., 2018), and Luk et al. (2010) supported the argument that the bilingual effect is observed in interference suppression but not response inhibition. Moreover, Thomas-Sunesson et al. (2018) obtained a bilingual (but not language proficiency) effect on flanker task and frog matrix but not

SSRT (although they did not recruit monolinguals and only compared bilinguals of varying L2 proficiency levels).

Yet, researchers do not always agree on the classification of tasks. Martin-Rhee and Bialystok (2008) argued that in the Simon task, one needs to inhibit conflicting stimuli, whereas other studies classified the Simon task as a stimulus-response conflict task (e.g., Blumenfeld & Marian, 2014; Valian, 2015). In addition, Blumenfeld and Marian (2014) and Verhagen et al. (2017) demonstrated a bilingual effect in the Stroop task, which arguably should not have occurred, according to Martin-Rhee and Bialystok's (2008) line of interpretation, because the Stroop task is a response inhibition task. It should be noted that inconsistencies in defining tasks are found even within the field of EF and are not unique to the bilingualism-cognition field. For instance, Peterson et al. (2002) found that both Simon and Stroop tasks activate similar brain regions (anterior cingulate cortex (ACC) and dorsal prefrontal cortex (DLPFC), among others) at similar time points, whereas another imaging study found that ACC is activated at different time points for flanker, Simon, and Stroop tasks (Fan et al., 2003). This variation illustrates the importance of considering task impurity when discussing the results obtained. As argued in Section 1.4.2, EF tasks do not tap exclusively on a single EF subcomponent. Hence, it is important to ensure that the tasks administered appropriately measure the intended cognitive ability. To further illustrate the importance of task validity, the heart-flower task used by Jaekel et al. (2019) did not yield any bilingual effect, leading to the conclusion that there is no bilingual effect in switching. Yet, this task also taps into response inhibition, which is purportedly not influenced by bilingualism. To address these issues, Paap and Sawi (2014) advised researchers to test a task's convergent validity through cross-task correlations that are assumed to measure the same component of EF. However, such methodology is typically resource intensive and time consuming. Researchers should balance the need to validate convergent validity and the need to answer specific research questions. Although addressing both is ideal, it may not always be practical.

It is also critical to take note of differences in task designs. First, the practice effect is an important aspect to consider, as Bialystok et al. (2004) demonstrated that after several test blocks, monolinguals' performance improves to be comparable to that of bilinguals. This effect is replicated in Costa et al. (2009), where a bilingual effect is only observed in the first test block but not in subsequent test blocks. Second, task difficulty impacts whether a bilingual effect will be observed. In terms of the ratio of congruent to incongruent trials, studies have shown that in high

monitoring conditions of the flanker task, that is, a roughly equal amount of congruent and incongruent trials, a bilingual effect is more likely to be observed (Costa et al., 2009). That the ratio of congruency affects task difficulty is also reported by Lehle and Hübner (2008), who found that a high number of congruent trials leads to less selective attention and thus a larger congruency effect. Furthermore, Keye et al. (2009) illustrated that switch trials lead to more conflict. Therefore, it is contended that the bilingual effect is best investigated using tasks that are cognitively demanding (Costa et al., 2009; Qu et al., 2015). In terms of the cue-target interval (CTI), a longer interval makes the task easier as it provides a longer preparation time for the participants, resulting in global improvements in reaction time (RT) that is unrelated to bilingualism (Hübner & Töbel, 2019; Stasenko et al., 2017). Introducing preparation time, such as forcing a delay between stimulus presentation and response, in response inhibition tasks allows the prepotent response to be more easily overridden (Coderre & van Heuven, 2014; Diamond et al., 2002; Ling et al. 2016; Simpson et al., 2012).

To this point, the bilingual effect has been discussed through a cross-sectional perspective, that is, those who are already bilingual are pitted against those who are already monolinguals in various executive function domains. Clear insights may be garnered from studies investigating the gradual emergence of the bilingual effect from a longitudinal perspective. Ljungberg et al. (2013) investigated the bilingual effect on episodic memory recall, verbal letter, and categorical fluency in 178 bilinguals and monolinguals from the Betula Prospective Cohort Study (age baseline between 35-70 years). Results indicate that bilinguals outperformed monolinguals in episodic memory recall and letter fluency across five test waves: T1 (1988–1990), T2 (1993–1995), T3 (1998–2000), T4 (2003–2005), T5 (2008–2010). However, there were no differences in performance in the categorical fluency task. Given the 20-year gap between T1 and T5, one could also argue that variation in lifestyles may have also contributed to portions of the variance in the outcomes (Judd et al., 2020; Olszewska et al., 2021; Peng & Kievit, 2020; Zhao et al., 2020). With a shorter timescale, the confounding effects of lifestyle choices may be minimized.

In Bak et al. (2016), 33 young to older adults (18–78 years) completed a three-week intensive Gaelic language learning course for non-Gaelic speakers (authors did not make it clear whether the participants were monolinguals or bilinguals to begin with). Results demonstrated that after the three-week language learning period, all participants demonstrated significant improvements in their attention-switching abilities from baseline compared to active and passive

control groups. A follow-up test after nine months on half of the original sample revealed that those who were actively learning Gaelic for more than 5 hours per week retained their improved performance. Those who learned for less than 4 hours per week showed an inconsistent pattern: some improved, whereas others deteriorated from baseline. The conclusion was that short-term language learning may modify EF. However, it is important to highlight that such improvements are not static and may disappear unless continually practiced.

Together, these studies imply that bilingualism, acquired either through short-term language learning or life-long experiences, can contribute to differences in cognition. Furthermore, the dynamic nature of cognitive improvements observed in Bak et al.'s (2016) study highlights the importance of sustained practice and engagement to maintain cognitive gains over time. Overall, these findings underscore the complicated interactions between bilingualism and cognitive functioning.

Although there is supporting evidence for a bilingual effect, there have been studies that also found no bilingual effects (Antón et al., 2019; Dick et al., 2019; Morton & Harper, 2007; Paap & Greenberg, 2013; Paap et al., 2015) and negative bilingual effects (Folke et al., 2016; Ratiu et al., 2017; Sadat et al., 2012). Lehtonen et al. (2018) and Paap et al. (2021) cite issues on power, sample size, and unclear theoretical frameworks as the reasons for contradictory evidence. They also concluded that a publication bias may have been the driving factor behind the disproportionate number of papers supporting a bilingual effect. This claim was based on the landmark meta-analysis by de Bruin et al. (2015), who found that conference abstracts that reported a bilingual effect were twice as likely to be published compared to those that reported null or negative effects. In addition, the meta-analysis by Lehtonen et al. (2018) suggests an observable, albeit small, evidence for a bilingual effect, but this effect disappeared when accounting for the abovementioned publication bias.

In response, Bialystok et al. (2015) highlighted three crucial errors that undermined the publication bias argument. First, the entire meta-analysis of de Bruin et al. (2015) was based on conference abstracts, which is not comparable to the actual publications or their peer-review processes. Concluding that there is a publication bias solely on conference abstracts is misleading. Second, null and negative effects were classified as the same kind of evidence. Such classification bias, combined with the exclusion bias (i.e., placing disproportionate emphasis on some types of literature at the expense of others), likely overinflated the estimations of publication bias (Leivada,

2023). Furthermore, according to Dentella et al. (2024), less than 50% of the studies that shows up with the search term "bilingual disadvantage" reported exclusively a disadvantage (i.e., bilinguals have worse task performance than monolinguals). Lastly, the decision to include studies using verbal tasks in the meta-analysis, given the reliable finding of a bilingual disadvantage on verbal tasks (Bialystok, 2009), was unjustified.

At this point, it is crucial to understand that while meta-analyses and systematic reviews provide a comprehensive overview of the bilingual effect debate, they do not represent the absolute truth. Implicit biases of the studies sampled in these reviews may be amplified, meaning that the results of the reviews are "objective" only to the extent that they accurately reflect the current biases of the literature. While their outcomes are certainly important to consider, they should not be treated as definitive truths (Bak, 2013).

1.3.2 Which Came First? The Multilingual Chicken or The Executive Function Egg?

Another crucial question that has not received sufficient attention in the literature is the direction of causality of the multilingual-cognition relationship. Like the chicken or egg dilemma, does multilingualism enhance EF or are high EF individuals more likely to be multilingual? Jackson (2020) argues that it is less about which comes first, and more about understanding the interconnectedness between multilingualism and EF as such variables are not mutually exclusive. Preliminary evidence suggests that the multilingual-cognitive relationship is positive and bidirectional (Cox et al., 2016; Grundy & Timmer, 2016). For example, the initial stages of L2 learning are supported by EF, such as working memory (WM). However, in later life, as cognitive functions decline due to ageing, the daily use of multiple languages has a facilitative effect on EF (Jackson, 2020). Such a pattern can be seen in how the effects of multilingualism on EF are usually more consistently found in older adults than in any other populations (Anderson et al., 2021; Bialystok & Craik, 2010; Bialystok et al., 2007; Bialystok et al., 2018; Kim et al., 2019).

Consider a child in a multilingual context, it remains a real possibility that the type of multilingual the child develops into may be a result of preexisting cognitive abilities. A child with outstanding preexisting EF abilities may seek highly multilingual situations which involve more intensive and diverse language exposure and code-switching (e.g., playing equally or more with multilingual than monolingual peers). Therefore, the trajectory of multilingual identity

development may still be subjected to the influence of preexisting cognitive abilities (Baum & Titone, 2014).

Jackson (2020) argues that WM, a subcomponent of EF, and second language acquisition (SLA) are intertwined (see also Shen & Park, 2020), with WM supporting SLA in the early stages of language learning but, as WM inevitably declines as a function of normal ageing, the daily use of two or more languages supports and strengthens WM. Here, WM is defined as the ability to hold a small amount of information in mind for a short time to facilitate, plan and direct behaviours related to comprehension, reasoning, and problem-solving (Cowan, 2014). Notably, both WM and EF abilities are not static throughout the lifespan and may be improved through training. For example, Wass et al. (2012) demonstrated that a 77-minute attention training implemented on typically developing 11-month-old infants resulted in improvements in attentional control compared to active controls. These training effects seem to be stronger the earlier the interventions are introduced, aligning with the narrative that bilingual effects are observed more strongly in early versus late bilinguals (López-Rojas et al., 2022; Luk et al., 2010). Wass et al. explained that this difference may be due to WM and attentional control sharing similar underlying neural networks, such as the DLPFC and the ACC. In early developmental stages, where functional neural networks are unspecialized and undifferentiated improvements in the strength of these neural networks may lead to improvements in attentional control and vice versa. However, whether bilingualism affects attentional control through WM, or a direct effect, remains to be investigated, representing a notable gap in the literature.

1.4 Theoretical Framework

In this section, relevant behavioural and neurophysiological theories outlining the relationship between multilingualism and cognition will be explored and discussed in detail. Frameworks of these theories as well as the specific predictions that can be made based on their parameters will also be discussed. Recent studies claim (DeLuca et al., 2019; Gullifer et al., 2020; Jared, 2015) that ambiguous theories on the multilingual effect have been impeding proper scientific progress. Vague theories, with unclear or unspecified predictions of when multilingual effects should and should not be seen, violate one of the core tenets of scientific pursuit: falsifiability. Unless theories can be empirically tested, data obtained may be reconfigured in an

infinite number of ways so that it fits with the biases of the researchers (see also Blanco-Elorrieta & Caramazza, 2021; de Bruin et al., 2021).

1.4.1 Behavioural Theories

Understanding the bilingual effect requires comprehension of how bilinguals manage multiple languages. Based on the well-known finding that both (or more) languages may be concurrently activated, to a certain extent, even if only exposed to one language (Kroll & Bialystok, 2013, 2015, 2017), this joint activation of languages leads to multiple lexical candidates competing for selection due to overlapping mappings of concepts and semantics between languages (Kroll, 2008). As an alternative to this joint-activation model, Chapter 7 (see Section 7.1.1) discusses the Bilingual Interactive Activation plus (BIA+) model and how that might impact the bilingual effect. Succinctly, the model provides an alternative narrative to the bilingual effect, one that allows for bilingual language control to function on automaticity rather than active and effortful cognition. This model may provide key insights into the mixed findings in the literature.

Colomé (2001) demonstrated a significant delay in response time by Catalan-Spanish bilinguals compared to Spanish monolinguals during a phoneme-monitoring task. The experiment required Spanish monolinguals and Catalan-Spanish bilinguals to determine if a target phoneme was present in the name of an object. For example, participants were shown a picture of a table and were required to determine if /t/ (present in TAULA, Catalan for table), /m/ (present in MESA, Spanish for table) and /f/ phonemes (not present in either Catalan or Spanish translations of “table”) were present in the Catalan word related to the picture. In this case, participants had to accept /t/, as it was present in the word TAULA, but reject the /m/ and /f/ phonemes. Interestingly, only for the bilingual participants did the RT to reject the /m/ phoneme take significantly longer than for the monolingual participants. This finding suggests that the bilingual L2 (Spanish) systems exerted an interfering effect on picture-word comprehension, which must be resolved before a decision can be made to accept or reject the phoneme, thereby resulting in longer RTs for the bilingual group but not the monolingual group. Importantly, the Spanish monolingual group displayed no significant difference in RT to accept or reject any phonemes.

Neurophysiological evidence also supports this phenomenon. In an electroencephalography (EEG) study by Thierry and Wu (2007), Chinese-English bilinguals were tasked to determine whether a pair of printed or orally presented English words were semantically related. In half of the trials, the English word pairs were carefully selected so that their Chinese

translations contained a repeated character. For example, the translation of the English word pair train-ham (火车 - 火腿) has a repeated Chinese character “火”. However, the repetition can only be seen if the English words were translated into Chinese and then written down, activities which could not practically be done within the constraints of their experimental procedure. The results indicated comparable RTs and accuracy for both congruent and incongruent trials, suggesting that participants performed the task with ease. Yet, interestingly, closer inspection of the N400 event-related potential (ERP) revealed a significant reduction in amplitude, indicating repetition priming (Kutas & Federmeier, 2011). This N400 amplitude reduction was only found for Chinese-English bilinguals, but not for English monolinguals. Importantly, the same N400 pattern was found also for Chinese monolinguals who performed the same semantic-relatedness task but in Chinese. Despite Chinese being irrelevant to the task, the language system was still actively processing words, resulting in the observed N400 repetition priming effect.

Although there seems to be overwhelming evidence in favor of the joint activation hypothesis, Costa et al. (2016) provide an alternative learning-based account based on the widely accepted spreading activation model of language processing (Dell, 1986; Levelt et al., 1999). They reasoned that during second language learning, associations between the two language structures are formed by virtue of semantic, phonological, or orthographical similarities. For example, in the train-ham example above (Thierry & Wu, 2007), a Chinese learner of English who is learning the word train for the first time would need to first activate its Chinese translation equivalent (火车) due to its semantic relatedness. The activation of 火车 would then spread to 火腿, by the shared phonological character “火”. Then, 火腿 would activate its English translation equivalent “ham” due to its semantic relatedness. This sequence of events demonstrates that languages may not always need to be jointly activated; rather, they can also be activated through learning and spreading activation. For successful communication to occur, these competing mappings need to be resolved, which alludes to the existence of an underlying cognitive mechanism responsible for the management of their bilinguals’ language. The earliest proposition for this cognitive mechanism is the Inhibitory Control (IC) Model (Green, 1998).

Inhibitory Control Model. This model (Green, 1998; see also the Bilingual Inhibitory Control Advantage (BICA) hypothesis, Hilchey & Klein, 2011) states that multilinguals recruit inhibitory control processes to suppress the activation of non-target languages at the lexical level,

preventing cross-linguistic interference. Consider a situation where a multilingual speaker is conversing with a monolingual. In this conversation, the multilingual speaker must be able to select the appropriate language while suppressing other languages to maintain a coherent conversation with the monolingual speaker. How, then, does the multilingual select the appropriate language? What cognitive mechanisms support this language selection process?

Green (1998) proposed that a conceptual top-down regulatory system, dubbed the Supervisory Attentional System (SAS), supports the language selection process through the recruitment of a variety of cognitive processes. Specifically, inhibitory control processes are heavily recruited to suppress task-irrelevant languages. The constant need to inhibit irrelevant languages strains and trains inhibitory control processes, leading to an inhibitory control advantage in multilinguals compared to monolinguals (Hilchey & Klein, 2011).

This transfer effect from the linguistic domain to inhibitory control abilities is possible through the Hebbian synapse (Bialystok, 2011; Hebb, 1949). Essentially, repeated management and resolution of linguistic interference strengthens and stabilizes synaptic pathways such that similar tasks that use the same pathway are also enhanced (Krizman et al., 2014), in other words, “cells that fire together, wire together” (Hebb, 1949). For example, the ability to inhibit the perception of misleading stimuli features in a flanker task (Eriksen & Eriksen, 1974) is also thought to tap into the same synaptic pathway of language control in bilinguals, specifically those located within the prefrontal cortex. For context, the flanker task requires participants to respond to the direction of the central arrow that is flanked by two arrows on both sides. On congruent trials, the direction of the flanking arrows is similar to the central arrow, on incongruent trials, the central arrow points in the opposite direction to the flankers. Based on Hebbian theory, bilinguals would be more resistant to flanker interference effects than monolinguals due to their strengthened prefrontal synaptic networks as a result of consistent multiple language management.

Based on this understanding, one should be able to predict experimental outcomes of multilinguals compared to monolinguals on tasks containing conflicting responses (i.e., Stroop, Simon, or flanker tasks). First, multilinguals should demonstrate shorter RTs and higher accuracy than monolinguals on incongruent trials. Second, balanced multilinguals (i.e., those who use their languages equally) should display greater inhibitory control than imbalanced multilinguals (i.e., those whose language use heavily favours one language). Lastly, early multilinguals (i.e., acquiring two or more languages at an early age) should display greater inhibitory control than late

multilinguals. However, von Bastian and Souza (2016) failed to find any convincing evidence in support of a bilingual effect in inhibitory control. The authors concluded that previous research supporting this framework might have been influenced by task-specific factors. This conclusion aligns with arguments made by Paap et al. (2015), suggesting that the bilingual effect may exist only under specific and yet undetermined circumstances.

Furthermore, predictions of such transfer effects often fail more than not (Paap et al., 2021; Studenica et al., 2022). This is not an isolated case, as even commercially available cognitive training (e.g., CogMed, Lumosity) that supposedly enhance working memory and attention fail to demonstrate any evidence for (far) transfer effects (Simons et al., 2016; Gobet & Sala, 2022). In a meta-analysis by Paap et al. (2024), bilingualism and various types of cognitive training likely only enhance performance in the specific cognitive domain required by the task (i.e., near transfer). The effects are unlikely to be transferred to broader cognitive domains. This conclusion is supported by Bialystok and Craik (2022), who argued that many purported far-transfer effects are likely near-transfer effects. Therefore, it is more likely that bilingualism improves inhibitory control than executive functions in general. This improvement is thought to occur because both bilingual language control and inhibitory control rely on inhibitory mechanisms.

Despite the lack of transfer effects, the inhibition account of the multilingual effect became a popular framework for explaining the relationship between language control and cognition (Bialystok et al., 2009). Its popularity was also compounded by the fact that it developed alongside the Unity and Diversity model of EF (Miyake et al., 2000), which proposes that EF contains three subcomponents: Inhibition (inhibition of prepotent responses), Switching (mental set shifting), and Monitoring (information updating and monitoring; see Miyake et al., 2000 for review). Therefore, the hypothesis that the relationship between bilingualism on cognition is, in part, due to inhibition, continues to frame bilingual research to date (Bialystok, 2017). Yet, there is an evident lack of explanation for the other components of EF. As indicated by Bialystok et al. (2012), the effects of bilingualism are simply too complex to be understood in the context of a single component of EF. In other words, both bilingualism and EF as a phenomenon are not monolithic constructs. Viewing either from the lens of a single, reductionistic definition leads to simplified and arguably inaccurate answers (Bialystok, 2021).

Therefore, despite providing an initial framework, the IC model does not hold up to closer scrutiny, especially when bilingual effects have been found in trials where inhibition was not

necessary. For example, bilinguals have been reported to have shorter global RT, which is the average RT across all conditions or trials in an experiment, compared to monolinguals in both congruent and incongruent trials on the flanker task (Costa et al., 2009) and Simon task (Bialystok et al., 2004). This result suggests a generally enhanced cognitive processing speed in bilinguals compared to monolinguals. Although bilingual effects on incongruent trials would make sense within the IC model parameters, effects on congruent trials go beyond its explanatory scope. Hence, it is unlikely that the bilingual effect lies solely in the inhibition domain. Hilchey and Klein (2011) proposed the bilingual executive-processing advantage (BEPA), which posits that the global RT effect in bilinguals may reflect a more domain-general EF effect rather than a domain-specific inhibitory control effect. However, evidence for a domain-general EF advantage remains contested, with some findings supporting BEPA and others supporting BICA (Hilchey et al., 2015).

It is also worth noting that recent advancements in understanding the bilingual effect have led researchers to move away from the IC Model approach. Bialystok and Craik (2022) discussed how the effects of bilingualism on inhibitory control appear to be more of a symptom rather than a root cause. They hypothesized that adaptations in the attentional control system may drive improvements in inhibition. While this idea is not entirely new, it aligns with the IC Model, where the SAS is an attentional system governing cognitive abilities. Therefore, the notion that attentional control underlies the bilingual effect is a perspective that integrates well with the existing framework. A more in-depth discussion on this attentional control system will be presented in Chapter 5, as the theory was specifically applied in that experiment.

Adaptive Control Hypothesis. Green and Abutalebi (2013) expanded on the inhibition approach by first outlining three interactional contexts in the Adaptive Control Hypothesis (ACH). Different interactional contexts place varying demands on the executive function system, specifically conflict monitoring and inhibitory control processes. In single-language contexts, a conversation is fixed in one language. Top-down cognitive control employs global inhibition to manage competition within or between language schemas. Dual-language contexts require speakers to monitor cues to determine which language to use and switch languages accordingly. This context involves conflict monitoring and bottom-up cognitive control to select the appropriate language, making it the most cognitively demanding context (Hartanto & Yang, 2019). In dense code-switching contexts, speakers freely engage in code-switching practices. According to the

typology of code-switching by Muysken (2000), there are three types of code-switching: 1) Congruent lexicalization, 2) Insertion, and 3) Alternation (see examples below).

Congruent lexicalization occurs when elements from two languages are used within a shared grammatical structure.

Next time 我去我invite你, then 我们一起去night market 找good food!

Next time I go, I'll invite you, then we'll go to the night market together to find good food!

English phrases are integrated into Chinese sentence structures.

Insertion occurs when well-defined chunks of a language are inserted into the grammatical structures belonging to another language

因为我爸爸是 English educated。

Because my father is English-educated.

The word “English educated” is inserted into a grammatically Mandarin sentence.

Alternation occurs when fragments of each language occur independently of each other's grammatical structure.

其实我很难判断, in my heart I think my Chinese is better but people tell me that my

English 比较好

Actually, it's hard for me to judge. In my heart, I think my Chinese is better, but people tell me that my English is better.

Alternating between full Mandarin sentences and English sentences, each retaining their grammatical structures

In these dense code-switching contexts, speakers are typically multilingual and may not require a specific language for mutual understanding. Hence, language selection is based on ease of access rather than restricted by top-down or bottom-up cognitive control, making it the least cognitively demanding context.

Further cementing the disparity in cognitive demands between interactional contexts is the presence of asymmetric switching costs when switching to a more dominant language (Meuter & Allport, 1999; Timmer et al., 2018). These findings illustrate that cognitive demands are greatest in dual-language contexts compared to single-language and dense code-switching contexts (see Table 1 for an overview of the cognitive demands). Together, these groups of cognitive processes are labelled as the bilingual language control (BLC) system (Calabria et al., 2018).

Table 1. *Cognitive Demands across Language Contexts*

<i>Control Processes</i>	<i>Interactional Contexts</i>		
	<i>Single-Language</i>	<i>Dual-Language</i>	<i>Dense-Code Switching</i>
Goal maintenance	+	+	=
Interference control	+	+	=
Salient cue detection	=	+	=
Selective response inhibition	=	+	=
Task disengagement	=	+	=
Task engagement	=	+	=
Opportunistic planning	=	=	+

Note: A “+” signifies an increase in cognitive demands (more so if bolded); “=” indicates that there is neither an increase nor decrease in that context. See Green & Abutalebi (2013) for an in-depth explanation of control processes. Table adapted from Green and Abutalebi, 2013

These variability in cognitive demands relative to the interactional contexts may adaptively enhance some aspects of EF in bilinguals (Bialystok, 2017; Hartanto & Yang, 2019). This parallels the conflict-monitoring hypothesis posited by Yeung (2013) which suggests that cognitively demanding tasks (i.e., dual-language contexts) signal the ACC to recruit greater cognitive resources, thereby enabling enhanced performance in subsequent trials requiring similar cognitive operations.

Individual and overarching sociological factors of bilinguals may also moderate the effects of interactional contexts on EF. For example, at the individual level, active bilinguals (see Section

1.1), may reap greater long-term EF benefits because they engage in both interference control and selective response inhibition in dual and dense code-switching environments.

On the other hand, macro-level factors (e.g., governmental policies and internet cultures) can exert a complex mix of direct and indirect influences on a community's language practices. In turn, these communities shape the language use of individuals. As a result, it is unlikely that every language community neatly follows the framework of single-, dual-, and dense-code-switching contexts. Instead, each community likely develops its own unique language practices that blend aspects of these three contexts. For instance, language communities in Malaysia and Singapore often do not exhibit clear-cut boundaries between single-, dual-, and dense-code-switching contexts (Hartanto & Yang, 2016, 2019; Lai & O'Brien, 2020; Ng & Yang, 2021). To date, no research has directly addressed this complexity. While the ACH provides valuable insights into different types of language contexts, it should not be viewed as a universally applicable framework across all language communities. The fluidity between these contexts may, in fact, contribute to the mixed findings in past research. Although this thesis does not directly tackle this issue, it explores how variations in language context fluidity may influence the manifestation of the bilingual effect.

1.4.2 Neurophysiological Theories

Advancements in neuroimaging techniques, as well as theoretical developments on the foundational mechanisms of multilinguals language control, have enabled researchers to further understand the nature of the elusive bilingual effect (Hervais-Adelman et al., 2011). This section intends to highlight the neurophysiological changes resulting from the BLC system. However, a critical issue here is whether the BLC and EF are separate domains or if BLC represents an instance of EF (Calabria et al., 2018). If the former is true, and BLC truly operates independently from EF, the lack of interaction between these domains would result in minimal cross-development. If the latter is true, then, the exertion and subsequent evolution of the BLC system should influence domain-general EF abilities. To determine which is true, a deeper understanding of the neural architecture of a bilingual brain is required.

A meta-analysis of the neuroimaging literature by Ware et al. (2020) suggests that bilingualism does indeed induce varying degrees of neuroplastic structural and functional changes in the brain. These changes are not surprising given how other aspects of one's lifestyle, such as

aerobic exercise (Zhao et al, 2020), academic performance (Peng & Kievit, 2020), musical training (Olszewska et al., 2021), and SES (Judd et al., 2020) are also related to neuroplastic changes in the brain.

Functional neuroimaging data provide strong evidence that language control is managed by a cortico-subcortico network that heavily overlaps with the neural architecture responsible for the EF system (Abutalebi & Green, 2007, 2016). Among these structures are the prefrontal cortex (PFC), basal ganglia, and inferior parietal lobe (IPL). Interestingly, individuals speaking more than one language activate areas outside the classical peri-sylvian network area. These extra-sylvian structures include the DLPFC, supplementary motor area (SMA), ACC, bilateral caudate, and cerebellum (Calabria et al., 2018). Overall, these patterns of activation suggest that multilinguals use an array of networks which coincide with both BLC and EF systems. Although there are many neurophysiological models that explore the association between bilingualism and the brain, a complete review of these models goes beyond the scope of this thesis (see DeLuca et al., 2020, for a review). For this thesis, only the Dynamic Restructuring Model (DRM) and the Bilingual Anterior-to-Posterior and Subcortical Shift (BAPSS) model are highlighted.

Dynamic Restructuring Model. This model, proposed by Pliatsikas (2020), has recently accrued much interest within the neuroscience field of multilingualism due to its specific predictions on the neuroplastic changes in the brain as a response to second language learning (see also Lövdén et al., 2020, for a parallel explanation on human skill learning). Such specificity has notably been missing in most bilingual effect theories (Blanco-Elorrieta & Caramazza, 2021; de Bruin et al., 2021).

Central to the DRM is the assumption that the quantity and quality of the bilingual experience dynamically modulates neuroplasticity in the peri-sylvian and extra-sylvian areas. For example, bilingual immersion is associated with increased neuroplasticity in the cerebellum (Pliatsikas, 2020). When faced with a mismatch between a goal (i.e., managing two or more languages) and existing cognitive capabilities, the brain initially undergoes rapid task-relevant cortical expansion. In this area, neural circuits are tested for efficiency in performing the required task. The most promising neural pathways are then chosen for further training, whilst ineffective neural pathways are pruned (Gallo et al., 2020). With continued bilingual exposure, cortical expansion almost renormalizes to pre-expansion levels without accompanying reductions in task performance (Lövdén et al., 2020; Pliastikas, 2019). The consensus is that bilinguals tend to

demonstrate greater grey matter density and white matter integrity in areas related to executive control (for review Abutalebi & Green, 2016; DeLuca & Voits, 2022; Hayakawa & Marian, 2019). In other words, straining BLC systems may induce neuroplastic changes in areas related to EF systems (see Abutalebi & Green, 2016; Coderre et al., 2015, for review, but see also Paap et al., 2021 for counterarguments). Yet, it is important to note that such bilingual experience-dependent neuroplasticity reverts to pre-exposure levels without continuous exposure or use (Pliatsikas, 2020).

Bilingual Anterior Posterior Subcortical Shift. To this point, it is well supported that specific bilingual experiences do induce neuroplastic changes in cortico-subcortico structures. However, another neural architecture unique to multilinguals is the shift in activation from frontal and executive regions to more posterior/subcortical regions to manage nonverbal executive tasks (Grundy et al., 2017). Specifically, compared to monolinguals, multilinguals rely less on frontal (e.g., DLPFC) and cortical regions (e.g., ACC) and instead rely more on posterior (e.g., perceptual/motor regions) and subcortical (e.g., basal ganglia) regions. Burgaleta et al. (2016) and Pliatsikas et al. (2016) found that bilingual experiences were significantly related to the expansion of subcortical regions, such as the basal ganglia and putamen. The anterior-to-posterior functional connectivity was also found to be stronger in bilinguals compared to monolinguals (Luk et al., 2011). Additionally, research on the N2, a negative-going ERP component that traditionally measures inhibitory control, indicates that bilinguals exhibit an earlier N2 response than monolinguals during go/no-go tasks (Barac et al., 2016; Sullivan et al., 2014). These findings suggest bilinguals resolve or detect the go/no-go conflict earlier than monolinguals. Gullifer et al. (2018) interpreted this early N2 as indicative of more efficient cognitive resource allocation. In support, Iragui et al. (2007) found that the N2/P3 component displays a linear increase in latency with age during an odd-ball paradigm. Therefore, this indicates that the inverse relationship between age and the latency of the N2 component is less pronounced in bilinguals when compared to monolinguals.

Similarly, the well-documented Posterior-Anterior Shift in Aging (PASA) effect (Davis et al., 2008; McCarthy et al., 2014) indicates a shift in cognitive processing among older adults from earlier, posterior, and less effortful processing during nonverbal Executive Control (EC) tasks to later, more effortful, top-down anterior processing (Jeon & Friederici, 2013) that was intriguingly not observed in elderly bilinguals when compared to elderly monolinguals. In Ansaldi et al.

(2015), elderly monolinguals and bilinguals completed a visuospatial interference control task while undergoing fMRI. Results showed that although both groups had equivalent behavioural performance, the elderly monolingual group relied on a different neural network compared to the elderly bilingual group. Specifically, the elderly monolingual group demonstrated the traditional PASA effect, whereas the elderly bilingual group only activated their posterior visuospatial processing networks. These results suggest that although bilingualism does not confer a cognitive advantage per se, given that both groups had equivalent task performance, the elderly bilingual group was able to manage interference control without the use of frontal executive networks which are vulnerable to the effects of ageing.

Alignment Problem. The validity of claims based on neurophysiological evidence, particularly ERP evidence, has come under increased scrutiny. This scepticism originates from multiple studies (Fernández et al., 2013, 2014; Jiao et al., 2020; Kałamała et al., 2022; Moreno et al., 2014; Timmer et al., 2021) that had reported non-significant behavioural differences between bilinguals and monolinguals, yet found significant ERP differences. This phenomenon has led researchers to identify what is termed the “Alignment Problem” (Paap et al., 2014, 2015, 2019; 2021; Vaughn et al., 2015). The crux of the alignment problem lies in the disconnect between neurophysiological and behavioural measures. For instance, when ERP measures do not correspond to behavioural outcomes, it remains unclear whether stronger or weaker ERP amplitudes are linked to better performance (i.e., valence ambiguity). As a result, an ERP finding in one study could be interpreted as an advantage, whereas in another study, the same finding may be interpreted as disadvantageous. This valence ambiguity contributes to the uncertainty in drawing firm conclusions based on ERP evidence alone.

Moreover, it is important to recognize that changes in brain architecture, as reflected by ERPs, do not always result in observable changes in cognitive performance (Duñabeitia & Carreiras, 2015; Paap et al., 2021). However, the mere existence of a mismatch between ERPs and behavioural outcomes does not invalidate the findings either, since no single cognitive task can serve as a pure measure of a narrowly defined cognitive process (Miyake et al., 2000). Instead, understanding these discrepancies requires a more nuanced interpretation of the data, taking into account the complexities of measuring both neurophysiological and cognitive processes.

Bialystok (2016) challenges this conclusion: “The clearest interpretation of brain data occurs when differences are found for equivalent performance, thereby ruling out competence or

ability as an explanation for the differences in brain results”. Both conclusions have merit. Although behavioural differences can indeed provide contextual data to interpret ERP differences, comparable behavioural performance does not immediately invalidate ERP findings. Likewise, neurophysiological data alone cannot confirm brain-behaviour links. The Neural Efficiency Hypothesis (NEH) argues that a more efficient brain (i.e., using fewer resources to achieve task goals) is correlated with improved RTs or accuracy on cognitive tasks (Dunst et al., 2014; Li & Smith, 2021; Neubauer & Fink, 2009). However, the backward association between performance and brain efficiency is not as straightforward because RTs and accuracy encompass the combined latency of all preceding cognitive processes (Beesley et al., 2019). Hence, it is difficult to associate improvements in performance with enhanced efficiency of a specific cognitive process unless there is complementary data.

Attempting to resolve the valence ambiguity problem (see the Alignment Problem above) is indeed complex (Grundy et al., 2017). It necessitates cross-disciplinary research from various labs to discern the signal amidst the noise (Bialystok, 2016). To bolster the robustness of such research, it is essential to ground findings in solid theoretical frameworks, apply triangulation of data sources to confirm results, and carefully tailor studies to the specific tasks and populations under investigation (Cespón, 2021; Cespón & Carreiras, 2020; de Bruin et al., 2021). Importantly, integrating complementary behavioural evidence alongside neurophysiological data offers crucial contextual insights, helping researchers link observed ERP differences with tangible behavioural outcomes (Paap, 2019).

Researchers should strive to collect both behavioural and neurophysiological data, which, together, provide more compelling evidence of brain-behavior relationships. Conclusions drawn solely from neurophysiological findings may be susceptible to confirmation bias, lacking the narrative that behavioural data offer. By incorporating behavioural results, the interpretation of whether larger or smaller ERPs signify stronger or weaker inhibitory control becomes more transparent. However, discrepancies between behavioural and neurophysiological data should not be viewed as problematic but rather as an invitation for deeper investigation.

1.4.3 Rapid Modulatory Effects of Language Contexts on The Language-Cognition Relationship

To this point, the fundamental behavioural and neurophysiological theories elucidating the connection between language and cognition were discussed. Most importantly, this relationship is

theorized to be modulated by the specific interactional context of the individual as outlined in the ACH. Interestingly, it appears that this effect may be rapid and transient. It is argued that rapid shifts in language contexts may lead to changes in EF. In Wu and Thierry (2013), the effects of immediate language context on the modulation of EF were investigated in a group of Welsh-English bilinguals. Here, they administered a modified flanker task, where the stimulus is interspersed with different word pairings based on three different interactional contexts; single-language contexts (Welsh or English) and dual-language context (Welsh-English). By experimentally switching language contexts, the authors intend to test whether the purported bilingual effect can be attributed to rapid changes in language contexts or not. They found that in dual-language contexts, error rates in incongruent flanker trials were significantly lower than in single-language contexts, indicating better inhibitory control in the dual-language context. ERP evidence also supports this finding as in incongruent trials compared to congruent trials, a smaller P3 component was found, which previous studies claim reflects better inhibitory control (Grundy et al., 2017; Neuhaus et al., 2010).

In a related study by Ye et al. (2017), L1-dominant Chinese-English bilinguals had better inhibitory control performance in the incongruent flanker trials during the dual-language context but not in the single-language context. Notably, the flanker task in Ye et al. (2017) differed from Wu and Thierry's (2013) study on the account of using intervening pictures of notable cultural places in China versus America or Britain (e.g., Great Wall of China vs. Statue of Liberty) instead of words to experimentally shift participants' control modes. Both authors suggested that as the intervening pictures and words were task-irrelevant, the incidental processing of stimuli from two language sets did not directly compete for cognitive resources with EF processes. The dual-language context was argued to have shifted bilinguals' EF system to a higher state of functioning (i.e., proactive control), allowing for enhanced performance in the modified flanker task.

Similarly, research by Jiao et al. (2020) also found improved performance of L1-dominant Chinese-English bilinguals in the dual-language context while completing a flanker task interleaved with a visual-word picture matching task. Participants were required to respond if a picture of an object matched the word below it as fast and as accurately as possible. Notably, results indicated improved performance in *both* congruent and incongruent flanker trials. This finding is important because based on the ACH, it is understood that inhibitory control processes are recruited during conflict, however, in this study, bilinguals perform better in the congruent

trials as well, where there are no conflicts. Therefore, Jiao et al. (2020) contend that a monitoring mechanism must be working alongside inhibitory control processes. In this unifying paradigm, monitoring processes are recruited to evaluate on a trial-by-trial basis before committing to a response, because there is always a probability that the next trial may be incongruent (Ye et al., 2017). Alongside this, during incongruent trials, inhibitory control processes are still recruited to resolve conflicts. Consequently, bilinguals, whose EF has been adaptively enhanced by the dual-language context, can rapidly detect and resolve conflicts, leading to improved performance in both congruent and incongruent trials. Collectively, these findings suggest that monitoring and inhibitory control processes are intricately related (Costa et al., 2009).

The findings from an ERP study on L1-dominant Chinese-English bilinguals by Jiao et al. (2019) further support this conclusion. They found a larger N2 amplitude and smaller P3 amplitude in both congruent and incongruent flanker trials in the dual-language context compared to single-language contexts. Past studies claim that larger N2 amplitudes reflect enhanced conflict monitoring abilities (Donkers & van Boxtel, 2004; Grundy et al., 2017), and smaller P3 amplitudes reflect better inhibitory control abilities (Grundy et al., 2017, Neuhaus et al., 2010, but see Huster et al., 2020 for alternative discussion). These findings indicate that bilinguals in the dual-language context demonstrate both behavioural and neurophysiological adaptation.

One crucial point that these studies did not discuss is the rapid language context priming mechanism. Although it is yet to be empirically studied within the context of the ACH, this language context priming is speculated to function on the Cascade Model of lexical access (López-Soto, 2010). In this model, bilinguals may easily retrieve words in the appropriate language if they were previously primed in the language because similar language words activate linked lexical and phonological words in a network. While cross-linguistic lexical access is possible in bilinguals, exposure to a specific language facilitates subsequent lexical access in that same language. This helps prevent between-language interference (Giavazzi et al., 2018; Martin & Nozari, 2020; Zhou, 2020). In situations where code-switching occurs, the code-switched words or sentences can prime access to multiple languages (Johns et al., 2021). Based on these findings, experimentally priming words (Wu & Thierry, 2013) and sentences (Adler et al., 2020) may shift bilinguals to be in single, dual or dense code-switching language contexts.

However, it is worth pointing out that the implicit assumption of the ACH, that bilinguals can be categorized solely based on the interactional context in which they were raised, overlooks

the potential for mobility between these contexts. This simplistic view fails to capture the complexity of actual bilingual experiences. Although this may be true in certain cultures where language use is relatively compartmentalized (e.g., first language at home, second language at work), this does not hold for every culture (Samuel et al., 2018). For a relatively compartmentalized culture, Hofweber et al. (2020a) found that German-English sequential bilinguals in the UK reported better performance in single-language contexts, but not in dense code-switching context, during a flanker task, which aligns with the ACH.

For integrated cultures such as Malaysia and Singapore, the distinctions between the interactional contexts (i.e., dual-language and dense code-switching contexts) are less apparent (Hartanto & Yang, 2016, 2019; Lai & O'Brien, 2020; Ng & Yang, 2021). Hence, the staircase patterns of cognitive adaptation elicited by different interactional contexts may express themselves differently in integrated than compartmentalized language cultures. For example, Hartanto and Yang (2019) measured 175 Singaporean bilingual participants' likelihood of being in either the single-language, dual-language or dense code-switching context through self-reports and administered a battery of 9 different EF tasks. Results suggest that the degree of dual-language experience predicted task-switching, but dense code-switching experience predicted both inhibitory control and goal maintenance. Based on the ACH, dense code-switching should not have facilitated any form of cognitive adaptation given that it has the least cognitive load compared to single- and dual-language contexts.

Furthermore, owing to differences in conceptualizing and categorizing bilinguals, comparing these studies becomes extremely difficult (Surrain & Luk, 2017). Yet, insights from Hofweber et al. (2020a) implicate the role of bilinguals' language dominance and immersion profiles in EF differences. For example, larger switch costs have been observed in L1-dominant bilinguals when code-switching from L2 to L1 compared to L1 to L2, suggesting that the dominant L1 language imposed greater inhibitory costs than the non-dominant L2 language (Filippi et al., 2014, Prior & MacWhinney, 2010; but see also Timmer et al., 2018, for an alternative view). Furthermore, Bosma and Pablos (2020) found a smaller P3 component in Dutch-English bilinguals when they encountered a code-switch from L1 to L2 compared to a code-switch from L2 to L1 while performing a flanker task, suggesting easier interference resolution when code-switching to a non-dominant language than to a dominant language. These studies suggest that L1-dominant bilinguals may be engaging EF processes to a greater extent than balanced bilinguals, leading to

group EF differences between different types of bilinguals. Indeed, bilinguals' language profile does not only differ in language dominance and immersion but also the age of language acquisition and proficiency (Beatty-Martínez et al., 2020; DeLuca et al., 2019; Kheder & Kaan, 2021; Luk & Bialystok, 2013; Treffers-Daller, 2015). For example, Yow and Li (2015) found that early second language acquisition is associated with smaller interference costs in the Stroop Task. Similarly, Ma et al. (2021) found that language proficiency had a non-linear influence on bilinguals on a cognitive task, where low and high-proficient bilinguals had better task performance than average proficient bilinguals. Hence, variations on these EBFs (DeLuca et al., 2019) might account for the contradicting findings in the literature.

1.5 Research Context

Situated in Southeast Asia, Malaysia is a Commonwealth country made up of Peninsular Malaysia and East Malaysia, with an estimated population of 32.7 million people made up of Malays (67%), Chinese (25%), and Indians (8%) (Department of Statistics, 2015). Officially, Bahasa Melayu is the de facto national language, but English, Mandarin, and Tamil are all widely spoken due to colonial influences. However, in major cities and towns, English is used above and beyond other languages owing to its position as an international trade language.

Furthermore, due to mandatory bilingual language policies (A'Beckett & Du Plessis, 2017, see also Chu & Le, 2020, for an in-depth sociopolitical review), most Malaysians are multilingual, speaking both Malay and English in addition to their heritage languages, often at a young age. For example, the use of Malay as the medium of instruction is enforced in all public schools except for vernacular schools. Furthermore, with the enactment of the Dual Language Program (DLP) in 2018 (Ministry of Education, 2018), selected public schools may use English as a medium of instruction for science, technology, engineering, and mathematics (STEM) subjects. For the other non-DLP schools, subjects are either taught in Malay (National Schools), Mandarin (National Chinese Schools), or Tamil (National Tamil Schools). Regardless of the type of school, English remains a mandatory language subject. Consequently, most graduates from public schools are communicatively competent in a minimum of two or three languages. This outcome is aligned with the Malaysian Education Blueprint 2013-2025 (Ministry of Education, 2013) which targets to have all Malaysian students be communicatively competent in at least Malay and English. For this reason, it is common to find Malaysians who speak more than three languages (e.g., Malay, English, and their heritage language).

It is also common to find individuals speaking a non-heritage language (e.g., Malaysian Indians speaking Mandarin), due to interracial mixing, either through formal education, immersion, or both. The presence of vernacular schools (Selvadurai et al., 2017) offers opportunities for students to be taught in either Tamil (Tamil vernacular schools) or Chinese (Chinese vernacular schools), regardless of race. However, most higher education institutions (HEIs) primarily use either English in private institutions or Malay in public institutions. As a result, some students may find themselves at a disadvantage. Whereas much of their high school education was conducted in Chinese or Tamil, many HEIs require proficiency in either English or Malay for admission.

Additionally, code-switching is extremely prevalent in daily conversations. For example, the usage of Malaysian English, a colloquial language derived from English with strong Malay, Chinese and Tamil influences, naturally requires interlocutors to engage in inter- or intra-sentential code-switching (Pillai & Ong, 2018; Tan, 2009).

Therefore, the existing sociolinguistic environment in Malaysia provides an excellent research opportunity to investigate the effects of multilingualism on different facets of cognition (see also Fernandez-Duque et al., 2023). Not only because past literature is saturated with studies on bilinguals, but also because it offers insights into the chicken-or-egg question (Cox et al., 2016; Grundy & Timmer, 2016) of whether multilingualism leads to greater EF, or whether individuals with higher EF are more likely to become multilinguals. Malaysians are multilinguals out of circumstance (e.g., environmental pressures) rather than choice. Hence, any cognitive benefits from multilingualism are unlikely to be a result of preexisting cognitive abilities (Baum & Titone, 2014). Investigating the Malaysian sample helps to better isolate the effects of pre-existing cognitive abilities from the literature.

However, an important question to investigate is the language attitudes of Malaysians. Sparse research suggests that Malaysians hold mixed attitudes towards multilingualism, depending on whether the question is posed to the language majority or minority. Language attitudes are an important but under-reported factor (Surrain & Luk, 2017) in the bilingualism-cognition literature, as language attitudes may predict language use choices (Kircher & Zipp, 2022). It has been argued that the exhibition of the bilingual effect is contingent on continuous exposure and use of bilingualism (Pliatsikas, 2020). A deeper discussion of the language attitudes of Malaysians will be presented in Chapter 3.

1.6 Research Questions and Hypotheses

The various theories discussed have provided sound theoretical grounds upon which the effects of multilingualism can occur. However, inconsistent evidence to support these theories has caused wavering support for the phenomenon, suggesting that if a multilingual effect does exist, they might only exist in specific, undetermined circumstances. There is simply too much variance in multilingualism and its constituents for researchers to consistently replicate the sample across studies and labs, making comparisons difficult. Hence, it remains uncertain if the findings in the literature stem from idiosyncratic conceptualizations of multilingualism or a true effect of multilingualism on cognition. Presently, researchers are collectively challenging the traditional binary perceptions of multilingualism and its constituents, which have served as the foundation for many theories. Consequently, it is now crucial to reevaluate past theories regarding the relationship between multilingualism and cognition to align them with modern methodologies.

Experiment 1 (Chapter 3) investigates the attitudes towards multilingualism and the language learning experience of the Malaysian public. It provides a general view of how Malaysians value multilingualism and their language learning experiences. This study serves two functions: 1) it characterizes the participants' language history in this thesis, thereby providing a necessary backdrop for interpreting experimental outcomes, and 2) it connects the thesis to public views on multilingualism and additional language learning, making the findings more applicable and impactful for societal and policy discussions.

Chapter 4 consists of three experiments, Experiment 2A investigates the rapid modulation of EF in multilinguals concerning language contexts using a modified Eriksen Flanker task. Additionally, it assesses whether the MLD score predicts EF outcomes above and beyond traditional measures of multilingualism. Experiment 2B explores the potential impact of linguistic distance between primed languages on Flanker task performance. Experiment 2C examines the potential influence of participants' language familiarity and proficiency on Flanker task performance. In Chapter 5, Experiment 3 assesses whether employing an eye-tracking methodology enhances sensitivity to cognitive adaptations in rapidly changing language contexts. In Chapter 6, Experiment 4 investigates whether increased task difficulty leads to more observable cognitive adaptations. It also aims to uncover novel insights into the cognitive adaptation processes

of multilinguals through ERPs. An overview of the methodology and materials used can be found in Chapter 2.

For experimental tasks, the first overarching hypothesis predicts that different language contexts (i.e., no language, single-language, dual-language) can adaptively modulate EF abilities in multilinguals. The second overarching hypothesis predicts that multilingualism measured as a continuum (i.e., indicated by higher values on the MLD score) can predict cognitive task performance over and above traditional measures of multilingualism. Each successive experiment, positioned within the larger context, addresses specific facets contributing to the understanding and testing of these two overarching hypotheses. Importantly, the experiments in this thesis rely on passive comprehension as a means to prime language contexts, rather than the active production of language. This decision was made in line with the past experimental studies that motivated this thesis. Hence, conclusions on the bilingual effect are constrained to this facet.

Chapter 2: General Methodology

In the coming chapters, the exact methodology of each experiment in this thesis will be outlined in greater detail. This chapter aims to provide a summary of the overlapping methodologies used across the experiments.

2.1 Covariates

2.1.1 *Socioeconomic Status*

The effects of SES on the cognitive outcomes of lifespan cognitive development have been well-documented (Vivas et al., 2020). SES, as a concept, can be measured by different proxies. Typically, these proxies fall under three overarching factors: 1) occupation, 2) income, and 3) education (Long & Renbarger, 2023). There is a degree of overlap between these factors; for example, higher educational attainment is associated with a more lucrative occupation (Hällsten & Thaning, 2021). SES itself does not directly influence cognitive outcomes but rather the experiences associated with low (or high) SES impact cognitive outcomes. For example, when children are brought up in impoverished settings, they are more likely to experience nutritional deficiencies, potential parental abuse, limited access to both physical and cognitive stimulation, and increased exposure to stress and trauma (Katz & Shah, 2017). It is these negative life experiences that are directly linked to poorer cognitive outcomes in developing children. Considering these associations, it was surprising to find that bilingual low-SES Turkish-Dutch children outperformed their respective monolingual counterparts on various WM tasks (Moriguchi & Shinohara, 2019; van den Noort et al., 2019). This bilingual effect seems to be present only in low-SES groups but disappears in high-SES groups (Naeem et al., 2018).

Traditionally, SES is measured based on proxy metrics, such as parental occupations (Ma et al., 2021) or parental education level (Wagner et al., 2023), but it can also be based on self-perceptions of income, education, and employment relative to their communities using the MacArthur Scale of Subjective Social Status (MacArthur SSS Scale; Adler et al., 2000). Hence, to account for the possibility of SES differences across participants, the experiments employed the adult version of the MacArthur SSS Scale to assess participants' self-perceived SES. This scale was chosen over self-reports of parental occupation or education level given that the scale demonstrated acceptable test-retest reliability even after 6 months ($\rho = 0.62$; Operario et al., 2004).

Participants complete a single-item assessment assessing their perceived SES relative to others in their group, using a scale from 1 to 10. They view a visual representation of a ladder with 10 rungs and are instructed to specify their position on the ladder (see Appendix 1). Higher positions denote a perceived higher SES compared to others in the group.

2.1.2 General Cognitive Ability

It is crucial to rule out the possibility that any effects found may be attributed to variances in participants' baseline general cognitive abilities. In the psychology of individual differences, general intelligence, *g*, is thought to be the main underlying factor across all cognitive processes (Spearman, 1904). Much evidence supports that *g* is strongly associated, among others, with executive function (Blair, 2016) and fluid reasoning abilities (Kyllonen et al., 2017). It stands to reason that cognitive task performance may be greatly affected by variations in a sample's *g* scores. Three methods of capturing *g* scores have typically been administered by past studies (Chung-Fat-Yim et al., 2021; Fyndanis et al., 2023; Vivas et al., 2020). The Wechsler's Intelligence Scale (WIS; Wechsler, 1958), Kaufman Brief Intelligence Test (K-BIT; Smith, 1999), or the Raven's Progressive Matrices (RPM; Raven, 1941).

Wechsler's Intelligence Scale is an extensive assessment instrument that evaluates a wide range of cognitive skills in both verbal and nonverbal domains. It includes variations for children (Wechsler's Intelligence Scale for Children; WISC) and adults (Wechsler Adult Intelligence Scale; WAIS). Its multifaceted method gives important insights into an individual's cognitive functioning, with independent indices for verbal comprehension, perceptual reasoning, working memory, and processing speed. Although this scale is considered the most comprehensive test of *g*, it is also the most time intensive.

The K-BIT stands out for its conciseness without compromising reliability and validity. Similarly, this test can assess children and adults on both verbal and non-verbal reasoning abilities but at a fraction of the time required. Walters and Weaver (2003) reported that the outcomes of the K-BIT are strongly correlated with the outcomes in the WAIS-III.

Lastly, RPM measures non-verbal fluid intelligence and non-verbal abstract reasoning across diverse populations and age groups. Unique to this test is its ability to assess these domains without relying on language skills, making it highly suited for cross-cultural assessments. In this case, given that bilinguals tend to score lower on verbal fluency measures (Bialystok, 2009;

Sandoval et al., 2010), RPM allows researchers to measure bilinguals' g without being confounded by variations in verbal fluency abilities. However, contrary to popular beliefs that RPM purely measures g (Martinez, 2013), large-scale studies suggest that RPM only accounts for about half of the variance of g (Gignac, 2015). This finding means that in comparison to K-BIT and WIS, RPM falls short in providing an accurate depiction of an individual's g factor. Because of its non-verbal nature and its ease of administration, RPM scores provide a good enough representation of the g factor in many multilingual-cognition studies.

Therefore, in the present experiments, Raven's Standard Progressive Matrices (SPM) were administered (Raven, 1941). In this task, participants are shown a series of matrices with missing elements. Participants must identify the missing element from a list of possibilities that logically completes the matrix. Typically, the interpretation of the SPM scores relies on normed values of a specific sample. However, Raven's SPM does not include normed values based on the Malaysian population. Given that the experiments are only concerned about their baseline general cognitive abilities, specificity and normed interpretations are less of a concern. Furthermore, as the participants were mainly young adults from an undergraduate program at a university, general cognitive ability was not expected to differ substantially between participants. Hence, to minimize participant fatigue and improve responses, only 6 of the 12 items from each set (Sets A to E) were used from the full Raven's SPM for a total of 30 items (see Appendix 1). The results of this test are interpreted based on the total number of correct responses, ranging from 0 to 30, where a higher total score indicates higher general cognitive ability.

2.2 Language Proficiency Measures

In the literature, vocabulary knowledge is generally taken as a good predictor of language proficiency, which spans four macro domains (i.e., Reading, Listening, Writing and Speaking). Reading and listening domains are considered receptive skills, whereas writing and speaking domains are considered productive skills. Although there is more than adequate evidence supporting this relationship (McLean et al., 2020; Rafique et al., 2023; Tong et al., 2023), Qian and Lin (2020) specified that vocabulary size itself does not equally measure all four domains. Where vocabulary size (i.e., the quantity of words known to a learner) is a good predictor of receptive skills, vocabulary use (e.g., the quality of use of a lexical item, such as overuse, underuse, or misuse) is a good indicator of productive skills (also termed as lexical richness). Ideally, when

measuring language proficiency, both vocabulary size and use are considered. However, there are much more readily available scales that measure receptive than productive skills, likely because it is much easier to document learner input than output.

Another issue is feasibility because where receptive skills can be easily measured using scales, like the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) or the Receptive One-Word Picture Vocabulary Test (ROWPVT; Martin & Brownell, 2011), which is a standardized instrument designed for quick assessment of a learner's receptive vocabulary, productive skills require a much more in-depth and time-consuming corpus and lexical analysis (Qian & Lin, 2020). In such cases, a learner's speech or written output is cross-examined for lexical features and patterns typical of the language's native speakers. Such assessment, although incredibly sensitive, may not be feasible for most researchers. In practice, it is more likely that researchers would measure vocabulary knowledge as a measure of general language proficiency. This approach is less concerning in view that both vocabulary size and use are significantly correlated with one another (Wang-Taylor & Clenton, 2022). Hence, measuring either one would sufficiently provide researchers with the information necessary to correctly understand an individual's language proficiency.

The experiments in this thesis will administer the Lexical Test for Advanced learners of English (LexTALE; Lemhöfer & Broersma, 2012) and its Mandarin variant (LexTALE_CH; Chan & Chang, 2018) as proxy measures for participants' language proficiency in English and Mandarin. LexTALE is a quick and objective measure of participant's general language proficiency in English (Lemhöfer & Broersma, 2012) and Mandarin (Chan & Chang, 2018). Participants are asked to indicate which words are real and which are fake in both tasks without external aid. This method of measuring language proficiency has been demonstrated to be superior to self-ratings and has substantial correlations with the participants' general English or Mandarin language proficiency.

English Version. Participants are given 60 English word stimuli, 20 of which are fake English words (see Appendix 1). To correct for the unequal proportion of words and non-words in the test, the average correct percentage was taken for these two item types.

$$\frac{\left(\frac{\text{number of words correct}}{40} \times 100\right) + \left(\frac{\text{number of nonwords correct}}{20} \times 100\right)}{2}$$

Mandarin Version. (Chan & Chang, 2018)². Similarly, the Mandarin version lists 90 single Mandarin characters, 30 of which are fake (see Appendix 1). Authors have formulated a calculation that penalizes guessing behaviour to allow for a more accurate measure of participants' Mandarin vocabulary proficiency. If participants respond randomly to the items (or respond "yes" to all items), their corrected accuracy scores would be 0. Likewise, if they incorrectly accept nonsense items as real Mandarin characters more frequently than the real items, their corrected accuracy would be negative, for as low as -30. The highest possible corrected accuracy score is 60, which requires participants to accurately determine all correct Mandarin characters and correctly reject all nonsense characters. However, to allow for ease of comparability with the English version of the LexTALE, the formula was modified to scale the results from 0 to 100%.

$$\frac{((\text{number of items correct}) - (2 \times \text{number of incorrectly accepted nonce items})) + 30}{90} \times 100$$

Note that another Mandarin variant called the LexCHI (Wen et al., 2023) is arguably more sensitive to the receptive vocabulary of Mandarin speakers as its items consist of two characters rather than the single character items in the LexTALE_CH. This would make it more reflective of the actual Chinese vocabulary which typically contains more than one character. For example, an analysis of a Chinese subtitle corpus (SUBTLEX-CH; Cai & Brysbaert, 2010) shows that among its 33.5-million-word database, 46.2% of the words contain two characters, whereas only 5.4% of the words consist of one character (Wen et al., 2023). However, at the time of the experiments, the LexCHI was not publicly available, hence, the LexTALE_CH was administered instead.

Furthermore, LexTALE was chosen instead of other industry standards, such as the aforementioned PPVT or ROWPVT, as it is free and less time-consuming to administer. Given

² The current thesis administered the LexTALE_CH rather than the LexCHI (Wen et al., 2023), which is arguably more sensitive to the receptive vocabulary of Mandarin speakers since its items consist of two characters rather than the single character items in the LexTALE_CH. This would make it more reflective of the actual Chinese vocabulary which typically contains more than one character. For example, an analysis of a Chinese subtitle corpus (SUBTLEX-CH; Cai & Brysbaert, 2010) shows that among its 33.5-million-word database, 46.2% of the words contains two characters, whereas only 5.4% of the words consist of one character (Wen et al., 2023). However, at the time of conceptualization, the LexCHI was not publicly available, hence, the use of the LexTALE_CH across all experiments to facilitate for cross-study comparison

that the experiments sample university students, extreme variations in language proficiency were not expected.

2.3 Language History Questionnaire 3.0

The LHQ (see Appendix 1) represents one of the three frequently used questionnaires to measure participants' degree of bilingualism that has been psychometrically validated (Anderson et al., 2018; Li et al., 2020). The current study used the LHQ over the LEAP-Q and LSBQ due to its appropriateness to be used with young adults and multilingual populations (Anderson et al., 2018). Based on the questionnaire, participants' AoA of languages, and Language Proficiency, Immersion and Dominance and Dominance Ratio are calculated. Additionally, an MLD Score was calculated, which provides a single value to determine participants' degree of multilingualism, allowing for a more objective between-subject comparison. Although language AoA, proficiency, immersion, and dominance are conventional indicators of multilingualism, the introduction of MLD scores represents a novel method to gauge the extent of multilingual capabilities in this thesis. The formulas required to calculate these dimensions will be briefly described in the following paragraphs (see Li et al., 2020, for review).

Language Proficiency. Based on Q11 of LHQ 3.0 (*Rate your current ability in terms of listening, speaking, reading and writing in each of the languages you have studied or learned*), participants' overall proficiency score of his i^{th} language is determined by the formula:

$$Proficiency_i = \frac{1}{7} \sum_{j=\{R,L,W,S\}} \omega_j P_{i,j}$$

Where $\{R, L, W, S\}$ stands for Reading, Listening, Writing and Speaking components of a language. $P_{i,j}$ indicates participants' self-rated proficiency level to the j^{th} component of the i^{th} language. As Q11 is based on a 7-point Likert scale, a scaling factor of 1/7 was used to normalize the output to a range of between 0 and 1 (1 indicates native-like proficiency). The variable w_j indicates the weight assigned to the j^{th} $\{R, L, W, S\}$ components which may be changed according to the discretion of the researcher. The default weightage of 25% was set for each component.

Language Immersion. Based on Q5 of LHQ 3.0 (*Indicate your native language(s) and any other languages you have studied or learned, the age at which you started using each language in*

terms of listening, speaking, reading, and writing, and the total number of years you have spent using each language), participants' immersion in the i^{th} language is determined by the formula:

$$Immersion_i = \frac{1}{2} \left[\sum_{j=\{R,L,W,S\}} \omega_j \left(\frac{Age - AOA_{i,j}}{Age} \right) + \left(\frac{YoU_i}{Age} \right) \right]$$

Here, Age is the participant's age in years. $AOA_{i,j}$ indicates participants' age of acquiring the i^{th} language in terms of the j^{th} {R, L, W, S} components. YoU_i indicates the total years of use (YoU) of the i^{th} language. This component serves to compensate for issues arising from non-continuous use of a language (e.g., learning and using a language but stopping using it for an extended period). A 1/2 scaling factor was used to ensure equal weightage of YoU and AoA on immersion scores, as well as to normalize the score to between 0 and 1 (1 indicates native-like immersion in a language)

Language Dominance. Based on Q11 (see Language Proficiency), Q14 and Q15 of LHQ 3.0 (14. Estimate how many hours per day you spend engaged in the following activities in each of the languages you have studied or learned; 15. Estimate how many hours per day you spend speaking with the following groups of people in each of the languages you have studied or learned), participants' language dominance is determined by the formula:

$$Dominance_i = \sum_{j=\{R,L,W,S\}} \omega_j \left[\frac{1}{2} \left(\frac{P_{i,j}}{7} \right) + \frac{1}{2} \left(\frac{H_{i,j}}{K} \right) \right]$$

Here, $H_{i,j}$ indicates the total estimated hours per day spent on the j^{th} {R, L, W, S} components of the i^{th} language. K is a constant, serving as a scaling factor (set to 16). Another 1/2 scaling factor is applied to ensure equal weightage of $P_{i,j}$ and $H_{i,j}$ on dominance scores.

Language Dominance Ratio. There is a large between-subject variability when estimating $H_{i,j}$, some may be more liberal in estimation, whereas others may be more conservative. To address this issue, Li et al. (2020) introduced the *Dominance Ratio*, which essentially converts Language Dominance scores into a standardized score, allowing for a more objective comparison between groups. This ratio is determined by the formula:

$$\text{Ratio}_{\text{Dominance}_i} = \frac{\text{Dominance}_i}{\text{Dominance}_1}$$

MLD Scores. Based on the Dominance score for the i th language, a temporary variable called Proportion of Dominance (PD_i) is formed, determined by the formula:

$$PD_i = \frac{\text{Dominance}_i}{\sum_{i=1}^n \text{Dominance}_i}$$

Where n indicates the total number of languages a participant has learned. This means that the MLD score accounts for as many languages as a participant has reported. This variable roughly indexes the proportion of the i^{th} language dominance in a given language environment/usage. Using PD_i , MLD scores can be calculated using the formula based on Shannon's Entropy:

$$MLD = - \sum_{i=1}^n PD_i \log_2(PD_i)$$

MLD scores range from 0 to 2, with 2 indicating highly multilingual.

2.4 Absence of Control Groups

The typical control group consisting of monolinguals will not be used in the following experiments because monolingualism is rare or even non-existent in Malaysian communities. These rare monolingual individuals in Malaysia are often from highly specific cultural or rural backgrounds, making their inclusion in the study unrepresentative of the broader Malaysian population. As a result, using monolinguals as a control group would not provide a meaningful comparison for this research. Although many past studies have used such methodology before, typically by participant self-identification (Lukasik et al., 2018) or through one of the psychometrically validated language questionnaires discussed in Chapter 1 (e.g., LSBQ; Chung-Fat-Yim et al., 2021), it is unclear what or why monolingual groups should be considered a suitable control for multilingualism.

As Rothman et al. (2022) discussed extensively, the notion of a control group implies the absence of the specific manipulation present in the experimental group, allowing researchers to “control” for confounding variables. For example, in typical drug trials, control groups are given placebos whereas test groups are given the experimental drug. This absence of manipulation is not achievable in a monolingual control group because language-related experiences and their subsequent linguistic and cognitive outcomes matter for all, even for those traditionally labelled as monolinguals. Although such control groups may be justifiable in certain experimental designs, in most cases, due to the lack of clear standards of what constitutes a typical control group of multilinguals, a monolingual “control” group fails its most essential role as it controls little or nothing at all. This is a hallmark of the false equivalence fallacy, where two things are argued to be the same despite important differences.

A similar argument can be made about the multilingual “experimental” group. As multilingualism is multifarious and dynamic, it is unwise to assume that any given group of multilinguals in a study is homogenous. Consequently, this approach presents a potential for erroneous or misleading conclusions, particularly when using popular statistical methods, such as the Analysis of Variances (ANOVA), which assumes group homogeneity. Although correlational analyses might mitigate this concern, their conclusions warrant careful interpretation to avoid falling into the correlation-as-causation fallacy.

Unfortunately, there are no straightforward answers to this dilemma as either method of classification would result in some level of noise. The best move would be to select the method that would generate the least amount of noise. Aligned with recent research (de Oliveira & Oliveira, 2022; Gosselin & Sabourin, 2023; Lai & O’Brien, 2020; Sorge et al., 2017; Sulpizio et al., 2020; Wagner et al., 2023), the present thesis opted to classify multilinguals on a continuum rather than to rely on dichotomous monolingual-multilingual groupings.

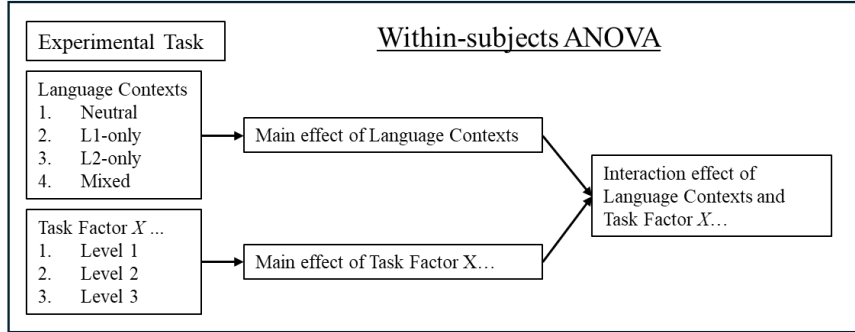
2.5 General Statistical Approach

Given that there are no monolingual control groups, the traditional ANOVA cannot be used to address the hypothesis about the enhanced performance of multilinguals. Instead, a regression analysis using each participant’s multilingualism variables (e.g., Language AoA, proficiency, MLD scores) provides insights into the continuous relationship between multilingualism and cognitive outcomes.

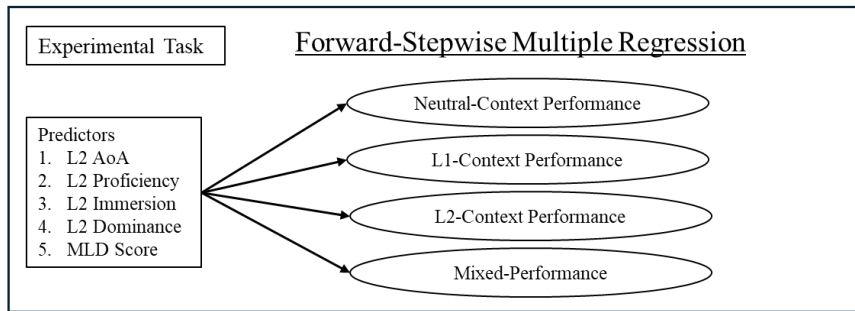
Within-subjects ANOVA was conducted without the inclusion of any multilingual experience in the models, instead, the analysis will solely focus on the effects of the manipulation of language priming of the task. This analysis would address the first hypothesis that different language contexts would rapidly modulate EF abilities. Multilingual experience variables are included as predictors of task performance in a regression model. This analysis would address the second hypothesis that MLD scores would predict task performance above and beyond other multilingual experience variables. Only L2 variables will be included in the modelling, because trilingualism, while common, may not be a characteristic of all Malaysians. L1 variables were excluded for two reasons: 1) traditionally, L2 variables are used as measures of bilingual expertise, since L1 would be the participant's native language and 2) including both L1 and L2 variables would lead to overfitting of the regression model. Figure 1 provides an overview of the statistical approach.

Figure 1: Overview of ANOVA and Regression Approach

Hypothesis 1: Different language contexts can adaptively modulate EF abilities in multilinguals



Hypothesis 2: MLD scores can predict task performance above and beyond traditional measures of multilingualism



2.5.1 Within-subjects ANOVA

The experimental tasks always consisted of a language context component that was strategically manipulated to prime participants to either be in a no language, single-language, or dual-language context. Depending on the experiment, task-specific factors (Factor *X*) may be included (e.g., trial type, cue type). The within-subjects ANOVA would offer insights into whether the experimental priming of language contexts rapidly influences task performance and whether this effect interacts with other task factors.

To improve confidence in the within-subjects ANOVA results, three core assumptions must be met or, if violated, corrected. First is the assumption of normality, that is, whether the residuals of the data follow a normal distribution. It is common knowledge that ANOVAs are resistant to violations of normality (Field, 2018; Tabachnick & Fidell, 2019). Simulation studies on various non-normal datasets have shown that violations of normality do not significantly influence Type-I error rates (Harwell et al., 1992; Lix et al., 1996). This is, in part, due to the central limit theorem, where the distribution of a sample would progressively become normally distributed when sample sizes are large enough. Sample sizes of more than 30 are generally

accepted as a “large enough” sample size (Field, 2018), although this value may change depending on various statistical parameters (e.g., skew and kurtosis).

Second is the assumption of sphericity, that is, whether the variance of all the differences across conditions is equal. Typically, this assumption is tested using Mauchly’s test of Sphericity, but Field (2018) argues that this is not a particularly informative test. Field argues that when sphericity is violated (i.e., Mauchly’s test is $p < .05$), a correction test (e.g., Greenhouse-Geisser, Huynh Feldt) is applied proportionate to the extent of the violation. Hence, when the data has no violations, no corrections are applied, for small violations, a small correction is applied, and for large violations, a large correction is applied. Therefore, Mauchly’s test may be ignored, and a correction test can be applied in any case. When choosing which correction test to apply, Field (2018) suggests that if the Greenhouse-Geisser epsilon (ϵ) value is above 0.75, the Huynh-Feldt correction should be used, otherwise apply the Greenhouse-Geisser correction.

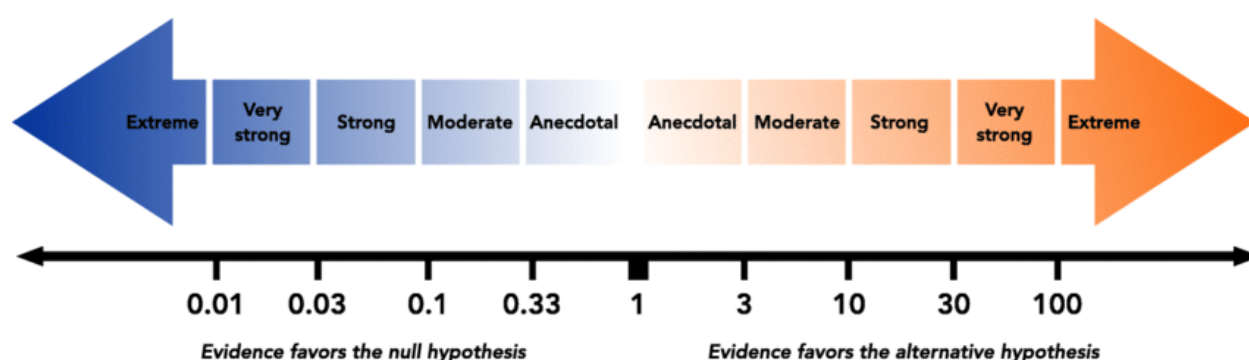
Lastly, there should be no outliers within the outcomes to allow for the best measure of central tendency (Friedman et al., 2011). There are many ways to deal with outliers, but specifically in the multilingualism-cognition literature (Hartanto & Yang, 2019; Lai & O’Brien, 2020; Paap et al., 2020; Zhou & Krott, 2016), RT data trimming procedures are common. These studies typically adopt the approach of removing false start trials (e.g., less than 200 ms) or outlier trials whose RT lies 2.5 standard deviations (SD) above the mean. Although there is merit to data trimming, this process must be theoretically informed or risk erasing important group differences at a minimum or drawing misleading conclusions at worst (Bialystok, 2016; Zhou & Krott, 2016).

2.5.2 Bayesian Within-subjects ANOVA

The null hypothesis significance testing (NHST) approach has been ubiquitously used in the field of psychology (Nickerson, 2000). It has been the driving force behind the popularity of the p -value as the be-all-end-all measure of statistical significance. However, there has been a growing chorus of criticisms against the exclusive use of the p -value in psychological hypothesis testing (Madjarova et al., 2022; Quintana & Williams, 2018). Recently, there has been a shift towards alternative measures of statistical significance. One such measure that has been garnering a large following is the Bayesian analysis (Halsey, 2019). This framework allows researchers to quantify, and continuously update, the probability of the observed data to support the null hypothesis (H_0) compared to the alternative hypothesis (H_1), given prior probabilities (Etz &

Vandekerckhove, 2017). Essentially, where p -values can only serve to accept or reject the null hypothesis, a Bayes Factor (BF) quantifies the support for the observed data as evidence for the null (BF_{01}) or alternative hypothesis (BF_{10}). This value is expressed on a scale of 0 to 100, if the value approaches 0, the value is typically better expressed as its reciprocal. For example, if an effect has a BF_{10} value of 0.01, it is much more intuitive to represent this value as a BF_{01} value of 100. Lee and Wagenmakers (2014) provided a classification scheme for the BF scale, from anecdotal to extreme evidence; a visualization of this scale can be seen in Figure 2.

Figure 2: Lee and Wagenmakers (2014) BF Scale Classification Scheme



Aligning the recent conversations surrounding the issues of the p -value, the current thesis incorporates a Bayesian analysis into the experimental chapters (Chapters 4, 5 and 6). The outcome of these analyses is presented alongside traditional NHST reports where appropriate. Ideally, it is expected that both the p -value and the Bayes Factor would align complementarily, with the Bayes Factor providing further information on the quality of support for either the null or alternative hypothesis. When the p -values are significant, but the BF values are anecdotal (e.g., less than 3), it could suggest that the statistical significance in the data may not strongly favour either the null or alternative hypothesis. This disconnect may signal methodological concerns, such as sample size or even the magnitude of the effect size, both of which have been documented to affect BF values (Stefan et al., 2019). Hence, contradictory evidence between p -values and BF values invites more detailed investigations.

2.5.3 Forward-stepwise Multiple Regression Analysis

From participants' language background data, five variables of interest in past research were extracted to be included as predictors in a forward-stepwise multiple regression model: L2

1) AoA (Kalamala et al., 2022), 2) Proficiency (Ma et al., 2021), 3) Immersion (Hannaway et al., 2017), 4) Dominance (Gathercole et al., 2014), and the MLD score (van den Berg et al., 2022). These variables are used to predict the outcomes of each of the four language contexts of the task. This analysis offers insights into which aspects of multilingual experience would significantly predict task performance.

Unlike simultaneous or sequential regression, which are theoretically guided, this form of regression uses computer algorithms to determine the order of variable entry into the model. The selection process is guided by statistical parameters like semi-partial correlation and p -values of the entered variables (Field, 2018). The algorithm automatically decides which variables to include and which to drop in the model based on these criteria, rather than being driven by specific theories or hypotheses. As such, this regression method is not able to explain any identified relationships but rather only assist in identifying the best predictors for predicting task performance (a comprehensive comparison between regression methods can be found in Keith, 2019). However, in the context of this thesis, which is focused on investigating the predictive capabilities of the multilingual experience variables, the absence of explanatory power is less problematic.

In addition, to improve confidence in the results, four key assumptions must be satisfied: normally distributed residuals, homoscedasticity, linear relationship between predictors and outcomes, and no multicollinearity. The first two assumptions are less critical for the same reasons that ANOVA is robust against violations of normality (Field, 2018; Tabachnick & Fidell, 2019). In general, with large sample sizes, residuals typically exhibit a normal distribution and homoscedasticity. As such, these two assumptions become particularly critical primarily in experiments with small sample sizes. It is the last two assumptions that are key to a valid outcome. If either of these assumptions is violated, it would bias all the obtained estimates (e.g., R^2 , regression coefficients, p -values; Keith, 2019).

Although a linear relationship may be assumed between the predictors and the outcomes based on relevant past research, multicollinearity can be determined by looking at the Variance Inflation Factor (VIF). If the largest VIF value is greater than 10, this outcome would indicate a serious multicollinearity problem (Field, 2018). Acceptable VIF values are, ideally, below 1, which indicates that the regression is unbiased.

2.6 Eye-tracking Methodology

Chapter 5 uses eye-tracking methodology to investigate the bilingual effect in an Attention Network Task (ANT) paradigm. This section introduces eye-tracking methodology in general and discusses how it has been used in past research, as well as how it will be used in the current thesis. A chapter-specific theory will be discussed in Chapter 5, the Attentional Control theory. This model highlights how bilingualism may potentially affect attentional processes, which then trickle down to other general cognitive areas.

Modern eye trackers use a technique called pupil centre corneal reflection (PCCR), using near-infrared illumination, to track eye movements when viewing visual stimuli. The near-infrared light creates reflections in the cornea and pupils, which an image sensor tracks and processes to obtain real-time spatial positioning of a subject's gaze. As the eyes move, the computed gaze position in space also moves, providing researchers with metrics, such as fixations, saccades, pursuits, and blinks (for review, see Skaramagkas et al., 2021).

These metrics are strongly coupled with a whole host of perceptual and cognitive processes (Hessels & Hooge, 2019; Pradhan et al., 2018). Researchers may use these metrics complementarily with traditional RT measures as a triangulation method. This conclusion is based on two core working assumptions (Cullipher et al., 2018; Hessels et al., 2019): 1) the immediacy assumption – the subject immediately processes the first object they fixate on before shifting their gaze to and processing the next object; and 2) the eye-mind assumption – whatever the eye sees, the mind processes. These assumptions allow researchers to measure implicit cognitive biases (e.g., facial recognition and racial prejudice), emotional recognition (e.g., stress response), and cognitive load (e.g., simple versus complex span working memory). Traditionally, the Inhibition-of-Return task (IOR), Antisaccade (AS) task, and ANT have been used to measure attentional control. A more detailed description of the ANT task is provided in Chapter 5.

In IOR tasks, participants respond to stimuli that are either cued or uncued. When stimuli appear in a cued location in less than 500ms (Satel et al., 2019), the response is slower compared to uncued trials. This inhibition of attention to a previously cued location encourages the orientation of attention toward novel stimuli. Mishra et al. (2012) found that highly proficient bilinguals show a greater IOR effect (occurring earlier and with slower responses) than less proficient bilinguals. The authors interpreted this finding as a more efficient disengagement of attention in highly proficient bilinguals, reflecting the frequent need for bilinguals to switch focus

between languages. In the context of the attentional model, this finding suggests that bilinguals may quickly orient their attention from one location to another, with this effect being modulated by their language proficiency. This interpretation was later supported by an ERP study by Grundy et al. (2020), who found that greater bilingual proficiency resulted in faster latencies of sensory perceptual processing (P1 and N1), later attentional control (P2 and N2), and stimulus categorization (P3) on IOR tasks. Grundy's findings also support the interpretation that IOR and EF are closely related. In a study with younger and older adults, age-related decline of cognitive functioning is associated with a delayed onset time of IOR (Li et al., 2018).

In AS tasks, participants must suppress the reflexive response to look at a peripheral stimulus and instead look in the opposite direction (antisaccade trials). In contrast, during prosaccade trials, participants are instructed to look directly at the stimulus as soon as it appears. In antisaccade trials, the stimulus location and the saccade goal are decoupled. Successful antisaccade performance requires top-down inhibition to prevent a reflexive saccade toward the stimulus and instead make a voluntary movement in the opposite direction (Munoz & Everling, 2004). Individuals with executive dysfunction or age-related cognitive decline typically perform worse on AS tasks, likely due to reduced top-down inhibitory control. Research on the effects of bilingualism on AS task performance is limited and has yielded mixed results. Liu et al. (2019) found that short-term, intensive training on monitoring and inhibitory control in bilinguals led to measurable improvements in the AS task. Bialystok et al. (2006) observed no significant differences between older monolingual and bilingual adults when saccades were used as the response trigger. However, when responses were tied to keypresses, bilingual older adults showed faster conflict resolution than monolinguals. This finding suggests that early, automatic components of attention may not differ between the two groups, but later stages requiring top-down executive control, such as inhibiting motor responses, may be affected by bilingualism. In contrast, Paap and Greenberg (2013) found no consistent evidence of a bilingual effect on the AS task.

In both tasks described above, saccadic eye movements were used as a measure of attentional control. Saccades are rapid, coordinated eye movements towards a point of interest, characterized by metrics, such as velocity, acceleration, and latency. Saccades are among the most understood ocular movements about cognition and memory that are easily measured (Leigh & Kennard, 2004).

Another metric of interest is fixations. According to Liu et al. (2022), fixation-related metrics, such as duration and frequency, may inform us of the perceptual and cognitive load of a particular stimulus. High perceptual loads are characterized by low fixation durations and high fixation frequencies, indicating that the complexity of stimuli determines the number of attended components. For instance, in visual search tasks, participants engage in visually sorting through numerous stimuli to locate the target. In contrast, high cognitive loads are marked by high fixation durations and low fixation frequencies. For instance, participants observe and mentally rotate geometric shapes to match a given model in mental rotation tasks.

One of the major limitations of eye-tracking methodology is the difficulty in interpreting eye movements, which do not always map clearly onto cognitive processes. Longer fixation durations may suggest higher cognitive loads, but they might also represent ambiguity or curiosity. However, a carefully designed experiment and attention to the psychophysical properties of the stimuli may help mitigate these concerns. Using both behavioral and eye-tracking metrics in the study in Chapter 5 allows for the triangulation of specific cognitive processes affected by bilingualism, which is difficult to achieve using either method on its own (Incera, 2018).

2.7 Electroencephalography (EEG) Methodology

Chapter 6 uses EEG methodology to investigate the bilingual effect in an N-back paradigm. Similar to the previous section, this section covers the methodology in general, how it has been used in past research, and how it will be used in the current thesis.

EEG is an imaging method that studies the brain's electrical activity using carefully positioned, highly sensitive silver/silver chloride (Ag/AgCl) electrodes on the scalp (Luck, 2014). These electrodes measure the absolute electrical potentials generated by pyramidal neuron bodies in the cerebral cortex within a roughly 10 cm² area. To facilitate this measurement, researchers use a conductive medium, such as a highly viscous gel, between the electrode and the scalp. This medium enhances the electrodes' sensitivity to electrical potentials in the cortex and reduces noise from surrounding electrical sources. Because the electrodes are placed on the scalp, which is not a good conductor, EEG is suitable for measuring electrical activity in the cortex but not in subcortical regions. Additionally, due to the 10 cm² area of electrical activity recorded, EEG has poor spatial localization. However, because electrical signals propagate almost instantaneously throughout the

brain, EEG offers excellent sub-millisecond temporal resolution, compensating for its lower spatial resolution.

For this reason, EEG is highly suited to measure EF abilities because specific segments of an EEG, called event-related potentials (ERPs), can be observed to understand specific cognitive processes and their time courses (Incera, 2018). ERPs may be linked to a response (response-locked ERPs) or exposure (stimulus-locked ERPs) to a specific stimulus. Typically, ERPs are isolated from the broader EEG signal by averaging across multiple trials to filter out electrical artifacts. The typical nomenclature of an ERP consists of two items: 1) the polarity (negative-going or positive-going waves) and 2) the timing of the component. For an N200 component, it is generally understood that it is a negative-going wave occurring around the 200ms mark. For the most part, the polarity of a component should not be interpreted beyond the organization of pyramidal cells under the electrode (Luck, 2014). If the sum of the pyramidal cells generates a net positive charge, the electrode will detect a positive-going wave; if they generate a net negative charge, a negative-going wave will be recorded. On the other hand, early ERPs occurring around 100 ms, are traditionally associated with early-stage visual information processing.

Language-based studies (for review, see Proverbio, 2022) traditionally investigate the N1 (Orthography), Mismatch Negativity (MMN; Phonology), N400 (Semantics), and P600 (Syntax). Generally speaking, the N170 component is linked to orthographic analysis, where real words produce a sharp increase in amplitude compared to non-words. The phonological MMN, peaking around 300ms after stimulus onset, is associated with phonological processing (Lewendon et al., 2020). The N400 component is related to semantic incongruence and violations of expectancy (Kutas & Federmeier, 2011). Lastly, the P600 component is associated with violations of syntactic rules (Gonda et al., 2020). Later ERPs typically reflect controlled, conscious, cognition (Luck et al., 1994). There are many types of ERPs, each associated with different cognitive processes, but a review of all ERPs is beyond the scope of this thesis. An overview of the ERPs related to EF will be provided here instead.

EF studies commonly investigate the N2 and P3 components, also referred to as the N2/P3 complex (Brydges et al., 2014; Downes et al., 2017; Kopp et al., 2020). The N2 has been associated with conflict monitoring in children and adults (Veen & Carter, 2002) and correlates with updating and manipulation of items in the working memory (López Zunini et al., 2016). For example, in the N-back task, a widely used measure of working memory (Warchoń & Zajac-Lamparska, 2023), a

large negative peak around 200ms has been observed when a stimulus is compared to another held in working memory for mismatches (López Zunini et al., 2016). The P3 component can be further delineated into two types: P3a and P3b. The P3a component is involved in the engagement of attentional control to orient or reorient attention, whereas the P3b component is involved in memory processing (Polich, 2007). Both of these subcomponents typically peak around 300ms and are generally found to be larger in more cognitively demanding EF tasks (Demirayak et al., 2023; Kok, 2001). By combining EEG/ERP data with traditional behavioral measures, researchers may gain a more comprehensive and insightful understanding of cognitive processes.

2.8 Summary

All chapters use the LHQ, MacArthur's SSS Scale, and Raven's SPM as measures of language history, SES, and non-verbal IQ, respectively. In addition, Chapter 3 includes the use of the Language Attitude Questionnaire for Language Learners (LAQ-LL), which will be explained in that chapter as it is only used once in this thesis. Similarly, the use of the Multilingual Switching Questionnaire (adapted from the Bilingual Switching Questionnaire; Rodriguez-Fornells et al., 2012) was only used once in Chapter 5 and will only be described in detail in that chapter. For Chapters 4, 5 and 6, behavioral (RT and accuracy) measurements were used, but only Chapter 5 used eye-tracking and only Chapter 6 used ERP methodology. The inclusion of eye-tracking and EEG methodologies within these chapters serves to provide general background information. Details are included to provide some basic details on how these methodologies may be used to measure cognitive processing in general. Related research on how these techniques can be used to measure EF is given in the relevant chapters.

Chapter 3: Understanding Attitudes towards Bi/Multilingualism and Language Learning Experiences in Malaysia

3.1 Introduction

Most of the research on attitudes towards multilingualism is heavily focused on Western countries (Sánchez-Pérez & Manzano-Agugliaro, 2021; Strani, 2020), this emphasis could be due to the West's greater emphasis on linguistic homogeneity (Pulinx et al., 2017). Additional language learning, either as the process itself or as a proxy of multilingualism, is affected by many sociological factors. These factors include, but are not limited to, sex, ethnicity, age, and socioeconomic status (SES; Ellis, 1994). These elements interact with one another in complicated ways, making it impossible to determine the precise role that each element plays. In the current study, the scope of the investigation will focus on the influence of age, socioeconomic status (SES), education, and language proficiency.

To establish proper contexts, it is important to first discuss the influence of sociopolitical factors before exploring individual factors that affect language attitudes. For example, it is well known that learning a second language offers numerous socioeconomic benefits to the individual. Ninety-five percent of research papers written between 2006 and 2016 supported the implementation of bilingual education in the United States (Lewis & Davies, 2018). Yet, despite the generally positive findings regarding second language learning, public attitudes towards bi/multilingualism in some Western countries (e.g., the United States) remain contested. For example, only 45% of media articles reporting on bilingualism cited support for learning an additional language in the United States between 2006 and 2016 (Lewis & Davies, 2018).

According to the Self-Determination Theory (SDT; Ryan & Deci, 2000), autonomy, defined as a sense of ownership over one's actions, can be undermined by external rewards or punishments (Jovanovic & Matejevic, 2014). When strict language policies are enforced, as can be seen in Taiwan (Chen, 2020) and Malaysia (Mostafa, 2016), even highly internally motivated learners may have their motivations suppressed (these case studies will be discussed in detail later). This shift occurs because compulsory language policies can change the learner's perception of the locus of control over their behaviour from internal to external (Verme, 2009). As a result, learners may experience a less enjoyable, externally motivated, language learning experience (McMillan & Rivers, 2011). Autonomy in language learning fosters positive attitudes towards the language,

whereas low internal motivation and autonomy are associated with poor language learning outcomes (Gottfried et al., 2008). When learners have poor language proficiency but are continually reinforced (through rewards or punishments) to learn it, they may become resentful towards the language (Wang et al., 2021). For example, Tan (2006, 2014) found that individuals who belong to the linguistic minority feel disadvantaged in a society speaking the majority language.

In a globalized world, bi/multilingualism³ is increasingly recognized as a valuable commodity (Ramírez-Esparza et al., 2020). As a result, the need to stay competitive globally has heightened many countries' focus on effective language planning and education to promote multilingualism (Bianco, 2014; Preece & Marshall, 2020). Effective language planning, however, is not a simple task. The following section highlights cases of less effective language planning in Taiwan and Malaysia and one case of effective language planning in Singapore.

In Taiwan, where the official language is Mandarin Chinese since 1946, Chen (2020) reported major language policy changes in the early 2000s. These changes included requiring English proficiency for promotion in government services and lowering the commencement of formal English learning from 10 to 9 years old. These implementations meant that learning English as a second language became necessary to secure a job in the government sector. These changes were implemented to equip its citizens in an increasingly globalized economy. Despite the good intentions, implementation faced challenges due to the lack of training for Taiwanese teachers. These teachers grew up learning Mandarin, being the medium of instruction. The sudden switch in requirements to deliver lessons in English or to teach the English language meant that many Taiwanese teachers were unprepared and under-equipped. The existing competitive spirit of Taiwanese education led to parents enrolling their children in English-only pre-schools at a younger age, further exacerbating the learning pressure that students already faced (Liu & Wei, 2023). In response, the government prohibited teaching English at the preschool level in 2004, citing that later exposure to English was better for the students. However, this decision did not align with the needs and wants of the parents. It was also not informed by current research, which

³ The term bilingualism and multilingualism can be used interchangeably. Where a “bilingual” is a term used to describe a person who speaks two languages, a “multilingual” is a term used to describe a person who speaks two or more languages.

agrees that, in general, the later the exposure to a language, the more difficult it is to achieve native-like proficiency (Jahrani & Listia, 2023).

In Malaysia, educational policies were implemented “top-down” without consulting key stakeholders, such as the Ministry of Education personnel, teachers, parents, or even students. According to Mostafa (2016), before the 2000s, all subjects in national schools were taught in Malay, the national language of Malaysia (see Liu & Ricks, 2012, for an extensive analysis of language politics in Malaysia). However, in 2002, the government introduced the PPSMI (*Pengajaran dan Pembelajaran Sains dan Matematik dalam Bahasa Inggeris*) policy, which changed the medium of instruction for Science and Math to English. This change caused a public backlash as it was seen as a threat to the status of Malay as the national language (Stephen, 2013; Yang & Ishak, 2012). Additionally, teachers were not equipped for the sudden change in language instruction, as many grew up learning and teaching Science and Math in Malay (Selamat, et al., 2010). This unpreparedness resulted in less than 5 percent of primary schools and 9 percent of secondary schools exercising the PPSMI policy (Yunus & Sukri, 2017). As it did not achieve its intended goals, the government reverted all teaching of Science and Math to Malay and officially abolished the PPSMI policy in 2014. In 2016, the government introduced the Dual Language Program (DLP) which allowed selected schools to teach science and mathematics in English.

In contrast, an excellent example of a “bottom-up” approach to policy implementation can be seen in Malaysia’s neighbouring country, Singapore. Ng (2012) outlines how policy implementation was closely monitored with feedback from students, parents, and teachers as seen in the *Report on the Ministry of Education 1978* (also known as the Goh Report; Goh, 1979). In summary, the Goh report identified three main shortcomings of Singapore’s education system: high education wastage, low levels of literacy, and ineffective bilingualism. These three issues stemmed largely from a peculiar situation brought about by the then-existing bilingual policy where the language of instruction (English and Mandarin) was not used at home by 85% of schoolchildren (Soon, 1988). Insights from this report were then used as a cornerstone for the establishment of the New Education System in 1980, which has since been recognized as a highly successful education regime globally (Dama, 2020). Among the major changes implemented by the Goh report was the streamlining of students into different courses based on their degree of language proficiency.

Insights from Singapore make it clear that, for successful language planning, it is important to understand the current sociolinguistic climate. This understanding ensures that resources are directed to fulfilling outcomes that will improve the receptibility of language policy changes. Understandably, this is a huge undertaking; however, a good first step is to understand the public's attitudes towards bi/multilingualism and their language learning experiences. Doing so provides a rich contextual background upon which larger-scale research can be built.

Building on Ecological Systems Theory (Bronfenbrenner, 1986), macrosystems and language policies play a role in shaping language attitudes at the grassroots level. While Chapter 1 highlights the significance of language use in demonstrating the bilingual effect on cognition, it is also essential to recognize that macrosystems influence language attitudes, which both shape and are shaped by language use. Within this socio-political and policy-driven context, discussions on the individual factors influencing language attitudes can be better contextualized. Additionally, this perspective provides a relevant sociolinguistic framework for interpreting measures of language use in the experimental chapters that follow.

3.1.1 The Influence of Age

Although age is an influential factor in the outcomes of the language learning process, with younger learners typically being better at acquiring a second language than older learners (Ellis, 1994), its relationship with language learning attitudes is not so straightforward. Furthermore, rather than as a continuum, age may be more appropriate as a categorical variable. For example, Zheng and Zhou (2021) analyzed population censuses from 2006 to 2011. They found that individuals who experienced the handover of Hong Kong's independence in 1997 during their schooling years were more likely to have a positive attitude towards the Mandarin language compared to other languages. This preference for Mandarin was not observed in older or younger generations who did not experience the handover during their schooling years. Such nuanced findings will likely be overlooked if age were to be measured on a continuum.

Similarly, Medeiros et al. (2019) discussed the antecedents of public support for bilingualism in Canada and Finland using each country's electoral survey data from 2015. The authors discovered two key findings. First, in general, the linguistic majority tends to show lower support for bilingualism compared to linguistic minorities. Second, demographic factors, such as age, may have contrasting effects depending on the country. In Canada, older Canadians are often

less supportive of bilingualism, but older Finns are typically more supportive of bilingualism. Based on the findings of Zheng and Zhou (2021), it is speculated that this difference may be due to the earlier adoption of bilingualism by Finland (see Language Act enacted in 1922; Ministry of Justice, Finland, 1996) compared to Canada (see Official Language Act enacted in 1969; Hudon, 2023). Hence, there is a compelling reason to examine the impact of cohorts on attitudes toward multilingualism and language learning. However, given that the categorization of cohorts can vary by country and involves a degree of subjectivity, it might be wise to assess cohorts as a continuous measure of age, especially when the exact parameters of a cohort are unclear. This approach is especially relevant when considering the back-and-forth language policy changes in Malaysia, which complicate the delineation of cohorts from the general population.

3.1.2 The Influence of Socioeconomic Status (SES), Language Proficiency, and Education Level

In Section 2.1.1, an overview of SES was provided. Although it is worthwhile investigating the impact of SES on language attitudes, a more accurate picture can be obtained by investigating the specific factors associated with high SES. In the context of the current study, language proficiency and education level were also investigated. This approach was inspired by Albury (2019), who found that Malaysian students had a positive attitude towards English as a second language because a good command of the language is required for work and international study (see also Albury, 2020). In many studies investigating language attitudes, the influence of SES has rarely been featured as an attitudinal variable (Gayton, 2010; Sicam & Lucas, 2016). This absence is unusual given that SES is well-documented to influence family language practices (Hoff, 2013; Hoff & Ribot, 2015; Schwab & Lew-Williams, 2016). Children from higher SES families typically demonstrate greater vocabulary growth and improved language and communication skills due to the higher quality and quantity of language input from their parents (Hoff et al., 2024). Yet, the relationship between SES and language attitudes is still under-researched.

Ladegaard (2000), using questionnaires, investigated attitudes and behaviours in language among Danish working- and middle-class adolescents. They found no significant difference in attitudes between the two SES groups. In contrast, Bokhorst-Heng and Caeon (2009), also using questionnaires, investigated the language attitudes of Singaporean youths and found that children from lower SES families held more positive attitudes towards other local languages than those

from middle and higher SES families. Sohidi and Rifai (2022), using qualitative interviews, examined the language attitudes among Indonesian university students and found that higher SES students tend to have more positive attitudes towards a second language compared to their lower SES counterparts.

These contradicting findings further underscore how SES should not be taken at face value, instead, it should be complemented with specific factors associated with higher SES that would influence language attitudes. A likely mediating factor could be the learner's exposure to different language cultures. In Gayton (2010), the SES-language attitude relationship was explained by investigating the learner's exposure to different language cultures. Through a series of interviews with high school teachers in Scotland, Germany, and France, the author found that all the teachers attributed higher language learning motivation and positive language attitudes in students to travelling experiences. They explained that students who have never travelled abroad often find it difficult to understand the value of learning a second language.

Education Level. Education level can either be conceptualized as parental education or the learner's education level (Long & Renbarger, 2023). Novak and Morvai (2019) found support for both forms of educational level and its relationship with language attitudes. First, a low level of maternal education was associated with extrinsic language learning motivations. Mothers from lower education backgrounds see language learning as only a requirement to obtain a degree or a job (Rahman & Singh, 2022), an attitude that their children share. Second, in terms of the learner's education level, it was found that higher education level changes the language learning motivation from extrinsic to intrinsic. Put differently, at lower levels of education, Degree students' main motivation for learning a language was to pass the course requirements. On the other hand, Master students' main motivation was more internally driven. They see second language learning as a means to connect with other cultures and to keep in touch with foreign students on campus.

Language Proficiency. In Singapore, Xie and Cavallaro (2016) found that the self-rated language proficiency of local youths is the best indicator of their attitudes towards Chinese-English bilingualism while accounting for current occupation, gender, and SES. Specifically, those who had equal language proficiency in both Chinese and English had the most positive language attitudes. This finding comes as no surprise as the socio-economy of Singapore highly values multilingual individuals (Cavallaro & Chin, 2014). Hence, equally proficient bilinguals reap the full rewards of being bilingual, which likely influences their positive outlook on languages.

3.2 Language Attitudes in Malaysia

Mostafa (2016) highlights that the language attitudes in Malaysia are intertwined with the discourse on Malay ethnonationalism. In Malaysia, the constitution dictates that the Malay language holds supremacy over other languages, hence, its exclusive use in national governance (Albury & Aye, 2016). The ethnic majority perceives the integration of the English language as a second language as a threat to the status of the Malay language. Hence, there is a long-standing belief that bilingualism harms Malaysian society in certain contexts, potentially polarizing the people and dividing the nation (Lee et al., 2010; Mostafa, 2016; Wong et al., 2012). Not all Malaysians likely hold the same belief. Considering that Malaysia is primarily made up of Malays (67%), Chinese (25%), and Indians (8%) (Department of Statistics, 2015), the linguistic majority may have a poorer attitude towards bilingualism compared to linguistic minorities. Similar trends have been observed in countries like Canada and Finland (Medeiros et al., 2019).

A deeper qualitative investigation aligns with these findings. Albury (2020) interviewed Malaysian university students on the sociolinguistic landscape in Malaysia using 25 semi-structured group discussions at five public and five private universities. Their findings provide valuable qualitative insights into Malaysian youth's language attitudes. The discussions were separated into three discourse analyses: Malay, Chinese, and Indian discussions. Although interethnic tensions were present in all three discourses, related to the position of the Malay language as the national language (Albury, 2019; Mostafa, 2016), the students frequently brought up the benefits of multilingualism on socioeconomic mobility. There was also a strong and positive support for societal multilingualism in the construction of their Malaysian identity.

When it comes to language learning motivations, Muftah and Rafic-Galea (2013) reported that university students ($N = 182$) were highly motivated to learn English as a second language. However, these motivations were primarily instrumental. For example, they reasoned that they are motivated to learn English because "The atmosphere in University requires us to be able to speak and communicate in English" (Muftah & Rafic-Galea, 2013, p. 97). This attitude is expected given Malaysia's mandatory bilingual education laws, which, as earlier argued by Verme (2009), would shift the learner's perspective of the locus of control over their behaviour from internal to external. As these are undergraduate students, these findings also align with Novak and Morvai (2019) and

Rahman and Singh (2022) who observed that degree students are more instrumentally motivated to learn a second language to pass the requirements set by the university.

On the other hand, Chan and Abdullah (2015) investigated a variety of teachers' perspectives ($N = 39$) on bilingualism. Results suggested that 54% of teachers perceive bilingual education as highly important; however, 41% perceive it as only moderately important. Similarly, 41% of the teachers reported that knowing more than one language was of no consequence to them, that is, they did not believe that bilingualism carries any significant advantages. The authors extrapolated this finding and concluded that the teachers' lack of enthusiasm towards bilingualism meant that they do not emphasize language learning, nor do they present themselves as strong role models for bilingual users. Consequently, the teachers' attitudes could detrimentally affect students' instrumental or integrative motivation to learn languages beyond the first.

When it comes to language teachers specifically, a qualitative case study (Too, 2023) found that all seven of the interviewed English language teachers engaged in the practice of translanguaging. This practice is defined as a "process of using language(s) to gain knowledge, to make sense, to articulate one's thought and to communicate about using language" (Li, 2011, p. 1223). Language is used as a scaffolding tool to help students in their learning. In practice, this process takes the form of using familiar home languages, such as the Kadazan language, to facilitate English language learning (Too, 2023). The increased use of bilingual pedagogies reflects a shifting perspective on the usefulness of multilingualism. This finding contrasts with the aforementioned study of Chan and Abdullah (2015), who found a significant portion of Malaysian teachers' perspectives on bilingualism to be mostly negative. Speculatively, the COVID-19 pandemic may have played an important role in this language attitude shift as it forced teachers to innovate their teaching practices (Schug & Sarnou, 2024).

What remains unclear is the factors associated with positive language attitude change. This would provide valuable insights for Malaysian policymakers advocating for multilingual education. A successful multilingual education program offers numerous benefits, including improving ethnic and national identities, promoting linguistic diversity, positively impacting cognitive and academic development, and easing sociopolitical tensions (Benson, 2002). Not to mention, language attitudes also affect language practices (Kircher & Zipp, 2022). If there are cognitive effects to multilingualism, it would only be afforded to those who actively practice the languages (Pliatsikas, 2020). Therefore, without clearly understanding what factors influence a

change in language attitudes, Malaysian policymakers may implement top-down policies that are detracted from the actual needs and wants at the grassroots level. For example, despite the laws that stipulate Malay to be the language of instruction, many publicly-funded universities use English as the language of instruction (Gill, 2013)

3.3 Measuring Attitudes Towards Multilingualism and Language Learning Experiences

There is a lack of psychometrically robust measures of language attitudes (Li & Wei, 2022). Recently, two measures have shown promise: the Language Attitude Questionnaire for Language Learners (LAQ-LL; Artamonova, 2020) and the Language Attitudes Scale-Student Form (LASS; Li & Wei, 2022). These surveys were developed based on the theoretical tripartite model of attitudes (Eagly & Chaiken, 1993), which states that attitudes may be divided into three components: affect (feelings), behavior (the influence of the attitude on behavior), and cognition (beliefs and knowledge).

LAQ-LL. Based on a Western demographic, the LAQ-LL sampled 127 students from an American college that majored or minored in Spanish (Artamonova, 2020). The initial questionnaire consisted of 43 items, which were grouped into five attitudinal areas: 1) Attitudes towards multilingualism, 2) attitudes towards language learning (in general), 3) attitudes towards a specific language (Spanish), 4) attitudes towards language classes (Spanish), and 5) attitudes towards the language group (Spanish speakers) and the target language culture (Spanish and Latin American cultures). Following an exploratory factor analysis (EFA) and the resulting principal component analysis (PCA), the author reduced the LAQ-LL to 28 items which were grouped into three subscales with high Cronbach's α : 1) Language Learning Experience ($\alpha = .880$), 2) Sociocultural Appeal (SA; $\alpha = .864$), and 3) Value of Multilingualism ($\alpha = .767$). Language Learning Experience reflects the behavioural component as it includes attitudes developed because of participating in language learning activities. SA refers to the affective component and includes the sentiments and emotional reactions elicited by the language, its speakers, and linked cultures. Value of Multilingualism may be regarded as expressing the cognitive component, which includes the logical knowledge of the benefits associated with multilingual ability (Artamonova, 2020).

LASS. Based on an Eastern demographic, the LASS sampled 5,237 primary ($M_{\text{age}} = 11.58$) to tertiary ($M_{\text{age}} = 19.60$) school students across 7 schools within the province of Hubei, China. Unlike the LAQ-LL, the authors conducted a confirmatory factor analysis (CFA) as their 10-item

questionnaire was developed based on 1) preliminary interview data with schoolteachers and (2) past instruments on language attitudes. Hence, the developed questionnaire had a strong theoretical foundation that warranted a CFA. Given the wider demographic, the analysis was split by ethnicity (i.e., Han, Tujia, and Miao) and language (i.e., Dialect, Putonghua, and English). The resulting CFA reported acceptable Cronbach's α levels (above .65, except for the Putonghua cognition subscale) across affect, behaviour, and cognition subscales across each language.

The present study chose to adapt the Language Learning Experience and Value of Multilingualism subscales from the LAQ-LL instead of the LASS, because the questions could be more easily adapted to reflect general language attitudes rather than attitudes towards a specific language or language culture. For example, in the LAQ-LL, questions that specified a language were rephrased as questions about languages in general. Thus, "Spanish classes are always among my favourite classes" became "Language classes are always among my favourite classes." However, in the LASS, similar affective questions like "Do you feel emotionally attached to [your ethnic language/dialect/Putonghua/English]?" became too ambiguous when phrased about languages in general, such as "Do you feel emotionally attached to other languages?".

The SA component was omitted because modifying the questions to reflect a general language attitude would similarly introduce ambiguity and reduce the clarity and accuracy of the responses. For example, changing "I find Latin American cultures interesting" to "I find multilingual cultures interesting" becomes too ambiguous, potentially causing items to cross-load onto multiple factors during an EFA.

3.4 Current Study

The present study aims to investigate which individual factors can predict changes in Value of Multilingualism and Language Learning Experience scores. It is anticipated that the public will generally hold positive sentiments towards multilingualism, given Malaysia's multicultural history. Understanding this is crucial because attitudes toward multilingualism influence language use practices within a society (Kircher & Zipp, 2022). As established in Section 1.2.1, the exhibition of the bilingual effect may be associated not with how many languages a person knows but with how they use them in daily life. However, it is hypothesized that age, SES, education level, and language proficiency will moderate this association. Only the first and second language proficiency was included as predictors because not all participants reported a third or fourth

language. The first hypothesis was that age would be a significant predictor of Value of Multilingualism and Language Learning Experience scores, with older participants predicted to have lower Value of Multilingualism and Language Learning Experience scores. The second hypothesis was that SES, education level, and language proficiency would be significant predictors of Value of Multilingualism and Language Learning Experience scores, with higher SES, education levels, and language proficiency predicting higher Value of Multilingualism and Language Learning Experience scores.

3.5 Method

3.5.1 Design

The present study adopted a correlational design, using multiple regression with three predictors and two outcome variables. The first predictor is age, measured on a continuum. The second predictor is SES, measured by the MacArthur Scale of Subjective Social Status (MacArthur SSS Scale) – Adult Version. The final predictor is the highest education level: No formal education, primary school, secondary school, diploma, bachelor’s degree, master’s degree, or doctoral degree. The answers to the educational attainment question were dummy-coded to a range of 0 (No formal education) to 6 (Doctoral degree). The outcome variables were the mean score of the Language Learning Experience and Value of Multilingualism subscales adapted from the LAQ-LL (Artamonova, 2020).

3.5.2 Materials

Language proficiency was calculated using the **LHQ** (Li et al., 2020). See Section 2.3 for more information. SES was measured using **MacArthur’s SSS Scale**, and non-verbal intelligence was measured using a short version of **Raven’s SPM**, see Chapter 2 for further details.

Modified LAQ-LL. Several questions were modified to ensure its suitability for the Malaysian context. This modification was done by changing all specific mentions of a language or a language culture into language in general. However, certain items were made redundant because of these changes. For example, in the Language Learning Experience subscale, “Learning Spanish is fun” (Item 3) would mean the same as “Studying languages is fun” (Item 4), and “I really enjoy learning Spanish” (Item 5) the same as “I like learning languages” (Item 6) after being rephrased. As such, Items 3 and 5 were removed from the current study to avoid cross-loading

(Costello & Osborne, 2019). The original Language Learning Experience and Value of Multilingualism subscales from the LAQ-LL have been demonstrated to be reliable at Cronbach's α of .88 and .77, respectively, which meets the requirement of reliability of .70 or greater. The items were rated on a 7-item Likert scale from Strongly Disagree to Strongly Agree. In the modified subscales, Language Learning Experience Item 5 and Value of Multilingualism Items 1, 2, 7, and 8 were reverse-coded. The questionnaire was provided in English and Malay (as they are the two most common languages in Malaysia). To ensure translation accuracy, two native English-Malay speakers performed a forward-backwards translation of all questionnaire items (see Appendix 1). Only three participants chose to complete the questionnaire in Malay.

3.5.3 *Participants*

A two-tailed a-priori power analysis was conducted with the following parameters: $f^2 = 0.08$, α error probability = 0.05, power = 0.80, and number of predictors = 3, arriving at a sample size of $N = 101$. Participant inclusion criteria are as follows: 1) above 18 years old as of 1st January 2023, 2) a Malaysian citizen, 3) reside in Malaysia since birth, and 4) never lived abroad for more than a month. To ensure that there was enough variability in sampling, the call for participants was distributed to both East and West Malaysia through social media and word of mouth. Advertising the call for participants directly in universities was avoided as typically, paid surveys are snowballed amongst university students, leading to biased sampling among young, educated, higher SES, adults. Despite this, due to the organic nature of survey distribution (i.e., via social media and word of mouth), these surveys eventually reached university students, as evidenced by a higher response rate from 18 – 24 years old. This study initially recruited 114 participants; however, after removing three duplicate submissions, the final total number of participants was 111 (72 females). Participants consisted of 77 Chinese, 18 Malay, 11 Indigenous Ethnic, 4 Indians, and 1 Eurasian.

The first language of participants includes dialects of Chinese (Mandarin: 41, Cantonese: 5, Hakka: 3, Hokkien: 2), English (37), and Malay (23). One-hundred seven participants reported knowing a second language, 71 participants reported knowing a third language, and 30 participants reported knowing a fourth language. During primary education, nine participants experienced a trilingual medium of instruction ($N = 8$ English-Chinese-Malay, $N = 1$ English-Malay-Arabic), 24 experienced a bilingual medium ($N = 5$ English-Chinese, $N = 18$ English-Malay, $N = 1$ Malay-

Chinese), and 78 experienced a monolingual medium of instruction (N = 8 English, N = 21 Malay, N = 49 Chinese). In secondary education, 12 participants had a trilingual medium (N = 11 English-Chinese-Malay, N = 1 English-Malay-Arabic), 38 had a bilingual medium (N = 6 English-Chinese, N = 28 English-Malay, N = 4 Chinese-Malay), and 61 had a monolingual medium (N = 24 English, N = 20 Malay, N = 17 Chinese). In tertiary education, five participants had a trilingual medium (English-Chinese-Malay), 29 had a bilingual medium (N = 7 English-Chinese, N = 22 English-Malay), and 76 had a monolingual medium (N = 68 English, N = 5 Malay, N = 3 Chinese). It was noticed that participants transitioned from a more diverse medium of instruction during primary education into a mostly English-centred medium of instruction during tertiary education. Further participant demographics can be found in Table 2 below.

Table 2: *Participant Demographic Measures*

Demographic	<i>Mean</i>	<i>SD</i>	<i>Lower Range</i>	<i>Upper Range</i>
Age (years)	28.51	8.18	18	53
SES	6.15	1.60	2	10
Education Level	4.01	0.84	2	6
L1 Proficiency	0.87	0.16	0.43	1.00
L2 Proficiency	0.78	0.16	0.36	1.00
L3 Proficiency	0.71	0.17	0.25	1.00
L4 Proficiency	0.70	0.18	0.25	1.00

Note: Answers for education level consist of no formal education, primary school, secondary school, diploma, bachelor's degree, master's degree, or doctoral degree. The answers were dummy-coded to a range of 0 (No formal education) to 6 (Doctoral degree). L1 – L4 proficiency here is the combined self-rated proficiency in reading, writing, listening, and speaking skills on a scale of 0 to 1 (see Materials).

3.5.4 Procedure

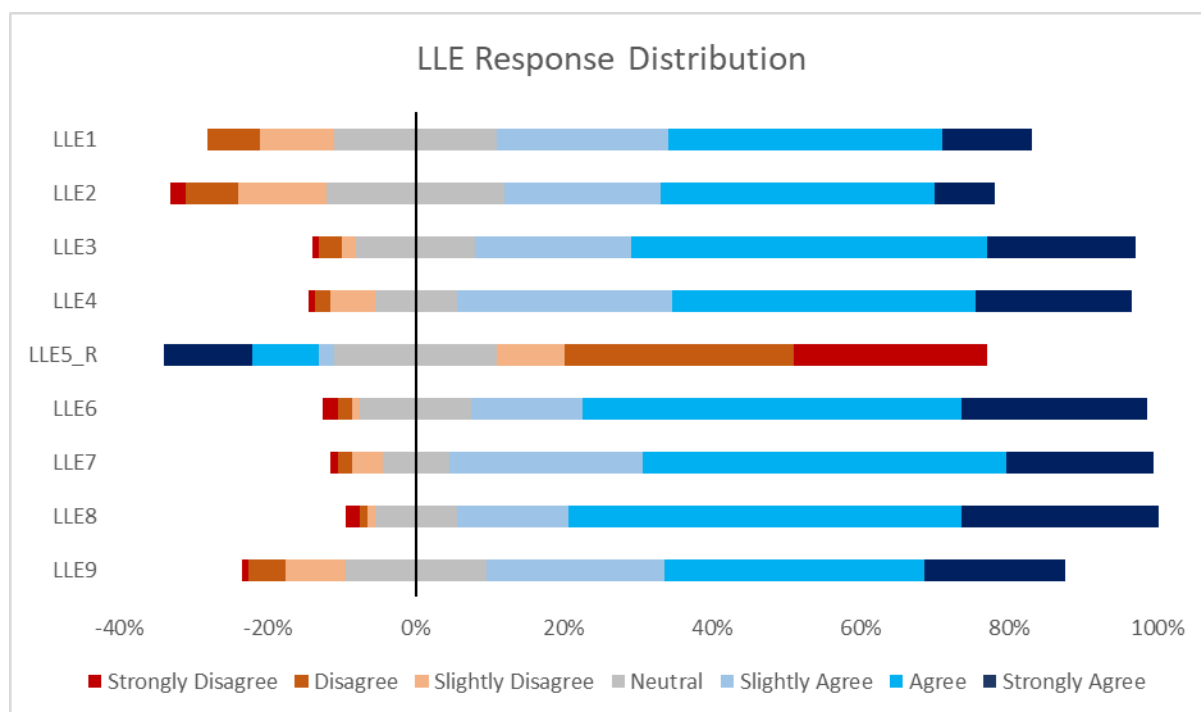
Before answering the questionnaires, all participants were given an informed consent sheet. Participants then completed a set of demographic questions, followed by the LHQ and the LAQ-LL questionnaire on Qualtrics (www.qualtrics.com), which could be completed in either English

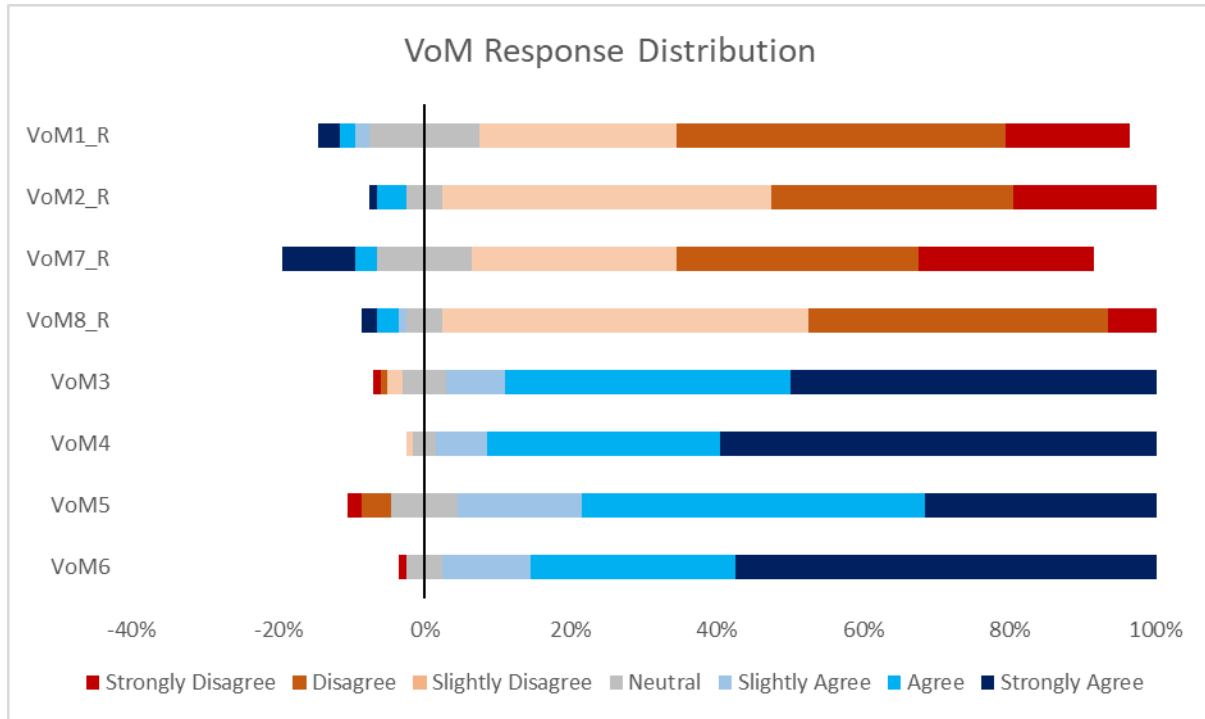
or Malay. The entire survey took approximately 10 minutes to complete. Participants were paid RM10 as compensation.

3.6 Results

The response distribution for both subscales can be seen in Figure 3. Overall, participants reported positive language learning experiences and highly valued multilingualism. However, responses for the Language Learning Experience subscale were less polarized than for the Value of Multilingualism subscale, as indicated by the difference in proportions of extreme responses (i.e., Strongly Agree). These findings mirror the findings from Artamonova (2020), who also found that students overwhelmingly agreed that speaking more than one language was beneficial but observed the largest variation in language learning experiences.

Figure 3: *Distribution of Responses on the Language Learning Experience and Value of Multilingualism Subscale.*





An EFA was conducted to ensure that the modified LAQ-LL used in our study has strong convergent validity with the original version, which was initially developed and validated for a Western population. It was also done to establish the reliability and validity of the modified questionnaire within the specific context of Malaysia.

3.6.1 Exploratory Factor Analysis

First, to ensure the internal reliability of the measure, Cronbach's α analysis was conducted. Results showed that all 17 items had a high global internal reliability at .89, with the Language Learning Experience subscale at .87 and the Value of Multilingualism subscale at .85. Based on Table 3 below, the removal of any items did not lead to a significant increase or decrease in internal reliability, therefore, all 17 items were retained.

Table 3: *Individual Item Reliability Statistics*

Item	If item dropped Cronbach's α
LLE_1: Language classes are always among my favorite classes	0.88

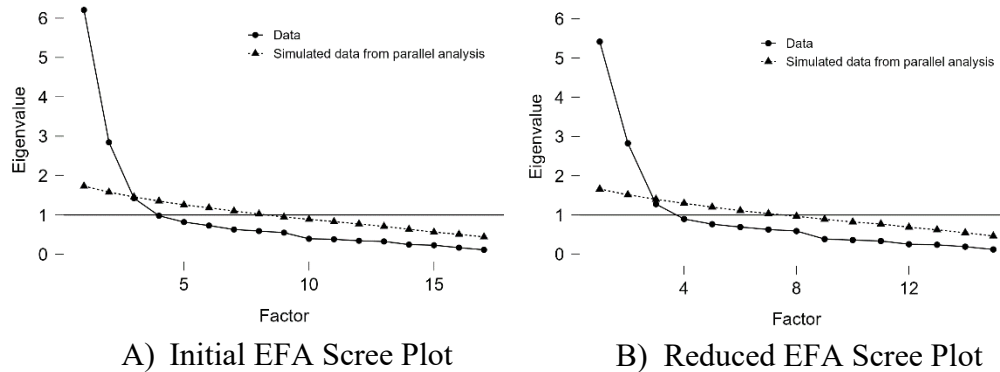
Item	If item dropped Cronbach's α
LLE_2: Language classes are more entertaining than other classes	0.89
LLE_3: Studying languages is fun	0.87
LLE_4: I like learning languages	0.87
LLE_5_R: Language classes are boring ^a	0.88
LLE_6: I like communicating with multilingual speakers	0.88
LLE_7: Studying languages is interesting	0.88
LLE_8: Learning languages is worth the effort that it takes	0.87
LLE_9: I would take language classes even if they were not required	0.88
VoM_1_R: I have no interest in learning languages	0.87
VoM_2_R: There is no need to learn other languages beyond my mother tongue ^a	0.88
VoM_3: I like the idea of being/becoming multilingual	0.88
VoM_4: Speaking more than one language is a valuable skill	0.88
VoM_5: Language study is a valuable pursuit	0.88
VoM_6: I admire people who speak multiple languages	0.88
VoM_7_R: Speaking just my mother tongue is enough for me ^a	0.88
VoM_8_R: Learning a second language is a waste of time ^a	0.88

Note: ^a indicates that the item was reverse-coded. LLE = Language Learning Experience; VoM = Value of Multilingualism

Next, an initial EFA was conducted based on PCA. Kaiser-Meyer-Olkin (KMO) test was .83, which is above the recommended value of .7 (Watkins, 2018), indicating that sampling was adequate. Bartlett's test of sphericity was significant ($\chi^2(136) = 1007.70, p < .001$), indicating that the items are not independent from one another and are thus suitable for factor analysis. Items with communalities above the recommended .4 value were retained. Oblique promax rotation was used as the items were expected to be associated with one another (Field, 2018). Factor extraction was based on the scree plot. Traditionally, the number of factors to be extracted depends on the position of the "elbow" of the scree plot, or where the line levels out. However, this is a problematic measure as evident in Figure 4A, the location of the "elbow" on the data plot is not immediately

clear (see Watkins, 2018). To address this, a parallel analysis was conducted by comparing the eigenvalues obtained from the actual data with the eigenvalues obtained from random data generated under the assumption of no underlying factors (see Figure 4B). The intersection between the data plot and the simulated data plot from the parallel analysis indicated that two factors should be extracted.

Figure 4: *EFA Scree Plot Comparison*



However, based on the resulting factor pattern matrix, a simple structure was not achieved as two items: LLE_6 and VoM_1_R were moderately loaded onto both extracted factors. As such, the EFA was conducted again by removing both items cross-loaded items. KMO test yielded a value of .82 and Bartlett's test of sphericity remained significant ($\chi^2(120) = 964.38, p < .001$). Again, the reduced EFA scree plot with parallel analysis (see Figure 4) confirms that two factors should be retained.

Consulting the structure matrix (see Table 4), a simple structure was observed with all items loading strongly and exclusively on either factor. The proportion of variance explained by both factors was 48.4%. Collectively, these results suggest the presence of two underlying factors labelled as "Language Learning Experience" and "Value of Multilingualism" respectively.

Table 4: *EFA Factor Pattern Matrix*

	Factor 1	Factor 2
LLE_1: Language classes are always among my favorite classes	0.77	
LLE_2: Language classes are more entertaining than other classes	0.59	
LLE_3: Studying languages is fun	0.80	

	Factor 1	Factor 2
LLE_4: I like learning languages	0.80	
LLE_5_R: Language classes are boring ^a	0.53	
LLE_7: Studying languages is interesting	0.77	
LLE_8: Learning languages is worth the effort that it takes	0.69	
LLE_9: I would take language classes even if they were not required	0.62	
VoM_2_R: There is no need to learn other languages beyond my mother tongue ^a		0.61
VoM_3: I like the idea of being/becoming multilingual		0.66
VoM_4: Speaking more than one language is a valuable skill		0.82
VoM_5: Language study is a valuable pursuit		0.62
VoM_6: I admire people who speak multiple languages		0.75
VoM_7_R: Speaking just my mother tongue is enough for me ^a		0.60
VoM_8_R: Learning a second language is a waste of time ^a		0.61

Note: The applied rotation method is oblique promax. Communalities below .4 were suppressed. LLE = Language Learning Experience; VoM = Value of Multilingualism. Item LLE_6 and VoM_1_R were removed for being cross-loaded onto both factors. ^a indicates that the item was reverse-coded

3.6.2 Regression Analyses

To understand how Age, SES, Education Level, and first and second language proficiency are related to the Language Learning Experience and Value of Multilingualism, a simultaneous multiple regression analysis was conducted for each outcome separately.⁴ The Language Learning Experience model was not significant ($F(5,105) = 0.90, p = .487$). The Value of Multilingualism model significantly predicted Value of Multilingualism scores, explaining 9.5% of the variance in the outcome ($R^2_{\text{adj}} = .10, F(5,105) = 3.37, p = .008$); Education Level significantly predicted an increase in Value of Multilingualism scores, where a one-unit increase resulted in a 2.11 unit increase in Value of Multilingualism scores ($B = 2.11, SE = 0.73, t(110) = 2.89, p = .005$). First

⁴ To evaluate the adequacy of these models, we examined the residual vs. predicted plot, QQ plots and Durbin-Watson statistics. For both models, the residual vs. predicted plots showed a random scatter of residuals around the zero line, indicating a satisfactory fit to the data. The QQ plots suggested acceptable normal distribution of the residuals. Finally, the Durbin-Watson tests for both models suggest no autocorrelation.

language proficiency also significantly predicted an increase in Value of Multilingualism scores, where a one-unit increase resulted in an 8.70 unit increase in Value of Multilingualism scores ($B = 8.70$, $SE = 4.19$, $t(110) = 2.08$, $p = .040$). The other factors did not predict Value of Multilingualism scores. The model had VIF below 1, suggesting no issues with multicollinearity.

3.7 Discussion

The current study is the first quantitative investigation into the factors affecting attitudes towards multilingualism and the language learning experience in Malaysia. Past studies have only investigated the perceptions of undergraduates (Albury, 2020) and teachers (Chan & Abdullah, 2015) on multilingualism, but not the factors that influence these perceptions. Hence, the present study investigated the relationship between age, SES, education level, and language proficiency on Language Learning Experience and Value of Multilingualism scores. Using simultaneous regression analyses, the study tested whether 1) age can predict lower Value of Multilingualism and Language Learning Experience scores, and 2) SES, education level and language proficiency can predict higher Value of Multilingualism and Language Learning Experience scores.

Overall, the sample demonstrated a generally positive language learning experience, that was consistent across age, SES, and educational level demographics. This outcome did not support the hypotheses on Language Learning Experience outcomes. A combination of national identity and intrinsic language learning motivation likely played a more influential role than age, SES, or education level. First, although language use has always been a contentious topic in Malaysian politics (Mostafa, 2016), the core identity of a Malaysian is rooted in a multicultural, multilingual identity (How et al., 2015; Rahman & Singh, 2021). Consequently, this identity shapes the preferences for language learning (Xie & Cavallaro, 2016). Second, the current sample demonstrated a high intrinsic motivation for language learning. This interpretation is evidenced by the fact that most participants responded positively to the question, “I would take language classes even if they were not required” (see LLE_9; Figure 3). During secondary school in Malaysia, language classes are compulsory due to the mandatory bilingual education policies (Liu & Ricks, 2012), which can also be observed in the current sample’s self-reported medium of instruction across education levels. However, in tertiary education and the workplace, language learning is optional, therefore, there are more variations in language learning motivation. Some may learn a language due to factors, such as having a personality that finds the learning experience enjoyable

for its own sake (Marszalek et al., 2022), having a growth mindset (Hu et al., 2022), or intrinsic motivations stemming from the ideal L2 self and the ought-to self (Mackay, 2019), whereas others may choose to learn another language due to external factors, such as fulfilling a job requirement.

Regarding the attitudes towards multilingualism, the results also do not fully support our hypotheses as no significant relationship was found between age and Value of Multilingualism scores or between SES and Value of Multilingualism scores. To understand these patterns of relationship, it is crucial to underscore Malaysia's colonial history from the 18th century (Stephen, 2013). During British rule, English was established as the language of education and trade. Being educated in English, often in prestigious schools established by the British, was a hallmark of the elite in the pre-independence era (Mostafa, 2016). After achieving independence, the English language was superseded by the Malay language as the medium of instruction in schools nationwide. However, the value of the English language in opening economic prospects could not be ignored, hence it remained a foreign language subject in educational institutions (Rahman & Singh, 2021). Most participants in the current study reported having English as the majority medium of instruction from primary to tertiary education. Over time, new generations internalize the importance of Malay as a national language, but also English, as a powerful economic, political, and scientific language (Wong et al., 2012). This conclusion is supported by the finding that multilingualism is highly valued regardless of age. Concerning SES, globalization, combined with the pre-independence context discussed earlier, has established multilingualism as a favourable attribute within Malaysia and across all SES strata (Kapong, 2022).

Importantly, there was a significant positive relationship between education level and Value of Multilingualism scores. At first glance, this relationship appears intuitive, given that higher education is associated with greater awareness of the benefits and importance of multilingualism in a globalized world (Surrain & Luk, 2019). However, this association oversimplifies the unique factors associated with a high education level that might, individually, better predict Value of Multilingualism scores. For example, Rahman and Singh (2022) discussed how English has been adopted as the medium of instruction for some programs in public HEIs and across all programs in private HEIs. This high level of spread necessitates a good command of English to be admitted entry in these institutes. Hence, many students value multilingualism because of the language requirements in many HEIs. These requirements align with the language learning motivations of undergraduate students (Novak & Morvai, 2019)

Another factor that is associated with higher education is the increased exposure to individuals from diverse cultural and linguistic backgrounds due to the internationalization of higher education (Teichler, 2017), that is, the process of integrating international, intercultural, or global dimensions into the delivery of tertiary education. This process is exemplified by student mobility across international borders due to twinning or transfer programs, which are typically offered by larger HEIs, fostering such movements. These programs offer students opportunities to complete their tertiary education partially in Malaysia and partially in another country, enabling students to connect firsthand with other cultures and language groups. Consequently, this exchange leads to a heightened awareness of the benefits and importance of multilingualism (see also Gayton, 2010).

The current results also support the hypothesis that language proficiency predicts positive changes in Value of Multilingualism scores. Language proficiency performed better than education level as a predictor of changes in Value of Multilingualism scores. This finding aligns with Xie and Cavallaro (2016) who found that, in Singapore, self-rated language proficiency was the best indicator of their language attitudes compared to gender, SES, and occupation. They noted that those who self-identified as equally proficient in English and Mandarin held the most positive outlook on bilingualism compared to those proficient in either language. This finding is because being multilingual is associated with the national Singaporean identity. However, the difference in language attitudes between language proficiency groups was minimal, meaning that all groups held generally positive language attitudes.

It is unclear why only first-language proficiency strongly predicted changes in Value of Multilingualism scores, but not second-language proficiency. It was speculated that the ethnicity of the participants may play a role in this finding. Half of the participants reported Chinese (or a dialect) was their first language. Due to the importance of Malay and English in the socioeconomic landscape in Malaysia, these participants may have realized that only knowing Chinese was not enough to remain socioeconomically competitive. Consequently, this realization may have fostered more positive language attitudes. This could also contribute to a sense of national identity, much like in Singapore, where the ideal Singaporean is perceived as someone who is multilingual (Xie & Cavallaro, 2016).

In the current sample, a rapid transition was observed from a linguistically diverse education in primary education to the dominance of English as the primary medium of instruction

in tertiary education. This transition is not an isolated phenomenon, as the position of the university as an influential social institution across the globalized world has established English as the academic *lingua franca* (Lin, 2020; Preece & Marshall, 2020). The issue here is not that English is the primary medium of instruction, rather, it is that linguistic diversity is being lost as students graduate from one education level to the next. This loss of diversity undermines the goals set out in Malaysia's Education Blueprint (2013–2025; Ministry of Education, 2013) which outlines that every Malaysian student should be bilingually proficient in at least Malay and English at the end of their formal education years. This goal has important implications for the overall aim of this thesis. If linguistic diversity is being lost, young adults may shift from a dual/dense code-switching context to primarily a single-language context. According to the ACH, such a shift may lead to systemic changes in cognitive outcomes for the young adult population.

Perhaps the way forward is less about choosing the appropriate medium of instruction, either Malay or English, and more about creating inclusive language policies that emphasize translanguaging (Lin, 2020; Phyak, 2023). There are concerns that mixing languages may be detrimental to learning outcomes, primarily because language switching is thought to increase the cognitive load (Meuter & Allport, 1999). Yet, these concerns are based on no empirical evidence (Antón et al., 2016). If anything, there is considerable evidence suggesting that children learn better in their mother tongue (Phyak, 2023; UNESCO, 2003). The question then arises, why not harness the vast linguistic capital in Malaysia to foster higher-quality learning in students? The prevailing hegemony of English in many educational institutions no longer holds up to today's multilingual, multicultural standards (Abrar-ul-Hassan, 2021; Taylor, 2021). Consequently, Malaysian policymakers should revisit and reconceptualize the education system to one that inspires both teachers and students to engage in translanguaging and multilingualism. Malaysian academics and state assemblies have recently advocated for teaching multilingualism and recognize that such a policy would require precise language planning (Jeevita, 2024). Embedding multilingualism early in education with consistent support can help future generations value multilingualism as a resource. This approach enables students to develop a diverse linguistic and communicative repertoire that can significantly benefit them across various facets of life (Kroll & Dussias, 2017; Pietrzyk-Kowalec, 2023).

Future research should consider investigating ethnic identities in relation to language attitudes. Ethnic identity has been documented to affect language attitudes (Trofimovich &

Turuševa, 2019; Zhang & Tsung, 2019). This difference in language attitudes is likely due to major ethnic groups constituting the linguistic majority, such as Malays speaking the Malay language. Based on an economic perspective, multilingualism may be perceived as a disadvantage to the linguistic majority, but an advantage to the linguistic minority (Bokhorst-Heng & Caleon, 2009). This perceived economic threat may be enough to form negative attitudes towards multilingualism and may disproportionately influence the linguistic/ethnic majority more than the linguistic/ethnic minority, but this has not yet been empirically explored (Medeiros et al., 2019).

The current research attempted to investigate this, however, because of the unbalanced sample sizes between the three main races (Chinese, Malay, Indians), a supplementary analysis was conducted by grouping the sample based on whether they were the linguistic majority (*bumiputra*-literally meaning sons of the soil, $N = 29$) or the linguistic minority (non-*bumiputra*, $N = 82$). A Mann-Whitney t-test revealed that linguistic minorities had significantly, but slightly less positive, language learning experience than linguistic majorities, however, this distinction does not diminish the overall positive view of language learning experience across all samples. Recall that *Bahasa Melayu* (BM) is designated as Malaysia's national language and primary language of instruction in national schools. In this context, *bumiputras* are typically native BM speakers and hence constitute the linguistic majority. Conversely, non-*bumiputras* must navigate the linguistic complexities of an education system tailored for the linguistic majority. For example, in order to obtain the *Sijil Pelajaran Malaysia* (SPM) – a national exam equivalent to the General Certificate of Secondary Education (GSCE), students must obtain at least a passing grade in *Bahasa Melayu* and *Sejarah* (i.e., Malaysian History) subjects. This requirement imposes an additional linguistic cost on students whose first language is not BM. This cost may have contributed to the observed negative relationship between ethnicities and Language Learning Experience scores. However, caution is warranted in interpreting this conclusion due to the unequal sample size between *bumiputras* and non-*bumiputras*, yet this highlights the possibility that ethnic identities may influence Language Learning Experience and Value of Multilingualism scores.

This study aimed to investigate which individual factors could predict changes in the value of multilingualism and language learning experience scores. The first key finding is that there is a generally positive value of multilingualism and language learning experience in the current sample. Second, higher education level and first language proficiency predicted higher value of

multilingualism scores, which aligns with past research. However, neither of these variables predicted changes in language learning experience. Additionally, changes in SES, age, and second language proficiency did not predict any changes in both subscales. These findings contradict past research that had found that SES and age can predict changes in language attitudes. Although these predictors may be associated with language attitude changes in some countries, their influence may be limited in others. Depending on the past and present social factors of a community, SES and age may exhibit different relationships to language attitudes.

It is important to note that the current results are biased toward the Malaysian Chinese population, which comprised roughly 70% of the sample. As such, generalizations to the wider population should be made with caution. Moreover, these findings may only be relevant to this specific subgroup of Malaysians (i.e., young, highly educated, higher SES individuals) rather than the general population. Further research with a larger and more representative sample could offer greater insights into the factors that lead to successful language planning policy in Malaysia. This conclusion warrants a psychometric investigation into developing a novel language attitude questionnaire for multilinguals. Both the LAQ-LL and the LASS questionnaires are not flexible enough to account for other combinations of language and language cultures.

As the world becomes increasingly globalized, multilingualism is becoming the rule rather than the exception. It is now more crucial than ever to ensure proper language planning. A successful language policy can immensely benefit the economy of a nation. Patterns of issues with language planning in several countries underscore the importance of grounding policymaking in empirical, bottom-up, research. This need warrants further local research to be conducted to understand how a community's unique sociopolitical landscape influences attitudes toward multilingualism and language learning experiences. The findings of this chapter indicate that Malaysians generally hold a positive attitude toward multilingualism and have a positive language learning experience. The results of the subsequent experiments are interpreted within this context.

Chapter 4: Relationship Between Language Context and Linguistic Distance on Executive Functioning in Multilinguals

4.1 Introduction

There have been studies demonstrating the adaptive effects of language contexts on EF (Jiao et al., 2020; Timmer et al., 2021; Wu & Thierry, 2013; Ye et al., 2017), there have also been studies that failed to find significant cognitive differences between the contexts (Kałamała et al., 2021; Paap et al., 2021). However, these studies typically assume group homogeneity when sampling bilinguals. As discussed in Section 1.2, there is a great deal of variation when it comes to bilinguals, meaning, this assumption is misleading, making their results unclear. In the present study, the experiment is a replication of Wu and Thierry (2013) with an important distinction. A holistic description of the sample's language history is provided, as well as the introduction of a language diversity score. This modification enables the study to also investigate how individual differences in multilingual experiences may affect task performance.

Additionally, Coderre and van Heuven (2014) argued that the cognitive outcomes may also be different depending on the similarity between the languages of a multilingual. This hypothesis is predicated on the notion that similar language pairs would require greater cognitive effort to keep the languages apart than dissimilar language pairs. Therefore, in dual-language or dense code-switching contexts, the cognitive outcomes may be modulated by the linguistic distance of the language pairs. Gallo et al. (2023) also suggested that this effect could be influenced by language learning, positing that early second-language learners benefit most from a dissimilar language pair because it results in weaker cross-linguistic transfer (Yang et al., 2017). In other words, cognitive load may be greater for early learners than for advanced language learners.

Conversely, advanced language learners can maximally benefit from studying similar languages, as the increased cognitive control required to separate them further strengthens the underlying neural substrates. Although some studies support this notion (Ramanujan, 2019; Cargnelutti et al., 2022; Radman et al., 2021), others contradict these findings, either reporting an opposite effect (Carthery-Goulart et al., 2023), a mixed effect (Laketa et al., 2021), no effect (Blom et al., 2017), or expressing concerns about methodological and theoretical framework of the studies (Paap et al., 2015).

Pinpointing the impact of linguistic distance on EF remains elusive without a universally accepted definition and method for measuring it. Past studies have used a phylogenetic approach to define Linguistic Distance (LD) as a measure of (dis)similarity between two languages (Ljungberg et al., 2019). For example, Beaufils and Tomin (2021) developed the eLinguistics genetic proximity software (<http://www.elinguistics.net>) based on a stochastic approach to classifying world languages. This tool uses an evolutionary tree model to assess genetic proximity across languages and is based on Swadesh Lists (Serva & Petroni, 2008; Swadesh, 1952). It uses a database of 220 different languages to generate a continuous numerical number. A larger value indicates that the LD between the two languages is greater. However, others have opted to use a feature-based approach (Olguin et al., 2019), matching languages based on similar linguistic features (e.g., phonology, orthography, and typology) within families or genera. Notably, both approaches lack objectivity, as the criteria for classifying similar versus dissimilar languages may differ between researchers based on what features they decide to focus on.

In addition, although a phylogenetic or feature-based approach for categorizing LD might offer greater reliability, these measures do not reflect the psychosocial experiences of language use. As discussed in de Oliveira and Oliveira (2022), the existence of named languages (e.g., English, Malay, Mandarin) as distinct categories only exists because it is shaped by social conventions, not because there is anything supposedly different between languages. Therefore, relying solely on historical relationships (i.e., phylogenetics) to measure linguistic distance may not provide a holistic picture. There needs to be consideration of the social and communicative aspects of languages as well.

The introduction of an indirect scalar measure of linguistic distance based on the learnability of another language relative to the L1 of a sample by Chiswick and Miller (2005) may address this concern. To them, LD is defined as the inverse of the language score obtained in a language proficiency test, such that the larger the LD, the greater the difficulty faced by a person speaking language A to learn dissimilar language C than similar language B. Yet, this measure could be confounded by individual variations in language learning aptitudes and working memory (Singleton, 2017). For example, if the sampled population has above-average WM abilities, the LD measure would be smaller (i.e., easier to learn the foreign language) compared to the actual scores of a sample with normally distributed WM abilities. However, because this scalar measure can potentially present a more ecologically valid assessment of LD compared to the phylogenetic

approach, this experiment adopts Chiswick and Miller's (2005) conceptualization of LD. This stance aligns with the understanding that multilingualism extends beyond units of knowledge of a language to encompass the active utilization of these languages.

4.2 The Current Study

One of the goals of the current study is to investigate whether MLD scores are a better predictor above and beyond the traditional measures of multilingualism, such as language AoA, proficiency, immersion, and dominance. The predictors are entered forward stepwise into a regression model to determine which of the predictors would best predict flanker task performance. The present study was the first to use Li et al.'s (2020) version of language entropy as operationalized in the LHQ. Other studies looking into language entropy used Gullifer and Titone's (2019) conceptualization and computation using the *languageEntropy* package in R (Kałamała et al., 2021; Li et al., 2021; van den Berg et al., 2022; Wagner et al., 2023). Furthermore, as discussed in Section 1.5, Malaysia's sociolinguistic diversity makes it a prime candidate to investigate the bilingual effect, as Malaysians tend to be multilingual out of necessity rather than choice, but, as shown in Chapter 3, Malaysians have a positive view of multilingualism and language learning experiences in general.

There are three experiments in this chapter. Experiment 2A investigates whether priming different language contexts can modulate cognitive control as measured by flanker task performance. Experiment 2B investigates the effects of linguistic distance on the relationship between language contexts and flanker task performance. Experiment 2C investigated the null hypothesis that there would be no effects of language contexts on flanker task performance by priming a familiar-non-familiar language pair. As participants are not familiar with the second primed language, no cognitive adaptation nor improvements in flanker task performance were expected. To date, no research has addressed this hypothesis, and its answers may provide novel insights into the complex multilingualism-cognition relationship.

4.3 Experiment 2A

Experiment 2A adapted the modified flanker task from Wu and Thierry (2013) by varying the language contexts between 1) no language, 2) single language context, and 3) mixed language context. Importantly, the current experiment used a control block with no words presented, this methodology was notably missing in other similarly designed studies (Beatty-Martínez et al.,

2020; Hofweber et al., 2020a; Jiao et al., 2020; Timmer et al., 2021). This block was designed to help identify a priming effect, if present. As no words are presented here, no priming effects were expected in this block. Accordingly, it is hypothesized that the mixed-language block, being more cognitively demanding, would signal the recruitment of greater cognitive resources, allowing for better flanker task performance compared to the single-language and no-language blocks. This hypothesis aligns with the ACH's predictions on the differences between dual and single-language contexts. Furthermore, it is hypothesized that the degree of multilingualism would be able to predict task performance; specifically, participants with higher MLD scores would have smaller flanker effects than those with lower MLD scores.

4.4 Method

4.4.1 Design

The current study used a 2 (Congruency: Congruent x Incongruent) x 4 (Language Context: NoLanguage x English x Mandarin x Mixed) within-subjects design with RT on the flanker task as the dependent variable. Non-verbal IQ, English and Mandarin vocabulary proficiency, and SES were measured as potential covariates using the Adapted Raven's Standard Progressive Matrices (SPM; Raven, 1941), the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012; LexTALE_CH; Chan & Chang, 2018) and the MacArthur Scale of Subjective Social Status (MacArthur SSS Scale; Adler et al., 2000), respectively. The current study also used a forward-stepwise hierarchical multiple regression analysis with second language AoA, proficiency, immersion, dominance, and MLD Score as predictor variables and the flanker effect of all four language blocks as the outcome variables, modelled separately.

4.4.2 Materials

Language proficiency was calculated using the **LHQ** (Li et al., 2020). See Section 2.3 for more information. SES was measured using **MacArthur's SSS Scale**, and non-verbal intelligence was measured using a short version of **Raven's SPM**, see Chapter 2 for further details.

Modified Flanker Task. Participants were instructed to respond to the direction of the target arrow (i.e., central arrow) flanked by two arrows on each side. In congruent trials, the flanker arrows point in the same direction as the central arrow (< < < <); in incongruent trials, the flanker arrows point to the opposite direction of the central arrow (< > < >). Thus, there were four possible

arrangements of flanker stimuli: (1) all left-facing arrows, (2) all right-facing arrows, (3) all left-facing arrows with central arrow facing right, and (4) all right-facing arrows with central arrow facing left.

One control block (i.e. NoLanguage block) and three experimental blocks (i.e. English block, Mandarin block, and Mixed-Language block) constitute this flanker task. NoLanguage and Experimental block sequences are outlined in Figure 5. All blocks contained 63 congruent trials and 27 incongruent trials for a total of 90 trials per block. The congruent to incongruent trials ratio mirrors Costa et al.'s (2009) study, which showed reduced conflict effects in bilinguals under high monitoring (75% or 50% congruent) compared to low monitoring (92% or 8% congruent) conditions. Unlike Wu and Thierry's experiment, the current flanker task version has faster stimulus timings and intervals, intending to increase task difficulty. This modification, combined with high monitoring conditions, maximizes cognitive demands, providing an ideal setting to explore potential bilingual effects.

In the English and Mandarin block, either English or Mandarin single words were displayed before every flanker trial. In the mixed-language block, half of the words were in English, and the other half were in Pinyin. The choice to use single words rather than full sentences or conversations was to attempt to replicate the findings of Wu and Thierry (2013) within the highly integrated and multilingual context of Malaysia so that a baseline could be established for the experiments moving forward. Word order was randomized within each block. These words were task-irrelevant, and participants were instructed to ignore them, aligned with Wu and Thierry's (2013) methodology⁵. This instruction also ensures that the cognitive load of the task is not affected by processing the rapid display of words in sequence. By varying the language context between blocks (i.e., single-language versus mixed-language), it is expected that participants will adaptively enhance EF processes according to the cognitive demands of the block elicited by the change in language context as predicted by the ACH (Green & Abutalebi, 2013).

A total of 270 words in English and Mandarin were selected. To control for the effects of unfamiliarity and emotionality, only high-frequency and neutral affective valence words were selected (see Appendix 2). First, the frequency of words was determined based on the Zipf scale

⁵ Even if participants are instructed to ignore the word stimuli, the immediacy and eye-mind assumptions (Cullipier et al., 2018; Hessels et al., 2019, see Section 2.6) contends that even ignored word stimuli would still be processed at a shallow level. A deeper discussion on the difference between shallow and deep processing of word stimuli and its effect on task performance will be presented in Section 7.1.1

of 1 - 7 (van Heuven et al., 2014), with higher values indicating higher frequencies. Second, neutral affective valence words were determined based on the Valence-Arousal (VA) dimension proposed by Russell (1980), which uses a 9-degree scale, with higher values denoting positive and excited words. Following the neutral-word semantic categorization of Ferré et al. (2011), neutral affective-valence words would fall between 4 and 6 on the VA scale. Therefore, word stimuli were selected based on two criteria: (1) above 4.5 on the Zipf scale, and (2) VA ratings between 4 and 6. Importantly, any words that might cause semantic interference with the flanker task (e.g., left, right, middle) were avoided (Wu & Thierry, 2013).

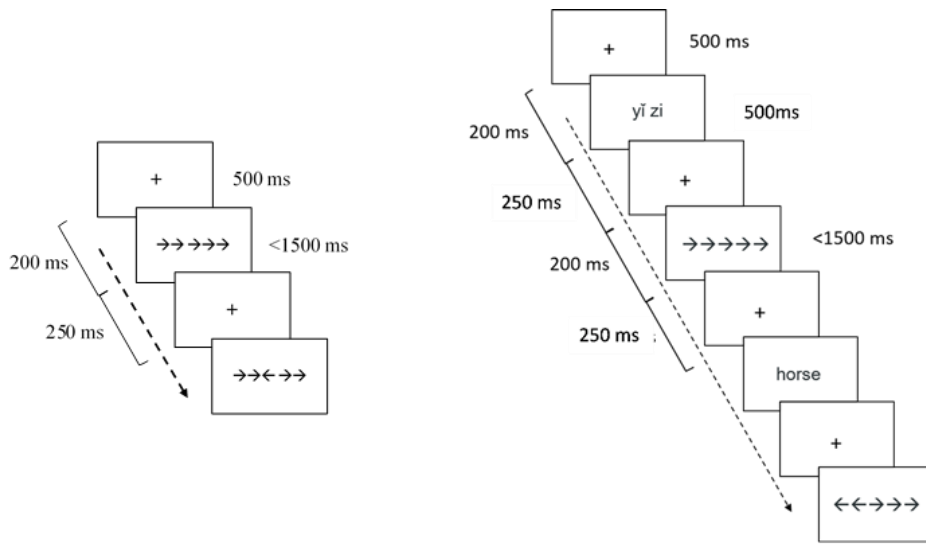
A total of 135 English words were sourced from Warriner et al.'s (2013) research which details 13,915 high-frequency English words with their corresponding VA data. These 135 words were cross-referenced with SUBTLEX-UK, a large corpus that identifies the frequency of words from film and television subtitles (van Heuven et al., 2014), to obtain their respective Zipf values.

Next, 135 Mandarin words, consisting of two characters, were selected from the Chinese Valence-Arousal Words 4.0 (CVAW 4.0; Yu et al., 2016) and 1,100 Chinese Words (Yao et al., 2017) corpora and Zipf data was obtained from SUBTLEX-CH (Cai & Brysbaert, 2010). To control for the possible confounding effects of different typography (Coderre & van Heuven, 2014; Linck et al., 2009) all Mandarin characters selected were transformed into Pinyin, a Romanized form of Mandarin, which uses Latin script similar to English⁶.

A pilot study was conducted before the current study to ensure the accuracy of the Mandarin to Pinyin translation. First, all selected Mandarin characters were translated into their Pinyin forms using an online translator (www.chineseconverter.com). Two highly proficient L1 Mandarin speakers reviewed the translations separately, if there were any incorrect translations, reviewers were asked to indicate the appropriate translation. If both reviewers independently agreed that a translation was incorrect, the suggested translation was used instead.

Figure 5: *Sequence of NoLanguage and Experimental Blocks*

⁶ Pinyin was also used so that when comparing with Experiment 2, the results cannot be ascribed to differences in typography



(A) NoLanguage

(B) Experimental Blocks

Note: One complete trial consists of a fixation point (denoted by “+”) and a word or flanker stimuli. Fixation points were presented for 500 ms followed by a 200-ms blank screen. Word stimuli were also presented for 500 ms. On flanker trials, participants had up to 1500 ms to respond to the target arrow. Intertrial interval (ITI) was fixed at 250 ms.

4.4.3 Participants

To estimate an appropriate sample size, an *a priori* power analysis was conducted using G*Power3 (Faul et al., 2007) to find a medium effect size of $f^2 = .15$. with α err prob = 0.05, Power (1- β) = 0.80, and number of predictors = 3. The output suggested that at least 77 participants were required in this study to detect an effect with 80% probability. Studies investigating similar variables included 60 participants or fewer (Adler et al., 2020; Hofweber et al., 2020a; Wu & Thierry, 2013). Furthermore, as reported by Paap et al. (2014), small-scale studies (average $N = 29$) usually report a bilingual effect, but larger-scale studies (average $N = 45$) often report null findings.

The current study recruited 85 participants based on the following criteria: (1) aged between 18-25 years old, (2) proficient and literate in both English and Mandarin, (3) local Malaysian-Chinese, and (4) no prior or current Mental Health diagnoses (see Table 5 for Participant Demographics).

The data preprocessing involves three stages, first, participants' results were screened for outliers. Three participants were removed at this stage due to having average RTs above 700ms across experimental blocks. Second, on a trial-by-trial basis, any trials clocking below 200 ms (i.e., false starts) or ± 2.5 SD above its experimental block average RT were removed from the analysis. Five participants were excluded at this stage due to the removal of more than 50% of block trials on one or more experimental blocks. Lastly, each participant's data integrity was calculated using the given formula ($\frac{\text{Number of Trials Retained}}{\text{Total Possible Number of Trials}}$), indexing the overall remaining useable trials after data preprocessing, ensuring that sufficient data remained for analysis. The cutoff score was arbitrarily set at 80% (i.e., at least 288 trials retained at the third stage of data cleaning). One participant whose data integrity fell below this cutoff score was excluded from the analysis. A total of 76 participants (65 females) were included in the final analysis.

Table 5: *Experiment 2A Language Background Statistics*

Demographic Measures	Mean	SD
Age (years)	21.71	2.03
English AoA (years)	3.66	1.61
English Proficiency (0-1)	0.79	0.14
English Immersion (0-1)	0.83	0.10
English Dominance (0-1)	0.59	0.17
Mandarin AoA (years)	3.43	1.74
Mandarin Proficiency (0-1)	0.84	0.12
Mandarin Immersion (0-1)	0.83	0.10
Mandarin Dominance (0-1)	0.57	0.13
L1 AoA (years)	2.80	1.20
L2 AoA (years)	4.30	1.80
L1 Proficiency (0-1)	0.86	0.12
L2 Proficiency (0-1)	0.77	0.12
L1 Immersion (0-1)	0.85	0.10
L2 Immersion (0-1)	0.82	0.09
L1 Dominance (0-1)	0.63	0.16

L2 Dominance (0-1)	0.53	0.13
MLD Score (0-2)	1.68	0.26
SES (0-10)	5.62	1.19
Adapted Raven's SPM (0-30)	25.19	2.71
LexTALE (100 %)	83.55	10.26
LexTALE-CH (100 %)	85.41	5.02

Note: N= 76; L1 and L2 languages differ across participants: L1 (English = 42; Mandarin =35), L2 (English = 35, Mandarin = 42). Between English and Mandarin languages, paired-sample t-tests revealed that only English and Mandarin Proficiency were significantly different from one another. No significant results were found for AoA, Immersion, and Dominance. For ease of comparability, LexTALE and LexTALE-CH scores were standardized to 100%.

Notably, 93% of participants ($N = 73$) reported knowing an L3 other than English and Mandarin. To investigate the effects of multilingualism on cognition, the following five variables will be used as indicators of multilingualism: (1) L2 AoA – the age when the second language was acquired, (2) L2 Proficiency – competency in the second language, (3) L2 Immersion – the extent that the surrounding sociolinguistic contexts use the second language, (4) L2 Dominance – the extent that the individual uses the second language and (5) MLD scores – the likelihood that one language is used over other languages based on Shannon's Entropy. Note that the effects of L1, L3, and L4 variables were not directly investigated, as L2 has been typically used as an indicator for bilingualism in past research. However, to account for the knowledge of multiple languages, the present study used a novel measure of multilingualism, the MLD scores, which provides a more accurate picture of participants' language profiles.

4.4.4 Procedure

Participants completed the MacArthur SSS Scale, Adapted Raven's SPM, and LexTALE (EN-CH) via Qualtrics (www.qualtrics.com). Once completed, a one-to-one video call was scheduled, where participants completed the LHQ 3.0 during the video call. As the LHQ is based on self-reports, participants would need to understand the question to be able to accurately complete the questionnaire (Li et al., 2020). Due to the ongoing COVID-19 pandemic during the

experiment, in-person supervision was not feasible; hence, a video call was used to allow participants to clarify any questions they might have.

Once completed, participants proceeded to complete the Modified Flanker Task via Pavlovia (www.pavlovia.org), in which they were asked to respond, with their index fingers, to the direction of the central arrow using the “F” key if the central arrow pointed left, or the “J” key if the central arrow pointed right. If no response was made within 1500 ms, the trial was coded as “incorrect”. Participants were reminded to ignore the intervening word stimuli (i.e., make no response to any of them). Before each block, participants were informed of what language the intervening words would be presented in. Participants were also provided with a practice block (N trials = 6) with feedback before the Experimental blocks to ensure that they understood the requirements of the task. The NoLanguage block was always presented first to get participants used to the task, whereas the presentation of the English, Mandarin, and Mixed-language blocks was fully counterbalanced between participants. An untimed break was provided between blocks.

4.5 Results

To investigate the interaction effects of Congruency (Congruent x Incongruent) and Language blocks (NoLanguage x English x Mandarin x Mixed), a two-way repeated-measures ANOVA was conducted (see Table 6 for mean estimates). Additionally, a BF analysis was conducted to provide complementary statistics to the traditional frequentist statistics.

Before analysis, a one-sampled t -test was conducted to check whether non-verbal IQ (Raven’s SPM) and SES were significantly different from the sample’s mean. Results indicate that neither SPM ($t(75) = -0.01, p = .996$) nor SES ($t(75) = -0.01, p = .991$) significantly differed from the sample’s mean (see Table 6). As such, neither of these variables were included as covariates in the analysis. Although error rate analyses were conducted, their outcomes did not yield any meaningful findings, likely because performance is near the ceiling (see Appendix 3). Therefore, the results section will focus on RT data.

Table 6: *Experiment 2A – Estimated Marginal Means by Congruency, by Language, and by Congruency * Language (with Standard Error)*

95% CI for Mean Difference

		Error Rates	Lower	Upper
By Congruency	Mean RT (ms) (%)			
Congruent	415.30 (5.29)	0.86	404.09	426.50
Incongruent	482.14 (5.96)	7.93	470.94	493.35
By Language				
NoLanguage	441.28 (5.60)	4.60	429.25	453.31
English	452.97 (5.81)	4.25	440.94	465.00
Mandarin	453.57 (5.80)	3.86	441.54	465.60
Mixed	447.05 (6.99)	4.86	435.02	459.08
By Congruency * Language				
Congruent				
NoLanguage	411.35 (5.79)	0.80	398.931	423.78
English	416.70 (5.48)	0.51	404.273	429.12
Mandarin	420.13 (5.67)	0.66	407.715	432.56
Mixed	413.00 (6.49)	1.48	400.581	425.43
Incongruent				
NoLanguage	471.21 (5.70)	8.40	458.79	483.64
English	489.25 (6.56)	8.00	476.82	501.67
Mandarin	487.01 (6.33)	7.06	474.59	499.43
Mixed	481.10 (7.87)	8.25	468.68	493.52

Note. $N = 76$.

As expected, there was a classical Flanker Congruency effect ($F(1,75) = 792.11, p < .001, \eta_p^2 = .91, BF_{10} = 2.88 \times 10^{38}$), with reaction times on Congruent trials ($M = 415.30, SE = 5.63$) being significantly faster than Incongruent trials ($M = 482.14, SE = 5.63$). The sphericity assumption was violated for the main effect of language blocks. Given that Greenhouse-Geisser epsilon values were above 0.75, the Huynh-Feldt correction was applied. A significant main effect of Language blocks was found ($F(2.42, 181.24) = 3.82, p = .017, \eta_p^2 = .05, BF_{10} = 3.47$). There was also a significant two-way interaction effect between Congruency and Language ($F(3, 228) = 4.53, p = .004, \eta_p^2 = .06, BF_{10} = 15.17$).

To investigate the significant main effect of Language blocks and the significant interaction effect between Congruency and Language blocks, a Bonferroni post-hoc correction test was conducted to adjust for multiple comparisons. Post-hoc analysis on the main effect of Language revealed that the NoLanguage block had significantly lower RTs compared to the English ($M_{\text{diff}} = -11.69, p = .033, \text{BF}_{10} = 54.53$) and Mandarin blocks ($M_{\text{diff}} = -12.29, p = .021, \text{BF}_{10} = 28.35$), but it was not significantly different than Mixed block ($M_{\text{diff}} = -5.77, p = 1.000, \text{BF}_{10} = 0.10$). There were no significant differences between English and Mandarin blocks ($M_{\text{diff}} = -0.60, p = 1.000, \text{BF}_{10} = 0.04$). Mixed language block was not significantly different than either of the single-language blocks (English: $M_{\text{diff}} = 5.92, p = .943, \text{BF}_{10} = 0.25$; Mandarin: $M_{\text{diff}} = 6.52, p = .715, \text{BF}_{10} = 0.26$).

A simple effects analysis on the interaction effect revealed that within the Congruent condition, there were no significant differences in RTs between all language blocks ($ps > .05$). Within the Incongruent condition, the NoLanguage block had significantly lower RTs than the English ($M_{\text{diff}} = -18.03, p = .002$) and Mandarin blocks ($M_{\text{diff}} = -15.80, p = .015$), but was not significantly different than Mixed block ($M_{\text{diff}} = -9.89, p = .826$). There were no significant differences between English and Mandarin blocks ($M_{\text{diff}} = 2.24, p = 1.000$). Mixed language block was not significantly different from either of the single-language blocks (English: $M_{\text{diff}} = 8.14, p = .943$; Mandarin: $M_{\text{diff}} = 5.91, p = 1.000$). The estimated marginal means can be found in Table 6.

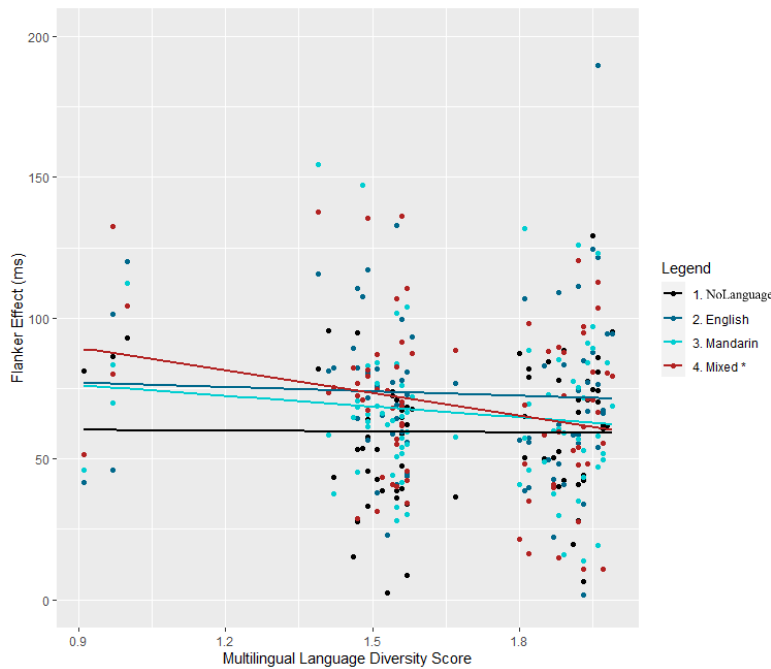
4.5.1 Forward-stepwise Multiple Regression

To investigate how multilingualism affects flanker task performance, a forward-stepwise hierarchical regression was conducted to identify which of the participants' multilingual experiences would emerge as a significant predictor of the task performance. For the current analysis, variables are only entered if they significantly predict the outcome at $p < .05$. As this regression method is criterion-driven rather than theory-driven, it cannot explain any identified relationships. Instead, it can only assist in identifying the best predictors for task performance (Keith, 2019). In the context of this study, which focuses on investigating the predictive capabilities of the MLD score compared to other traditional multilingual experience variables, the absence of explanatory power is less problematic. Data was visualized using the *ggplot2* package (Wickham, 2016) in R (R Core Team, 2021).

Results only converged for the NoLanguage and mixed block flanker effect. In the NoLanguage block, only L2 immersion significantly predicted the flanker effect, explaining 5.7% of the variance in the outcome ($F(1, 74) = 4.43, p = .039, R^2_{\text{adj}} = .04$). As a predictor, a unit increase in L2 immersion results in a 59.43-ms increase in the NoLanguage-block flanker effect ($\beta = -59.43, SE = 28.23, t = 2.11, p = .039$). In the mixed block, only MLD scores significantly predicted the flanker effect, explaining 7.5% of the variance in the outcome ($F(1, 74) = 6.02, p = .016, \text{Adjusted } R^2_{\text{adj}} = .06$). As a predictor, a unit increase in MLD results in a 32.5-ms decrease in the mixed block flanker effect ($\beta = -32.50, SE = 13.25, t = -2.45, p = .016$). Given only a single predictor was entered into both models, multicollinearity was not a concern in these models.

When MLD scores were plotted against flanker effects across language blocks (see Figure 6), the relationship between MLD and flanker effects appeared to be modulated by linguistic context, with higher MLD scores significantly predicting smaller flanker effects in the Mixed block. However, MLD scores did not significantly predict the flanker effect in the NoLanguage, English, and Mandarin blocks.

Figure 6: *Experiment 2A - Scatterplot of MLD Scores Against Flanker Effect Across Language Blocks*



Note: * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

4.6 Discussion

Experiment 2A aimed to investigate the effects of different linguistic contexts on inhibitory control in multilinguals. First, it was hypothesized that performance in the mixed-language context would be better compared to single-language contexts. Second, it was hypothesized that participants with higher MLD scores would perform better on the flanker task than those with lower MLD scores.

Contrary to the first hypothesis, no significant mean RT differences were found between the language blocks in either the congruent or incongruent trials. This absence could imply that there are indeed no effects of the rapid modulation of language contexts on flanker task performance. Yet, it is plausible that the non-significant result stems from Mandarin's dissimilarity to English. As per Gallo et al. (2023), such dissimilarity might only impact cognitive outcomes during the early stages of second language learning. It is suggested that a more similar language pair would be required to affect cognitive outcomes for highly proficient bilingual participants.

However, other studies using the same English-Mandarin language combination on highly proficient bilinguals, albeit with different EF tasks, found a significant difference between mixed-language and single-language contexts in performance (Han et al., 2022; Yang et al., 2023), which does not support Gallo's et al. (2023) interpretation. A possible explanation for these contradictory findings may be related to the use of Pinyin rather than Mandarin characters. In a meaning-judgement task by Chen et al. (2017), advanced learners of Chinese took significantly longer than intermediate learners to determine the meaning of Pinyin words compared to Mandarin characters. The authors interpreted this difference as intermediate learners needing to derive meaning directly from Pinyin, whereas advanced learners needed to activate the character representation of the Pinyin word before accessing its meaning, thus leading to increased RTs. Therefore, given the high Mandarin proficiency in the present sample, it is speculated that their lexical access is slower. It is possible that the cognitive adaptations resulting from language context priming could have concealed the negative effects of slower lexical access on flanker task performance. However, as RTs did not show significant differences between language blocks, this interpretation must be approached with caution.

Although there were statistically no RT differences between the incongruent trials of the language blocks, single-language blocks had significantly slower mean RT than NoLanguage blocks. The mean RT between the NoLanguage and the mixed language block was not

significantly different. Bayesian analysis strongly supports the difference between NoLanguage and single language blocks, but the evidence for the difference between NoLanguage and mixed language blocks is anecdotal. This finding aligns with the ACH, which predicts that a mixed-language context would elicit greater cognitive adaptation than single-language contexts. However, the non-significance between the language blocks prevents any further conclusions beyond speculation.

Next, consistent with our second hypothesis, we found that higher MLD scores significantly predicted smaller flanker effects (i.e., better flanker task performance) compared to the other multilingualism experience variables. Importantly, this relationship is moderated by language contexts, as the relationship was only significant in the mixed-language context but not in the single-language or NoLanguage contexts. This finding is echoed by van den Berg et al. (2022), who found a more efficient performance on the color-shape switching task (CSST), where those with greater language entropy had smaller mixing RT costs than those with smaller language entropy. Similarly, Gullifer and Titone (2021) found that language entropy was related to the engagement of proactive control – a form of cognitive control that is much more resource efficient than reactive control, on the AX-Continuous Performance Task (CPT). However, they noted that this language entropy-cognitive control relationship is likely small and that an equally predictive model can also be achieved by removing language entropy.

4.7 Experiment 2B

Based on the findings from Experiment 2A, this experiment considered the possibility that linguistic distance may differentially affect the outcomes by replacing the Mandarin words with Malay words while retaining the same design and procedures. As discussed in earlier sections, the conceptualization of LD follows that of Chiswick and Miller (2005), defined as a function of the inverse test scores of English speakers learning other languages. Based on this metric, the LD for the language pairs in Experiment 2A (English-Mandarin) is more dissimilar ($LD = 0.67$) than in Experiment 2B (English-Malay; $LD = 0.36$).

Therefore, due to the smaller linguistic distance between English and Malay, more inhibitory control should be engaged to handle cross-linguistic interference. Similar to Experiment 2A, the mixed block is anticipated to exhibit significantly better flanker task performance

compared to the single language blocks. Also, higher MLD scores should predict better flanker task performance compared to the other multilingual experience variables.

4.8 Method

4.8.1 Design

Experiment 2B used the same 2x4 within-subjects design as with Experiment 2A. However, language contexts were importantly modified to include Malay interleaved words instead of Mandarin.

4.8.2 Materials

In addition to the materials listed in Section 4.4.3, Experiment 2B uses the following materials in replacement of the LexTALE-CH and Mandarin word stimuli. The English word stimuli used were different than Experiment 2A (see Appendix 2, English Word Priming List Set Experiment 2B)

Malay Cloze Test. The cloze test used here was used as a proxy for Malay vocabulary proficiency (see Appendix 1). To cover a wide range of difficulty levels, a total of 20 Malay cloze questions were selected from past national examination papers at different education levels. Participants' performance on the cloze test is scored between 0 to 20, with higher scores indicating better Malay vocabulary proficiency. To facilitate comparisons with LexTALE scores, its raw scores are converted into percentages. Lee et al. (2023) suggests that results on this cloze test were able to distinguish proficient Malay speakers from non-proficient Malay speakers, with proficient speakers significantly scoring higher than non-proficient speakers.

Malay Word Stimuli. There is a lack of Malay word corpora that provides valence and arousal data. Therefore, this study obtained VA values for Malay words by proxy via translation to English and sourced the VA values from English word corpora (see Appendix 2). First, word frequency (words per million) data was obtained from the Malay Lexicon Project (Yap et al., 2010), which was subsequently converted into the Zipf standard using logarithmic functions (van Heuven et al., 2014). Second, a set of Malay-English forward-translation data (Lee et al., 2022) was used to obtain the dominant English word translations of the data in the Malay Lexicon Project. Third, VA data was cross-referenced from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999) and the 13915-word corpora (Warriner et al., 2013). Finally, the

words were selected based on the following criteria: (1) word length between 3 and 8 characters, (2) above 3 Zipf values, and (3) VA values between 3-7. This change in filtering criteria, compared to Experiment 2A, was implemented because there would be insufficient Malay words to be used if the threshold criteria were to be maintained at above 4.5 Zipf and between 4-6 VA values. Due to this change, English word stimuli were also reevaluated and replaced using these new criteria.

4.8.3 *Participants*

Eighty participants were recruited based on the following criteria: 1) aged between 18-25 years, 2) proficient in English and Malay, with Malay as the first or second language, 3) local Malaysians, and 4) no prior or current mental health diagnoses. These participants were a different set of participants from Experiment 2A.

Applying the same data preprocessing steps in Experiment 2A, a total of eight participants were removed: two participants were removed for missing demographic data, one participant was removed due to having a 100% error rate in the NoLanguage block, and five participants were removed for having suboptimal data integrity below 80%, thereby leaving 72 participants in the final data analysis. Notably, 81% of participants reported an L3 other than English and Malay. Further descriptions of participants can be found in Table 7.

Table 7: *Experiment 2B - Language Background Statistics*

Demographic Measures	Mean	SD
Age (years)	21.06	2.05
EN AoA (years)	3.44	1.68
EN Proficiency (0-1)	0.89	0.10
EN Immersion (0-1)	0.79	0.12
EN Dominance (0-1)	0.77	0.23
BM AoA (years)	4.80	2.02
BM Proficiency (0-1)	0.76	0.12
BM Immersion (0-1)	0.77	0.11
BM Dominance (0-1)	0.47	0.15
L1 AoA (years)	3.02	1.49
L2 AoA (years)	4.34	1.96

L1 Proficiency (0-1)	0.84	0.17
L2 Proficiency (0-1)	0.75	0.17
L1 Immersion (0-1)	0.82	0.14
L2 Immersion (0-1)	0.77	0.14
L1 Dominance (0-1)	0.68	0.22
L2 Dominance (0-1)	0.54	0.25
MLD Score (0-2)	1.50	0.30
SES (0-10)	5.86	1.26
Adapted Raven's SPM (0-30)	24.86	3.28
LexTALE (100 %)	87.59	8.82
BM Cloze Test (100 %)	55.70	4.20

Note: N= 72; L1 and L2 languages differ across participants: L1 (English = 47; Malay = 7; Mandarin = 14; Tamil = 4), L2 (English = 22; Malay = 36; Mandarin = 12; Japanese = 1; Burmese = 1). Between English and Malay languages, paired-sample t-tests revealed that AoA, Proficiency, Immersion, and Dominance were significantly different from one another. For ease of comparability, LexTALE and the BM Cloze Test scores were standardized to 100%.

4.9 Results

Before analysis, a one-sample t-test was conducted to check whether non-verbal IQ (Raven's SPM) and SES were significantly different from the sample's mean. Results indicate that neither SPM ($t(71) = 0.00, p = .998$) nor SES ($t(71) = 0.01, p = .994$) significantly differed from the sample's mean. Neither variable was included as a covariate in the analysis, and as the error analysis was not significant, the results will focus solely on RT data (see Appendix 3) for error analysis).

The expected classical Flanker Congruency effect was found ($F(1,71) = 1004.18, p < .001, \eta_p^2 = .93, BF_{10} = 1.28 \times 10^{14}$), with significantly faster reaction times in Congruent ($M = 417.87, SE = 5.01$) than Incongruent trials ($M = 489.57, SE = 5.01$). The sphericity assumption was violated for the main effect of language blocks. Given that Greenhouse-Geisser epsilon values were above 0.75, the Huynh-Feldt correction was applied. Results indicate a significant main effect of Language ($F(2.69, 191.06) = 3.50, p = .020, \eta_p^2 = .05, BF_{10} = 2.54$). There was also a significant

two-way interaction effect between Congruency and Language ($F(3, 213) = 8.28, p < .001, \eta_p^2 = .10, BF_{10} = 1495.48$).

A Bonferroni-corrected, post-hoc analysis on the main effect of Language revealed that NoLanguage block had significantly lower RTs compared to English ($M_{\text{diff}} = -12.60, p = .031, BF_{10} = 6.39$) but marginally significant compared to Malay blocks ($M_{\text{diff}} = -11.58, p = .060, BF_{10} = 2.24$), but not significantly different than Mixed block ($M_{\text{diff}} = -5.13, p = 1.000, BF_{10} = 0.10$). There were no significant differences between English and Malay blocks ($M_{\text{diff}} = 1.12, p = 1.000, BF_{10} = 0.04$). Mixed language block was not significantly different than either of the single-language blocks (English: $M_{\text{diff}} = 7.47, p = .574, BF_{10} = 1.00$; Malay: $M_{\text{diff}} = 6.45, p = .899, BF_{10} = 0.33$).

On the interaction effect, a simple main effects analysis revealed no significant differences in RTs between all language blocks ($ps > .05$) within the Congruent condition. Within the Incongruent condition, the NoLanguage block had significantly lower RTs than English ($M_{\text{diff}} = -19.21, p = .009$) and Malay blocks ($M_{\text{diff}} = -20.50, p = .005$). Mixed block's RT was not significantly different than NoLanguage ($M_{\text{diff}} = 11.05, p = .226$), English ($M_{\text{diff}} = -8.17, p = .371$), and Malay blocks ($M_{\text{diff}} = -9.45, p = .268$). Further details on the estimated mean RTs can be found in Table 8.

Table 8: *Experiment 2B – Estimated Marginal Means by Congruency, by Language, and by Congruency * Language (with Standard Error)*

			95% CI for Mean Difference	
By Congruency	Mean RT (ms)	Error Rate (%)	Lower	Upper
Congruent	417.87 (4.67)	1.19	408.55	427.18
Incongruent	489.57 (5.32)	10.85	478.96	500.18
By Language				
NoLanguage	446.39 (5.08)	5.47	436.26	456.52
English	458.99 (6.03)	6.04	446.96	471.02
Malay	457.97 (5.74)	6.17	446.53	469.41
Mixed	451.52 (5.47)	6.40	440.63	462.42
By Congruency * Language				
Congruent				

NoLanguage	415.90 (5.17)	1.21	405.60	426.20
English	421.89 (5.78)	0.82	410.37	433.41
Malay	418.56 (5.37)	1.40	407.85	429.28
Mixed	415.12 (5.19)	1.32	404.78	425.46
Incongruent				
NoLanguage	476.88 (5.45)	9.72	466.01	487.75
English	496.09 (6.74)	11.27	482.66	509.53
Malay	497.38 (6.53)	10.96	484.36	510.40
Mixed	487.93 (6.09)	11.47	475.79	500.07

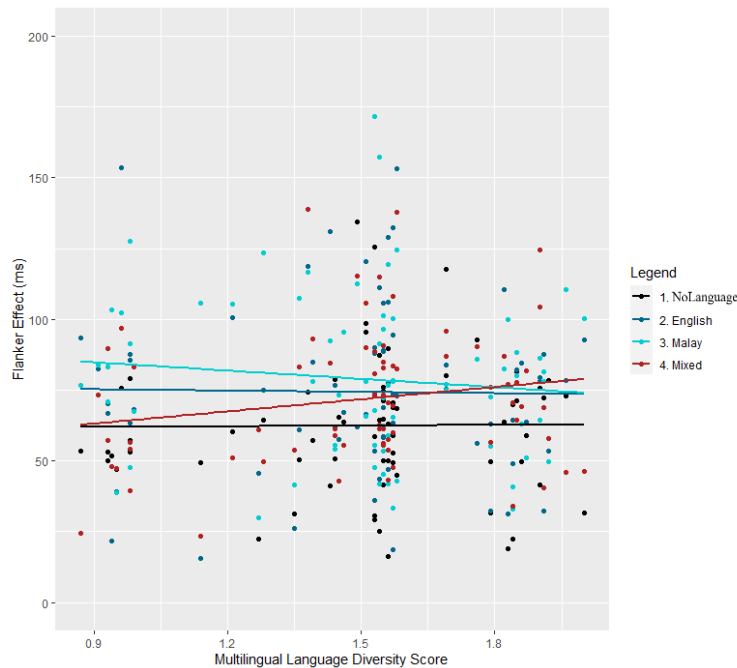
Note. $N = 72$.

4.9.1 Forward-stepwise Multiple Regression

The regression was modelled similarly to Experiment 2A. To reiterate, a forward-stepwise hierarchical regression was conducted using L2 AoA, proficiency, immersion, dominance, and MLD scores on the flanker effect of each language block, modelled separately. Results indicate that only models for the Malay and Mixed block flanker effect converged successfully. The model for the Malay block flanker effect only includes L2 proficiency, which explained 5.9% of the variance in outcome ($F(1, 70) = 4.36, p = .031, R^2_{\text{adj}} = .05$). As a predictor, a unit increase in L2 Proficiency results in a 41-ms decrease in the Malay block flanker effect ($\beta = -41.08, SE = 19.67, t = -2.09, p = .040$). Furthermore, the model for the Mixed block flanker effect only includes L2 AoA, explaining 5.9% of the variance in the outcome ($F(1, 70) = 4.41, p = .039, R^2_{\text{adj}} = .05$). As a predictor, a unit increase in L2 AoA results in a 3-ms decrease in the Mixed block flanker effect ($\beta = -3.00, SE = 1.43, t = -2.1, p = .039$). Interestingly, no models converged with MLD scores as a predictor in any language block flanker effect ($ps > .05$). Given only a single predictor was entered into both models, multicollinearity was not a concern in these models.

Plotting MLD scores against each block's flanker effect (see Figure 7), there appears to be a non-linear relationship between MLD scores and the flanker effect across language blocks, although this relationship is non-significant.

Figure 7: Experiment 2B - Scatterplot of MLD Scores Against Flanker Effect Across Language Blocks



4.10 Discussion

It was hypothesized that, as compared to English and Mandarin (Experiment 2A), the closer LD between English and Malay would require greater inhibitory control to maintain separation between the languages. This requirement, in turn, would predict improved performance on the flanker task in the mixed-language block compared to the single-language blocks. However, the results showed no significant mean RT differences between the language blocks. Bayesian analysis supports this finding, as between the language blocks, there was no evidence for the alternative hypothesis.

This finding does not support the first hypothesis, suggesting that the role of LD in the multilingual effect is limited. The results are aligned with recent research by Wolfsturn et al. (2023), supporting the null hypothesis. They found no significant differences in Stroop task performance between participants with highly similar languages (Italian-Spanish) and those with highly dissimilar languages (Dutch-Spanish). Blom et al. (2017) also support this claim. They argued that although there is evidence supporting the idea that similar languages lead to larger bilingual effects, there is also contradictory evidence that found that distant languages also lead to larger bilingual effects (Carthery-Goulart et al., 2023). These contradictory findings suggest that

factors (e.g., proficiency in the primed languages, language switching practices) other than LD could be the deciding factor in eliciting the bilingual effect.

Regarding the second hypothesis, it was predicted that higher MLD scores would predict smaller flanker effects. However, inconsistent with this hypothesis, none of the multilingual variables, including MLD scores, were able to predict any language blocks' flanker effects. This finding aligns with Kałamała et al.'s (2021) study in which language entropy was unrelated to the performance of four response-inhibition tasks. Together, the findings suggest that language entropy is not linked to improvements in inhibitory control, adding to the ongoing discussion about the multilingual effect.

4.11 Experiment 2C

This experiment aimed to determine if the adaptation of cognitive control, as outlined in the ACH, could be linked to the presence of a second language irrespective of proficiency or familiarity. To assess this hypothesis, the same methodology as in previous experiments was employed, but this time, participants were exposed to English and Welsh. The Welsh language was chosen for two reasons: (1) it is extremely unlikely that the current sample of Malaysian undergraduate students would have any exposure or proficiency in the Welsh language and (2) Welsh uses the Latin script in their written language, which is similar to the language pairs used in Experiment 2A (English-Pinyin) and 2B (English-Malay), thereby ruling out the possibility that any effects found here would be a result of different typography.

No prior research has explored the effects of a familiar-unfamiliar language pairing on the cognitive outcomes outlined in the ACH. To reiterate, the ACH predicts that single-language contexts should demonstrate weaker cognitive adaptations than mixed-language contexts. As participants are only familiar with English in this experiment, participants should theoretically only be primed in either no context (NoLanguage or Welsh-only) or single-language contexts (English-only or English-Welsh). It is expected that Welsh-only would be slower than NoLanguage and English-Welsh would be slower than English-only. These expectations are made because if priming an unfamiliar language only increases cognitive load without any associated enhancement in cognitive abilities, adverse effects, rather than a facilitatory effect, should be observed. This prediction is in line with Mueller's (2009) findings that reading words in an unfamiliar language led to increased processing costs. Therefore, it is hypothesized that the

NoLanguage block would have the fastest RT, followed by Welsh-only, English-only, and finally English-Welsh. Additionally, Experiment 2A found that MLD only significantly predicted better flanker task performance in the mixed-language block, and because the mixed-language block in this experiment is theoretically comparable to a single-language block, the second hypothesis predicts that MLD scores will not predict better flanker task performance.

4.12 Method

4.12.1 Design

Experiment 2C used the same 2x4 within-subjects design as Experiments 2A and 2B. However, all interleaved words were replaced with English and/or Welsh words.

4.12.2 Materials

In addition to the materials listed in Section 4.4.3, Experiment 2C uses the following materials in replacement of the Mandarin word stimuli. A different set of English priming stimuli (see Appendix 2, English Word Priming List Set Experiment 2C) was used with the same criteria for valence, arousal, and word length.

Welsh Word Stimuli. 135 high-frequency noun Welsh words (see Appendix 2) were taken from the *Corpws Cenedlaethol Cymraeg Cyfoes* (CorCenCC) corpus (Knight et al., 2020). Here, unlike the past two experiments, frequency (zipf), valence, and arousal values were not important factors to control for, as the present sample did not report any familiarity nor proficiency in Welsh in their language history data. Therefore, regardless of the frequency of the words, or where the word falls on the valence-arousal space, the Welsh stimuli are unlikely to meaningfully affect the present sample as participants would not understand the words. However, the length of each Welsh stimulus was limited to 3 to 8 characters to ensure consistency with the English stimuli. Further, interlingual homographs like “bore” and “pen” in the Welsh corpus, meaning “morning” and “head” in English respectively, were excluded.

4.12.3 Participants

In this experiment, 82 participants were based on the following criteria: (1) between 18-25 years old, (2) must be proficient in English, (3) must be a local Malaysian, and (4) no prior or

current mental health diagnoses. These participants were a different set of participants than Experiment 2A and 2B. No participants reported knowing Welsh as their L2, L3, or L4.

Applying the same data preprocessing steps as in Experiment 2A, 6 participants' data were excluded for having suboptimal data integrity below 80%, leaving 75 (54 female) participants in the final data analysis. Notable, 94% of participants reported knowing an L3 other than English. Further descriptions of the participants can be found in Table 9.

Table 9: *Experiment 2C Participant Demographics*

Demographic Measures	Mean	SD
Age (years)	21.31	1.63
English AoA (years)	3.56	1.72
English Proficiency (0-1)	0.85	0.14
English Immersion (0-1)	0.84	0.11
English Dominance (0-1)	0.71	0.29
L1 AoA (years)	2.77	1.26
L2 AoA (years)	4.13	2.05
L1 Proficiency (0-1)	0.86	0.18
L2 Proficiency (0-1)	0.75	0.21
L1 Immersion (0-1)	0.87	0.14
L2 Immersion (0-1)	0.81	0.13
L1 Dominance (0-1)	0.68	0.30
L2 Dominance (0-1)	0.55	0.18
MLD Score (0-2)	1.60	0.23
SES (0-10)	6.01	1.26
Adapted Raven's SPM (0-30)	24.53	4.17
LexTALE (100 %)	84.23	10.22

Note: N= 75; L1 and L2 languages differ across participants: L1(English = 30; Mandarin = 36; Cantonese = 5; Malay = 2; Tamil = 2), L2 (English = 42; Malay = 12; Mandarin = 17; Tamil = 1; Telugu = 1; Teo Chew = 1; Vietnamese = 1). For ease of comparability to Experiments 2A and 2B, LexTALE scores were standardized to 100%.

4.13 Results

Before analysis, a one-sample t-test was conducted to check whether non-verbal IQ (Raven's SPM) and SES were significantly different from the sample's mean. Results indicate that neither SPM ($t(75) = 0.01, p = .994$) nor SES ($t(75) = 0.02, p = .982$) significantly differed from the sample's mean (see Table 9). Neither variable was included as a covariate in the analysis, and as the error analysis was not significant, the results will focus solely on RT data (see Appendix 3) for error analysis).

There was the expected classical Flanker Congruency effect ($F(1, 74) = 615.78, p < .001, \eta_p^2 = .89$), with significantly faster Congruent ($M = 420.59, SE = 6.52$) than Incongruent trials ($M = 489.46, SE = 6.52$). The sphericity assumption was violated for the main effect of language blocks. Given that Greenhouse-Geisser epsilon values were above 0.75, the Huynh-Feldt correction was applied. Results indicate a significant main effect of Language ($F(2.69, 191.06) = 3.50, p = .020, \eta_p^2 = .05, BF_{10} = 102.58$) and a significant two-way interaction effect between Congruency and Language ($F(3, 213) = 8.28, p < .001, \eta_p^2 = .10$).

A Bonferroni-corrected, post-hoc analysis of the main effect of language revealed significant differences between the NoLanguage and English blocks ($M_{diff} = -15.66, p < .001, BF_{10} = 221.07$) and the NoLanguage and Welsh blocks ($M_{diff} = -15.48, p < .001, BF_{10} = 316.45$), but there was no significant difference between the NoLanguage and Mixed blocks ($M_{diff} = -9.81, p = .087, BF_{10} = 1.64$). Average RT on English blocks were not significantly different than average RT on Welsh ($M_{diff} = 0.18, p = 1.00, BF_{10} = 0.04$) or Mixed blocks ($M_{diff} = 5.85, p = .858, BF_{10} = 0.25$). Welsh blocks were also not significantly different than Mixed blocks ($M_{diff} = 5.67, p = .934, BF_{10} = 0.23$).

On the interaction effect, there were no significant differences in RTs between all language blocks ($ps > .05$) within the Congruent condition. Within the Incongruent condition, the NoLanguage block had significantly lower RTs than English ($M_{diff} = -19.76, p < .001$), Welsh blocks ($M_{diff} = -20.44, p < .001$) and Mixed block ($M_{diff} = -16.10, p = .015$). Average RT was not significantly different between English and Welsh ($M_{diff} = -0.68, p = 1.00$) or Mixed ($M_{diff} = 3.67, p = 1.00$) blocks. The Welsh block was also not significantly different than the Mixed block ($M_{diff} = 4.34, p = 1.00$). Further details on the estimated mean RTs can be found in Table 10.

Table 10: *Experiment 2C – Estimated Marginal Means by Congruency, by Language, and by Congruency * Language (with Standard Error)*

			95% CI for Mean Difference	
By Congruency	Mean RT (ms)	Error Rate (%)	Lower	Upper
Congruent	420.59 (6.52)	1.20	407.62	433.56
Incongruent	489.46 (6.52)	9.64	476.49	502.43
By Language				
NoLanguage	444.79 (6.82)	4.78	431.25	458.33
English	460.45 (6.82)	5.44	446.91	473.99
Welsh	460.27 (6.82)	5.93	446.73	473.81
Mixed	454.60 (6.82)	5.52	441.06	468.14
By Congruency * Language				
Congruent				
NoLanguage	414.19 (6.75)	1.12	400.12	428.26
English	425.75 (7.15)	0.91	411.68	439.82
Welsh	424.71 (7.51)	1.55	410.65	438.78
Mixed	417.72 (6.71)	1.21	403.65	431.79
Incongruent				
NoLanguage	475.39 (7.01)	8.44	461.32	489.46
English	495.15 (6.70)	9.98	481.08	509.22
Welsh	495.83 (7.41)	10.32	481.76	509.89
Mixed	491.49 (7.52)	9.83	477.42	505.55

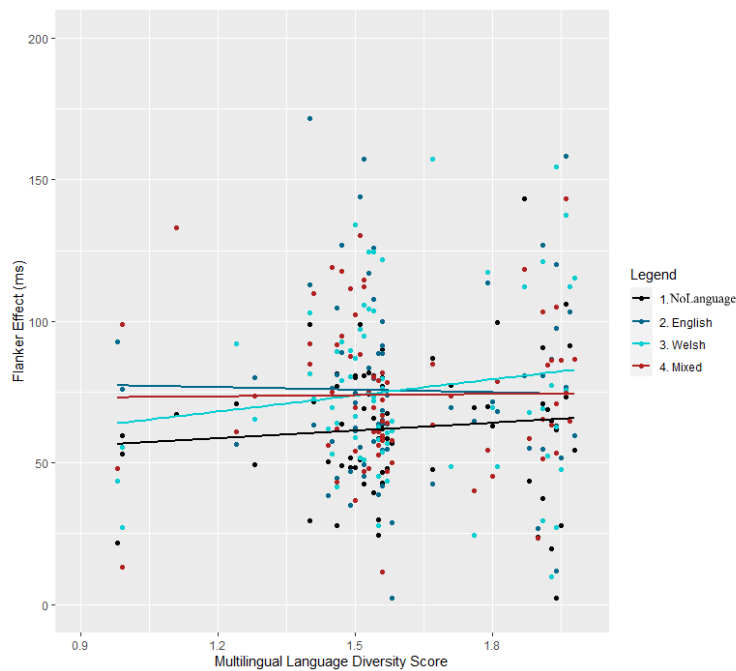
Note. $N = 75$.

4.13.1 Forward-stepwise Multiple Regression

The same regression modelling was conducted as with the previous experiments. To reiterate, a forward-stepwise hierarchical regression was conducted using L2 AoA, proficiency, immersion, dominance, and MLD scores on the flanker effect of each language block, modelled separately. Results indicated that only the model for the Welsh block flanker effect converged successfully. The model consists only of L2 proficiency, which explained 5.5% of the variance in the outcome ($F(1, 73) = 4.21, p = .044, R^2_{\text{adj}} = .04$). As a predictor, a unit increase in L2 Proficiency

results in a 41.21-ms increase in the Welsh block flanker effect ($\beta = 41.21$, $SE = 20.08$, $t = 2.05$, $p = .044$). Notably, MLD scores were also not a significant predictor of any language block flanker effects ($ps > .05$). Given only a single predictor was entered, multicollinearity was not a concern in the model. Figure 8 depicts the relationship between MLD scores and language block flanker effects.

Figure 8: *Experiment 2C - Scatterplot of MLD Scores Against Flanker Effect Across Language Blocks*



4.14 Discussion

The current study assessed whether an unfamiliar language would have any impact on the cognitive implications stipulated in the ACH. First, it was predicted that the Welsh-only block would have the poorest flanker task performance, followed by the mixed language and the English-only block. Results did not support the hypothesis as there were no significant differences in mean RT between the language blocks for both congruent and incongruent trials. However, within the incongruent trials, the mean RT on the single language blocks was significantly slower than the NoLanguage control block, but only slightly slower than the mixed language block. Together, these findings suggest that mixed-block incongruent trial performance was more similar to control than single-language blocks incongruent trial performance. However, Bayesian analysis suggests

that this difference is likely anecdotal, as compared to the decisive evidence in favour of the alternative hypothesis for the differences between NoLanguage blocks and single language blocks. Therefore, it is unlikely that familiar-unfamiliar language pairing can induce cognitive adaptation.

Alternatively, the results may be explained via language learning mechanisms. As the languages primed were explicitly communicated, the participants understood that the unfamiliar words were Welsh and not meaningless words. Hence, Welsh words could have potentially activated relevant language-learning neural pathways. This conclusion is in line with the discovery that even short-term exposure to another language can result in cognitive changes (Linck et al., 2009). Bermúdez-Margaretto et al. (2021) found that although unfamiliar words with a similar script may induce slower naming latencies, by the end of the experiment, its latency matched the fast latency of familiar words. The authors argued that repeated exposure (roughly 10 exposures) to unfamiliar, but similar script words would result in the formation of new memory representations of the word. Subsequent access to these novel word forms is faster, leading to faster naming latencies at the end of the experiment, suggesting rapid orthographic learning. Speculatively, this rapid word-learning mechanism might explain the similarities between Experiment 2A and 2C. If participants are creating novel word representations of Welsh, then, it stands to reason that some degree of language learning is occurring. As demonstrated by both Bak et al. (2016) and Ma et al. (2021), language learning, even as brief as three weeks, can lead to noticeable differences in EF abilities.

When examining the neural patterns of native Chinese speakers as they learned an artificial language using an fMRI, Li et al. (2018) found that merely reading words in the artificial language elicited a similar pattern of neural activity when reading Chinese (see also Grey et al., 2017; Perfetti & Liu, 2005; Perfetti et al., 2007). This artificial language was constructed by combining the visual forms and sounds of Korean Hangul words and matching them with the visual complexity of Chinese words. In a sense, their experiment likely reflected the English-Welsh language pairing in Experiment 2, as both languages are written in the Latin script and therefore share similar visual complexity. If this is the case, then, the unexpected findings in Experiment 2 are less likely due to rapid cognitive adaptations, but rather the effects of language learning.

For the second hypothesis, it was found that MLD scores were not related to flanker effects, which supports the hypothesis. If the effects of language entropy on cognition were spurious, there would be cases where an effect is observed even when none should exist. To date, research to date

has only addressed the relationship in contexts where an effect was expected (Kałamała et al., 2021, 2022; Li et al., 2021; van den Berg et al., 2022; Wagner et al., 2023;), but no one has investigated in contexts where the effect would not have been expected. The present results contribute to the discussion that the relationship between language entropy and cognition is not random.

4.15 Comparison Between Experiment 2A, 2B and 2C

To investigate deeper, the study compared flanker task performance between the three experiments. The datasets from the three experiments were combined ($N = 224$) and a three-way mixed ANOVA, with Experiment (Exp2A x Exp2B x Exp2C) as the between-subjects variable and Congruency (Congruent x Incongruent) and Language blocks (NoLanguage x L1 x L2 x Mixed) as the within-subjects variables were used to analyze it. Neither SES nor SPM was significantly different across experiments, and thus, was not included as covariates in the model.

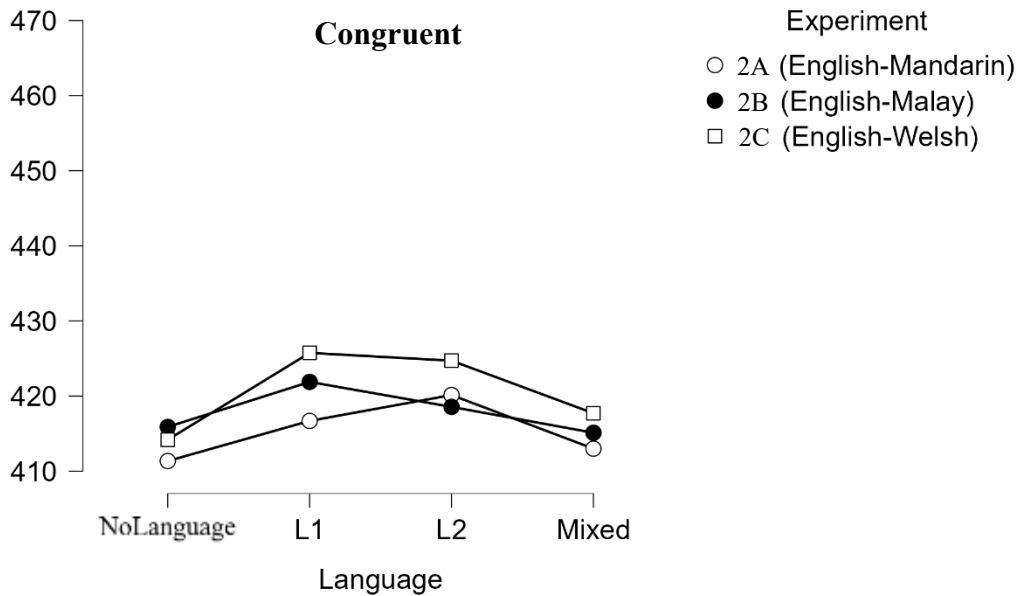
Results revealed that there was the expected main effect of congruency on RT ($F(1, 220) = 2319.02$, $p < .001$, $\eta_p^2 = .913$, $BF_{10} = 6.70 \times 10^{115}$), with significantly faster Congruent ($M = 417.92$, $SE = 3.22$) than Incongruent trials ($M = 487.06$, $SE = 3.45$). The sphericity assumption was violated for the main effect of language blocks. Given that Greenhouse-Geisser epsilon values were above 0.75, the Huynh-Feldt correction was applied. After correction, there was a main effect of language blocks on RT ($F(2.64, 580.38) = 13.50$, $p < .001$, $\eta_p^2 = .06$, $BF_{10} = 8583396.05$). However, there were no main effects of experiment group on RT ($F(2, 220) = 0.35$, $p = .702$, $\eta_p^2 = .00$, $BF_{10} = 0.27$).

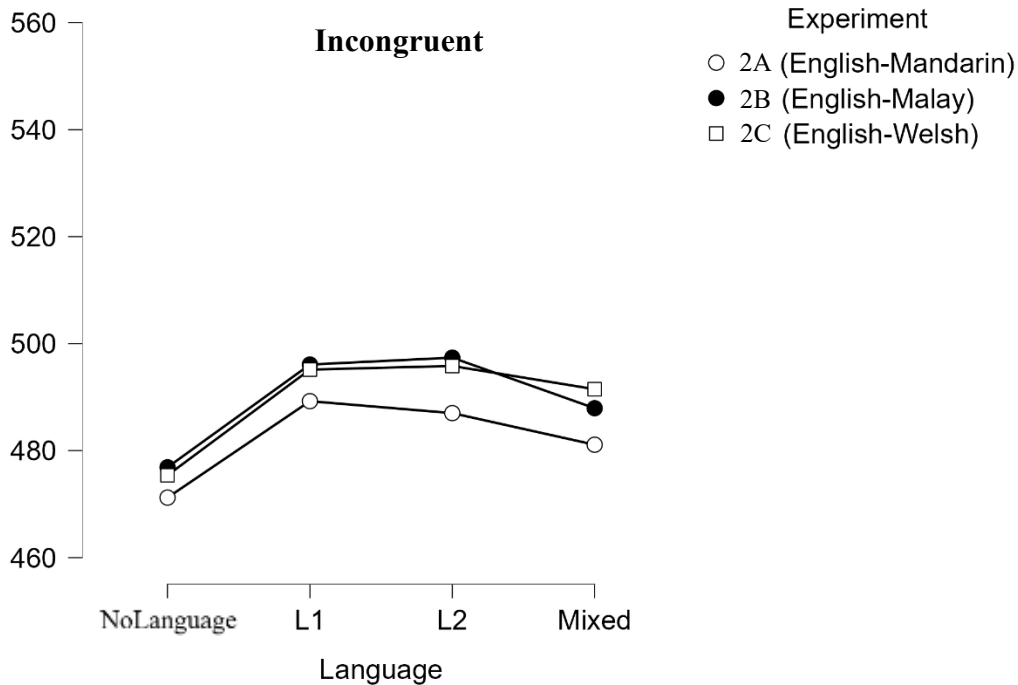
A Bonferroni-corrected, post-hoc analysis on the main effect of Language revealed that NoLanguage block had significantly lower RTs compared to L1 ($M_{diff} = -13.32$, $p < .001$, $BF_{10} = 2.68 \times 10^7$) and L2 ($M_{diff} = -13.12$, $p < .001$, $BF_{10} = 5.22 \times 10^6$), and Mixed block ($M_{diff} = -6.91$, $p = .027$, $BF_{10} = 3.08$). There were no significant differences between L1 and L2 blocks ($M_{diff} = 0.20$, $p = 1.000$, $BF_{10} = 0.02$). There was a marginally significant difference between L1 and Mixed block ($M_{diff} = 6.41$, $p = .051$, $BF_{10} = 23.04$), but no significant difference between L2 and Mixed block ($M_{diff} = 6.21$, $p = .064$, $BF_{10} = 8.53$).

There were no significant interaction effects of experiment group and congruency ($F(2, 220) = 0.96$, $p = .702$, $\eta_p^2 = .00$, $BF_{10} = 0.02$; see Figure 9) or experiment group and language ($F(5.28, 580.38) = 0.16$, $p = .981$, $\eta_p^2 = .00$, $BF_{10} = 0.00$). However, there was a significant interaction effect between congruency and language ($F(3, 660) = 12.16$, $p < .001$, $\eta_p^2 = .05$, BF_{10}

= 214390.44). A simple main effects analysis revealed no significant differences in RTs between all language blocks ($p > .05$) within the Congruent condition. Within the Incongruent condition, NoLanguage block had significantly lower RTs than L1 ($M_{\text{diff}} = -19.01, p < .001$) and L2 ($M_{\text{diff}} = -18.91, p < .001$) and Mixed ($M_{\text{diff}} = -12.34, p < .001$). There was no significant difference between L1 and L2 ($M_{\text{diff}} = 0.09, p = 1.000$) and Mixed ($M_{\text{diff}} = 6.66, p = .372$). There was also no significant difference between L2 and Mixed ($M_{\text{diff}} = 6.57, p = .409$). Further details on the estimated mean RTs can be found in Tables 6, 8, and 10.

Figure 9: Comparison of Congruent and Incongruent Condition Mean RT (ms) Against Language Contexts Across Experiments 2A and 2B and 2C,





4.16 General Discussion

4.16.1 Language Context Rapidly Modulates Executive Functions

Experiments 2A, 2B, and 2C revealed a remarkably similar pattern of results. As seen in Figure 9, mean RT is the fastest in the NoLanguage block, followed by the mixed and single-language blocks. In all three experiments, the mean RT for the mixed-language block is slightly attenuated compared to the single-language blocks, but this difference is never significant. Although a single instance of this pattern could be attributed to random chance, the same pattern occurring in three separate experiments suggests that some form of cognitive adaptation must have taken place but was not strong enough to lead to any measurable behavioural differences between the language blocks. This interpretation gains support from the consistently high performance across all experiments. Essentially, due to the low task difficulty, any cognitive adaptations that did take place might not be substantial enough to be discernible.

The present findings do not fully align with a similar study by Kałamała et al. (2021), who found that in a large sample of Polish-English bilinguals ($N = 195$), language entropy and language mixing (combined as a measure of dual-language context intensity) were unrelated to the performance of four response-inhibition tasks (i.e., Antisaccade, Stroop, Go/no-go, and Stop-signal tasks). Their findings support the argument that the intensity of bilingual experience (i.e.,

diversity in language use and language mixing habits) is unrelated to the efficiency of response inhibition. However, unlike the present experiments, Kałamała et al. (2021) did not experimentally manipulate different language contexts.

Curiously, the large disparity in mean RT between the NoLanguage control and the Language blocks suggests that the inclusion of a non-target word increased task-cognitive demands. Despite being instructed to ignore the intervening words, participants will still automatically and unconsciously process lexical stimuli (Shtyrov et al., 2013). This additional reaction time means that the inclusion of intervening word stimuli taxes an already limited supply of cognitive resources, which must be shared between language processing, inhibitory control, and task-goal maintenance. If this is the case, then the addition of another language in the mixed block should have increased the cognitive load of that block. Not only would participants be expected to maintain task goals, but they would also need to juggle the processing of words in two languages. If there were no cognitive adaptations, this inclusion would have led to the mixed block having the slowest mean RT compared to the single language blocks. As the data from all three experiments suggest otherwise, there is evidence to support that language contexts do indeed rapidly modulate EF, thereby supporting the ACH.

4.16.2 Linguistic Distance and the Multilingual Effect

Another aim of this study was to investigate the effects of LD on the multilingualism-cognition relationship. The distance between English and Mandarin was 0.67, whereas the distance between English and Malay was 0.36. Although Welsh was primed in Experiment 2C, the lack of familiarity and proficiency of participants makes its inclusion in a linguistic distance analysis unnecessary. The results demonstrated no significant differences between the mean RT of Experiment 2A and 2B, which contradicted the hypothesis that the English-Malay participants would have better performance than the English-Mandarin participants. A plausible explanation for this finding could be attributed to how the Malay language was initially categorized as linguistically similar to English and Mandarin as linguistically dissimilar to English (Chiswick & Miller, 2005). Yet, according to a phylogenetic approach (Beaufils & Tomin, 2021), it appears that both Malay (86.7) and Mandarin (84.9) may be equally linguistically dissimilar to English.

Currently, it is unclear whether the more traditional approach to evaluating LD via phylogenetics is better or worse than the scalar measure via relative test scores of language

learning. Although a phylogenetic approach attempts to conceptualize linguistic distance in the form of evolutionary histories and (dis)similarities on various lexical features, such as phonology and script, it is unclear how, or if, these matter to a multilingual's language control ability the way that traditional multilingual variables (e.g., proficiency, immersion, dominance, AoA) do. On the other hand, the scalar measure's shortcoming is in its specificity. As its outcomes are based on the language learning ability of a particular group, the results can be biased by a host of individual factors (e.g., working memory capacity, language popularity).

To follow the scalar approach is to conclude that there are no effects of LD on the multilingualism-cognition relationship. However, two critical issues undermine confidence in this conclusion. First, the data from Chiswick and Miller was based on a 1993 study (Hart-Gonzalez & Lindemann, 1993), and second, the sample used was English-speaking Americans. Based on psychometric principles, test scoring and standardization (i.e., norming) must be continuous (Kyriazos & Stalikas, 2018). The resulting normed scores would then only be valid and reliable to that specific group. This process would reduce error and bias arising from different cultures, cognitive abilities, and age groups. Following this logic, the LD outcomes of Chiswick and Miller (2005) could be considered outdated, highlighting the need for contemporary research.

A more likely explanation for the results, according to the phylogenetic approach, is that the second languages primed (i.e., Mandarin and Malay) in this study were equally dissimilar to English. This interpretation would explain why the mean RT was not significantly different between Experiment 2A and 2B. Furthermore, the findings of past studies align with this conclusion. Table 11 summarizes various studies and tasks, detailing the language pairs employed, their phylogenetic distance indicated by the Genetic Proximity Value (calculated post-hoc via www.elinguistics.net), and the accuracy of the author's LD classification compared to the genetic proximity value. In hindsight, it is unsurprising to note that most of these studies' categorizations of language pairs as either similar or distant align closely with their genetic proximity values. This alignment is to be expected, given that they employ the same phylogenetic approach.

Table 11: *Summary of Past Studies and Their LD Classification*

Author	Task	Results	Language Pairs (LD classification)	Genetic Proximity Value	Correctly Classified?

Coderre & van Heuven (2014)	Stroop task	<i>sig.</i>	English-German	31.3	True
		Similar	(Similar)	(Related)	
		> Distant	English-Polish (Moderate)	54.1 (Remotely Related)	True
Ljungberg et al. (2019)	Episodic memory recall tasks, Verbal letter fluency, Verbal categorical fluency, Mini-Mental State Examination	<i>sig.</i>	English-Arabic (Distant)	85.5 (Unrecognized)	True
		Similar	Swedish-English (Similar)	31.0 (Related)	True
		> Distant	Swedish-Finnish (Distant)	84.6 (Unrecognized)	True
Ramanujan (2019)	Translation and Reading tasks	<i>sig.</i>	English-Dutch	21.8	True
		Similar	(Similar)	(Highly related)	
		> Distant	English-Hindi (Moderate)	68.9 (Related)	True
Olguin et al. (2019)	Dichotic-listening task	<i>sig.</i>	English-Cantonese (Distant)	87.5 (Unrecognized)	True
		Similar	English-Dutch	21.8	True
		> Distant		(Highly related)	
Radman et al. (2021)	Picture-naming tasks	<i>sig.</i>	English-Spanish	59.3 (Related)	False
		Similar	English-French (Similar)	46.9 (Related)	True
		> Distant	English-Persian (Distant)	77.3 (Very remotely related)	True
Wolfsturn et al. (2023)	Stroop task	<i>n.s.</i>	Spanish-Italian	14.0	True
		Similar	(Similar)	(Highly related)	
		= Distant	Spanish-Dutch (Distant)	58.4 (Related)	False

Note. *sig.* = significant; $p < .05$, *n.s.* = non-significant, $p > .05$. The genetic proximity is expressed as a value between 0 (same language) and 100 (biggest possible distance). Values between 1 - 30 = Highly related; 30 - 50 = Related; 50 - 70 = Remotely related; 70 - 78 = Very remotely related; 78 - 100 = Unrecognized.

It appears that if a study correctly classifies their language pairs as similar vs distant according to the genetic proximity values, it is more likely to find a significant group difference. The exceptions to this rule are Olguin et al. (2019) and Wolfsturn et al. (2023), whose distant language classifications did not match the genetic proximity, yet, where Wolfsturn et al. found no significant group differences, Olguin et al. found significant differences. Unsurprisingly, the different tasks used in these studies make it difficult for clear comparisons. However, until more research is done, it seems that differences in LD may lead to measurable differences in cognitive abilities. A caveat to this conclusion is that the languages must be sufficiently distant from one another for a bilingual effect to be observable. If the LD is not distinct enough, such as in the current study, or Olguin et al. and Wolfsturn et al., the effects would not be clear.

Furthermore, given that only two recent papers (Gallo et al., 2023; Sörman et al., 2019) have tackled the impact of LD on cognitive outcomes in older adults, there is a potential that these conflicting findings stem from the peak cognitive abilities of younger adults. As proposed by Valian (2015) and Gallo et al. (2020), when other cognitively demanding life experiences, such as sports, social interactions, or education, reach a certain intensity, any potential effects of bilingualism on cognition might be overshadowed.

The rationale behind Experiment 2C was that if there is a true effect of language context priming on cognitive adaptation, then, by priming a familiar-non-familiar language pair, the mixed language context would mirror the cognitive requirements of a single-language context. This should have led to similar flanker task performance between the English-only and the English-Welsh context. Essentially, Experiment 2C investigates whether knowing both primed languages is a unique requirement for cognitive adaptation to occur, or could the mere exposure to any second language, familiar or not, be enough to elicit cognitive adaptation?

The results of Experiment 2C reveal that cognitive adaptation is not quantitatively unique to knowing both primed languages in a mixed context. Qualitatively, however, Experiment 2C had the slowest mean RT on incongruent trials followed by 2B and 2A. Together, these findings suggest that there may be an effect of language familiarity (or proficiency in the case of 2A and 2B), but the effect is not noticeably different given the observed ceiling effect in accuracy. Further research on this using more cognitively demanding tasks is required to confirm such conclusions.

4.16.3 Predicting Flanker Task Performance with MLD Scores

The present study also investigated the utility of MLD scores in the investigations of the multilingual effect. The results provided mixed support for MLD as a superior measure of multilingualism compared to other traditional multilingualism variables (e.g., L2 AoA, Proficiency). Although MLD predicted flanker task performance on the mixed-block English-Mandarin incongruent trials, this relationship was not found for the other blocks, nor in the English-Malay or English-Welsh experiments. It also did not predict task performance in single-language contexts, suggesting that the predictive power of MLD may be limited in scope. Importantly, MLD could not predict performance on all the experiments' NoLanguage block, which is essentially a standard flanker task without language context priming, similar to Kalamala et al. (2020) findings. Synthesizing these findings, it is argued that the effects of language context in modulating EF are transient, existing only when a mixed-language context is primed.

One potential source of noise may be the variations in the second primed language proficiency (i.e., Mandarin, Malay, and Welsh). In all three experiments, English proficiency was high. However, in Experiment 2A, participants were highly proficient in Mandarin; in Experiment 2B, participants were moderately proficient in Malay; in Experiment 2C, participants reported no knowledge of Welsh. Therefore, it might be misleading to conclude that the MLD score is a superior measure compared to other traditional measures of multilingualism. It is plausible that although MLD can predict task performance under certain conditions, it likely interacts with other variables related to multilingual experience to create a complex relationship with cognitive control.

Other studies using language entropy support the utility of the MLD score in predicting EF. For example, Li et al. (2021) found that for English-Mandarin multilinguals, higher language entropy explained the greatest variance in higher brain network specialization and segregation and lower signal variability, which were associated with better task switching, goal maintenance, and monitoring abilities, but worse conflict effect in inhibitory control (but see also Grundy et al., 2017). In Kang et al. (2023), fMRI of Cantonese-Mandarin multilinguals during a semantic Go/No-Go picture naming task, demonstrated that as task demands increased, multilinguals performance decreased. However, higher language entropy was related to higher accuracy in naming and greater involvement of domain-general, compared to domain-specific, neural networks. Importantly, differences in calculating language entropy across studies complicate direct comparisons,

underscoring the importance of a standardized measure for accurate multilingual effect investigations (Kalamala et al., 2021; Kang et al., 2023; Li et al., 2021; Wagner et al., 2023).

There is also a possibility that instructing participants to ignore the word stimuli may have led them to look away from the words, preventing any language context priming. However, because the words appear immediately after and directly on top of central fixation, it is unlikely that participants averted their gaze. If participants had looked away from central fixation, it would have adversely affected the RT of the flanker task. To ensure a more robust methodology, researchers could either implement a simple memory word-recall test after the experiment or use eye-tracking methodology to confirm that participants are looking at the word stimuli (this is addressed in Chapter 5).

Furthermore, given the ceiling effect in accuracy across the three experiments, the next logical step would be to test the sample with a similar but more complex experimental task. For example, the Attention Network Test (ANT; Fan et al., 2002), combines a Posner Cueing task (Posner, 1980) with the flanker task. This combined task could potentially introduce a higher level of cognitive demand and promote more variability in task performance. To enhance the sensitivity of the results to changes in cognitive abilities, the following chapter will employ an eye-tracking methodology. By using both eye-tracking and behavioral data, it will be possible to triangulate the changes in cognitive abilities resulting from rapid shifts in language contexts.

Lastly, because there were no significant differences in the results between the Malaysian Chinese and Malaysian Malay samples, the rest of the thesis will only recruit Malaysian Chinese participants. This restriction also has the added benefit of allowing for more direct study-to-study comparisons, as there are significantly more studies involving Mandarin speakers than Malay speakers (see Hartanto & Yang, 2016; Ma et al., 2021; Yang et al., 2023).

Chapter 5: Does Linguistic Contexts Affect Multilinguals’ Performance on the Attention Network Task? An Eye-Tracking Study

5.1 Introduction

Chapter 4 focused on whether a bilingual effect can be found in inhibitory control. Succinctly, the IC/BICA Model used in Chapter 4 (see also Section 1.4.1) suggests that the constant recruitment of inhibitory control to manage cross-language activation strains and trains the brain, resulting in better EF. However, the results of Chapter 4 did not find any effect. In line with recent discussions within the literature, inhibitory control may not be the primary mechanism underlying the bilingual effect. Bialystok and Craik (2022) argued that because many predictions regarding inhibitory control effects fail (see also Paap & Sawi, 2014), the IC model provides an incomplete account of the findings of past research. Bialystok (2017) proposed, instead, that “lifelong bilingualism impacts a set of processes subsumed under the category of executive attention” (p. 250). This proposal means that bilingualism leads to improvements in attentional control, which confers domain-general benefits across different aspects of cognition across the lifespan. This section describes the case for attention as the mechanism underlying the bilingual effect.

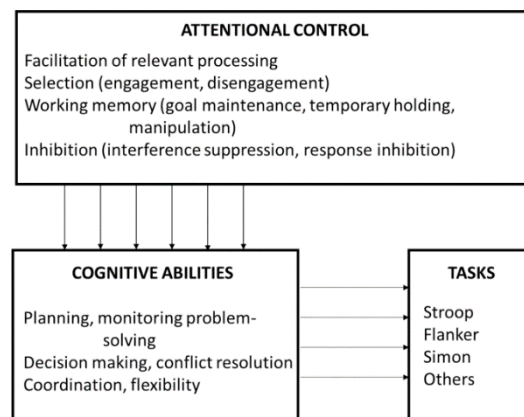
Defining attention has been challenging (Bialystok & Craik, 2022), as many researchers tend to interchangeably describe it as both a finite processing resource (Wickens, 1980) and a top-down control process (Chun et al., 2011). Depending on the research, attention can either be considered a resource or a system (Oberauer, 2019). In the case for attentional control proposed by Bialystok and Craik (2022; see also Bialystok, 2017), attention is considered a system that is influenced by bilingual experiences.

The general attentional network model by Posner and Petersen (1990) describes how attention, as a system, is a multifaceted construct that can be delineated into three distinct, yet interconnected components, which are Alerting, Orienting, and Executive Control. Decades of neuroimaging data have provided strong empirical support for the existence of these components and have been validated in a variety of populations (for review, see Petersen & Posner, 2012; Posner, 1980, 2008; Posner & Petersen, 1990; Posner & Rothbart, 2007). Although there is no clear one-to-one mapping between Posner and Petersen’s general attention network model and

Bialystok and Craik's attentional control framework (see Figure 10), it may be speculated that the alerting and orienting components share similarities with selective attention engagement and disengagement, whereas the executive control component is likely mapped onto both working memory and inhibitory control.

Hence, attention, or more specifically attentional control (see Figure 10), will be henceforth defined as a broad hierarchical set of top-down control processes that manage and mobilize processing resources for cognitive processing, goal maintenance, interference control, and cognitive switching to fulfil specific task goals (Bialystok & Craik, 2022). Burgoyne and Engle (2020) also agree that attentional control is the common thread linking the performance of many higher-order cognitive abilities. This view aligns with how language control functions because multilingual language control does not only involve inhibitory control but also monitoring, working memory, and language (dis)engagement (Green & Abutalebi, 2013). While bilingual language control primarily operates in the auditory modality, and the concept of far-transfer (i.e., from the auditory domain to domain-general executive function) has been criticized (Paap et al., 2021), this attentional control model is based on adaptation rather than far-transfer. In this model, an operation or set of operations, along with their underlying neural networks, is modified through bilingual experience so that all involved domains are affected (Bialystok & Craik, 2022). Thus, by enhancing the cognitive operations, individual task processes become more efficient.

Figure 10: *Overview of The Proposed Attentional Control Hierarchy. Adapted from Bialystok and Craik (2022)*



The alerting component serves to achieve and maintain a heightened state of vigilance toward an oncoming stimulus. Fan et al. (2009) importantly delineates the different components of alertness, specifically the difference between tonic and phasic alertness. Where tonic alertness is implicated in sustained attention (e.g., wakefulness and vigilance), phasic alertness is the ability to increase one's response readiness to a target following a warning cue. The orienting component serves to selectively attend to specific information from sensory inputs. Here, orienting may either be exogenous, as in when an external cue draws attention to its location, or endogenous, as in when voluntarily focusing attentional fields. Orienting may also be either overt, as in the movement of the head or eyes towards an external cue, or covert, as in the shifting of attentional fields without the accompanying head or eye movements. The executive control component, compared to the other two attentional components, represents a higher-order cognitive process that is heavily involved in processing, monitoring, and resolving information from different brain areas.

The Attention Network Task (ANT; Fan et al., 2002) allows for the simultaneous measurement of all three components: alerting, orienting, and executive control. The ANT is a combination of the Posner cueing task (Posner, 1980) and the Flanker task (Eriksen & Eriksen, 1974). It enables researchers to examine alerting and orienting (measured by the Posner cueing task) and executive control (measured by the Flanker task). In this task, participants are given on-screen cues that appear either at 1) the center (center cue), 2) either above or below the central fixation (spatial cue), 3) both above and below the fixation (double cue), or 4) not at all (no cue). Spatial cues provide both temporal and spatial information about the appearance of the flanker stimuli, whereas center and double cues only provide temporal information. Participants are instructed to make a saccade towards the cues whenever they appear on screen. While there is only one cue presented during spatial cues, thus only one possible saccade direction, participants may saccade towards either cue during double cue presentation. During no cue trials, the flankers appear without cue and participants would continue fixating on the fixation cross. No saccades should happen here as no other stimuli would be presented on screen.

The flanker stimuli are either congruent (all arrows pointing in the same direction) or incongruent (the target arrow pointing in the opposite direction to the flankers). The alerting component is calculated by subtracting RT of double cue conditions from no cue conditions. Orienting is calculated by subtracting the RT of spatial cue conditions from center cue conditions. Executive control is calculated by subtracting the RT of congruent flankers from incongruent

flankers. Participants are required to respond to the direction of the target arrow while ignoring the directions of the surrounding flanker arrows.

Bilingual effects on the ANT have been observed across various age groups, including bilingual infants (Arredondo et al., 2022), young adults (Costa et al., 2008), and older adults (Dash et al., 2019), when compared to their monolingual counterparts. However, it appears that bilingualism may affect some, but not all, attentional components. For example, Costa et al. (2008) and Tao et al. (2011) are large-scale studies ($N > 100$) that found significant differences only in the alerting and executive control, but not orienting, subcomponents.

Similarly, Marzecová et al. (2012; $N = 35$) investigated the attentional networks in bilinguals and monolinguals using the LANT, where the stimuli were presented vertically in the right and left visual fields, which measures participants' hemispheric lateralization. Results similarly demonstrated a significant difference in alerting and executive control, but not orienting, between bilinguals and monolinguals, neither was there a difference in hemispheric lateralization.

Timmer et al. (2021) found that although no behavioural differences between bilinguals and monolinguals were detected, they did find a significant difference in the ERPs, suggesting that bilingualism does impact attentional control by heightening the overall attention state to deal with conflicts more readily, but this effect could not be seen behaviourally due to the facilitatory effects of the alerting cue on RT performance.

5.2 Current Study

The present study aimed to investigate whether attentional networks (Alerting and Executive Control) are modulated by language contexts. Specifically, whether language contexts can rapidly modulate attentional networks. Studies typically find that bilingual adults are either better or no different in attentional abilities than monolinguals, but never worse (Bialystok, 2017; Bialystok & Craik, 2010; Bialystok & Viswanathan, 2009; Hartanto & Yang, 2019; Janic et al., 2020; Marian & Shook, 2012; Sörman et al., 2019). Yet, it is unclear how this relationship is modulated by existing language contexts. Timmer et al.'s (2021) experiment suggests that bilingual contexts demonstrated enhanced target-P3 amplitudes compared to monolingual contexts in an ANT. Notably, no RT differences were found. Studies using an inhibition-of-return (IOR; Wang et al., 2015) task and the AS task (Riek et al., 2023) demonstrated that eye-tracking paradigms may reveal important and complementary insights to behavioural data related to changes in EF and

attentional control. The study additionally aims to investigate whether MLD scores can positively predict attentional networks above and beyond traditional measures of multilingualism. Finally, the use of an eye-tracker means a comparison can be made between traditional RT measures and eye-tracking metrics, to see which is more sensitive to the effects of the primed language contexts.

Although eye-tracking is commonly employed in language-based research to explore linguistic processing during reading or listening tasks (Hayakawa & Marian, 2020; Incera, 2018), its application in investigating the bilingualism effects has been limited (Bialystok et al., 2006; Blumenfeld & Marian, 2011; Blumenfeld et al., 2016; Mercier et al., 2014). For example, the combination of the ANT and eye-tracking methodology has received little attention in the bilingual-cognitive literature, despite the significance of attention in this context (Hilchey et al., 2014). No previous studies examining the influence of bilingualism on cognitive control using the ANT have employed an eye-tracking methodology (Incera, 2018). For the studies that do use the ANT to investigate the bilingual effect, they typically rely on behavioural measures (Ooi et al., 2018; Yang & Yang, 2016) or neurophysiological measures, such as fMRI and EEG (Dash et al., 2019; Timmer et al., 2021).

Incorporating eye-tracking methodology into an ANT paradigm may offer novel and complementary insights into the cognitive processes and temporal dynamics of the bilingual effect, since RTs can be dissociated from cognitive control (Wong et al., 2017). For example, in tasks like the ANT, where a specific response is required per trial, responses in similar subsequent trials can become habitually initiated. This pre-initiation of movement has been documented to be independent of executive planning (Haith et al., 2016). In other words, RTs may be faster or slower without necessarily indicating changes in underlying cognitive control. Beesley et al. (2019) further support this by arguing that RTs represent the cumulative latency of all preceding cognitive processes, making it difficult to pinpoint which specific process is affected by bilingualism or changing language contexts. Early, quicker autonomous processing may be overshadowed by later, slower controlled processing. Therefore, it is essential to complement RT data with other sources to capture nuances that might otherwise be missed.

In the ANT task, given its low perceptual load, saccadic latency and fixation duration were selected to measure Alerting and Executive Control respectively. Here, saccadic latency refers to the time it takes for a saccade to be initiated in response to the cue. It can be affected by the saliency of the stimulus/cue, and or the individual's alerting ability. Shorter saccadic latency is typically

associated with better alerting abilities (Posner & Rothbart, 2007) and working memory (Hutton, 2008). This relationship is more evident during double cue trials compared to spatial cue trials, as the former provides only temporal information, isolating the alerting effect from the orienting effect. Few studies found a facilitative effect of saccades on executive control (Edlin & Lyle, 2013; Lyle & Martin, 2010; Ries et al., 2019), but no previous studies have examined saccades as a measure of the Alerting component. On the other hand, fixation-duration refers to the time spent engaging with a stimulus, in this case, the flanker stimuli. The flankers have low perceptual complexity, but specifically during incongruent trials, the flankers have a higher degree of cognitive complexity. The faster resolution of incongruent trials may be reflected in shorter fixation durations (Liu et al., 2022). Hence, it is speculated that shorter fixation durations would indicate better Executive Control.

Therefore, the current study hypothesizes that, based on RT data, the mixed language block is expected to exhibit the fastest RT compared to the NoLanguage and single-language contexts. Additionally, a larger Alerting component and a smaller Executive control component (i.e., smaller interference effect) are anticipated in the mixed language block compared to the NoLanguage and single-language blocks. Next, based on eye-tracking measures, fixation durations (henceforth flanker-fixation duration) were taken as a measure of Executive control. In addition, the first saccade latencies (henceforth first cue-saccade latency) were taken as a measure of Alerting. Based on both metrics, the mixed language block is expected to show a smaller flanker fixation duration component and shorter first cue-saccade latency compared to the NoLanguage and single-language blocks.

The current study also predicts that participants' MLD scores are positively associated with better performance in both Alerting and Executive control components, as assessed by RT, flanker-fixation duration, and first cue-saccade latency. It is important to note that no specific hypotheses were formulated for the orienting component, as previous studies have not reported a consistent relationship between bilinguals and monolinguals in this regard (Costa et al., 2008; Marzecová et al., 2012; Tao et al., 2011).

5.3 Method

5.3.1 Design

The present study used a 2 (Congruency: Congruent vs. Incongruent) x 4 (Language: NoLanguage vs. English vs. Mandarin vs. Mixed) x 4 (Cue Type: None vs. Centre vs. Double vs. Spatial) within-subjects design.

5.3.2 Materials

Language proficiency was calculated using the **LHQ** (Li et al., 2020). See Section 2.3 for more information. SES was measured using **MacArthur's SSS Scale**, and non-verbal intelligence was measured using a short version of **Raven's SPM**, see Chapter 2 for further details. The present experiment also included a Multilingual Switching Questionnaire (MSWQ), adapted from the BSWQ (Rodriguez-Fornells et al., 2012) and a modified ANT (Fan et al., 2002).

Multilingual Switching Questionnaire. The decision to include this questionnaire is informed by the findings that code-switching may shape EF in bilinguals (Hofweber et al., 2020b). Although Hofweber et al. suggested that Alternation has the strongest training effects on inhibitory components of EFs, Kuzyk et al. (2019) suggested that general code-switching practices can also lead to better inhibitory control. As such, the BSWQ was adapted to measure code-switching practices in general. In the MSWQ, languages beyond L2 are accounted for, unlike the BSWQ. The questionnaire similarly measures the same language-switching constructs as the original questionnaire: (1) L1-switch (L1S: tendencies to switch to L1 from other languages); (2) L2-switch (L2S: tendencies to switch to L2 from other languages); (3) L3-switch (L3S: tendencies to switch to L3 from other languages); (4) contextual switch (CS: frequency of language switches that are prompted by a specific event, topic, or setting); and (5) unintended switch (US: indexes the lack of intention and awareness of the language switches). Questions are based on a 5-point Likert scale, where “Never” represents 1 point, and “Always” represents 5 points. Each component is measured by three questions, therefore, the range of scores of each component is between 3 – Never engage in this type of switching to 15 – Very frequently engage in this type of switching.

Modified ANT. Attention is a multifaceted construct that can be delineated into three distinct, yet interconnected, components, which are Alerting, Orienting, and Executive control (Dash et al., 2019; Petersen & Posner, 2012). These components can be measured using the ANT (Fan et al., 2009; see Macleod et al., 2010, for a psychometric review of the ANT task). The present

study used the standard version of the ANT (Fan et al., 2002) but was modified to include interleaved words in English, Mandarin, or both, in prime language contexts (Timmer et al., 2021; Wu & Thierry, 2013).

The English word priming stimuli from Experiment 2A were used in this study (see Appendix 2, English Word Priming List from Experiment 2A). For the Mandarin stimuli, they were the same as those used in Experiment 2A, but Mandarin characters were displayed rather than pinyin. This was done because the current sample consisted of highly proficient Chinese speakers. As discussed in Section 4.6, Chen et al. (2017) found that only low proficient Chinese speakers access meaning through pinyin. Displaying pinyin to highly proficient Chinese speakers would result in orthography interference effects.

As a NoLanguage (control) context, a noise-textured mask was presented instead of words. This address concerns that the NoLanguage block in the experiments in Chapter 4 was not appropriate as it had shorter ITIs compared to the language blocks. The use of these noise-textured masks ensures that the ITI between the NoLanguage block and the language blocks are the same. These priming stimuli were presented at the start of every trial on top of the fixation cross in the center. Participants were instructed to ignore these priming stimuli. The experiment was programmed using Experiment Builder v. 2.3.38 (SR Research Ltd.).

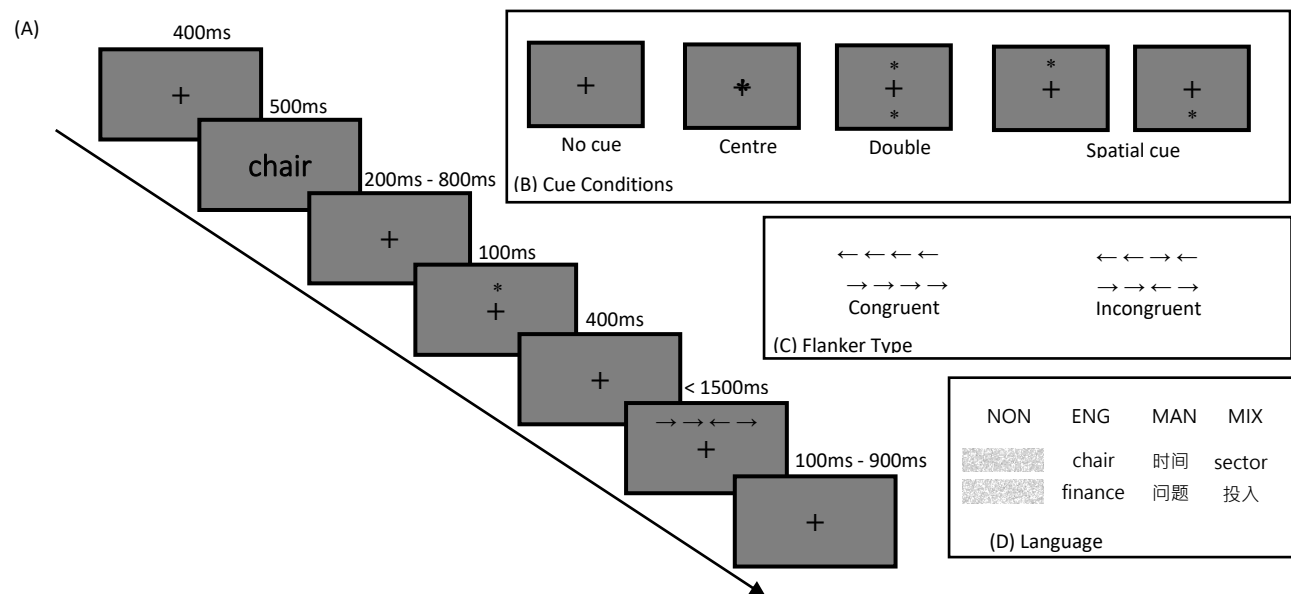
In this task, a central arrow, flanked by two arrows on each side, was presented to the participants. In 50% of the trials, the central arrow points in the same direction as the flanking arrows (Congruent), otherwise, it points in the opposite direction (Incongruent). The task required participants to respond to the direction of the central arrow while ignoring the directions of the flankers. A fixation cross was shown continuously at the center of the screen throughout each trial, disappearing only briefly at the end of the trial. Preceding each flanker stimulus, an asterisk (*) cue may be present or absent (No Cue). If the cue is present, it will either appear above or below (Spatial Cue), both (Double Cue), or directly on top of the fixation cross (Centre Cue). The spatial cues always correctly predict the spatial location of the flanker stimuli. Overall, it took 40 minutes, on average, for participants to complete the modified ANT. An overview of the task procedures can be seen in Figure 11.

Flanker stimuli dimensions remained similar to the original measures provided by Fan et al. (2002). Each flanker arrow subtended to 0.60° of visual angle, where the contours of each arrow were separated by 0.06° of visual angle. The combination of all five arrows subtended a total of

3.27 visual angle. However, distances between stimuli were modified to accommodate the eye-tracking methodology. In every trial where the cues were present, the location where the asterisk appears was presented at 5.49° of visual angle above and/or below the fixation. The fixation cross was presented at the center of the screen, subtended to 0.48° of visual angle. The cue stimuli subtended to 0.8° of visual angle. A grey background (RGBA; 125, 125, 125, 255) was used throughout the experiment and was displayed on a 1920 x 1080, 60-Hz refresh rate, monitor.

The experiment consists of one practice block with 12 trials, and 4 counterbalanced experimental blocks, each with 96 trials for a total of 384 experimental trials. A nine-point calibration and validation procedure were conducted before the start of each experimental block to ensure optimum eye-gaze tracking.

Figure 11: *Overview of the Modified ANT*



Note: (A) An example of the sequence of events of the Modified ANT with English word stimuli, spatial cues, and incongruent flanker stimuli. (B) Overview of the four possible cue conditions. (C) Overview of the two types of flankers. (D) Example of Language Context priming, where noise masks are presented in the NON block, English words are presented in the ENG block, Mandarin characters are presented in the MAN block, and both English words and Mandarin characters are presented in the MIX block.

Typically, the ANT has a standard main effect of cues and congruency on RTs, and a two-way interaction effect between cues and congruency on RTs (Costa et al., 2008; Fan et al., 2002; Fan et al., 2009; MacLeod et al., 2010). Incongruent flankers result in the slowest RT. This slowing effect is greatly enhanced when no cues are presented. However, the effect is attenuated when the center, double, and spatial cues are provided, with spatial cues showing the greatest attenuation.

RT Metrics. Note that, as past studies suggest no effect of bilingualism on the Orienting component (Costa et al., 2008; Dash et al., 2020; Marzecová et al., 2012; Tao et al., 2011), the present study will only report on the Alerting and Executive control components. However, the indices of the Orienting component (i.e., spatial or center cue; Fan et al., 2009) remain in the ANT task design to improve comparability with past ANT studies. Attention and Executive control components are operationally defined as a function of the correct RT performance between one condition and an appropriate reference condition (Fan et al., 2009):

1. *RT_Alerting.* The double cues provide only temporal information about the appearance of the flanker stimuli. Therefore, a facilitative effect of double cues on target response speed about no-cue trials was expected (evaluated across congruency type). This facilitating effect can be calculated by subtracting the mean RT of all double-cue conditions from the mean RT of all no-cue conditions. Larger positive values here are indexed by a more efficient alerting component. This formula is expressed as:

$$RT\_Alerting = RT_{no\ cue} - RT_{double\ cue}$$

2. *RT_Executive control.* The flanker conflict effect indexes executive control (specifically, response inhibition). This effect can be calculated by subtracting the mean RT (evaluated across all cue types) of all congruent flanker conditions from the mean RT of all incongruent flanker conditions. Smaller values here index a more efficient executive control component. This formula is expressed as:

$$RT\_Executive\ control = RT_{flanker\ incongruent} - RT_{flanker\ congruent}$$

Eye-tracking Metrics. The EyeLink 1000 Plus system employed an online parser using a “saccade picking” approach. A sample is classified as a saccade if its velocity and/or acceleration exceeds 30°/s and 8000°/s² respectively. Conversely, if a sample’s velocity and/or acceleration falls below either threshold, it is categorized as a fixation. To ensure consistency and facilitate

comparisons with the RT data, the same calculation method was used for determining the Alerting and Executive control components. This approach enhances parity and interpretability across the different measurements:

1. *Sacc_Alerting*. The first cue-saccade latency during cue presentation was used as a measure of Alerting (see Section 5.2). Note that saccades occurring during no-cue trials, which are centered on the fixation cross, serve as a baseline and are not meaningful when evaluated independently. They are not meaningful, because they do not represent pro-saccadic movements towards any external cues. Hence, as a measure of Alerting, the difference between the mean saccadic latency in no-cue trials and double-cue trials was calculated. To capture only relevant saccades, only saccade samples occurring 200 – 800 ms before the end of cue presentation and within 3 degrees of visual angle to the left and right from central fixation were considered. These saccades were baseline corrected by deducting the fixation (both static and variable) and word display durations. Smaller latencies index a more efficient alerting component. This formula is expressed as

$$\text{Sacc\_Alerting} = \text{Sacc\_Latency}_{\text{double cue}} - \text{Sacc\_Latency}_{\text{no cue}}$$

2. *Fix_Executive control*. Flanker-fixation duration was used as a measure of Executive control, with shorter durations indicating better performance (see Section 5.2). Like RT_Executive control, the difference between mean fixation durations in incongruent and congruent trials was taken as an indicator of Fix_Executive control. To capture only relevant fixations, fixations occurring within 4.09 horizontal degrees of visual angle and 2.45 vertical degrees of visual angle parameter centered on the flanker stimuli during stimulus presentation was considered. This formula is expressed as

$$\text{Fix\_Executive control} = \text{Fix\_Duration}_{\text{flanker incongruent}} - \text{Fix\_Duration}_{\text{flanker congruent}}$$

5.3.3 Participants

To arrive at an appropriate sample size, an a priori analysis for a regression analysis was conducted using G*Power3 (Faul et al., 2007). The output predicted that at least 77 participants are required to detect an effect with 80% probability, similar to Chapter 4. The current study recruited 71 young adult students from the University of Nottingham Malaysia based on the

following criteria: (1) aged between 18 to 25 years old, (2) proficient and literate in both English and Mandarin, (3) local Malaysian Chinese, and (4) no prior or current mental health diagnoses. Specifically, only local Malaysian Chinese were recruited to control for the confounding effects of culture on the bilingual effect (Samuel et al., 2018; Ye et al., 2017).

Before analysis, incorrect responses, and outlier trials with response latencies 2.5 SD above or below the participant's mean per condition were removed from the data. Based on participants' data integrity score ($\frac{\text{Number of Trials Retained}}{\text{Total Possible Number of Trials}}$), nine participants with scores below 80% were subsequently removed from the analysis. This exclusion brought the final number of participants to 62, with 46 females and a mean data integrity score of 97.20%.

Six participants had scores that exceeded the maximum value of 1 on their L1 dominance scores. It was found that they were very liberal in the way that they rated their L1 dominance. For example, one participant listed "24 hours" as the number of hours spent using L1 for reading, writing, speaking, and listening. To address this problem, these six participants had their L1 dominance data retroactively capped at 1 to indicate a high L1 dominance compared to their other languages. Furthermore, all participants reported learning an L2, but 40 participants reported learning an L3 or L4 language, which was anticipated given Malaysia's status as a multilingual community (Eberhard et al., 2022). As the languages of participants differed across L1 to L4, measures for English and Mandarin languages were computed separately. All descriptive data can be found in Table 12 below.

Table 12: *Participant Demographic Measures (Range)*

Demographic	<i>M</i>	<i>SD</i>
Age (years)	19.78	1.82
English AoA (years)	3.55	1.71
English Proficiency (0 – 1)	0.82	0.14
English Immersion (0 – 1)	0.82	0.13
English Dominance (0 – 1)	0.68	0.24
Mandarin AoA (years)	3.61	2.06
Mandarin Proficiency (0 – 1)	0.82	0.15
Mandarin Immersion (0 – 1)	0.85	0.12

Mandarin Dominance (0 – 1)	0.61	0.17
L1 AoA (years)	3.05	1.54
L2 AoA (years)	4.08	1.87
L1 Proficiency (0 – 1)	0.86	0.13
L2 Proficiency (0 – 1)	0.80	0.15
L1 Immersion (0 – 1)	0.87	0.09
L2 Immersion (0 – 1)	0.80	0.12
L1 Dominance (0 – 1)	0.67	0.18
L2 Dominance (0 – 1)	0.59	0.16
L1S (1 – 15)	9.82	1.88
L2S (1 – 15)	9.37	1.58
CS (1 – 15)	10.73	2.54
US (1 – 15)	8.18	1.76
MLD Score (0 – 2)	1.49	0.32
SES (1 – 10)	5.95	1.26
Adapted Raven’s SPM (1 – 30)	24.19	3.38
LexTALE (100%)	82.22	10.12
LexTALE-CH (100%)	77.08	9.23

Note: $N = 62$; L1 and L2 languages differ across participants, L1: 43 Mandarin, 18 English, 1 Cantonese; L2: 18 Mandarin, 40 English, 1 Korean, 2 Cantonese, 1 Malay. AoA = Age of acquisition, SES = Social Economic Status, MLD = Multilingual Language Diversity score

5.3.4 Procedure

Participants answered the MacArthur’s SSS Scale, Adapted Raven’s SPM, LexTALE (EN & CH versions), LHQ 3.0, and the MSWQ online via Qualtrics before completing a Modified ANT in the lab. Participants rested their heads on a chin and forehead support to stabilize their position, 78 cm away from the monitor, and completed a 12-trial practice of the Attention Network Test (ANT). Before attempting the experimental blocks, an EyeLink 1000 Plus eye tracker sampling at 1000 Hz was set up to record gaze data to an accuracy of ± 0.5 degrees of visual angle. Calibration and validation procedures were conducted before the start of every new experimental block. The experimental stimuli were presented using SR Research’s Experiment

Builder (v2.3.38). The order of experimental blocks (NoLanguage, English, Mandarin, and Mixed blocks) was fully randomized between participants. If the central arrow pointed to the left, participants were instructed to press “F” using their left index finger. Similarly, if the central arrow is pointed right, participants must press “J” using their right index finger. An untimed break (5 minutes on average) was provided between experimental blocks.

5.4 Results

The results will be first broken down into the behavioural data (RT & Accuracy), the attentional network data, and eye-tracking using ANOVAs to investigate the modulatory effects of language contexts on task performance⁷. No specific hypothesis was made for Orienting, but a simple ANOVA was conducted to verify this point⁸.

First, two one-sample t-tests were conducted to check whether non-verbal IQ (Raven’s SPM) and SES were significantly different from the sample’s mean. Results indicate that neither SPM ($t(61) = 0.01, p = .993$) nor SES ($t(61) = 0.01, p = .992$) significantly differed from the sample’s mean (see Table 12). As such, neither of these variables were included as covariates in the analysis.

Then, a multiple linear regression analysis was conducted to investigate whether individual differences in multilingual experiences can predict RT or attentional network changes. The multiple regression analysis focuses on the modulatory effects of individual differences in bilingual experiences on task performance.

5.4.1 Behavioural Results – RT & Accuracy

Reaction Time. A three-way within-subjects ANOVA (Analysis of Variance) was conducted with Congruency (Congruent vs. Incongruent), Language (NON vs. ENG vs. MAN vs. MIX), and Cue Type (None vs. Centre vs. Double vs. Spatial). Descriptives can be found in Table

⁷ As the RTs, Fixation durations, and Saccades followed a non-normal distribution, a Friedman’s test (i.e., a non-parametric equivalent to a within-subjects ANOVA) was used to run the analysis. However, because a Friedman’s test is only meant for a one factor within-subjects ANOVA, a separate Friedman’s test was conducted for each of the main factors. The Friedman outcomes was not significantly different than the traditional three-way within-subjects ANOVA. Because running multiple Friedman’s increases type 1 error, and ANOVA is generally robust against violations of normality, the data analysis proceeded with using the traditional three-way within-subjects ANOVA.

⁸ While an Orienting effect was not hypothesized/expected, an ANOVA was nevertheless conducted on the differences between the spatial and no cue data, and as expected, there were no significant differences between the contexts on the indices of Orienting

13. Analysis revealed the typical main effects of Congruency ($F(1,61) = 339.94, p < .001, \eta_p^2 = .848$) and Cue Type ($F(1.88, 114.46) = 243.46, p < .001, \eta_p^2 = .800$) and Cue Type ($F(1.82, 111.21) = 243.36, p < .001, \eta_p^2 = .800$). Congruent trials had significantly faster mean RT compared to incongruent trials ($M_{\text{diff}} = -102.08 \text{ ms}, SE = 5.54, t(61) = -18.44, p < .001, d = -1.45$). Bonferroni-corrected pairwise comparisons on Cue Type showed a significant difference between several pairs of Cue Type levels. There was a significant difference between No Cue and Centre ($M_{\text{diff}} = 17.81 \text{ ms}, SE = 4.12, t(56) = 4.32, d = 0.254, p < .001$), No Cue with Double ($M_{\text{diff}} = 27.59 \text{ ms}, SE = 4.12, t(56) = 6.69, d = 0.393, p < .001$), No Cue with Spatial ($M_{\text{diff}} = 103.18 \text{ ms}, SE = 4.12, t(56) = 25.03, d = 1.470, p < .001$), Centre and Spatial ($M_{\text{diff}} = 85.37 \text{ ms}, SE = 4.12, t(56) = 20.71, d = 1.216, p < .001$), and Double with Spatial ($M_{\text{diff}} = 75.58 \text{ ms}, SE = 4.12, t(56) = 18.33, d = 1.077, p < .001$). However, there was no significant difference between Centre and Double ($M_{\text{diff}} = 9.78, SE = 4.12, t(56) = 2.37, p = .112, d = .139$). The fastest RT was observed for spatial cues ($M = 575.99, SE = 9.70$) followed by double cues ($M = 653.62, SE = 7.62$), center cues ($M = 663.28, SE = 8.08$), and no cues ($M = 681.43, SE = 8.12$).

There was no main effect of Language ($F(3, 183) = 1.56, p = .200, BF_{10} = 0.14$). Curiously, the typically reported interaction effect between Congruency and Cue Type (e.g., MacLeod et al., 2010) was not found ($F(2.54, 155.01) = 0.95, p = .408, BF_{10} = 0.03$). There were also no significant interaction effects between Congruency and Language ($F(2.66, 155.01) = 0.86, p = .450$), Language and Cue Type ($F(7.85, 478.88) = 0.86, p = .550$), or Congruency, Language and Cue Type ($F(9, 549) = 0.52, p = .863$).

Accuracy. The same three-way within-subjects ANOVA was used to analyze the accuracy data. As previous Bayesian analyses yielded no additional insights beyond those already provided by traditional frequentist methods (see Chapter 3), no further Bayesian analyses were conducted for the sake of conciseness. Further descriptives can be found in Table 13. Curiously, there was no typical main effect of Congruency ($F(1, 61) = 0.32, p = .576$) and Cue Type ($F(3, 183) = 0.79, p = .501$). Yet, a significant main effect of Language ($F(1.73, 105.64) = 65.47, p < .001, \eta_p^2 = .518$) was observed. However, this is likely due to the cueing effect masking the effect of Congruency (Timmer et al., 2021). To confirm this possibility, the analysis was rerun using a two-way within-subjects ANOVA with Congruency and Language. This analysis confirmed that when the results partials out the cueing effect, there was a significant main effect of Congruency ($F(1, 61) = 96.64, p < .001, \eta_p^2 = .613$), where congruent trials had better accuracy than incongruent trials ($M_{\text{diff}} =$

4.65, $SE = 0.47$, $t(61) = 9.83$, $p < .001$, $d = 1.247$), but the main effect of Language disappeared ($F(3, 183) = 0.16$, $p = .925$).

Table 13 Descriptives of RT and Accuracy (with Standard Deviations) separated by Language, Congruency, and Cue Type

		No Cue		Centre Cue		Double Cue		Spatial Cue	
		CON	INCON	CON	INCON	CON	INCON	CON	INCON
RT (ms)	NON	636.55	732.12	621.06	719.66	607.37	707.76	527.43	632.77
		(57.02)	(88.38)	(75.05)	(77.74)	(58.79)	(70.98)	(72.33)	(90.38)
	ENG	630.95	720.96	611.37	712.19	594.33	700.37	521.30	624.88
		(53.04)	(82.19)	(49.11)	(80.37)	(44.56)	(80.73)	(57.59)	(76.74)
	MAN	623.19	728.91	605.40	708.45	594.43	701.06	520.42	628.79
		(48.97)	(74.72)	(48.19)	(73.61)	(45.60)	(81.82)	(79.11)	(90.82)
	MIX	631.57	733.21	606.90	709.97	606.14	705.27	525.62	630.85
		(48.93)	(75.05)	(40.85)	(80.31)	(44.57)	(79.59)	(61.20)	(91.51)
Accuracy (%)	NON	99.60	99.87	99.87	99.60	99.19	99.87	99.33	99.46
		(1.80)	(1.06)	(1.06)	(1.80)	(2.91)	(1.06)	(2.29)	(1.06)
	ENG	95.03	93.15	95.16	94.62	96.51	95.83	93.82	95.03
		(7.51)	(8.61)	(6.67)	(7.86)	(5.95)	(6.36)	(8.66)	(7.91)
	MAN	99.87	99.60	99.19	99.06	99.46	99.60	99.60	99.33
		(1.06)	(1.80)	(2.91)	(3.06)	(2.56)	(1.80)	(2.35)	(2.74)
	MIX	94.09	94.76	94.09	95.16	94.89	94.62	96.51	94.76
		(9.35)	(7.73)	(7.60)	(7.78)	(7.75)	(7.71)	(6.67)	(7.88)

Note: 384 total trials per participant across 4 experimental blocks. Each block contains 96 trials with 48 congruent and 48 incongruent trials. There are 12 trials of each of the four cue types. CON = Congruent trials, INCON = Incongruent trials, NON = NoLanguage block, ENG = English block, MAN = Mandarin block, MIX = Mixed block

5.4.2 Behavioural Results - the Attentional Networks

RT_Alerting. Based on RT data, an analysis of Cue Type and Language revealed a significant main effect only for Cue Type ($F(1, 61) = 92.01$, $p < .001$, $\eta_p^2 = .601$). Bonferroni-corrected pairwise comparisons show that no-cue trials had slower mean RTs than double-cue trials ($M_{\text{diff}} = 27.61$ ms, $SE = 2.88$, $t(61) = 9.59$, $p < .001$, $d = 0.469$). There was no main effect of Language ($F(2.82, 171.97) = 2.08$, $p = .108$), nor an interaction of Cue Type and Language ($F(2.81, 171.11) = 0.02$, $p = .994$).

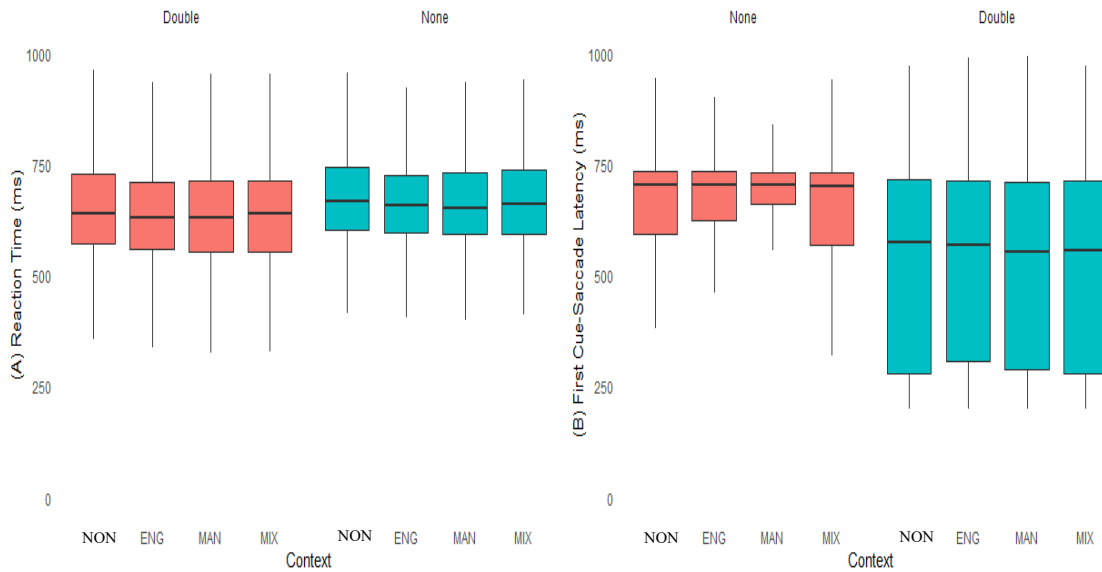
RT_Executive Control. Based on RT data, an analysis of Congruency and Language revealed a significant main effect only for Congruency ($F(1, 61) = 336.27$, $p < .001$, $\eta_p^2 = .846$).

Bonferroni-corrected pairwise comparisons show that congruent trials had faster mean RTs than incongruent trials ($M_{\text{diff}} = -101.79$ ms, $SE = 5.55$, $t(61) = -18.34$, $p < .001$, $d = -1.63$). There was no main effect of Language ($F(3, 183) = 1.55$, $p = .204$), nor an interaction of Congruency and Language ($F(2.62, 159.97) = 0.93$, $p = .419$).

5.4.3 Eye-Tracking Results – the Attentional Networks

Sacc_Alerting. Based on the first cue-saccade latency data (see Figure 12), Cue Type and Language revealed a significant main effect for Cue Type ($F(1, 61) = 209.80$, $p < .001$, $\eta_p^2 = .775$). Double-cue trials had faster mean saccade latencies than no-cue trials ($M_{\text{diff}} = 130.03$ ms, $SE = 8.98$, $t(61) = 14.49$, $p < .001$, $d = 1.37$). There was no significant main effect of Language ($F(3, 183) = 0.62$, $p = .600$, $\eta_p^2 = .010$). The interaction of Cue Type and Language was also not significant ($F(3, 187.86) = 0.40$, $p = .752$).

Figure 12: Boxplot of No-Cue and Double-Cue RT (ms) and First Cue-Saccade Latency (ms) across Language.

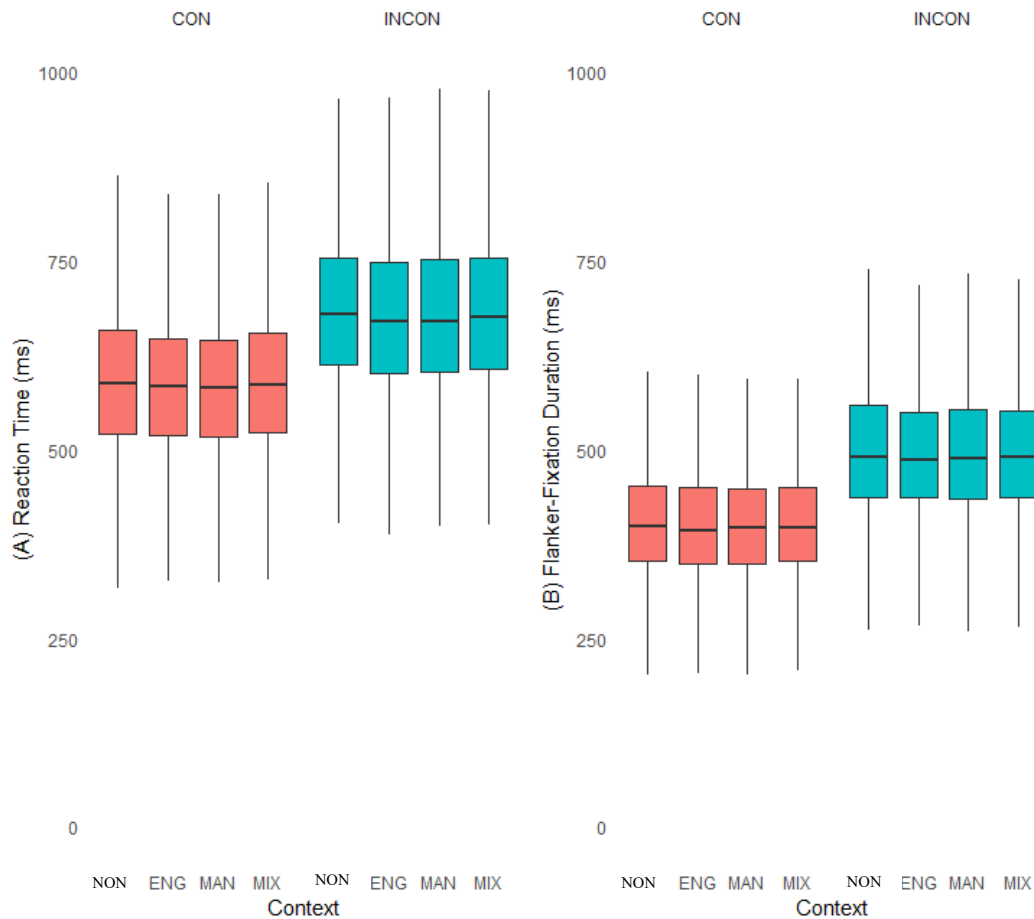


Note: (A) Boxplot of RT across Language contexts. (B) Boxplot of First Cue-Saccade Latency across Language contexts.

Fix_Executive Control. Based on flanker-fixation duration data (see Figure 13), there was a significant main effect only for Congruency ($F(1,61) = 316.95$, $p < .001$, $\eta_p^2 = .839$). Bonferroni-

corrected pairwise comparisons show that Congruent trials had shorter mean fixation durations than incongruent trials ($M_{\text{diff}} = -101.79$ ms, $SE = 5.63$, $t(61) = -17.80$, $p < .001$, $d = -1.759$). There was no main effect of Language ($F(3, 183) = 0.96$, $p = .415$), nor an interaction of Congruency and Language ($F(2.67, 152.85) = 0.09$, $p = .956$).

Figure 13: Boxplot of Congruent and Incongruent trials in (A) RT (ms) and (B) Flanker-Fixation Duration (ms) across Language



Note: (A) Boxplot of RT across Language contexts. (B) Boxplot of Flanker-Fixation Duration across Language contexts.

5.4.4 Regression analysis

The present regression includes the following variables: MLD, L2 AoA, L2 Proficiency, L2 Immersion, L2 Dominance, English AoA, English Proficiency, English Immersion, English

Dominance, Mandarin AoA, Mandarin Proficiency, Mandarin Immersion, and Mandarin Dominance. L1S, L2S, CS, and US variables measured by the MSWQ were not found to correlate, nor predict any outcomes. For brevity, these variables were not reported. L3 and L4 variables were also omitted as not all participants were reported to be multilingual. This study examined RT_Alert, Sacc_Alert, RT_Executive control, and Fix_Executive control across the four language contexts: NoLanguage, English, Mandarin, and Mixed. Separate models were created for each combination of language context and outcome measure.

The analysis did not identify any noteworthy variables that could effectively and reliably predict the outcomes under investigation. The following models had VIF below 1, suggesting no issues with multicollinearity. In summary, L2 Immersion and Mandarin proficiency predicted English block RT_Alerting, explaining 14.4% of the outcome ($F(2, 59) = 6.15, p = .004, R^2_{\text{adj}} = .14$). As a predictor, a unit increase in L2 Immersion resulted in a 72.49-ms increase in the English block RT_Alerting effect ($\beta = 72.49, SE = 34.08, t = 2.13, p = .038$). A unit increase in Mandarin proficiency results in a 55.07-ms increase in the English block RT_Alerting effect ($\beta = 55.07, SE = 25.90, t = 2.13, p = .038$).

In another model, only Mandarin proficiency predicted Mandarin RT_Alerting, explaining 9.4% of the outcome ($F(1, 60) = 7.32, p = .009, R^2_{\text{adj}} = .09$). As a predictor, a unit increase in Mandarin proficiency results in a 98.64-ms increase in the Mandarin block RT_Alerting effect ($\beta = 98.64, SE = 36.45, t = 2.71, p = .009$). No other outcomes in any other language blocks successfully converged into a model.

5.5 Discussion

The aims of the present study were threefold: 1) to investigate whether language contexts can rapidly modulate attentional networks; 2) to investigate whether MLD scores can positively predict attentional networks above and beyond traditional measures of multilingualism; and 3) to investigate the difference between traditional RT measures and eye-tracking metrics. To explore these areas, English-Mandarin dominant multilingual young adults were recruited to perform a modified ANT while an eye-tracker was employed to measure gaze and saccade metrics. Using an eye tracker enabled gaze verification which ensured that participants were overtly looking at the word, flanker stimuli and cues using central vision rather than relying on peripheral vision. This verification meant that all gaze-related metrics obtained (i.e., fixations and saccades) were

confirmed to be confined within the designated regions of interest (i.e., word, flanker stimuli and cues).

The mixed-language block was expected to exhibit the fastest RT compared to the NoLanguage and single-language blocks. Second, a larger RT_Alerting component and a smaller RT_Executive control component (i.e., smaller interference effect) were expected in the mixed language block compared to the NoLanguage and single-language blocks. Third, a smaller Sacc_Alerting and Fix_Executive control component was expected in the mixed-language block compared to the NoLanguage and single-language blocks. Lastly, participants' MLD score was expected to be the best predictor of RT_Alerting, RT_Executive control, Sacc_Alerting, and Fix_Executive control components across all the language contexts above and beyond the other traditional measures of multilingualism. No specific hypotheses for the orienting component were made, as past studies typically report a null effect of multilingualism on orienting abilities (Costa et al., 2008; Marzecová et al., 2012; Tao et al., 2011).⁹

This is the first study to investigate the effects of language contexts on ANT task performance using eye-tracking methodology. In summary, for the RT data, other than the expected main effects of congruency and cue types, language contexts did not affect RT performance either by itself or through interactions with congruency and cue types. In the analysis of eye-tracking metrics, no significant main or interaction effects with the language contexts were found in either the flanker-fixation duration data or the first cue-saccade latency was observed.

As for the regression analysis, although some significant models were found in predicting various attentional components, different predictors lead to different outcomes, without any clear patterns making it difficult to interpret. L2 Immersion and Mandarin Proficiency predicted RT_Alerting in the English block but had no influence elsewhere. Similarly, Mandarin Immersion predicted RT_Alerting in the Mandarin block but not in any other contexts. If the linear relationship between predictors and outcomes were genuinely influenced by language contexts, we should have seen a clear and consistent prediction pattern, rather than these one-off models. Based on the ACH, mixed-language contexts should have shown a stronger predictive relationship than single-language contexts. Furthermore, the use of step-wise regression prevents any definitive

⁹ ANOVA analysis of Orienting measures confirmed that there were no significant differences between language contexts. For brevity, only results of Alerting and Executive Control were reported to better align with the study's hypotheses

conclusions from being drawn. Hence, despite the convergence of some models, it is unlikely that the measured multilingual variables are associated with any improvements in attentional networks.

5.5.1 Language Contexts and Cognitive Control

There were no main effects of Language contexts on RT, accuracy, fixation times, or first-cue saccade latencies. This could be because of the strong facilitatory effect of the cues. Given that these cues always correctly predict the presentation timing and location of the flanker stimuli, participants may have relied exclusively on the cues to optimize their performance on the ANT. Therefore, with the cues providing such strong facilitatory effects, there was no space for any cognitive adaptation to arise through the language contexts.

Past studies using ERP methodology provide compelling evidence that the effects of language contexts on ANT performance are likely observable only with highly sensitive measures. For example, Timmer et al. (2021) investigated the lateralized readiness potential (LRP), an ERP that reflects the preparation and execution of motor responses (Eimer, 1998), within the ANT. The LRP is commonly observed in experiments with left/right response requirements and provides information about the pre-motor activation of the correct or incorrect hand. Results indicated that incongruent, but not congruent, trials of the correct hand exhibited faster activation in mixed-language blocks compared to single-language blocks, as indicated by larger negative LRP amplitudes. However, despite this faster activation, there was no observable difference in RT between mixed and single-language blocks. This finding is interesting because it implies that the correct response is evaluated earlier in the cognitive processing pipeline when participants were primed in a mixed-language compared to a single-language context. Except, this speeded evaluation did not translate to observable RT differences in ANT performance. Similarly, Nayak et al. (2020) discovered that multilingual children exhibited faster activation of the correct hand on incongruent trials in a non-verbal Stroop task compared to monolingual children, but like Timmer et al. (2021), this speeded activation did not translate to improvements in RT in the Stroop task.

Research suggests that bilingualism is unlikely to have increased the amount of cognitive resources, instead it likely impacts the efficiency of resource management, where fewer resources are used to achieve baseline outcomes (Calvo et al., 2023). This interpretation is grounded in the concept that automatic processing, which occurs early in the cognitive processing pipeline, is

autonomous and unconscious, and requires much less effort compared to controlled processing. Controlled processing, on the other hand, occurs later in the cognitive processing pipeline, involving the deliberate retrieval of information from memory, and is conscious and slow (Fabio et al., 2019). Results of both Timmer et al. (2021) and Nayak et al. (2020) could indicate that although bilingualism and bilingual contexts may have improved the automaticity of response processing on the ANT and Stroop task compared to monolingualism and monolingual contexts, no discernable behavioural differences were observed.

It is also important to address that the closer an arbitrary alerting cue (i.e., words, cues) is to the target (i.e., smaller CTI), the stronger the alerting effect (Timmer et al., 2021). In the current study, the priming words and asterisks (cues) can both be considered as alerting cues as they provide consistent and reliable temporal information to the upcoming presentation of the flanker stimuli. As the asterisks had a smaller CTI of 500 ms compared to the CTI of the priming words (ranging from 1200 to 1800 ms) to the flanker stimuli, the facilitatory effects of the asterisks likely overshadowed those of the language context priming words. Therefore, participants may have exclusively relied on the facilitatory effects of the cues to enhance task performance, rather than an actual cognitive adaptation related to language contexts.

Fixation duration was used here because it has been traditionally used as an indicator of perceptual and cognitive load (Liu et al., 2022). However, such a metric would only be viable with a scene presenting multiple objects of interest. As each trial in the ANT only ever projects one or two objects at a time, either a word, a fixation cross, cues, or the flankers, using fixation duration to index cognitive load is less appropriate. In addition, each scene (i.e., trials) terminates upon receiving a response or when 1500 ms lapses. As the flanker fixation duration is measured from the initial presentation of the flanker stimuli until the termination of the trial, this metric is non-distinguishable from behavioural RT. For a more accurate assessment of changes in cognitive load between language contexts, physiological measures like pupil dilation and ERPs could serve as more valid and reliable indicators (Kuipers & Thierry, 2013).

Could the use of single-word priming, both in the current experiment and in Chapter 4, potentially confound the results? Although there is no clear evidence to support whether priming using single words is any different than priming using dialogues or in different modalities (e.g., listening to conversations or spoken responses), related research indicates that this is an important factor to consider. For example, Mandarin monolinguals, Hakka-Mandarin bilinguals, and Hakka-

Mandarin-Minnan trilinguals completed two versions of a read-aloud task (Hsu, 2014). In the first version, the reading materials were presented on a computer screen, one word at a time, whereas in the second version, the materials were presented on an A4 paper. Results indicated that across all the measured outcomes, there was a clear bilingual and trilingual effect over the monolinguals during the first version of the read-aloud task. However, in the second version, trilinguals demonstrated better attentional control than bilinguals nor monolinguals.

Additionally, in Kałamała et al. (2022), authors primed different language contexts through natural conversations while EEG was being recorded. Here, participants were always tasked to play the role of a host in a map task (Berriós et al., 2023) with two bilingual confederates acting as players. In the control condition, the confederates were instructed to use only Polish (i.e., the participants' native language). In the second condition, the confederates were instructed to use only English (i.e., the participants' second language) to create a single-language environment. In the third condition, one confederate used Polish and the other English to create a dual-language environment. Results indicated that although there were ERP differences between the conditions, there were no behavioural differences.

Hsu (2014) indirectly demonstrated that the task requirements (i.e., reading one word at a time vs self-paced reading of a text) may affect the exhibition of a bilingual effect, whereas Kałamała et al. (2022) demonstrated that the use of natural conversations in priming language contexts could affect the neural mechanisms of the bilingual effect. Therefore, it is reasonable to conclude that the use of single words to prime language contexts in the current study may have been insufficient to elicit meaningful cognitive changes.

In the present study, MLD scores were predicted to be associated with better performance in both Alerting and Executive control components, as assessed by RT, flanker-fixation duration, and first cue-saccade latency. However, the results demonstrated that L2 immersion and Mandarin proficiency predicted English block RT_Alert and Mandarin immersion predicted Mandarin block RT_Alert. As a stepwise methodology was used, these results are not driven by theory, but by arbitrary statistical values. Combined with the fact that there does not seem to be any consistent patterns in the results, it is unlikely that this result is due to any genuine associations between multilingual experiences and attentional control. Furthermore, given that there was near-ceiling performance, as seen by the high accuracy scores across language contexts on both congruent and incongruent trials (see Table 13), combined with the facilitative effects of the alerting cue, it may

be speculated that the cognitive demand of the current ANT is not intensive enough to allow for meaningful differences to arise as a result of variations in multilingual experiences.

Borrowing an excellent analogy from Bialystok and Craik (2022), if a researcher would aim to compare the fitness levels between runners and non-runners by assessing their walking, jogging, or running performance, larger group differences are expected to only emerge during higher aerobic demands like jogging or running. The absence of group differences within walking conditions may not provide significant diagnostic information. In the current study, the near-ceiling ANT performance is analogous to the walking condition. Should the task demands be greater, as evidenced by greater variance in RT or accuracy, the effects of multilingual experience may be observed more clearly (Bialystok, 2009; Bialystok & Craik, 2022). For example, Janic et al. (2020) specifically investigated the effects of task difficulty on performance in young adult monolinguals and bilinguals using the Multiple Object Tracking (MOT) task. Participants were required to complete the MOT task twice. During the first instance, they performed the MOT task without any distractions. During the second instance, participants had to concurrently engage in a distractor task, which involved counting backward in 1s from a randomly given number. Results indicated that performance was only significantly different between groups during the task with distraction instance, whereby bilinguals' performance decreased by 15%, whereas monolinguals' performance decreased by three times as much (see also Comishen & Bialystok, 2021). This finding suggests that the presence of a bilingual effect may be influenced by task difficulty. It is possible that the lack of observed language context differences in this study was due to the task being too simple. Introducing shorter ITIs or adding a concurrent distractor task during the ANT could increase cognitive load. This increase in difficulty may result in higher task performance for individuals with greater multilingual experience, compared to those with less multilingual experience.

Although this narrative is a compelling explanation for the null results, insights from Costa et al. (2009) provide an alternative explanation. First, the current study ANT follows the "high monitoring requirements" of Costa et al. (2009). In their experiments, authors demonstrated that bilinguals had significantly faster mean RT on the ANT compared to monolinguals when task demands were high (i.e., 50% - 75% incongruent trials) than when task demands were low (i.e., 8% incongruent trials). This difference means that the task demands of the current ANT should have been sufficient to elicit a bilingual effect. Second, the current study's task performance, in

terms of mean RT and accuracy, was comparable to the bilingual sample's performance in Costa et al. This comparison means that any bilingual effects found in Costa et al. were similarly observed in the current study, but they can only be seen when they are directly compared to a monolingual group¹⁰. Therefore, although a ceiling effect might contribute to some level of noise, these two key observations suggest that bilingual effects may still be found despite ceiling effects in the form of speeded RT.

Alternatively, to sidestep the ceiling effect issue, a significantly harder task could be used. For example, the N-Back Task (Owen et al., 2005), a highly taxing task of working memory, requires participants to memorize a sequence of digits presented one at a time. Participants are required to correctly identify if the on-screen digit matches the digit presented N positions ago. It is recommended to use either the 2-back or 3-back versions of the task, as demonstrated in studies by Barker and Bialystok (2019), Morrison et al. (2018), and Comishen and Bialystok (2021). These versions have been shown to produce results that do not reach ceiling performance. If the effect of multilingualism becomes observable only when cognitive resources are exhausted, an effect should only be seen under cognitively strenuous tasks, which may provide novel findings on the effect of multilingualism on cognition, given the over-saturation of research utilizing Simon, Flanker, and Stroop tasks (Paap et al., 2020).

In Section 5.3, an a-priori analysis suggested that 77 participants were required to achieve 80% power to detect an effect. Since only 62 participants' data was included in the final analysis, the study is likely underpowered. However, the ceiling performance of the current sample suggests that the addition of more data sets into the analysis may not significantly change the study's outcomes. The ceiling performance also suggests that the ANT may be too simple and thus, not suitable to investigate the bilingual effect hypothesis. A much more difficult task, like the N-Back task discussed above, may be more appropriate.

In the next study, a dialogue reading task is introduced to prime the different language contexts, potentially creating a stronger language context priming effect. Although Chapter 4 used Pinyin and Chapter 5 used Mandarin characters, a mixed three-way ANOVA comparing the behavioural results of Experiment 2A and Experiment 3 demonstrated no significant difference in using either Pinyin or characters. Therefore, the next study will use Mandarin characters in the

¹⁰ The issues of using a monolingual control group and the dichotomic categorization of bilingualism were outlined in Section 2.4

dialogue priming so that it better aligns with the actual reading preferences of Mandarin native speakers. Furthermore, the 2-back task is used, as studies have demonstrated that it is more difficult than the 1-back task but not as difficult as the 3-back task (Comishen & Bialystok, 2021; Morrison et al., 2018).

Chapter 6: Can Language Entropy Predict ERP Differences in Multilinguals During a Working Memory Task?

6.1 Introduction

WM is a cognitive ability that enables the manipulation of memory to facilitate, plan, and direct behaviours related to comprehension, reasoning, and problem-solving (Cowan, 2014). The relationships between WM and language and attention make WM a prime candidate to investigate the bilingual effect.

Jackson (2020) describes a mutually beneficial relationship between WM and language: WM facilitates language learning early in life, while language abilities sustain WM over the years. Additionally, WM is intimately linked to attention (Oberauer, 2019) as strong attentional control enables individuals to effectively exclude irrelevant content from WM (Hasher & Zacks, 1988). This ability is done by either filtering them out during the encoding process (Vogel et al., 2005) or by discarding them once they are no longer needed (Oberauer et al., 2012). Studies have found that bilingualism was associated with improvements in WM (Antón et al., 2019; Blom et al., 2014; Grundy & Timmer, 2016; Liu & Liu, 2021; Lukasik et al., 2018; Morales et al., 2013; Yang, 2017; Monnier et al., 2021; Monnier et al., 2021), whereas other studies have found no such improvements (Calvo et al., 2016; Lehtonen et al., 2018; Namazi & Thordardottir, 2010; Ratiu & Azuma, 2014).

According to Lukasik (2018), the majority of studies in this area primarily employ simple span tasks (i.e., tasks that use short-term memory rather than WM). Although these tasks are useful in their own rights to investigate short-term memory abilities, the innate simplicity of these tasks means that any effects arising from bilingual experiences would not be used, thereby leading to null differences between groups. This conclusion is not surprising, given that the brain is an incredibly expensive metabolic organ relative to the size of the body. Therefore, how the brain chooses to organize itself can be simplified to a drive to minimize energy costs and maximize efficiency (Bullmore & Sporns, 2012; Marsman et al., 2017).

If there is an effect in WM in multilinguals, it may not be observable in just any type of task (Monnier et al., 2021). Rather, a certain level of cognitive load must be imposed for the brain to recognize the need for the cognitive adaptations afforded by multilingualism. Task difficulty

plays an important role in determining whether the bilingual effect will be observed or not (Antón et al., 2019; Bialystok & Craik, 2022; Jiao et al., 2017). Simple span tasks lack discriminant validity, as the nature of their simplistic response could remove the need for any of the proposed bilingual adaptations to be recruited and used. As the cognitive load of simple-span tasks is low (Mathy et al., 2018), both monolinguals and bilinguals were able to maintain existing task goals without tapping into WM resources.

A widely used measure of WM capacity is the N-back task (Owen et al., 2005). As briefly described in Section 5.5.1, the N-back task typically requires participants to hold a string of digits in their memory to reference whether a displayed number matches the number presented N positions ago. Common versions are the 2- and 3-back tasks, whereas the 0- and 1-back tasks typically serve as control. Although this task has been championed as a “pure” measure of WM, investigations into its construct validity suggest otherwise (Jaeggi et al., 2010). For example, other notable tasks of WM include the reading span task (RST) or an operation span (Ospan) task. Jaeggi et al. (2010) found weak intercorrelations between the N-back tasks and the RST/Ospan task, ranging between $r = .10$ and $r = .24$. These small associations could mean one of two things. First, the N-back task may not be as pure as a measure of WM as once thought, suggesting that it may be inadequate to measure individual differences in WM. Second, and more likely, is that the N-back, RST, and Ospan tasks are all valid measures of WM in their own regard, except that they measure different facets of WM (see the Unity and Diversity model by Miyake et al., 2000). Kane et al. (2007) suggested that the N-back task may rely more on recognition abilities, whereas both the RST and Ospan may require more active recall of working memory (Jaeggi et al., 2010). Nevertheless, Jaeggi et al. (2010) concluded that the N-Back task is useful for experimental research in WM as it can predict individual differences in other higher cognitive functions, especially at higher levels of N (i.e., 2 or 3-back).

Further support for the use of the N-back task comes from EEG studies. Scharinger et al. (2017) compared the WM-load between simple (digit span task) and complex span tasks (N-back and operation span tasks). They hypothesized that increasing WM load would generally result in decreased alpha power and increased theta band power. Their results indeed indicated that complex span tasks showed decreased alpha power compared to simple span tasks, suggesting that complex span tasks require greater WM than simple span tasks. The N-back task demonstrated the highest theta band power, suggesting specific demands on cognitive control that are not present in the

simple span and Ospan tasks. It also measured the P3 amplitude across the three tasks. The P3 is a positive-going ERP component that is maximal over centroparietal regions and occurs approximately 300–500 ms post-stimulus. Traditionally, its amplitude reflects the intensity of cognitive processing (Kok, 2001) and resource allocation (Polich, 1996), while its latency reflects the speed of cognitive processing (Polich, 2007). The sparse research on the P3 component on WM load seems to suggest a significant association between its amplitude and task difficulty, however, its latency does not seem to be modulated by task difficulty (Watter et al., 2001; Scharinger et al., 2017).

In Scharinger and colleagues's (2017) experiment, it was observed that P3 amplitude decreased as WM load increased across the three WM tasks. Specifically, the N-back task exhibited the most significant decrease in P3 amplitude, with the 0-back task showing the highest P3 amplitude compared to the 3-back task. The negative correlation observed between P3 amplitude and WM load is likely attributable to the diminishing availability of WM resources necessary to complete increasingly difficult WM tasks (Kok, 2001; Scharinger et al., 2017). A decrease in P3 amplitudes reflects the reallocation of resources from simple stimulus matching to more executive processing of rising memory demands (Watter et al., 2001). This association was also found between high-scoring and low-scoring individuals on the 2-back task, with high-scoring individuals generally having higher P3 amplitudes suggesting increased efficiency in WM resource allocation (Daffner et al., 2011).

In addition to the P3 component, studies have also investigated the N2 component, typically referred together as the N2/P3 complex as both components are intricately involved in EF (Warchoř & Zajac-Lamparska, 2023). This ERP complex has been frequently examined in the bilingualism-cognitive literature and several studies found significant differences in the N2/P3 complex between bilinguals and monolinguals (Jiao et al., 2020; Jiao et al., 2021; Kařamařa et al., 2022; Timmer et al., 2021; Wu & Thierry, 2013). The N2 is a negative-going component that is maximal over centro-frontal regions which typically occurs between 200 to 350 milliseconds after stimulus presentation (Folstein & Van Petten, 2008; Luck, 2014). Its amplitudes traditionally indicate the ability to discriminate between the presented stimulus and one held in WM (Bennys et al., 2007; Daffner et al., 2011; Patel & Azzam, 2005). On the other hand, its latency traditionally reflects the speed of stimulus evaluation and categorization (Howe, 2014; Lee et al., 2020).

Importantly, increasing task difficulty has been found to reduce N2 amplitude (Comishen & Bialystok, 2021; Morrison et al., 2018) and prolong its latency (Morrison et al., 2019).

The N2 and P3 components have also been traditionally investigated in studies using the Flanker task (Bringas et al., 2020). The P3 is also relevant to the ANT (Galvao-Carmona et al., 2014), although because of the attentional nature of the task, both early N1 and P1 components are also investigated. Therefore, to a certain degree, the N-back task functions on similar underlying neural mechanisms responsible for EF as the Flanker task and ANT, establishing task-convergent validity between Experiment 2 (Chapter 4), Experiment 3 (Chapter 5) and Experiment 4 (the current chapter).

Lastly, another frequently investigated ERP component related to WM is the P2, which is a positive-going component that is maximal over centro-frontal regions, typically occurring 150–275 ms following stimulus presentation (Luck, 2014). It is thought to be associated with WM encoding (Lijffijt et al., 2009; O’Connell et al., 2011), where larger amplitudes traditionally indicate larger WM capacity in individuals (Bourisly & Shuaib, 2018) and shorter latencies indicate faster stimulus classification (Breznitz & Meyler, 2003).

The work by Comishen and Bialystok (2021) and Morrison et al. (2018) systematically investigated the influence of bilingualism on N-Back task performance using an EEG paradigm. Morrison et al. (2018) recorded the brain activity of young adults (23 English monolinguals and 21 English-French bilinguals) while they performed the 0-back, 1-back, and 2-back digit tasks. Both accuracy and RT data did not differ significantly between groups. However, monolinguals exhibited smaller P3 amplitudes compared to bilinguals, suggesting less available cognitive resources to complete the n-back tasks. No differences were found for the N2 and P2 components. On the other hand, Comishen and Bialystok (2021) sampled 33 English monolinguals and 31 English-Non-English bilinguals while they performed a 0-back (symbols), 1-back (letter), 2-back (shape), and 3-back (digit). Accuracy and RT performance decreased more drastically for monolinguals than for bilinguals. P3 latencies were faster for bilinguals compared to monolinguals. No group differences were found for the N2 and P2 components.

6.2 Current Study

The aims of the current experiment are threefold: 1) to investigate whether rapidly changing language contexts can modulate WM in multilinguals, 2) to determine whether changes in

language contexts or differences in multilingual experiences are associated with changes in the N2, P2, and P3 ERP components of WM, and 3) to examine whether multilingual experiences can predict WM. To measure WM, the current study employed the N-back task, specifically, the 2-back version used by Morrison et al. (2018), because of the 2-back's significant increase in WM load compared to the 0-back and 1-back versions. The 0-back and 1-back accuracies were close to ceiling at 95% and 90% mean accuracy respectively, in contrast to the 2-back's mean accuracy of 67% (Comishen & Bialystok, 2021; Morrison et al., 2018). The task involves presenting a series of numbers one at a time. Participants are required to identify and match the current number displayed on-screen with the number in working memory that was presented 2 trials ago.

Importantly, to address whether rapidly changing language contexts can modulate WM and its associated ERPs, this task was modified to include an interleaved language priming task that would prime participants to be in a no-language, single-language, or mixed-language context. The purpose of this modification was to test the parameters of the ACH (Green & Abutalebi, 2013). Until now, no research has explored the ACH using an N-Back task. Using single words rather than sentences for language priming may have led to smaller cognitive adaptation effects which could explain the patterns found in the previous experimental chapters. Neurobiologically, single words activate limited orthographic-semantic networks, whereas sentences engage broader networks, including the dorsolateral prefrontal cortex, requiring more cognitive resources (Aboud et al., 2016). For example, where the current experiments did not strongly support the rapid modulation of EF through single-word language context priming, Adler et al. (2020) achieved this using sentence priming. In their study, participants perform a modified Flanker task interleaved with code-switched or non-code-switched (i.e. monolingual) sentences. Their results found that participants had significantly faster mean RTs in incongruent code-switched conditions compared to non-code-switched conditions. Unfortunately, no effect size was reported in their study, making it difficult to ascertain the exact effect size difference between sentences versus words. This approach also better mimics real-world language use, potentially offering more relevant insights into cognitive processing during language switching.

Following past studies (Barker & Bialystok, 2019; Comishen & Bialystok, 2021; Fernández et al., 2013; 2014; Moreno et al., 2014; Morrison et al., 2018), it is expected that a bilingual effect should only be observable within ERP but not behavioural, outcomes. Given the lack of research on the N-Back task and the ACH, the directions of the outcomes are unclear. Yet,

according to the ACH, the mixed language context should exhibit the smallest N2, largest P2, and largest P3 peak amplitude followed by single and no language contexts. Further, like Experiments 2 and 3, it is expected that higher MLD scores will predict smaller N2, larger P2, and larger P3 amplitudes.

No specific hypothesis was made for the latency domain, because past research has mostly focused on amplitudes. However, given that interpretations for differences in latencies are more straightforward than amplitude (Cespón & Carreiras, 2020), if multilinguals and multilingual contexts enhance EF, then it should be expected that the N2, P2, and P3 latency should be faster in mixed contexts and high MLD participants. A faster ERP latency may be simply interpreted as “faster ERP in group X compared to group Y suggests more efficient processing by group X” (Cespón & Carreiras, 2020).

6.3 Method

6.3.1 Design

The present study used a one-way within-subjects design to investigate the effects of Language (NoLanguage vs. English vs. Mandarin vs. Mixed) on RT and Accuracy.

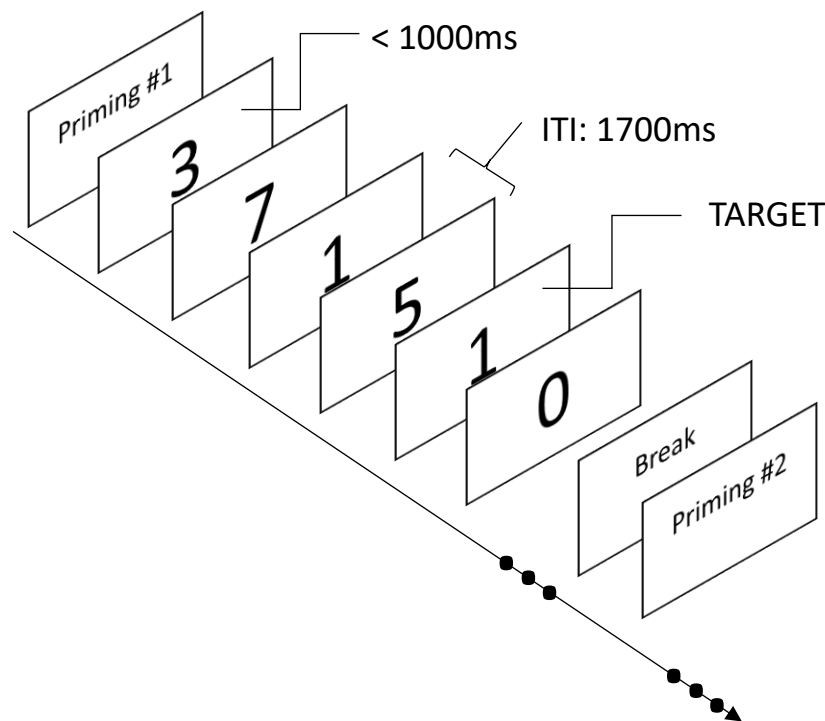
6.3.2 Materials

Language proficiency was calculated using the **LHQ** (Li et al., 2020). See Section 2.3 for more information. SES was measured using **MacArthur’s SSS Scale**, and non-verbal intelligence was measured using a short version of **Raven’s SPM**, see Chapter 2 for further details. The MSWQ was omitted from this study, because no significant findings emerged in Chapter 5. Therefore, it was not expected to yield any interesting results in the present study either.

Modified 2-Back Task. Participants were presented with single digits one at a time at the centre of the computer screen for a maximum of 1000 ms with an ITI of 1700 ms. In this task, participants must determine if the digit on-screen matches the digit displayed two trials previously. Matching digits are labelled as targets, whereas non-matching digits are labelled as non-targets. This 2-back version of the task was chosen over the 0-back or 1-back task due to its steeper cognitive load, and its ability to discriminate between monolinguals and bilinguals (Comishen & Bialystok, 2021). Participants have up to 1000 ms to respond to targets. The first three sequences of digits presented will never be targets. There were four blocks to this experiment, each block

consists of 52 targets and 92 non-targets for a total of 144 trials per block. Each block was split into two equal parts (26 targets and 46 non-targets per split) with a timed 20-second break. These breaks were added to ensure participants did not experience fatigue, which could negatively impact the outcomes of the n-back task (Chen et al., 2021). Language priming tasks were presented at the start of the block and after the timed break, which will be discussed below. After each block, a short, untimed break was provided before the next block started. Each block lasted approximately 10 minutes, for a total of 40 minutes. For an overview of the task, see Figure 14 below.

Figure 14: *Overview of the 2-Back Task*



Language Priming Tasks. The 2-back task was split into four blocks, each primed with different language contexts: NoLanguage, English-only, Mandarin-only, and Mixed (see Appendix 2 for the dialogue priming stimuli). The NoLanguage block, acting as a control block, presents non-verbal IQ questions from Raven’s SPM (Raven, 1941) as a filler task. In the English-only, Mandarin-only, and Mixed blocks, participants were required to read a short conversation in either English-only, Mandarin-only, or an English-Mandarin code-switched language respectively.

These conversations are self-contained vignettes, sourced from the Southeast Asia Mandarin-English (SEAME) corpus (Lyu et al., 2010), which is a large repository of natural and spontaneous English-Mandarin code-switched conversations between two Singaporean adults on a wide range of topics. The conversations were adapted to the language contexts outlined in the ACH, such that for single-language contexts, the conversations were translated to either English-only or Mandarin-only conversations. Three native English-Mandarin Malaysian speakers were hired separately to ensure that these conversations and their respective translations were accurate and culturally appropriate to Malaysia. Notably, the dual-language context was not explicitly primed here because recent evidence suggests that in naturalistic conversations, especially in multilingual countries such as Malaysia and Singapore, dual-language and dense-code-switching contexts are practically indistinguishable from one another (Lai & O'Brien, 2020). As a manipulation check, a simple true/false comprehension question appeared at the bottom of the screen after 15 seconds in each priming section.

Although the effects of reading-based priming in engaging domain-general cognitive control compared to the presentation of single words, as seen in Wu and Thierry (2013), remain ambiguous, previous studies, such as Bosma and Pablos (2020), have shown that reading code-switch sentences engages domain-general cognitive control. The use of reading-based priming might be considered a more natural methodology than single-word priming, potentially reflecting the actual cognitive adaptations resulting from changes in the interactional contexts outlined in the ACH. The decision to use written passages rather than auditory stimuli is based on research by Loh (2021), who replicated Experiment 2A. Rather than written words appearing on the screen, participants listened to single words, but this did not seem to impact the performance as there were no significant differences between language contexts. Furthermore, the mean RT differences between the language contexts were not in the expected directions, with mixed language contexts having RTs faster than Mandarin-only, but slower than English-only.

The language of instructions within each block reflected the current primed language, except for the NoLanguage block, which replaced all texts with X masks, in this case, "Get Ready!" became "XXX XXXXX".

6.3.3 Participants

A total of 44 young adults (28 females) English-Mandarin multilinguals were recruited from the University of Nottingham Malaysia. Past studies in this area suggested sample sizes of around 20, but as the number of trials per condition was half as small in the present study compared to past studies, double the number of participants was deemed appropriate (Barker & Bialystok, 2019; Comishen & Bialystok, 2021; Morrison et al., 2018). Participants were recruited based on the following criteria: 1) Age 18-26 years old, 2) No history of psychiatric or neurological diagnoses, 3) Local Malaysian Chinese, and 4) Proficient and literate in English and Mandarin.

Before analysis, incorrect responses, and outlier trials with response latencies 2.5 SD above or below the participant's mean per condition were removed from the data. In Experiments 2 and 3, participants with less than 80% data integrity ($\frac{\text{Number of Trials Retained}}{\text{Total Possible Number of Trials}}$) were excluded from the analysis. However, the same restriction was not applied to the current study because the N-Back task is significantly more difficult than the ANT or flanker task and would have led to the exclusion of more trials due to errors or missing responses. Therefore, the current study determined that participants with data integrity below 50% would be omitted from the analysis. Notably, three participants fell below this threshold. Two of them had accuracy scores below 10%, indicating that they did not perform the task correctly. The third had an accuracy score of 48%. Given that signal-to-noise ratio improves with larger datasets (Luck, 2014), this participant was retained for the final data analysis. A follow-up analysis confirmed that excluding this 48% accuracy dataset did not significantly alter the results. Additionally, three other participants were removed due to equipment failure. The final sample size for data analysis was 39 participants, with an average data integrity score of 80.02%. Descriptive data of the participants are reported in Table 14 below.

Table 14: *Experiment 3 Participant Demographics*

<i>Demographic Measures</i>	<i>Mean</i>	<i>SD</i>
Age (years)	20.46	1.83
English AoA (years)	3.66	1.83
English Proficiency (0-1)	0.83	0.14
English Immersion (0-1)	0.81	0.15
English Dominance (0-1)	0.79	0.53

Mandarin AoA (years)	3.42	1.60
Mandarin Proficiency (0-1)	0.80	0.17
Mandarin Immersion (0-1)	0.82	0.15
Mandarin Dominance (0-1)	0.60	0.30
L1 AoA (years)	2.79	1.26
L2 AoA (years)	4.40	2.05
L1 Proficiency (0-1)	0.87	0.14
L2 Proficiency (0-1)	0.76	0.15
L1 Immersion (0-1)	0.87	0.12
L2 Immersion (0-1)	0.78	0.15
L1 Dominance (0-1)	0.72	0.36
L2 Dominance (0-1)	0.68	0.58
MLD Score (0-2)	1.58	0.31
SES (0-10)	5.59	1.41
Adapted Raven's SPM (0-30)	24.31	3.50
LexTALE (100 %)	81.96	13.28
LexTALE-CH (100%)	74.70	10.16

Note: N = 39; L1 and L2 languages differ across participants: L1 (English = 13; Mandarin = 24; Hakka = 2), L2 (English = 24, Mandarin = 12; Malay = 3). For ease of comparability, LexTALE and LexTALE-CH scores were standardized to 100%.

6.3.4 Procedure

Participants completed an online questionnaire consisting of the MacArthur Scale of Subjective Social Status (MacArthur SSS Scale; Adler et al., 2000), Adapted Raven's Standard Progressive Matrices (Raven's SPM; Raven, 1941), Lexical Test for Advance Learners of English (LexTALE; Lemhöfer & Broersma, 2012), in addition to its Mandarin version (LexTALE-CH; Chan & Chang, 2018), and the LHQ 3.0 (Li et al., 2020).

In the experimental session (lasting approximately 1.5 hours), participants completed the modified 2-back digit task while EEG was recorded. Participants were instructed to memorize the sequence of numbers presented on-screen, one at a time. When the on-screen number matched, the number presented two numbers ago (i.e., targets), participants were required to press the "G" key

using their right index finger. If the numbers did not match (i.e., non-targets), no response should have been made. Making a response during non-target trials resulted in the trial being listed as “false alarm”.

6.3.5 EEG Preprocessing

EEG data were recorded continuously using 32 active Ag/AgCl electrodes attached to an electrode cap (www.brainproducts.com) placed according to the international 10–20 system at a sampling rate of 512 Hz. FP1 was used to capture blinks. All electrodes were referenced to Cz and grounded using the Common Mode Sense (CMS) electrode. A criterion of 20 k Ω was used for electrode impedance. Electrodes that surpassed this threshold were interpolated.

EEG data was processed offline using Brain Vision Analyser 2.2. The data was first visually inspected for bad electrode sites. Topographic interpolation was applied to sites with high-frequency noises and was down-sampled to 250Hz and re-referenced to the average of both mastoids. An infinite impulse response (IIR) filter was applied, consisting of a low pass filter of 0.01 Hz, a high pass filter of 30 Hz, and a notch at 50 Hz. The filter order was set to 2 to ensure effective suppression of frequencies outside the desired range. All channels were then visually inspected for high levels of noise, these channels were replaced by interpolating the data of surrounding electrode sites (Perrin et al., 1989). Automated rejection criteria were also employed, excluding channels with amplitudes exceeding a threshold of 200 μ V. An Ocular Independent Components Analysis (ICA) was used to identify and correct eye movements and blinks that were independent of the EEG activity. The continuous EEG was manually reconstructed into four segments corresponding to the start and end markers of each language block. Each block segments were further segmented into 1200-ms ERPs that comprised a 200-ms baseline (for baseline correction) and 1000-ms event response window locked to target and non-target stimuli presentation. Finally, ERPs on correct trials were averaged for each condition.

6.4 Results

6.4.1 2-Back Behavioural Performance

Trials, where RT was below 200ms (i.e., false starts) or exceeded the threshold of ± 2.5 SD from the mean by participant and language context, were excluded as outliers, which constituted less than 1.7% of the trials in each language context. RT was analysed using a one-way repeated-

measures ANOVA with Language (NoLanguage, English, Mandarin, and Mixed) as the within-subjects factor. Accuracy was analysed using a separate two-way repeated-measures ANOVA with Language (NoLanguage, English, Mandarin, and Mixed) and Target (Target and Non-Target) as the within-subjects factors. Average RT and Accuracy by Language are shown in Table 15 below.

Table 15: *Descriptives of RT and Accuracy (with Standard Deviations)*

	<i>NoLanguage</i>	<i>English</i>	<i>Mandarin</i>	<i>Mixed</i>
RT (ms)	569.64 (74.72)	563.30 (84.30)	552.67 (75.17)	562.04 (84.51)
Target Accuracy (%)	80.28 (13.34)	78.30 (15.74)	80.38 (13.93)	81.11 (16.38)
Non-Target Accuracy (%)	96.88 (2.81)	97.85 (1.86)	97.38 (1.98)	97.19 (1.86)

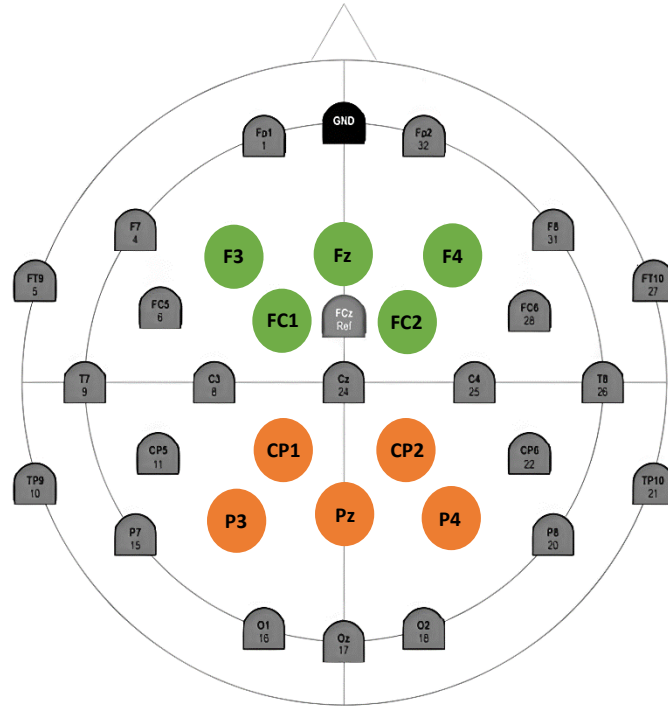
Note: Each language context block contains 52 targets for a total of 208 targets per participant.

The one-way within-subjects ANOVA revealed no significant main effect of Language on RT ($F(3, 114) = 1.02, p = .387, \eta_p^2 = .026, BF_{10} = 0.11$). Similarly, there was no significant main effect of Language on Accuracy ($F(3, 114) = 0.64, p = .592, \eta_p^2 = .017, BF_{10} = 0.13$).

6.4.2 ERP Analyses

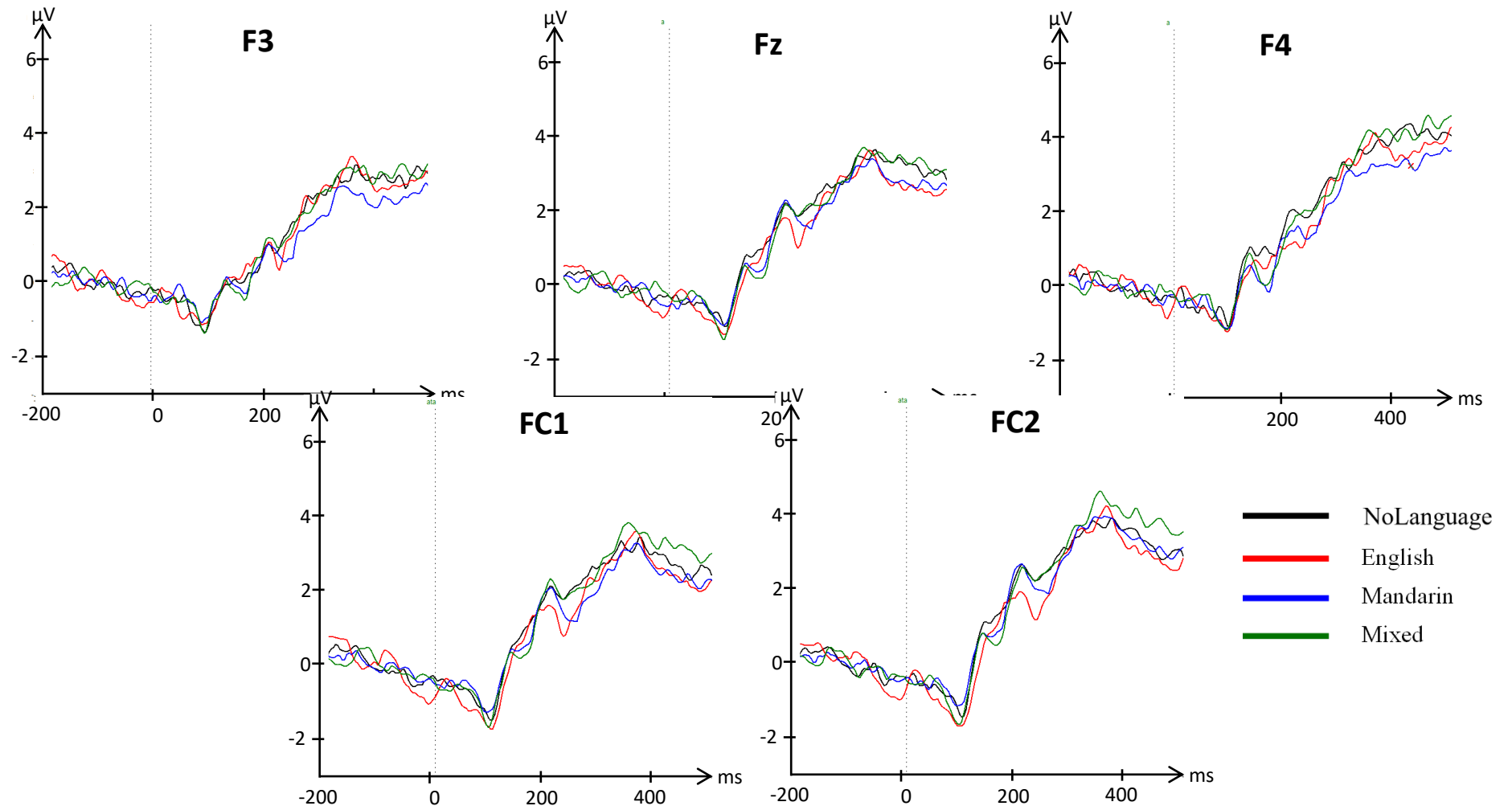
Mean amplitudes and peak latencies of the P2, N2, and P3 components were analyzed using a one-way repeated-measures ANOVA with Language (NoLanguage, English, Mandarin, and Mixed) as the within-subjects factor. According to Luck (2014), the P2 and N2 components are maximal over the midline anterior scalp sites, hence, they were analyzed by taking the average waveforms recorded at electrodes: F3, Fz, F4, FC1, and FC2. Time windows of 140–170 ms and 170–210 ms were selected to analyze the P2 and N2 components respectively. The P3 component is typically maximal over midline centro-parietal scalp sites (Polich, 2012), hence, it was analyzed by taking the average waveforms recorded at electrodes: CP1, CP2, P3, Pz, and P4 with a time window of 325–425 ms (Luck, 2014). Figure 15 visualizes the electrodes included in the anterior and posterior regions of interest.

Figure 15: *Visual Representation of Anterior (Green) and Posterior (Orange) Regions of Interest and Their Electrode Makeup*



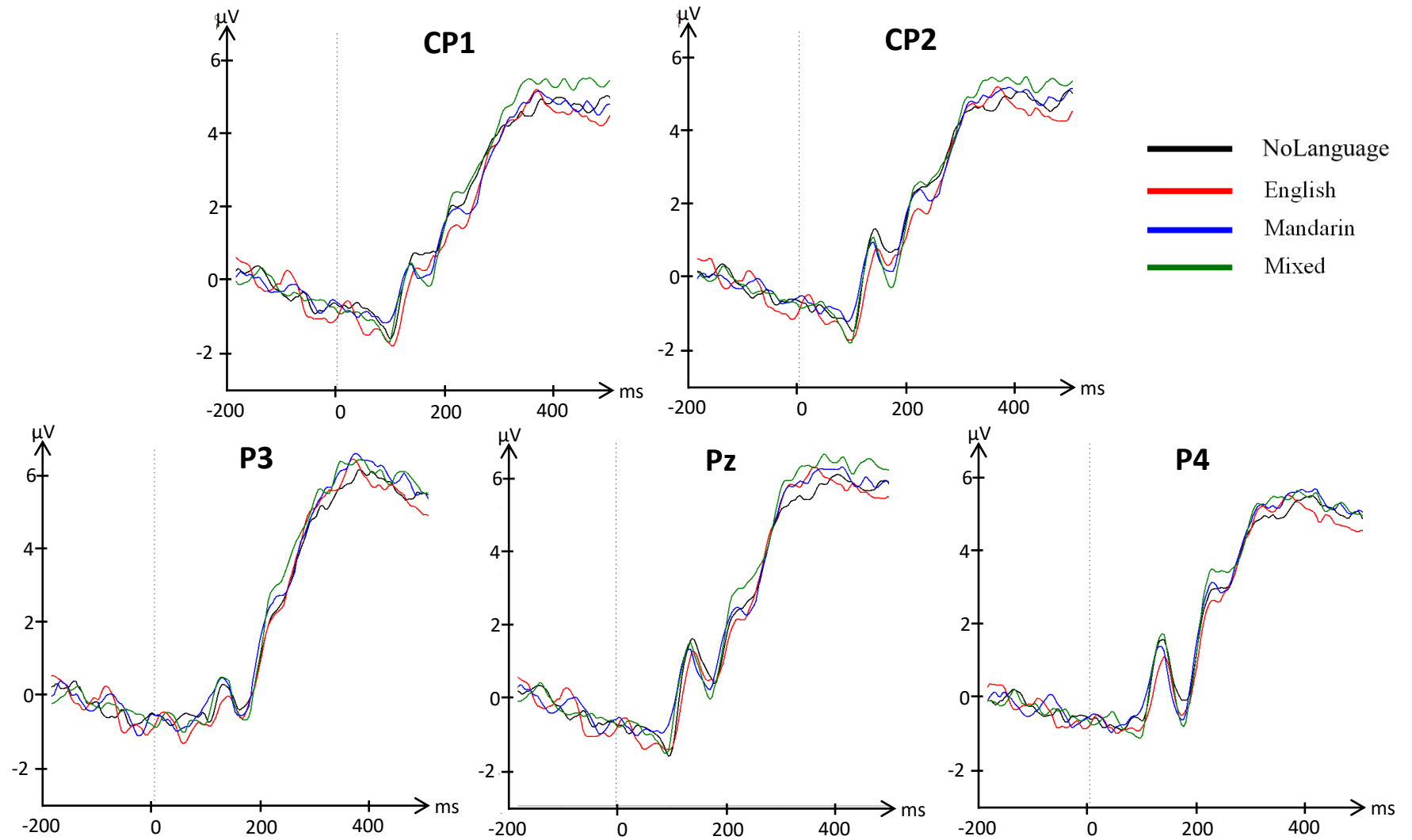
There was no significant main effect of Language on the N2 amplitude ($F(3, 114) = 0.44$, $p = .723$, $\eta_p^2 = .012$, $BF_{10} = 0.06$) nor on its latency ($F(3, 114) = 1.19$, $p = .315$, $\eta_p^2 = .030$, $BF_{10} = 0.14$). Similarly, no significant main effect of Language was found on the P2 amplitude ($F(3, 114) = 0.55$, $p = .652$, $\eta_p^2 = .014$, $BF_{10} = 0.05$), nor on its latency ($F(3, 114) = 0.75$, $p = .526$, $\eta_p^2 = .019$, $BF_{10} = 0.08$). Additionally, there was no significant main effect of Language on the P3 amplitude ($F(3, 114) = 0.39$, $p = .758$, $\eta_p^2 = .010$, $BF_{10} = 0.05$), nor on its latency ($F(3, 114) = 0.40$, $p = .750$, $\eta_p^2 = .011$, $BF_{10} = 0.05$). Figures 16 through 19 visualize the ERP scalp distribution at the corresponding sites.

Figure 16: ERP Distribution of Target Stimuli at Anterior Sites.



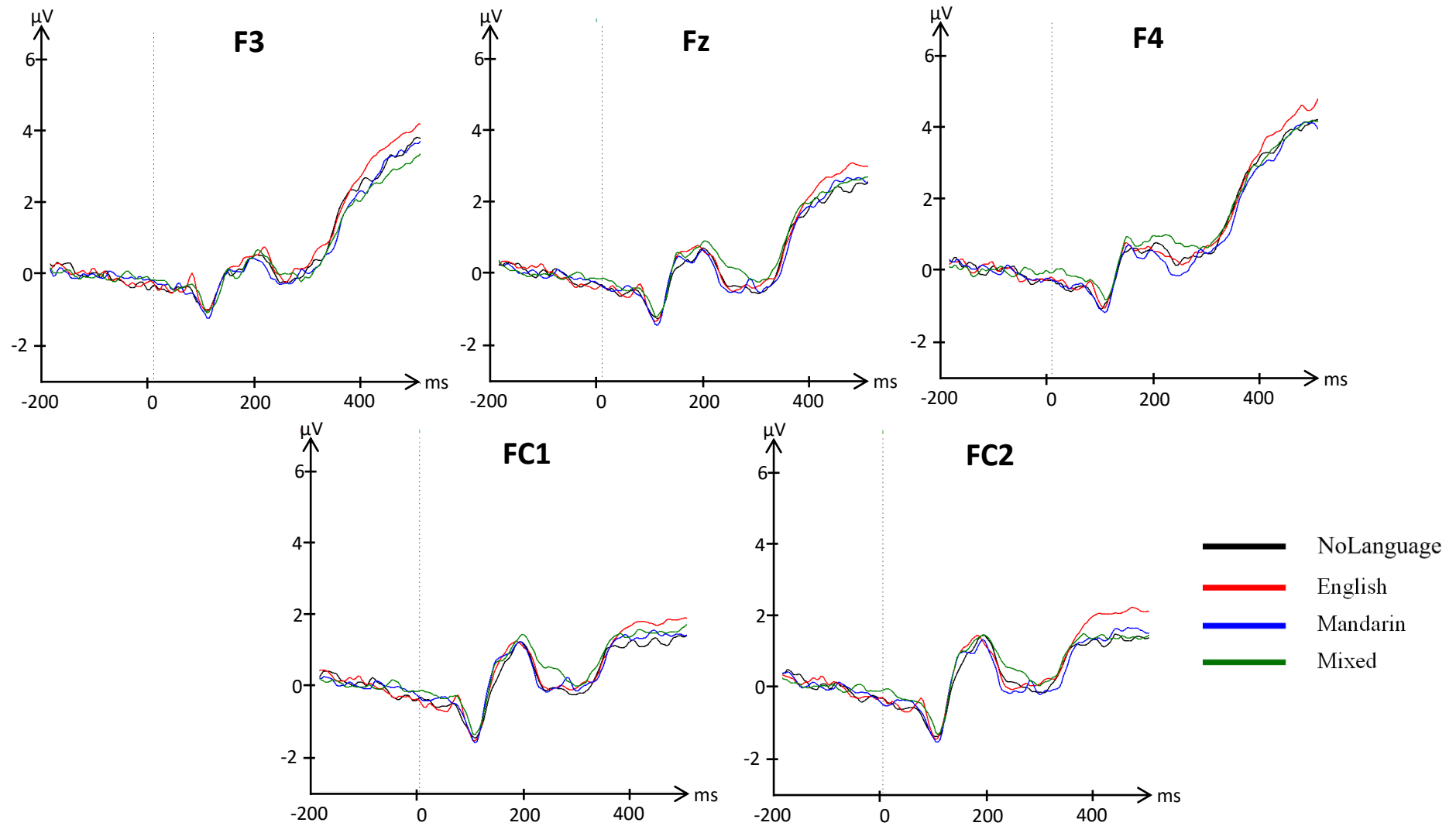
Note. Negative is plotted downwards

Figure 17: ERP Distribution of Target Stimuli at Posterior Sites



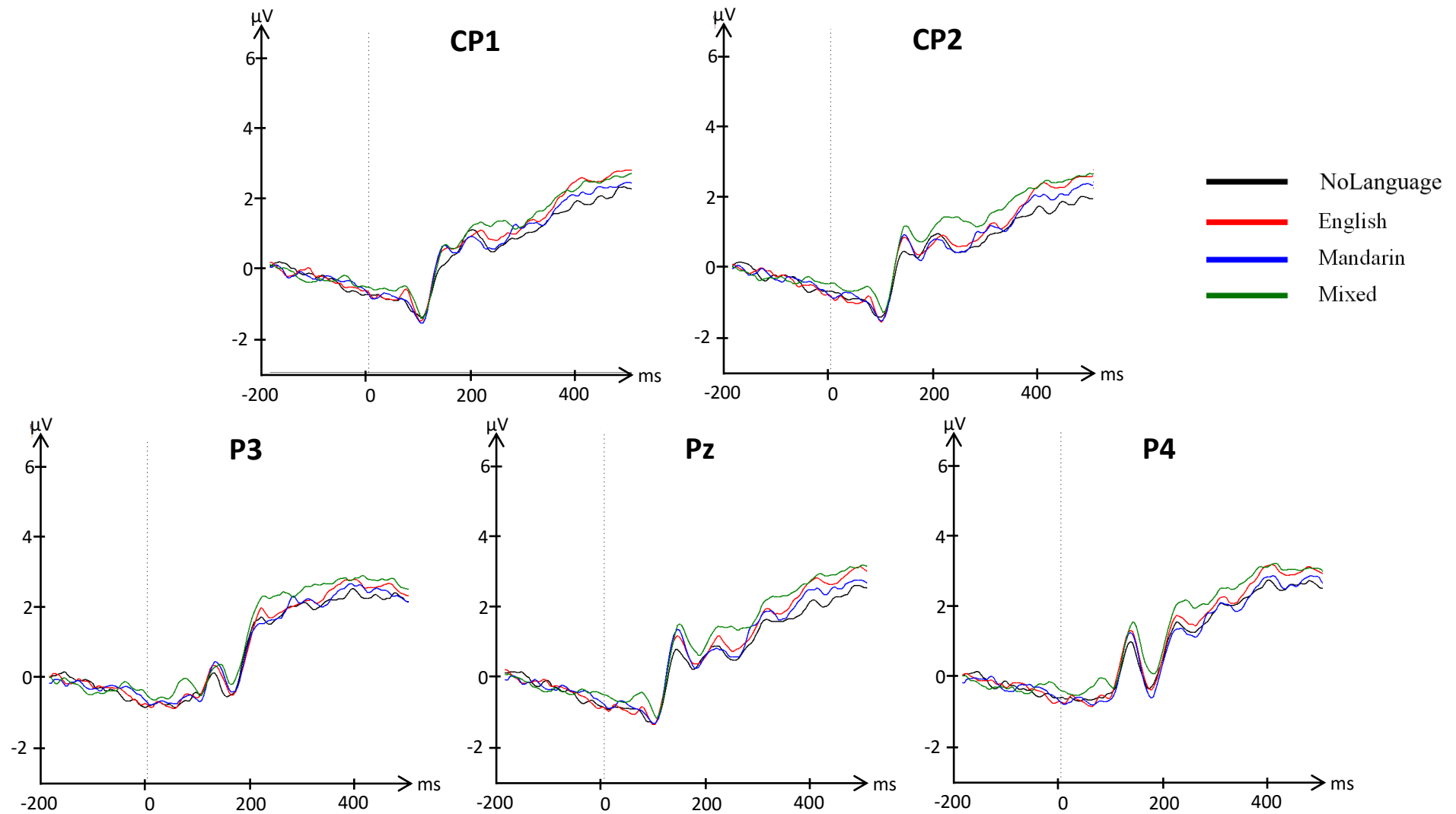
Note. Negative is plotted downwards

Figure 18: ERP Distribution of Non-Target Stimuli at Anterior Sites



Note. Negative is plotted downwards

Figure 19: ERP Distribution of Non-Target Stimuli At Posterior Site

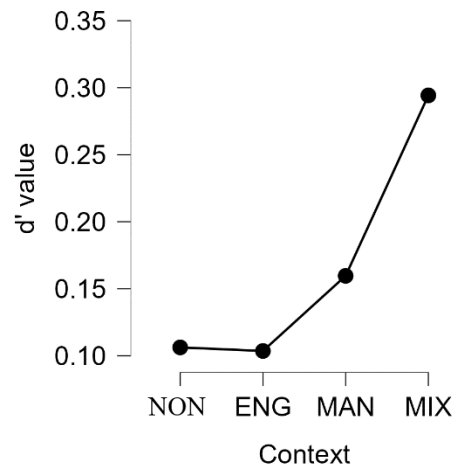


Note. Negative is plotted downwards

D Prime Analysis. To further analyze the data, a sensitivity analysis using D Prime (d')¹¹ was conducted based on the Signal Detection Theory (Green & Swets, 2000). This analysis aims to provide insights into the changes in the ability to discriminate targets from non-targets in the 2-back task. Furthermore, d' analysis is typically applicable to memory tasks, like the N-back task (Mikhalishchina & Utochkin, 2022). A high d' indicates a stronger ability to discriminate targets from non-targets, whereas a d' of 0 indicates the inability to discriminate targets from non-targets (Stanislaw & Todorov, 1999).

It was observed that the Mixed-language block had the highest d' ($M = 0.29$, $SD = 0.77$) followed by Mandarin-only ($M = 0.16$, $SD = 0.74$), English-only ($M = 0.10$, $SD = 0.78$), and NoLanguage ($M = 0.11$, $SD = 0.68$). A one-way repeated-measures ANOVA was conducted with language contexts being the within-subjects factor, and the d' scores being the dependent variable. Results revealed no significant main effect of language context on d' scores ($F(1.71, 64.96) = 0.77$, $p = .449$, $\eta_p^2 = .020$). See Figure 20 for a visualization of the d' values across language contexts.

Figure 20: Plot of d' Values Across Language Contexts



¹¹ A Snodgrass and Corwin (1988) correction factor was applied because there are instances where participants did not commit any false alarms, which would bias the calculation of d' . This correction factor applies a 0.5 frequency to all data. Using the formula: $\frac{(\text{number of hits/false alarm}) + 0.5}{(\text{total number of trials with signal present/absent}) + 1}$, the hit/false alarm rates were calculated. These values were then transformed into a Z-score. Finally, d' is calculated using the formula $Z_{Hit} - Z_{FalseAlarm}$.

6.4.3 *Forward-stepwise Multiple Regression*

The present regression included the following 13 variables: MLD, L2 AoA, L2 Proficiency, L2 Immersion, L2 Dominance, English AoA, English Proficiency, English Immersion, English Dominance, Mandarin AoA, Mandarin Proficiency, Mandarin Immersion, and Mandarin Dominance. L3 and L4 variables were omitted as not all participants were reported to be multilingual. This study examined the N2, P2, and P3 peak amplitudes across the four language contexts: NoLanguage, English, Mandarin, and Mixed¹². Twelve separate models were created, one for each combination of language context and ERP component.

However, the analyses did not identify any noteworthy variables that could effectively and reliably predict changes in N2, P2, or P3 peak amplitudes. For N2 peak amplitudes, two models successfully converged; one model with only MLD scores predicted NoLanguage block peak amplitudes ($F(1, 37) = 4.37, p = .044, R^2_{\text{adj}} = .08$), another model with only L2 Immersion predicted Mandarin-only block peak amplitudes ($F(1, 37) = 4.22, p = .047, R^2_{\text{adj}} = .08$). For P2 peak amplitudes, no successful models converged. For P3 peak amplitudes, two models successfully converged; one model with only L2 AoA scores predicted English-only block peak amplitudes ($F(1, 37) = 4.30, p = .045, R^2_{\text{adj}} = .08$), another model with only L2 Proficiency predicted Mandarin-only block peak amplitudes ($F(1, 37) = 7.64, p = .009, R^2_{\text{adj}} = .15$). Given only a single predictor was entered into both models, multicollinearity was not a concern in these models.

6.5 Discussion

The present study investigated the effects of different language contexts on the 2-Back task performance in multilinguals using ERPs. First, it was hypothesized that language context differences would only be observable in the ERP data but not behavioural data. Second, it was expected that the mixed-language context would demonstrate the smallest N2, largest P2, and largest P3 components compared to single-language and NoLanguage contexts. Third, it was expected that higher MLD scores would be associated with smaller N2 amplitudes and larger P2 and P3 amplitudes.

¹² Regression analyses on RTs and accuracy did not yield meaningful results. As it was hypothesized that there would be no significant behavioural differences, the regression analysis results on RTs and accuracy were not reported for the sake of brevity.

Contrary to the first hypothesis, neither behavioural nor ERP data demonstrated any significant differences between language contexts. Past studies traditionally have found significant differences in ERP data but not behavioural data (Barker & Bialystok, 2019; Comishen & Bialystok, 2021; Fernández et al., 2013, 2014; Moreno et al., 2014; Morrison et al., 2018). According to Paap et al. (2015), because both ERP and behavioural data align, there is strong evidence opposing the cognitive adaptation effect proposed by the ACH.

For the second hypothesis, although the results do not support the predictions as there were no significant differences between amplitudes (or latencies) of the N2, P2, and P3 components, the amplitudes of the N2 and P3 components were in the expected directions (Comishen & Bialystok, 2021; Morrison et al., 2018). The P2 component did not demonstrate any notable patterns of results.

First, consider the behavioural results in more detail (see Table 15). The RT and accuracy scores are similar across language conditions suggesting that language priming did not directly influence performance. This finding aligns with Comishen and Bialystok (2021) and Morrison et al. (2018) who also found no significant differences in either RT or accuracy measures between monolinguals and bilinguals across the 0-, 1-, 2-, and 3-back versions. However, both did find significant differences in P3 amplitudes between monolinguals and bilinguals. There was no monolingual control group in the current study. This absence highlights that detecting differences within multilinguals is much more challenging than comparing monolinguals and bilinguals, as seen in previous research. Notably, the present study's accuracy scores were similar to the previous studies, albeit slightly higher. The accuracy score in the current study across all language conditions averages approximately 80%. In comparison, on the 2-back task with bilinguals, Comishen and Bialystok achieved an accuracy of around 71%, whereas Morrison and colleagues achieved 75%.

Could this increase in average target accuracy scores be an indicator that perhaps language context priming improved WM abilities in general and not via a rapid, block-by-block, cognitive adaptation? To answer this question, the current study is compared to Comishen and Bialystok (2021) and Morrison et al. (2018), both of which did not feature a rapidly changing language context. To facilitate cross-study comparisons, the RTs and accuracy scores of these studies were

standardized using the inverse efficiency score¹³ (IES; Bruyer & Brysbaert, 2011). If language context priming improved WM abilities in general, the current study should have demonstrated the lowest IES scores. However, it appears that on the 2-back task scores of bilinguals, the study by Morrison and colleagues had the lowest IES scores at 628.20 ms, followed by the current study at 704.61 ms (averaged across language contexts), and finally Comishen and Bialystok at 906.94 ms (see Appendix 4).

Furthermore, there are no discernible differences between the language contexts on the amplitude or latency of the ERP components (see Figures 16–19). At the minimum, there should have been an observable difference between the language and NoLanguage contexts. There are two possible interpretations. First, there may not have been enough code-switching stimuli to necessitate the recruitment of additional WM resources. Green and Wei (2014) and Hofweber et al. (2016, 2020b) both suggested that congruent lexicalization recruits the most monitoring resources (i.e., working memory, Miyake et al., 2000), followed by insertions, and alternations, which require the least monitoring resources but primarily involve inhibitory control (see Section 1.4.1 and Muysken, 2000, for a review of the different types of code-switching). A code-switching analysis of the English-Mandarin code-switched dialogues (see Appendix 2) suggests that it is made up of mostly congruent lexicalization and alternation, with insertions being the least frequent. As a result, WM and inhibitory control were both recruited, leading to ambiguous findings in the present study. A clearer interpretation should be observable if the code-switches were mostly congruent lexicalizations. However, this explanation cannot explain why behavioural, or ERP data of the NoLanguage block was indistinguishable from the language blocks.

Another possible explanation could be that carryover effects from the language priming may have spilled over into other contexts. Although there have not been any empirical investigations into the difference in effect size or longevity of the priming effect between using single words (e.g., Wu & Thierry, 2013) versus sentences (e.g., Adler et al., 2020), there is evidence that sentence priming is stronger than single word priming (as discussed in Section 6.2). Therefore, unlike studies that used single-word priming during an ANT (Timmer et al., 2021) and observed a significantly larger P3 component in mixed-language contexts than in single-language

¹³ The IES score is a measure of performance that considers speed-accuracy tradeoffs (SATs), calculated by the formula $\frac{RT}{1-ERR}$. By accounting for SATs, which can be highly variable and challenging to measure, the IES score provides a more accurate indicator of performance. A higher IES score indicates lower efficiency, meaning more time is taken to determine whether a given stimulus is a target or non-target independent of the overall accuracy.

contexts, the use of dialogue-reading priming in this study might have produced a larger priming effect that persisted across blocks. This possibility could explain why even the control block had comparable RT and ERP performance with the language blocks. Overall, both behavioural and ERP data suggests that the predictions of the ACH does not hold up in multilingual integrated communities. Language contexts did not appear to modulate behavioural and ERP outcomes, which goes against the predictions of the ACH. For a more holistic understanding of the bilingual effect, it is vital for the ACH to be reviewed and revised according to multilingual integrated communities

Regarding the P2 component, the results do not support the initial hypothesis, neither quantitatively nor qualitatively. Although this finding does not align with Comishen and Bialystok (2021), it does align with Morrison et al. (2018). In hindsight, this finding may not be surprising. Although the P2 amplitude may be associated with WM encoding, the amplitude is likely an indicator of object feature encoding rather than WM per se (Mecklinger & Pfeifer, 1996; Shalchy et al., 2020). Comishen and Bialystok found a significant P2 amplitude difference because they experimentally manipulated the N-back stimulus type (i.e., icon, letter, shape, and number). The current study and Morrison's experiment did not find P2 amplitude differences because the N-back stimuli were presented as numbers only.

The lack of significant differences observed in the N2 and P3 components could be attributed to the less distinct contrasts between single, dual, and dense code-switching contexts, especially in highly integrated language cultures like Malaysia (Lai & O'Brien, 2020). As language contexts are blended and indistinguishable in such cultures, priming in a single-language context could potentially evoke similar adaptive effects on cognition as in mixed-language contexts. Likely, the relationships between language contexts and cognitive adaptations may not be as straightforward as initially proposed in the ACH. Instead, the relationship is likely influenced by the language practices of a particular culture. This has been implicated before (Samuel et al., 2018; Karimi & Rad, 2021; Xie et al., 2022; Gosselin & Sabourin, 2023; Beatty-Martínez et al., 2020). Concurrent with the current study's results, there is strong evidence to suggest that the manifestation of a bilingual effect is contingent on the macro culture of the sample (Xie et al., 2022). For example, Gosselin and Sabourin (2023) concluded that where a rapid modulation of cognitive abilities may take place for certain bilinguals (such as those living in predominantly dual-language contexts), it may not occur for those living in predominantly dense code-switching

contexts. Likewise, bilinguals predominantly in integrated language contexts, where the boundaries between single, dual, and dense code-switching contexts are blurred, might exhibit cognitive adaptation patterns that are difficult to distinguish compared to those predominantly in segregated language contexts. Interestingly, Raisman-Carlovich et al. (2024) found that bilinguals in an integrated context exhibited more efficient performance on the flanker task compared to those in the segregated context. The studies discussed here underscore the complexity of the relationship between bilingualism and cognition. They suggest that the inconsistent results observed in past research may be partially due to the unaccounted influence of participants' macro culture.

In Chapter 1, the valence ambiguity problem was briefly introduced. In essence, it highlights the challenge of interpreting ERP differences when no significant behavioural effects (e.g., RTs, accuracy) are observed. This can lead to misinterpretations, where lower or higher ERP amplitudes may be interpreted to support a research hypothesis. However, in the current experiment, behavioural and neurophysiological results were aligned, eliminating any valence ambiguity. Both ERP amplitudes and latencies remained consistent across language contexts, mirroring the pattern observed in RTs.

Finally, with d' values below 1 across language contexts, the current sample demonstrated poor discriminability between targets and non-targets. However, qualitatively, it seems that participants displayed relatively greater sensitivity to targets when primed in a mixed-language context compared to single-language and NoLanguage contexts (see Figure 20). Although this difference is not statistically significant, it aligns with the expected direction and complements the ERP results. This finding suggests that there might be a slight EF adaptation occurring when participants are primed in the mixed-language context, but for some underlying reasons, the adaptation was not observable in the behavioural or ERP data.

6.5.1 Role Of Multilingual Experience on ERP Components

No specific hypotheses were made for RT data as no significant differences were expected in that domain. However, it was expected that higher MLD scores will be associated with smaller N2 amplitudes, and larger P2 and P3 amplitudes. Contrary to expectations, MLD scores did not predict changes in N2, P2, or P3 amplitudes. Although MLD scores did predict P2 amplitudes only in the NoLanguage block, the lack of a consistent pattern across the other blocks and ERP components undermines confidence in this outcome. As a stepwise regression analysis was

conducted, interpretations of the results must be conservative. Hence, in the absence of discernable patterns, these significant models are more likely to be the result of random chances rather than indicating a true effect. Further discussion and ideas for future research will be explored in the following General Discussion chapter.

Chapter 7: General Discussion

The primary objective of this thesis was to explore how language use and language context might influence executive function. A survey was first conducted to explore the public's sentiment towards language attitudes was conducted to understand attitudes to language use in Malaysia. Then, focus was placed on whether different language contexts (e.g., no language, single-language, and dual-language) can adaptively modulate EF abilities in multilingual individuals and whether the degree of multilingualism (i.e., MLD scores) can predict cognitive task performance beyond traditional measures of multilingualism. In this chapter, the key findings of the four experimental chapters (Chapters 3 to 6) will be summarized, and potential interpretations of these results will be discussed, followed by a discussion of the implications of the findings for future research in this field.

Chapter 3 starts off the thesis by investigating the current sociolinguistic background of Malaysia and uses the LAQ-LL to explore Malaysian's attitudes towards multilingualism and language learning experiences. This enabled an empirical investigation into the specific individual factors that may be used to predict changes in these language attitudes. In general, Malaysians appear to have a strong positive sentiment towards multilingualism and their language learning experience. Although there was almost an unanimously positive response to the value of multilingualism, there was more variability in the language learning experience. This pattern was similarly reflected in Artamonova (2020). Among the individual factors measured: Age, SES, education level, and first and second language proficiency, only higher education level and first language proficiency predicted a higher value of multilingualism. These findings are consistent with previous research, higher education levels are typically linked to greater awareness of the benefits and importance of multilingualism (Surrain & Luk, 2019), increased exposure to people from diverse cultural and linguistic backgrounds (Teichler, 2017), and stronger instrumental motivation for language learning (Novak & Morvai, 2019). While language proficiency was found to be the best predictor of language attitudes (Xie & Cavallaro, 2016), likely because it aligns one's actual self with Malaysia's multilingual identity. Finally, no factors were significant predictors of language learning experience.

In Chapter 4, three successive experiments were conducted: Experiment 2A (English-Mandarin proficient multilinguals), Experiment 2B (English-Malay proficient multilinguals), and Experiment 2C (English proficient multilinguals). Each experiment used a modified flanker task designed to prime participants into different language contexts. Across all three sub-experiments, mixed-language context performance was not significantly different from single-language context performance. However, qualitatively, there appears to be a consistent pattern, where the fastest RT was in NoLanguage contexts, followed by mixed-language contexts with single-language contexts typically showing the slowest RT. Qualitatively, this pattern seems to be influenced by language proficiency. RT was slightly faster in mixed-language contexts than in single-language contexts when participants were primed in a second language that was both familiar and highly proficient (i.e., Mandarin). This difference diminished when participants were primed in a second language that was familiar with but only moderately proficient (i.e., Malay), and showed almost no difference when the second language primed was unfamiliar (i.e., Welsh).

Additionally, MLD scores significantly predicted smaller flanker effects, but only in the mixed-language block and only when primed in familiar and highly proficient languages. This relationship disappeared when participants were not highly proficient in the primed languages. There were several spurious significant regression models, these models consisted of single predictors such as L2 AoA, L2 Immersion, and L2 Proficiency and predicted flanker effects. Considering that a forward-stepwise regression was used, caution was warranted in interpreting these results as empirical support for a bilingual effect. Overall, results did not convincingly demonstrate that MLD scores were a superior predictor of flanker task performance over traditional multilingual experience variables.

Chapter 5 investigated whether the rapid cognitive adaptation, brought on by different language contexts, can be more readily observed using an eye-tracking paradigm, on a group of highly proficient English-Mandarin multilingual young adults. To accommodate the new paradigm, an ANT task was used. No significant main effect on language was found on RT, Accuracy, Flanker Fixation duration or the First Cue-Saccade Latency. These results might be due to the prominent facilitatory effects of the asterisk cues on RT. As these cues provided the most reliable and valid indication of the flanker stimuli, participants may have relied solely on them to enhance task performance, eliminating the need for cognitive adaptation related to language contexts.

Chapter 6 extends the investigation by implementing a more challenging task of EF, the 2-back task, while EEG was recorded. This task was chosen as there was a ceiling effect on the accuracy of the task performance in Chapters 4 and 5. This introduces more variability in performance, allowing for more opportunities to observe the effects of rapid cognitive adaptation, brought on by different language contexts. Unfortunately, neither behavioural nor ERP data demonstrated any evidence that mixed-language contexts adaptively improved EF any better than single-language or NoLanguage contexts. While there may have been expected to be no significant differences within the behavioural data, it was surprising to find no significant differences in the N2, P2, and P3 amplitudes across language contexts. This contradicts the ACH which proposes that mixed-language contexts should have induced greater cognitive adaptations than single-language contexts. Furthermore, multilingual experiences did not predict behavioural or ERP differences. This result could suggest that differences are not discernible because everyone in the sampled population is already experiencing the effects of being multilingual.

7.1 Language Contexts and Cognitive Control

While it is appealing to reduce the complexity of multilingual interactions into three distinct categories: single-language, dual-language, and dense code-switching contexts—the reality is that multilingual interactions often shift fluidly between these contexts. For example, while some multilinguals are predominantly from a dual-language community (i.e., where no code-switching occurs, and each language is used in different social contexts), code-switching may still occur (intentional or not). How, then, would cognitive adaptation appear for these multilinguals who transition between different language contexts? This multilingual feature is important, but commonly overlooked. As previously discussed, in cultures where language use is highly integrated in all social contexts (e.g., Malaysia and Singapore), distinctions between dual-language and dense code-switching contexts are less clear (Lai & O'Brien, 2020; Hartanto & Yang, 2016, 2019; Ng & Yang, 2021). As a result, cognitive adaptations may be expressed differently in such countries.

Wu and Thierry's (2013) findings on how rapidly changing language contexts may influence cognitive control closely resemble the actual experiences of multilinguals. Yet, within the context of the ACH, it remains unclear how the intensity of exposure to different interactional contexts relates to the engagement of cognitive adaptation. The intensity of exposure can be thought of as passive (comprehension) or active (production) exposure. For example, a multilingual on a tour

could be considered passively exposed to multiple languages through silent reading and listening. On the other hand, a multilingual interpreter could be considered actively exposed to multiple languages given their active production through speech or writing. Understanding how these types of exposure would or would not lead to the engagement of domain-general cognitive control may shed light on the mixed findings in past literature.

In the current thesis, participants were passively exposed to rapidly changing language contexts, either through reading single-word priming or a dialogue. Statistically, neither behavioural nor ERP data showed any significant differences between language contexts. However, throughout the experiments, there was qualitative evidence to suggest that there may be differences in cognitive adaptation depending on language contexts. Figures 9, 12, and 13 demonstrate that differences in task performances were in the predicted directions. Specifically, task performance appeared to be slightly improved in the mixed-language contexts compared to single-language contexts, although these differences were not statistically significant.

Kałamala et al. (2022) actively primed different language contexts through the clever use of confederates. This method of priming is more natural, as it better reflects the nature of bilingual interactions compared to the forced language priming methods used in the current thesis. Using the Map Task (Brown et al., 1985), authors set up a scenario where participants were the hosts, and confederates were the players. The confederates were instructed to communicate with the host using L1-only, L2-only, or both L1 and L2. After 2 hours of the game, participants completed the Stop-signal and Stroop tasks to measure their response inhibition abilities. Results indicated that while there were no behavioural differences between the language context blocks, there was a reduction of P3 and N450 amplitudes after the L2-only and L1-L2 games, reflecting better response inhibition abilities in the stop-signal and Stroop tasks respectively. This finding suggests that the method of priming language contexts could influence whether cognitive differences between language contexts are observed. However, evidence for this conclusion is mixed, as there is no direct research investigating the differences between active and passive priming on the observation of the bilingual effect. Thus, whether there is a difference between passive priming (e.g., listening or silent reading) or active priming (i.e., production and comprehension) remains undetermined and presents a viable area for future research.

Another growing concern about the ACH model is whether the cognitive adaptations seen in past studies were an artifact of artificial priming methods. The basic assumption is that switching

languages is difficult, hence, when required to switch between two languages (e.g., dual-language contexts), additional cognitive resources are recruited to facilitate the switching, leading to domain-general cognitive improvement. However, there is growing evidence to support that free-and-voluntary language switching is not effortful (Kleinman & Gollan, 2016; Blanco-Elorrieta & Pylkkänen, 2017; Gollan et al., 2014; Gollan & Ferreira, 2009; Zhang et al., 2015; Reverberi et al., 2018). When participants engage in naturalistic language switching (i.e., not cued, prompted, or constrained), the traditionally observed language switch costs (Prior & MacWhinney, 2010; Filippi et al., 2014) disappeared (Blanco-Elorrieta & Pylkkänen, 2018). Even the study by Kałamała et al. (2022), which claims to have used naturalistic conversation priming, may paradoxically not be truly naturalistic. The confederates may also implicitly and artificially constrained the use of languages to either single- or dual-language contexts by their presence alone.

In summary, the current thesis does not provide support for the ACH model in its current iteration. More research needs to be conducted to extend the parameters of the ACH within highly multilingual integrated language communities. This finding cannot be attributed to order effects, as the order of experimental blocks was fully randomized across Experiments 2, 3, and 4. A supplementary analysis confirmed that outcomes were consistent regardless of block order (i.e., Range of p-values between .197 and .980). Similarly, Kalamala et al. (2020) found that across other EF tasks (i.e., Antisaccade, Stroop, Go/No-Go, and Stop-signal reaction time tasks), no effects were found for bilinguals who were immersed in single-language versus dual-language contexts. Further, past evidence in support of the ACH is called into question against the backdrop of the issue of artificial versus naturalistic language context priming. This also suggests that the ACH model may not fully reflect the actual multilingual reality.

7.1.1 Does Bilingual Language Control Necessarily Require Executive Control?

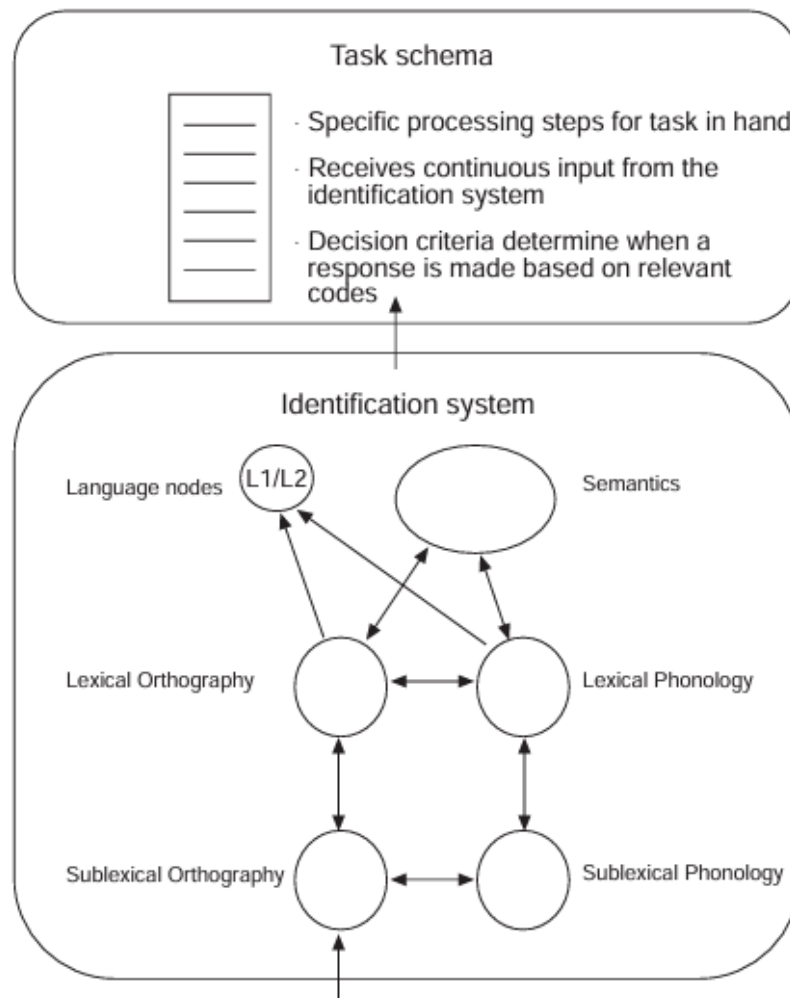
Is EF necessary during bilingual language control if naturalistic language switching is not inherently effortful? According to Ayres et al. (2021), fMRI could be used to investigate the relationship between intrinsic cognitive effort and EF. Typically, fMRI uses blood oxygen level-dependent (BOLD) contrast, which measures the changes in blood flow that are associated with the amount of energy used in brain cells. Increased energy use, especially in fronto-parietal regions, can be associated with increased cognitive load (Mäki-Marttunen et al., 2019). Hence, in continuation of this thesis, the next logical step would be to investigate the relationship between naturalistic versus artificial language switching manipulation and EF using fMRI.

Zhang et al. (2015) used fMRI to study brain activation differences between free-language and forced-language switching and reported that both engage non-overlapping prefrontal-parietal-subcortical regions. During free-language switching, the middle frontal cortex (MFC) was engaged. During forced-language switching, bilateral frontoparietal regions were engaged. This finding suggests that executive control is recruited differently during free versus forced language switching. This was supported in another study by Reverberi et al. (2018), who also reported that when bilinguals were given free will to choose which language to speak, the medial prefrontal cortex (BA10) was activated before language execution, which activated a separate network of brain regions (i.e., left Broca's area, right supramarginal gyrus, right inferior temporal gyrus). Zhu et al. (2022) found that recruitment of the medial frontal cortex (MFC) during free language choice was linked to smaller mixing costs and no significant switch costs in a picture-naming task. However, when participants were forced to switch languages, they exhibited increased brain activity in the right inferior frontal gyrus, consistent with findings by Reverberi et al. (2018).

To understand why this is the case, it is important to understand how bilingual language switching works. According to the Bilingual Interactive-Activation Plus (BIA+; see Figure 21) model (Dijkstra & van Heuven, 2002), language switching functions on the interplay between two systems, the Word Identification System (WIS) and the Task/Decision System. Effectively, the WIS is a bottom-up control system that functions by activating a network of orthographic, phonological, and semantic representations from a given word stimulus. On the other hand, the Task/Decision system is a top-down control system that incorporates an individual's expectations, strategies, and contexts into understanding and acting on the output from the WIS (Paap et al., 2019). In this configuration, the authors argue that the WIS operates independently of the Task/Decision system. In other words, bilingual language control does not necessarily require top-down cognitive control if a WIS, bottom-up control selects the language before its output reaches the Task-Decision system (Kleinman & Gollan, 2016).

Within the context of the ACH, language selection in any interactional contexts do not require the breadth or depth of cognitive resources as initially proposed (see Table 1, Section 1.4.1). Consequently, the cognitive adaptations arising from the different interactional contexts would not be cognitive stimulating enough to strain and train the EF system.

Figure 21. *Overview of the Word Recognition and Processing System as Discussed in The BIA+ Model.*



Note: Figure adapted from (Dijkstra & van Heuven, 2002)

Additionally, this configuration offers greater insight into the language context priming method used in the experiments (i.e., Single-word vs dialog-comprehension priming). According to Pollack (2022), there are two types of word processing: appearance/sound-based (shallow) processing and meaning-based (deep) processing. Shallow processing occurs when word identification stops at the level of lexical orthography (as in Experiments 2 and 3), whereas deep processing occurs when word identification reaches the semantic system (as in Experiment 4, see Figure 21). Currently, there is no literature on whether shallow or deep processing leads to stronger language context priming. However, according to the BIA+ model, semantic processing is not a prerequisite for language selection, meaning that language selection may still occur even during shallow processing.

When a language is artificially primed, the Task/Decision system overrides language selection by imposing a top-down language constraint on the WIS and this may lead to multilinguals engaging in inefficient strategies by mixing both top-down and bottom-up control. Without language constraints, multilinguals can engage exclusively in bottom-up control (Kleinman & Gollan, 2016). Consequently, it is only during artificial language priming that multilinguals experience a more difficult language switching, resulting in the need for the recruitment of more cognitive control. It is likely that naturalistic language switching is not effortful and thus, requires no adaptation in cognitive control.

This poses an interesting question. Considering artificial priming was used in Experiments 2, 3, and 4, why were there no task performance differences between the language contexts? The non-significant outcomes observed across all three experiments could be due to the invariance of multilingual experiences in the sample recruited. To verify this claim, a supplementary K-means cluster analysis was conducted (See Appendix 5). Confirming suspicions, the best model was the model with only two clusters. In the first cluster, participants acquire L2 at a later age and experience less L2 immersion. In the second cluster, participants acquire L2 earlier and experience more L2 immersion. Other multilingual experience variables show little deviations within and between clusters. A three-way mixed ANOVA was conducted to confirm the differences between the clusters. The results suggest no differences between the clusters, further supporting the conclusion that the samples recruited with this thesis did not significantly differ on all the measured language experience variables. It is reasonable to conclude that the samples recruited across the experiments would likely be traditionally labelled as "highly proficient, balanced, simultaneous bilinguals". As such, the current sample is well-suited to examine the effects of language contexts on executive functioning within a relatively homogeneous group, but not to explore the impact of the intensity of multilingual experiences on EF.

This limitation is an unavoidable consequence of conducting research in a highly integrated multilingual society. In Malaysia, being multilingual is the norm rather than the exception, this can be seen from the survey data in Chapter 3. Therefore, it is challenging to find research participants who are Malaysian but can speak only one language (i.e., those with low MLD scores). Instead, it may be more valuable to explore the differences between individuals stereotypically labelled as "bilinguals" and those who are "trilinguals" or beyond. This approach leverages existing sociocultural norms rather than searching for niche samples of monolinguals in Malaysia.

Additionally, it offers the opportunity to investigate polyglotism and how it fits within the bilingualism-cognition framework.

Past research is mixed on the difference between simultaneous bilinguals (i.e., learned L1 and L2 at the same time) compared to sequential bilinguals (i.e., learning L2 sometime after L1). The samples recruited in this thesis can be considered as simultaneous bilinguals (see Appendix 5). It is important to note here that the concept of simultaneous and sequential is not well defined as it is unclear at what age learning a second language transitions from being considered simultaneous to sequential. However, clearer insights may be obtained, if one considers the ends of the extreme, that is, those who simultaneously acquired both languages around birth versus those who acquire an L2 in adulthood. D'Souza et al. (2021) report that late sequential bilinguals had poorer change detection ability compared to early simultaneous bilinguals. Kousaie et al. (2017) found that early simultaneous bilinguals had more efficient neurofunctional connectivity and cognitive functioning than late sequential bilinguals, but, when compared to behavioural measures, there were no significant differences between the groups (Sabourin & Vinerte, 2014; Gathercole et al., 2014; van den Noort et al., 2019). Against this backdrop, it can be concluded that the differences in L2 AoA and L2 immersion in the samples of the current thesis are too small to observe significant group differences. While the k-means analysis (see Appendix 5) identified two clusters of participants across all experiments, the differences in their multilingual experiences are too minor to be meaningfully distinct.

Lastly, several alternative views should also be considered. Paap et al. (2019) discuss three other explanations of bilingual language control that do not require top-down cognitive control. First, the target language could be upregulated through repeated stimulus-response associations. In this scenario, the representation of the target language lexical candidates is greatly enhanced relative to the non-target candidates, leading to easier language selection and inhibition. Second, there may be no conflict resolution mechanism to resolve cross-linguistic interference. In this scenario, bilinguals live with the language competition and the risk of an occasional, but unintended, non-target language intrusion. Third, bilingual language control may also depend on the specialized mechanism underlying articulatory suppression. In this scenario, multilinguals' speech in a specific language suppresses any occurrence of a non-target language subvocal rehearsal, allowing for easier language selection. These scenarios are likely to occur in combination. Importantly, none of these scenarios implicate a transfer effect from bilingual

language control to domain-general abilities. Unless future research rules out these scenarios, the theoretical basis for the bilingual effect is ambiguous.

7.1.2 Reconciling The Opposing Findings: Is There Cognitive Control in Language Control?

At this juncture, two schools of thought were discussed. In the original stance of the thesis, EF is required to manage bilinguals' constant language switching, and this was based on the idea that an attentional system – the SAS (IC Model, Green, 1998; Attentional Control Model, Bialystok & Craik, 2021) – monitors for language intrusions and apply necessary cognitive control. The other school of thought, presented in Section 7.1.1, challenges this idea and suggests that language control is effortless because language selection happens outside the purview of EF, rather, it is based off bottom-up control instead. So, which is it? Is there cognitive control in language control?

Fyndanis and Lehtonen (2022) discuss the possibility that language control might rely on domain-general cognitive control, particularly in cases of pathological language-switching/mixing. Pathological language-switching/mixing, defined as the inability to control language switching or mixing (Fabbro et al., 2000), is often associated with aphasia (Kong et al., 2014). Aphasia offers a unique opportunity to investigate the relationship between language control and EF. People with Aphasia (PWA) typically exhibit deficits across a range of EF tasks (Purdy, 2002), suggesting a strong connection between language control and cognitive control mechanisms.

For example, Kong et al. (2014) report a case study of Dr. T, a trilingual speaker of Cantonese, English, and Mandarin, who developed fluent aphasia following a brain injury to the left frontal and temporoparietal areas. Dr. T exhibited pathological language switching, particularly during an unstructured discourse production task, where she involuntarily switched between her three languages without intent or control. This behavior was less frequent during structured naming tasks. Dr. T's broader EF deficits, as evidenced by her performance on tasks like the Stroop color-word and Wisconsin Card Sorting Task, further suggest that her impaired EF system played a role in her pathological language switching. This case aligns with Fabbro et al. (2000), who similarly found that brain injury to the frontal lobe led to uncontrolled language switching/mixing in another bilingual patient.

These findings led Fyndanis and Lehtonen (2022) to conclude that a deficient EF is partly responsible for the pathological language switching/mixing observed in multilingual PWAs. Their

review corroborates the IC model (Green, 1998), which argues that cognitive control is necessary for language switching. However, the BIA+ model (Dijkstra & van Heuven, 2002) suggests that language control and switching/mixing happens independently of EF. Paap et al. (2019) also provide alternative explanations, positing that language switching is not reliant on EF. How can these seemingly contradictory models coexist?

Kleinman and Gollan (2016) offer a potential reconciliation by suggesting that cognitive control in bilingual language processing is optional. Bilinguals may engage in either a more effortful, top-down control strategy or a less effortful, bottom-up control strategy, depending on the context and their language proficiency. This idea of flexible control is supported by the notion of automaticity in cognitive processes. In the field of psychology, automaticity is defined as the ability to perform actions without conscious effort or recruiting cognitive and attentional resources that are typically required for information processing (Schneider & Shiffrin, 1977). Segalowitz (2003) contends that automatization—the gradual process by which an action becomes automatic—is more important than the state of automaticity itself. The changes that occur when a conscious activity transforms into an unconscious process are more significant than the resulting automaticity. When language proficiency is low, language processing is conscious and effortful. Here, executive control is recruited extensively, as per the IC model. As proficiency increases, language processing automatizes, gradually becoming more unconscious and effortless, more automatic (Tavakoli, 2019). At this point, executive control becomes optional, recruited only when required, as per the BIA+ model. This perspective might explain the effortless nature of natural language switching (Kleinman & Gollan, 2016; Kałamała et al., 2022), while also accounting for the need for EF in controlling language switches in less proficient speakers or in PWAs.

Thus, it seems possible that both schools of thought - those that propose EF is essential for language control and those that argue for an independent, automatic language control process—could be valid depending on the context and the proficiency of the language user. Language control may involve a dynamic interaction between cognitive control and language proficiency, where automaticity and executive control are recruited flexibly based on the demands of the situation.

7.2 The Complexities of Measuring Multilingualism

Sections 1.1 and 1.2 discussed the issues of defining and measuring multilingualism and it is clear there is no one agreed upon measurement of bi/multilingualism. Some research only measures AoA or proficiency of the acquired languages, while others consider contextual code-

switching practices. It has been argued that such a reductionistic approach—the oversimplification of the richness and diversity of multilingual experiences to a single factor—has contributed to the mixed findings in the literature on the cognitive implications of multilingualism (Bialystok, 2021). Further concerns about the use of experiential variables of multilingualism are that it only reflects the individual’s experience with multilingualism, such as their proficiency in their languages or how dominant they are in the language. Yet, this is only one facet of multilingualism. First described by Haugen (1972), the use of language by individuals involves a dynamic interplay with social and cultural forces, termed “language ecology”. This understanding is evident even in modern psycholinguistic research, such as the ACH (Green & Abutalebi, 2013).

Therefore, it is not enough to measure multilingualism on only the experience variables (i.e., Proficiency, AoA, Immersion, Dominance) because multilingualism is not only an internal psychological phenomenon but also a sociological one. However, it is unclear what precisely constitutes a “language ecology”, and more importantly, how can it be quantified in a way that produces clean, empirical data to be used in research. The use of language entropy (Gullifer & Titone, 2019; Li et al., 2020) attempts to address this. This measure provides a single numerical factor that quantifies the “randomness” of language use. Low language entropy suggests that language use is compartmentalized, that is, languages are kept separate in distinct social contexts. High language entropy suggests that language use is integrated, that is, languages are used regardless of the communicative context and code-switching is common. It is important to note that the effectiveness of this measure requires further empirical scrutiny. Incorporating such a complex variable into relatively straightforward empirical models of bilingual language experiences challenges the principle of parsimony, which has long guided scientific research.

In the current thesis, the utility of the MLD score is compared with traditional multilingual experience variables in predicting task performance. The findings did not provide exceptional evidence to suggest that MLD scores were superior to traditional measures. Hence, while MLD scores represent a valuable step forward in assessing the ecology of multilingual experience, it may not be suitable for highly homogenous samples. Thus, this finding offers important insights for future research designs, suggesting that the utility of language entropy is contingent on the language use diversity within the sample being studied.

It would be easy to conclude that the experiments in this thesis are also indicative of the replication crisis of the bilingual effect. However, this conclusion should not be taken at face value

because it is based on “the assumption that the same experiment should produce the same results independent of the environment in which it was conducted” (Bak, 2016). Given the diversity of multilingual experiences, it is unlikely that an experiment would yield the same results even if the methodology was replicated to the letter. While there may be truly no cognitive effects to bilingualism, the replication crisis may also stem from the failure to adequately characterize the complexity of the multilingual experience (Beatty-Martinez & Titone, 2021; Bak, 2016).

Navarro-Torres et al. (2021) describe this characterization in terms of bilingual "phenotyping" - the assessment of language experiences and their interactions with the environment. It is argued that some, but not all, multilingual phenotypes would demonstrate a bilingual effect (Bak, 2016, Navarro-Torres et al., 2021; Beatty-Martinez & Titone, 2021). Both Hofweber et al. (2020a) and Bosma and Pablos (2020) observed that L1-dominant bilinguals showed an effect on a flanker task during an L2-only condition, but not during the code-switching condition. More specifically, Goral et al. (2015) found that L1-dominant older adult bilinguals showed little to no age-related inhibition decline on the Simon task, but balanced bilinguals showed significant age-related inhibition decline. Hofweber et al. (2020a) argued that L1-dominant bilinguals require greater inhibitory control in L2 contexts due to the saliency of their dominant language, and therefore, may experience greater inhibitory control training than balanced bilinguals. Based on these findings, it could be concluded that it is the interactions between language experiences and the environment that will lead to differences in cognition. Hence, the question now is to identify which "phenotypes" of bilinguals would lead to the expression of a bilingual effect. Such efforts require an interdisciplinary approach, as they involve examining bilingual language experiences through linguistic competencies, the broader social contexts in which these languages are used, and the neurobiological factors underlying language control.

7.3 Future Research

Several questions remain unanswered at present, this section highlights a few avenues for future research. First, in testing the ACH, what constitutes an appropriate control block? NoLanguage blocks in the experimental chapters contained no interleaved stimuli (Experiment 2), a Gaussian mask (Experiment 3), and an “X” mask (Experiment 4). These types of control blocks are non-linguistic, are they an appropriate control for the language context blocks? Perhaps a more convincing control could come from artificial languages, such as the Korean-Chinese artificial language demonstrated by Li et al. (2018). Applying this type of control block could resolve the

conflicting mechanisms of bilingual language control as discussed in section 7.1.1. Specifically, it could help researchers understand whether semantic processing is necessarily required in language selection or can language selection occur entirely independent of semantic knowledge (for examples of cognitive research using artificial languages, see Batterink & Neville, 2013; Friederici et al., 2002). As the artificial language resembles the known languages of a participant, it rules out the influence of differences in phonology, morphology, and grammar.

Second, the priming mechanism must be explored further for investigations on the ACH to be more aligned with psycholinguistic research. One critique of Wu and Thierry (2013), the pioneering paper on the investigations of rapid modulations of EF by language context, is that priming single words may not fully reflect the mechanisms underlying real-world dual-language or dense code-switching contexts. Code-switching is not just a receptive process (Bultena et al., 2014), but also a productive one (Meuter & Allport, 1999). It is unclear whether this variable is influential to the expression of a bilingual effect. Studies that required participants to produce language switches (or no switches for single-language contexts) found a significant difference between single-language and dual/dense code-switching contexts (Jiao et al., 2020, Kalamala et al., 2021, Lai & O'Brien, 2020, Ng & Yang, 2021). Yet, studies requiring only receptive comprehension of the priming stimuli also produced similar significant differences in the same direction (Timmer et al., 2021; Han et al., 2022). Future research should consider exploring the differences between comprehension-only, production-only, and both comprehension-production to better understand whether such a variable can affect experimental outcomes. In comprehension-only blocks, participants could be required to read/listen to the priming stimuli and answer a comprehension question later. In production-only blocks, participants could be instructed to read out loud as many words on-screen as possible within a given time limit. In comprehension-production blocks, participants could be instructed to discuss, out loud, pre-determined NoLanguage affective topics with a confederate. Such methodology, combined with triangulation methods from eye-tracking and ERP techniques, would help shed light on the nuanced effects of language priming medium on the ACH.

Third, future studies should consider sampling populations whose cognitive processing is either developing (i.e., children) or declining (i.e., older adults) because the current sample is made up of mostly healthy young adults at their prime cognitive functions (Bialystok, 2016; 2017; Ferguson et al., 2021). Hence, the current sample's baseline inhibitory control abilities may already

be at their peak. When accuracy scores are highly clustered around the maximum value, it results in minimal variance, constraining the range of values that a variable may take (Arslan & Benke, 2023). This leaves little room to observe any potential benefits of multilingualism.

Fourth, it is crucial to investigate how code-switching practices can improve EF, given past findings (Hofweber et al., 2016; 2020b; Treffers-Daller et al., 2020). In Chapter 5, the MSWQ (an adapted version of the BSWQ; Rodriguez-Fornells et al., 2012) was used to measure participants' code-switching practices. None of the code-switching variables measured significantly predicted any of the outcomes of the experiment. However, only some forms of code-switching practices likely influence EF. Hofweber et al. (2016; 2020b) found that only congruent lexicalization and alternation practices improve EF, but not insertion. This finding could be due to the higher need for monitoring in congruent lexicalization, and inhibitory control in alternation (see Section 1.4.1; Green & Wei, 2014). Insertion practices may not be cognitively taxing enough to elicit adaptive changes in EF. The MSWQ does not measure the frequency of congruent lexicalization and alternation, and thus may not be sensitive to the changes in EF. Additionally, there are some methodological concerns with the MSWQ. According to Rodriguez-Fornells et al. (2012), Unintended Switching (US) is a measure of the lack of intention and awareness of the language switches. However, asking participants to rate the frequency of engaging in a practice that is defined as “lack of intention and awareness” is paradoxical. How can a measure be valid and reliable if it occurs subconsciously? Therefore, it is contended that the MSWQ may not be suitable for bilingual-cognition research. Future studies should consider using the bilingual discourse completion task (DCT) or the code-switching frequency judgement task (CSFT; Hofweber et al., 2018) to investigate the influence of code-switching practices instead of the BSWQ. Both the bilingual DCT and the CSFT provide a basic frequency measure of CS practices. In the bilingual DCT, participants are tasked to write a bilingual email, and the frequency of their code-switches within the text is taken as a conservative estimate of their code-switching in real speech. In the CSFT, participants are tasked to rate how frequently they have encountered alternations, insertions, or congruent lexicalization code-switches. Either of these measures may provide a better measure of code-switching measures that aligns with Muysken's (2000) typology of code-switching.

Finally, future studies should closely review the use of eye-tracking and EEG in experiments so that it correctly aligns with key research questions. For research using the ANT with an eye-tracking paradigm, it is recommended to use saccadic movements to log responses

rather than manual keystrokes (Gorina et al., 2021). This recommendation is to address the issue of RTs not being a pure measure of changes in cognitive processing (See Section 5.2). Next, instead of relying on fixation duration, it is recommended to use valid and invalid cues that require the execution of pro- and anti-saccades. The cognitive mechanisms involved in performing pro- and anti-saccades, fortunately, align with executive control, specifically the need for inhibitory control.

For research using EEG, given that both Comishen and Bialystok (2021) and Morrison et al. (2018) have established that the P3 component is sensitive to N-back task difficulty, it would be interesting to see whether changing the method of language priming used and incorporating 1 and 3-back conditions would have any effects on the outcome. However, it is recommended to carefully consider the number of trials in the task and the sample size recruited (Luck, 2014). A within-subjects design that manipulates both task difficulty and language contexts may require an extremely high number of trials per block. This may induce participant fatigue that negatively impacts the quality of the data collected. In such cases, researchers may consider reducing the number of trials by increasing the sample size or by introducing longer breaks in-between experiments.

7.4 Implications and Significance

The findings from this thesis have important implications for the field. First, the ACH requires revision to bring it up to the modern understanding of multilingualism. While highly controlled lab experiments could prime different interactional contexts, this does not align with how multilinguals switch between contexts naturally. Even if we measured the "dominant" language contexts of multilinguals (Kalamala et al., 2020)—whether they mostly engage in single- or dual-language contexts—it remains unclear at what point one crosses the threshold from being a single-language dominant bilingual to a dual-language dominant bilingual. For instance, if a single utterance is spoken in another language, should that context then be labelled as dual language? Language dominance may change over time, and this factor of time is missing in the current version of the ACH. These changes may occur quickly or gradually and may shift from improvement to decline. They might affect only one language, both languages equally, or both languages in an orthogonal manner—where increased use of one language coincides with decreased use of the other (Schmid & Yilmaz, 2018). Additionally, the ACH does not account for inner speech. Could multilingual inner speech (Resnik, 2018) contribute to the shift between

interactional contexts? While the ACH may be relevant to cultures with compartmentalized language use (e.g., L1 at home, L2 at work; Gullifer & Titone, 2021), the current thesis argues that the predictions of ACH may not be visible in highly integrated language cultures such as Malaysia (see also Wagner et al., 2023).

Second, MLD scores are a great first step in providing a holistic understanding of multilingual experiences, but it is still subject to the pitfalls of self-reported data. Both variations of language entropy (Gullifer & Titone, 2019; Li et al., 2020) are contingent on the ability of multilinguals to accurately estimate how much each language is used in a social context per day (see Section 2.3). However, how can this measure be accurate if languages are code-switched frequently? Take the LHQ language dominance question “Estimate how many hours per day you spend engaged in the following activities in each of the languages you have studied or learned”. How should one estimate the hours spent engaged in a certain language if code-switching occurs within the activity? This question is important for multilingual countries like Malaysia because much psycholinguistic research depends on self-reported data (Olson, 2023). The current thesis demonstrated that while MLD scores are not a superior measure of multilingualism above and beyond traditional measures, it does provide a good enough snapshot of an individual’s multilingual experiences.

Third, these insights empower policymakers to make more informed language policy decisions. While the effects of multilingualism on cognition remain undetermined, there is clear evidence that multilingualism is not disadvantageous. Scores on the cognitive tasks throughout the experiments demonstrate the equal or better performance of Malaysian young-adult multilinguals compared to their peers in other countries. The data on the Malaysian public’s language attitudes revealed that multilingualism is seen as a commodity, aligning with past literature. Public sentiment on multilingualism is positive, the results from Chapter 3 indicate that this is the case. If there is strong empirical evidence in the literature that multilingualism offers numerous benefits at both the individual and societal levels and no cognitive drawbacks to multilingualism are found, then language policies that emphasize monolingual education need to be revised. Integrating multilingual practices and values (e.g., translanguaging) into early education could create a domino effect that strengthens ethnic and national identities, promotes linguistic diversity, enhances cognitive and academic development, and helps ease sociopolitical tensions (Benson, 2002).

Fourth, understanding the consequences of multilingualism on cognition has far-reaching practical implications in the healthcare sector. It is well understood that EF abilities are a fundamental set of cognitive skills responsible for intentional, volitional self-directed control of behavior (Blair, 2016). Impairments of the EF systems are strongly correlated with a broad range of psychological disorders including depression (Nuño et al., 2021), (ADHD; Roselló et al., 2020), bipolar disorder (Dickinson et al., 2017), obsessive-compulsive disorder (OCD; Yun et al., 2017), and schizophrenia (Kamal et al., 2016). Conversely, neurodegenerative diseases such as Alzheimer's (Guarino et al., 2019) and Parkinson's disease (Vlagsma et al., 2018), result in impairments of EF systems. Therefore, understanding how to sustain or enhance EF has important implications for diagnosing and treating these disorders and pathologies. Furthermore, given that many of these conditions are life-long with no functional cure (Kousaie & Taler, 2015; Pliatsikas et al., 2021), even small or moderate effects of non-pharmacological interventions on the rate or onset of cognitive decline at the individual level would have a massive global impact on healthy ageing (del Maschio et al., 2018).

Frequent evidence supporting the relationship between multilingualism and cognition is the finding that bilinguals are associated with a significant delay in the onset of dementia by up to 4 years compared to monolinguals (Bialystok et al., 2007). Since then, there has been increasing evidence demonstrating similar neuroprotective effects of multilingualism (Anderson et al., 2021; Bialystok et al., 2018; Kim et al., 2019). For example, Aveledo et al. (2020) found preliminary evidence indicating that bilingualism provides a neuroprotective buffer against multiple sclerosis. In the same vein, Paplikar et al. (2018) found decreased severity of post-stroke aphasia in bilinguals compared to monolingual counterparts. DeLuca et al. (2019) argue that specific multilingual experiences could positively act against neurocognitive decline and prospective dementia, merely knowing multiple languages does not produce a neuroprotective effect against cognitive decline, pathological or otherwise. Rather, it is the active use of multiple languages that is hypothesized to create a buffer against neurocognitive decline (Hofweber et al., 2020a; 2020b; Kałamała et al., 2022). This pattern of effect also aligns with how other life experiences affect cognitive integrity. For example, enhanced neuroplasticity and cognition have been linked to musical training and aerobic fitness, not merely musical knowledge, and physical exercise (del Maschio et al., 2018). Importantly, multilingualism has never been depicted as a cognitive panacea for age- or pathology-related cognitive decline. Rather, what past research alludes to is that multi-

modal interventions on cognitive development or decline can be improved with the inclusion of multilingualism (Pliatsikas et al., 2021).

The implications of enhancing or sustaining EF are not only limited to healthcare as it has been reported that high EF in early childhood predicts better academic (Pascual et al., 2019), psychosocial (Luerssen & Ayduk, 2017), and occupational (Miller et al., 2012) outcomes in later life. Therefore, an argument can be made whereby fostering EF growth in childhood, especially in low socioeconomic status (SES) contexts (Blom et al., 2014), via multilingual education may create a cascading effect in other developmental areas. That being said, this field is relatively new, and any conclusions from existing research must be scrutinized before they can be integrated into clinical or educational intervention models. The existing hurdle is in understanding the mechanisms and parameters by which the multilingual-cognitive relationship functions.

7.5 Conclusion

Returning to the initial question of this thesis, it is evident that 1) Malaysians have a generally positive language attitude, 2) there is no simple “yes/no” answer to whether language contexts rapidly modulate executive functions in multilinguals, and 3) the measure of language entropy may not fully capture the richness and variety of multilingualism. The present studies found that when a culture practices multilingualism daily, is highly proficient in language use and has an integrated language culture, the predictions of the ACH model may not be relevant. This discrepancy likely arises from a mismatch between the model and the actual experiences of multilingual language use. While the model suggests a step-like cognitive adaptation to different language contexts, multilinguals seem to navigate these contexts fluidly, indicating that cognitive changes are likely gradual and non-linear. It would be misleading to claim cognitive advantages to multilingualism based on these findings, but the indirect evidence suggests that there are, at the least, no cognitive disadvantages. However, this outcome might also be due to the invariance of multilingual experiences in the samples. To gain a deeper understanding of how multilingualism influences cognition, future research should explore the ACH model using distinct bilingual phenotypes. Overall, these studies highlight the need for more representative research from highly multilingual countries. Such research would not only provide new insights into the literature but also help local policymakers create language policies grounded in scientific evidence.

References

- A'Beckett, L., & Du Plessis, T. (2017). In pursuit of societal harmony. In *SUN MeDIA eBooks*.
<https://doi.org/10.18820/9781928424130>
- Aboud, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental Science*, 19(4), 632–656. <https://doi.org/10.1111/desc.12422>
- Abrar-ul-Hassan, S. (2021). Linguistic capital in the university and the hegemony of English: Medieval origins and future directions. *SAGE Open*, 11(2), 215824402110218. <https://doi.org/10.1177/21582440211021842>
- Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20, 3. <https://doi.org/10.1016/j.jneuroling.2006.10.003>
- Abutalebi, J., & Green, D. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition*, 19, 4. <https://doi.org/10.1017/s1366728916000225>
- Abutalebi, J., Rosa, P., Green, D., Hernandez, M., Scifo, P., Keim, R., Cappa, S., & Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, 22, 9. <https://doi.org/10.1093/cercor/bhr287>
- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology*, 19, 6. <https://doi.org/10.1037/0278-6133.19.6.586>
- Adler, R. M., Kroff, J. R., & Novick, J. M. (2020). Does integrating a code-switch during comprehension engage cognitive control? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46, 4. <https://doi.org/10.1037/xlm0000755>
- Albury, N. J. (2019). Beyond economy and culture: language-in-education preferences of Malaysian youth. *Current Issues in Language Planning*, 21(3), 301–319. <https://doi.org/10.1080/14664208.2019.1680161>
- Albury, N. J. (2020). Language policy, ideological clarification and theory of mind. *Language Policy*, 20(2), 193–214. <https://doi.org/10.1007/s10993-020-09547-z>

- Albury, N. J., & Aye, K. K. (2016). Malaysia's national language policy in international theoretical context. *Journal of Nusantara Studies*, 1(1), 71.
<https://doi.org/10.24200/jonus.vol1iss1pp71-84>
- Anderson, J. A. E., Grundy, J. G., Grady, C. L., Craik, F. I. M., & Bialystok, E. (2021). Bilingualism contributes to reserve and working memory efficiency: Evidence from structural and functional neuroimaging. *Neuropsychologia*, 163, 108071.
<https://doi.org/10.1016/j.neuropsychologia.2021.108071>
- Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2017). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, 50(1), 250–263. <https://doi.org/10.3758/s13428-017-0867-9>
- Anderson, J., Hawrylewicz, K., & Bialystok, E. (2018). Who is bilingual? Snapshots across the lifespan. *Bilingualism: Language and Cognition*, 23, 5.
<https://doi.org/10.1017/s1366728918000950>
- Ansaldi, A. I., Ghazi-Saidi, L., & Adrover-Roig, D. (2015). Interference control in elderly bilinguals: Appearances can be misleading. *Journal of Clinical and Experimental Neuropsychology*, 37(5), 455–470. <https://doi.org/10.1080/13803395.2014.990359>
- Antón, E., Carreiras, M., & Duñabeitia, J. A. (2019). The impact of bilingualism on executive functions and working memory in young adults. *PLOS ONE*, 14(2), e0206770.
<https://doi.org/10.1371/journal.pone.0206770>
- Antón, E., Thierry, G., Gborov, A., Anasagasti, J., & Duñabeitia, J. A. (2016). Testing bilingual educational methods: A plea to end the language-mixing taboo. *Language Learning*, 66(S2), 29–50. <https://doi.org/10.1111/lang.12173>
- Antoniou, K., Grohmann, K. K., Kambanaros, M., & Katsos, N. (2016). The effect of childhood bilingualism and multilingualism on executive control. *Cognition*, 149, 18–30.
<https://doi.org/10.1016/j.cognition.2015.12.002>
- Arredondo, M. M., Aslin, R. N., Zhang, M., & Werker, J. F. (2022). Attentional orienting abilities in bilinguals: Evidence from a large infant sample. *Infant Behavior and Development*, 66, 101683. <https://doi.org/10.1016/j.infbeh.2021.101683>
- Arslan, J., & Benke, K. (2023). Statistical analysis of ceiling and floor effects in medical trials. *Applied Biosciences*, 2(4), 668–681. <https://doi.org/10.3390/applbiosci2040042>

- Artamonova, T. (2020). L2 learners' language attitudes and their assessment. *Foreign Language Annals*, 53(4), 807–826. <https://doi.org/10.1111/flan.12498>
- Aveledo, F., Higuera, Y., Marinis, T., Bose, A., Pliatsikas, C., Meldaña-Rivera, A., Martínez-Ginés, María Luisa, García-Domínguez, José Manuel, Lozano-Ros, A., Cuello, J. P., & Goicochea-Briceño, H. (2020). Multiple sclerosis and bilingualism. *Bi-/Multilingualism and the Declining Brain*, 11, 4. <https://doi.org/10.1075/lab.18037.ave>
- Ayres, P., Lee, J. Y., Paas, F., & van Merriënboer, J. J. G. (2021). The validity of physiological measures to identify differences in intrinsic cognitive load. *Frontiers in Psychology*, 12, e702538. <https://doi.org/10.3389/fpsyg.2021.702538>
- Bak, T. (2016). Cooking pasta in La Paz. *Linguistic Approaches to Bilingualism*, 6, 5. <https://doi.org/10.1075/lab.16002.bak>
- Bak, T. H. (2013). The importance of looking in dark places. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 14(1), 1–2. <https://doi.org/10.3109/21678421.2013.760150>
- Bak, T. H. (2016). The impact of bilingualism on cognitive ageing and dementia: Finding a path through a forest of confounding variables. *Linguistic Approaches to Bilingualism*, 6(1-2), 205–226. <https://doi.org/10.1075/lab.15002.bak>
- Bak, T. H., Long, M. R., Vega-Mendoza, M., & Sorace, A. (2016). Novelty, challenge, and practice: The impact of intensive language learning on attentional functions. *PLOS ONE*, 11(4), e0153485. <https://doi.org/10.1371/journal.pone.0153485>
- Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*, 75(6), 959–963. <https://doi.org/10.1002/ana.24158>
- Bak, T., & Mehmedbegovic, D. (2017). 'Healthy linguistic diet: The value of linguistic diversity and language learning across the lifespan. *Languages, Society & Policy*, 1, 1–9. <https://doi.org/10.17863/CAM.9854>
- Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th ed.). Multilingual Matters.
- Bar-yam, Y. (1997). *Dynamics Of Complex Systems*. CRC Press.
- Barac, R., Moreno, S., & Bialystok, E. (2016). Behavioral and electrophysiological differences in executive control between monolingual and bilingual children. *Child Development*, 87(4), 1277–1290. <https://doi.org/10.1111/cdev.12538>

- Barke, E. M., & Williams, D. E. P. (1938). A further study of the comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *British Journal of Educational Psychology*, 8, 1. <https://doi.org/10.1111/j.2044-8279.1938.tb03183.x>
- Barker, R. M., & Bialystok, E. (2019). Processing differences between monolingual and bilingual young adults on an emotion n-back task. *Brain and Cognition*, 134, 29–43. <https://doi.org/10.1016/j.bandc.2019.05.004>
- Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience*, 25(6), 936–951. https://doi.org/10.1162/jocn_a_00354
- Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics*, 35, 5. <https://doi.org/10.1017/s0142716414000174>
- Baumgart, C. Q., & Billick, S. B. (2018). Positive cognitive effects of bilingualism and multilingualism on cerebral function: A review. *Psychiatric Quarterly*, 89(2), 273–283. <https://doi.org/10.1007/s11126-017-9532-9>
- Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(6), 1022–1047. <https://doi.org/10.1037/xlm0000770>
- Beatty-Martínez, A. L., & Titone, D. A. (2021). The quest for signals in noise: Leveraging experiential variation to identify bilingual phenotypes. *Languages*, 6(4), 4. <https://doi.org/10.3390/languages6040168>
- Beaufils, V., & Tomin, J. (2021). Stochastic approach to worldwide language classification: The signals and the noise towards long-range exploration. SocArXiv <https://doi.org/10.31235/osf.io/5swba>
- Beesley, T., Pearson, D., & Le Pelley, M. (2019). Chapter 1 - Eye tracking as a tool for examining cognitive processes. *Biophysical Measurement in Experimental Social Science Research*, 1–30. <https://doi.org/10.1016/b978-0-12-813092-6.00002-2>
- Ben-Gal, I., & Kagan, E. (2021). Information theory: Deep ideas, wide perspectives, and various applications. *Entropy*, 23(2), 232. <https://doi.org/10.3390/e23020232>

- Bennys, K., Portet, F., Touchon, J., & Rondouin, G. (2007). Diagnostic value of event-related evoked potentials N200 and P300 subcomponents in early diagnosis of Alzheimer's Disease and mild cognitive impairment. *Journal of Clinical Neurophysiology*, 24(5), 405–412. <https://doi.org/10.1097/wnp.0b013e31815068d5>
- Benson, C. J. (2002). Real and potential benefits of bilingual programmes in developing countries. *International Journal of Bilingual Education and Bilingualism*, 5(6), 303–317. <https://doi.org/10.1080/13670050208667764>
- Bermúdez-Margaretto, B., Kopytin, G., Myachykov, A., & Shtyrov, Y. (2021). Behavioral and neurophysiological correlates of orthographic learning in L1 and L2 alphabets. In *Advances in intelligent systems and computing*, pp. 345–358. https://doi.org/10.1007/978-3-030-71637-0_40
- Berrios, J., Swain, A. K., & Fricke, M. (2023). Implementing the map task in applied linguistics research: What, how, and why. *Research Methods in Applied Linguistics*, 2(3), e100081. <https://doi.org/10.1016/j.rmal.2023.100081>
- Berroir, P., Ghazi-Saidi, L., Dash, T., Adrover-Roig, D., Benali, H., & Ansaldi, A. I. (2017). Interference control at the response level: Functional networks reveal higher efficiency in the bilingual brain. *Journal of Neurolinguistics*, 43, 4–16. <https://doi.org/10.1016/j.jneuroling.2016.09.007>
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3–11. <https://doi.org/10.1017/s1366728908003477>
- Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 65(4), 229–235. <https://doi.org/10.1037/a0025406>
- Bialystok, E. (2016). The signal and the noise. *Linguistic Approaches to Bilingualism*, 6, 5. <https://doi.org/10.1075/lab.15040.bia>
- Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, 143, 3. <https://doi.org/10.1037/bul0000099>
- Bialystok, E. (2021). Bilingualism as a slice of Swiss cheese. *Frontiers in Psychology*, 12, e769323. <https://doi.org/10.3389/fpsyg.2021.769323>

- Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science*, 19(1), 19–23.
<https://doi.org/10.1177/0963721409358571>
- Bialystok, E., & Craik, F. I. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review*, 29(4), 1246–1269.
<https://doi.org/10.3758/s13423-022-02057-5>
- Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, 45, 2.
<https://doi.org/10.1016/j.neuropsychologia.2006.10.009>
- 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.
<https://doi.org/10.1037/0882-7974.19.2.290>
- Bialystok, E., Craik, F. I. M., & Ryan, J. (2006a). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1341–1354. <https://doi.org/10.1037/0278-7393.32.6.1341>
- Bialystok, E., Craik, F., Green, D., & Gollan, T. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10, 3. <https://doi.org/10.1177/1529100610387084>
- Bialystok, E., Craik, F., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16, 4. <https://doi.org/10.1016/j.tics.2012.03.001>
- Bialystok, E., Anderson, J. a. E., & Grundy, J. G. (2018). Interpreting cognitive decline in the face of cognitive reserve. *Linguistic Approaches to Bilingualism*, 11(4), 484–504.
<https://doi.org/10.1075/lab.18040.bia>
- Bialystok, E., Kroll, J. F., Green, D. W., MacWhinney, B., & Craik, F. I. M. (2015). Publication bias and the validity of evidence. *Psychological Science*, 26(6), 944–946.
<https://doi.org/10.1177/0956797615573759>
- Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, 112(3), 494–500.
<https://doi.org/10.1016/j.cognition.2009.06.014>
- Bianco, J. L. (2014). Domesticating the foreign: Globalization's effects on the place/s of languages. *The Modern Language Journal*, 98(1), 312–325.
<https://www.jstor.org/stable/43651760>

- Blair, C. (2016). Developmental science and executive function. *Current Directions in Psychological Science*, 25, 1. <https://doi.org/10.1177/0963721415622634>
- Blanco-Elorrieta, E., & Caramazza, A. (2021). On the Need for Theoretically Guided Approaches to Possible Bilingual Advantages: An Evaluation of the Potential Loci in the Language and Executive Control Systems. *Neurobiology of Language*, 2(4), 452–463. https://doi.org/10.1162/nol_a_00041
- Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *The Journal of Neuroscience*, 37(37), 9022–9036. <https://doi.org/10.1523/JNEUROSCI.0553-17.2017>
- Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends in Cognitive Sciences*, 22(12), 1117–1126. <https://doi.org/10.1016/j.tics.2018.10.001>
- Blom, E., Boerma, T., Bosma, E., Cornips, L., & Everaert, E. (2017). Cognitive advantages of bilingual children in different sociolinguistic contexts. *Frontiers in Psychology*, 8, e552. <https://doi.org/10.3389/fpsyg.2017.00552>
- Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology*, 128, 105–119. <https://doi.org/10.1016/j.jecp.2014.06.007>
- Blommaert, J., Creve, L., & Willaert, E. (2006). On being declared illiterate: Language-ideological disqualification in Dutch classes for immigrants in Belgium. *Language & Communication*, 26, 1. <https://doi.org/10.1016/j.langcom.2005.03.004>
- Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition*, 118, 2. <https://doi.org/10.1016/j.cognition.2010.10.012>
- Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in stimulus–stimulus inhibition. *Bilingualism: Language and Cognition*, 17(03), 610–629. <https://doi.org/10.1017/s1366728913000564>
- Blumenfeld, H. K., Schroeder, S. R., Bobb, S. C., Freeman, M. R., & Marian, V. (2016). Auditory word recognition across the lifespan: Links between linguistic and nonlinguistic

- inhibitory control in bilinguals and monolinguals. *Linguistic Approaches to Bilingualism*, 6(1-2), 119–146. <https://doi.org/10.1075/lab.14030.blm>
- Bokhorst-Heng, W. D., & Caleon, I. S. (2009). The language attitudes of bilingual youth in multilingual Singapore. *Journal of Multilingual and Multicultural Development*, 30(3), 235–251. <https://doi.org/10.1080/01434630802510121>
- Bosma, E., & Pablos, L. (2020). Switching direction modulates the engagement of cognitive control in bilingual reading comprehension: An ERP study. *Journal of Neurolinguistics*, 55, 100894. <https://doi.org/10.1016/j.jneuroling.2020.100894>
- Bourisly, A. K., & Shuaib, A. (2018). Neurophysiological effects of aging: A P200 ERP study. *Translational Neuroscience*, 9(1), 61–66. <https://doi.org/10.1515/tnsci-2018-0011>
- Bradley, M., & Lang, P. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1. <http://www.uvm.edu/>
- Breznitz, Z., & Meyler, A. (2003). Speed of lower-level auditory and visual processing as a basic factor in dyslexia: Electrophysiological evidence. *Brain and Language*, 85(2), 166–184. [https://doi.org/10.1016/s0093-934x\(02\)00513-8](https://doi.org/10.1016/s0093-934x(02)00513-8)
- Bringas, M. L., Liu, S., Zhang, M., Ibañez, I. P. I., Morales, L. M., Garcia, L. G., Bocourt, V. P., Jahanshahi, M., & Valdes-Sosa, P. A. (2020). Flanker task-elicited event-related potential sources reflect human recombinant erythropoietin differential effects on Parkinson's patients. *Parkinson's Disease*, 2020, 1–10. <https://doi.org/10.1155/2020/8625794>
- Broer, T., Pickersgill, M., & Deary, I. J. (2017). The movement of research from the laboratory to the living room: A case study of public engagement with cognitive science. *Neuroethics*, 9, 2. <https://doi.org/10.1007/s12152-016-9259-6>
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*, 22(6), 723–742. <https://doi.org/10.1037/0012-1649.22.6.723>
- Brown, G., Anderson, A., Shillcock, R., & Yule, G. (1985). *Teaching Talk*. Cambridge University Press.
- Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, 51(1), 5. <https://doi.org/10.5334/pb-51-1-5>

- Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). Predictive validity of the N2 and P3 ERP components to executive functioning in children: A latent-variable analysis. *Frontiers in Human Neuroscience*, 8, e80. <https://doi.org/10.3389/fnhum.2014.00080>
- Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5), 336–349. <https://doi.org/10.1038/nrn3214>
- Bultena, S., Dijkstra, T., & Van Hell, J. G. (2014). Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition*, 18(3), 453–469. <https://doi.org/10.1017/s1366728914000145>
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children. *Neuron*, 33(2), 301–311. [https://doi.org/10.1016/s0896-6273\(01\)00583-9](https://doi.org/10.1016/s0896-6273(01)00583-9)
- Burgaleta, M., Sanjuán, A., Ventura-Campos, N., Sebastian-Galles, N., & Ávila, C. (2016). Bilingualism at the core of the brain. Structural differences between bilinguals and monolinguals revealed by subcortical shape analysis. *NeuroImage*, 125, 437–445. <https://doi.org/10.1016/j.neuroimage.2015.09.073>
- Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, 29(6), 624–630. <https://doi.org/10.1177/0963721420969371>
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, 5, e10729. <https://doi.org/10.1371/journal.pone.0010729>
- Calabria, M., Costa, A., Green, D. W., & Abutalebi, J. (2018). Neural basis of bilingual language control. *Annals of the New York Academy of Sciences*, 1426(1), 221–235. <https://doi.org/10.1111/nyas.13879>
- Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, 130(3), 278–288. <https://doi.org/10.1016/j.cognition.2013.11.015>
- Calvo, N., Ibáñez, A., & García, A. M. (2016). The impact of bilingualism on working memory: A null effect on the whole may not be so on the parts. *Frontiers in Psychology*, 7, e265. <https://doi.org/10.3389/fpsyg.2016.00265>

- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2–14). London: Longman.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
<https://doi.org/10.1093/applin/i.1.1>
- Cargnelutti, E., Tomasino, B., & Fabbro, F. (2022). Effects of linguistic distance on second language brain activations in bilinguals: An exploratory coordinate-based meta-analysis. *Frontiers in Human Neuroscience*, 15, e744489.
<https://doi.org/10.3389/fnhum.2021.744489>
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11(2), 282–298. <https://doi.org/10.1111/j.1467-7687.2008.00675.x>
- Carrow, S. M. A. (1957). Linguistic functioning of bilingual and monolingual children. *Journal of Speech and Hearing Disorders*, 22(3), 371–380. <https://doi.org/10.1044/jshd.2203.371>
- Carthery-Goulart, M. T., Privitera, A. J., & Weekes, B. S. (2023). Does language distance modulate the contribution of bilingualism to cognitive reserve in seniors? A systematic review. *American Journal of Alzheimer's Disease and Other Dementias*, 38, 153331752311672-153331752311672. <https://doi.org/10.1177/15333175231167223>
- Cavallaro, F., & Chin, N. B. (2014). 2. Language in Singapore: From multilingualism to English plus. In *Multilingual Matters* (pp. 33–48). <https://doi.org/10.21832/9781783092529-005>
- Cenoz, J. (2013a). Defining multilingualism. *Annual Review of Applied Linguistics*, 33(33), 3–18. <https://doi.org/10.1017/s026719051300007x>
- Cenoz, J. (2013b). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, 46(1), 71–86.
<https://doi.org/10.1017/s0261444811000218>
- Cespón, J. (2021). Neural processing underlying executive functions in bilinguals: “Heads I win, tails you lose.” *Frontiers in Human Neuroscience*, 15, e710905.
<https://doi.org/10.3389/fnhum.2021.710905>

- Cespón, J., & Carreiras, M. (2020). Is there electrophysiological evidence for a bilingual advantage in neural processes related to executive functions? *Neuroscience & Biobehavioral Reviews*, 118, 315–330. <https://doi.org/10.1016/j.neubiorev.2020.07.030>
- Chan, I. L. & Chang, C. B. (2018). LEXTALE_CH: A quick, character-based proficiency test for Mandarin Chinese. In A. B. Bertolini, & M. J. Kaplan (Eds.), *Proceedings of the 42nd Annual Boston University Conference on Language Development* (pp. 114–130). Cascadilla Press.
- Chan, S. H., & Abdullah, A. N. (2015). Bilingualism in Malaysia: Language education policy and local needs. *Pertanika Journals of Social Sciences and Humanities*, 23, 55–70.
- Chen, C. (2020). Public health messages about COVID-19 prevention in multilingual Taiwan. *Multilingua*, 39(5), 597–606. <https://doi.org/10.1515/multi-2020-0092>
- Chen, L., Perfetti, C. A., & Leng, Y. (2017). Reading Pinyin activates character orthography for highly experienced learners of Chinese. *Bilingualism: Language and Cognition*, 22, 1. <https://doi.org/10.1017/s136672891700058x>
- Chen, Y., Fang, W., Guo, B., & Bao, H. (2021). Fatigue-related effects in the process of task interruption on working memory. *Frontiers in Human Neuroscience*, 15, 703422. <https://doi.org/10.3389/fnhum.2021.703422>
- Chiswick, B., & Miller, P. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development*, 26, 1. <https://doi.org/10.1080/14790710508668395>
- Chrysochoou, E., Vivas, A. B., Cana, L., & Demetriou, A. (2021). The bilingual effect on cognitive development: not an executive function advantage, but a differentiation of mental abilities. *Journal of Cognitive Psychology*, 34(4), 470–484. <https://doi.org/10.1080/20445911.2021.2002875>
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62(1), 73–101. <https://doi.org/10.1146/annurev.psych.093008.100427>
- Chung-Fat-Yim, A., Calvo, N., & Grundy, J. G. (2022). The multifaceted nature of bilingualism and attention. *Frontiers in Psychology*, 13, 910382. <https://doi.org/10.3389/fpsyg.2022.910382>

- Chung-Fat-Yim, A., Poarch, G. J., Comishen, K. J., & Bialystok, E. (2021). Does language context impact the neural correlates of executive control in monolingual and multilingual young adults? *Brain and Language*, 222, 105011.
<https://doi.org/10.1016/j.bandl.2021.105011>
- Cockcroft, K., Wigdorowitz, M., & Liversage, L. (2017). A multilingual advantage in the components of working memory. *Bilingualism: Language and Cognition*, 22, 1.
<https://doi.org/10.1017/s1366728917000475>
- Coderre, E. L., & van Heuven, W. J. B. (2014). The effect of script similarity on executive control in bilinguals. *Frontiers in Psychology*, 5, e1070.
<https://doi.org/10.3389/fpsyg.2014.01070>
- Coderre, E. L., Smith, J. F., van Heuven, W. J. B., & Horwitz, B. (2015). The functional overlap of executive control and language processing in bilinguals. *Bilingualism: Language and Cognition*, 19, 3. <https://doi.org/10.1017/s1366728915000188>
- Comishen, K. J., & Bialystok, E. (2021). Increases in attentional demands are associated with language group differences in working memory performance. *Brain and Cognition*, 147, 105658. <https://doi.org/10.1016/j.bandc.2020.105658>
- Comishen, K. J., Bialystok, E., & Adler, S. A. (2019). The impact of bilingual environments on selective attention in infancy. *Developmental Science*, 22(4), e12797.
<https://doi.org/10.1111/desc.12797>
- 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.
<https://doi.org/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.
<https://doi.org/10.1016/j.cognition.2006.12.013>
- Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2016). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*, 41(6), 1629–1644. <https://doi.org/10.1111/cogs.12434>
- Costello, A., & Osborne, J. (2019). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(7), 7. <https://doi.org/10.7275/jyj1-4868>

- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, 26(2), 197–223. <https://doi.org/10.1007/s10648-013-9246-y>
- Cox, S. R., Bak, T. H., Allerhand, M., Redmond, P., Starr, J. M., Deary, I. J., & MacPherson, S. E. (2016). Bilingualism, social cognition and executive functions: A tale of chickens and eggs. *Neuropsychologia*, 91, 299–306. <https://doi.org/10.1016/j.neuropsychologia.2016.08.029>
- Cullipher, S., Hansen, S. J. R., & VandenPlas, J. R. (2018). Eye Tracking as a Research tool: An Introduction. In *ACS symposium series* (pp. 1–9). <https://doi.org/10.1021/bk-2018-1292.ch001>
- D’Souza, D., Brady, D., Haensel, J. X., & D’Souza, H. (2021). Early bilingual experience is associated with change detection ability in adults. *Scientific Reports*, 11(1), 2068. <https://doi.org/10.1038/s41598-021-81545-5>
- D’Souza, D., Brady, D., Haensel, J., & D’Souza, H. (2020). Is mere exposure enough? The effects of bilingual environments on infant cognitive development. *Royal Society Open Science*, 7, 180191. <https://doi.org/10.1098/rsos.180191>
- Daffner, K. R., Chong, H., Sun, X., Tarbi, E. C., Riis, J. L., McGinnis, S. M., & Holcomb, P. J. (2011). Mechanisms underlying age- and performance-related differences in working memory. *Journal of Cognitive Neuroscience*, 23(6), 1298–1314. <https://doi.org/10.1162/jocn.2010.21540>
- Dama, S. (2020). Singapore educational system as synthesis of τέχνη, humanities and social sciences. Is it a pathway for future education in a globalised world? *Colloquium Pedagogika*, 4(40), 1–26. <https://doi.org/10.34813/31coll2020>
- Dash, T., Berroir, P., Joannette, Y., & Ansaldo, A. I. (2019). Alerting, orienting, and executive control: The effect of bilingualism and age on the subcomponents of attention. *Frontiers in Neurology*, 10, e1122. <https://doi.org/10.3389/fneur.2019.01122>
- Dash, T., Masson-Trottier, M., & Ansaldo, A. I. (2020). Efficiency of attentional processes in bilingual speakers with aphasia. *Aphasiology*, 1–25. <https://doi.org/10.1080/02687038.2020.1719970>

- Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cerebral Cortex*, 18(5), 1201–1209. <https://doi.org/10.1093/cercor/bhm155>
- de Bruin, A., Dick, A. S., & Carreiras, M. (2021). Clear theories are needed to interpret differences: Perspectives on the Bilingual Advantage debate. *Neurobiology of Language*, 2(4), 433–451. https://doi.org/10.1162/nol_a_00038
- de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, 26(1), 99–107. <https://doi.org/10.1177/0956797614557866>
- De Oliveira, G. S., & Oliveira, M. B. (2022). Bilingualism is always cognitively advantageous, but this doesn't mean what you think it means. *Frontiers in Psychology*, 13, 867166. <https://doi.org/10.3389/fpsyg.2022.867166>
- del Maschio, N., Fedeli, D., & Abutalebi, J. (2018). Bilingualism and aging. *Bi-/Multilingualism and the Declining Brain*, 11, 505–519. <https://doi.org/10.1075/lab.18032.del>
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283–321. <https://doi.org/10.1037//0033-295x.93.3.283>
- DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, 116, 7565–7574. <https://doi.org/10.1073/pnas.1811513116>
- DeLuca, V., Segaert, K., Mazaheri, A., & Krott, A. (2020). Understanding bilingual brain function and structure changes? U bet! A unified bilingual experience trajectory model. *Journal of Neurolinguistics*, 56, 100930. <https://doi.org/10.1016/j.jneuroling.2020.100930>
- DeLuca, V., & Voits, T. (2022). Bilingual experience affects white matter integrity across the lifespan. *Neuropsychologia*, 169, 108191. <https://doi.org/10.1016/j.neuropsychologia.2022.108191>
- Demirayak, P., Kıyı, İ., İşbitiren, Y. Ö., & Yener, G. (2023). Cognitive load associates prolonged P300 latency during target stimulus processing in individuals with mild cognitive impairment. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-43132-8>

- Dentella, V., Masullo, C., & Leivada, E. (2024). Bilingual disadvantages are systematically compensated by bilingual advantages across tasks and populations. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-52417-5>
- Department of Statistics. (2015, May). *Department of statistics Malaysia official portal*. https://www.dosm.gov.my/v1/index.php?r=column/cthem&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09&bul_id=MDMxdHZjWtk1SjFzTzNkRXYzcVZjdz09
- Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, 38(3), 352–362. <https://doi.org/10.1037/0012-1649.38.3.352>
- Dick, A., Garcia, N., Pruden, S., Thompson, W., Hawes, S., Sutherland, M., Riedel, M., Laird, A., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour*, 3, 692–701. <https://doi.org/10.1038/s41562-019-0609-3>
- Dickinson, T., Becerra, R., & Coombes, J. (2017). Executive functioning deficits among adults with Bipolar Disorder (types I and II): A systematic review and meta-analysis. *Journal of Affective Disorders*, 218, 407–427. <https://doi.org/10.1016/j.jad.2017.04.010>
- 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(3), 175–197. <https://doi.org/10.1017/s1366728902003012>
- Donkers, F., & Boxtel, van. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition*, 56, 165–176. <https://doi.org/10.1016/j.bandc.2004.04.005>
- Downes, M., Bathelt, J., & De Haan, M. (2017). Event-related potential measures of executive functioning from preschool to adolescence. *Developmental Medicine & Child Neurology*, 59(6), 581–590. <https://doi.org/10.1111/dmcn.13395>
- Duñabeitia, J. A., & Carreiras, M. (2015). The bilingual advantage: Acta est fabula? *Cortex*, 73, 371–372. <https://doi.org/10.1016/j.cortex.2015.06.009>
- Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test--Third Edition. *PsycTESTS Dataset*. <https://doi.org/10.1037/t15145-000>
- Dunst, B., Benedek, M., Jauk, E., Bergner, S., Koschutnig, K., Sommer, M., Ischebeck, A., Spinath, B., Arendasy, M., Böhner, M., Freudenthaler, H., & Neubauer, A. C. (2014).

- Neural efficiency as a function of task demands. *Intelligence*, 42, 22–30.
<https://doi.org/10.1016/j.intell.2013.09.005>
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Wadsworth Cengage Learning.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (eds.). 2022. *Ethnologue: Languages of the World*. Twenty-fifth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>
- Edlin, J. M., & Lyle, K. B. (2013). The effect of repetitive saccade execution on the attention network test: Enhancing executive function with a flick of the eyes. *Brain and Cognition*, 81(3), 345–351. <https://doi.org/10.1016/j.bandc.2012.12.006>
- Eimer, M. (1998). The lateralized readiness potential as an on-line measure of central response activation processes. *Behavior Research Methods, Instruments, & Computers*, 30(1), 146–156. <https://doi.org/10.3758/bf03209424>
- Elliott, M. R., Mungas, D. M., Rentería, M. A., Whitmer, R. A., DeCarli, C., & Fletcher, E. M. (2024). Bilingualism reduces associations between cognition and the brain at baseline, but does not show evidence of cognitive reserve over time. *Bilingualism: Language and Cognition*, 1–9. <https://doi.org/10.1017/S1366728924000105>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149.
<https://doi.org/10.3758/bf03203267>
- Escobar, G. P., Kalashnikova, M., & Escudero, P. (2018). Vocabulary matters! The relationship between verbal fluency and measures of inhibitory control in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 170, 177–189.
<https://doi.org/10.1016/j.jecp.2018.01.012>
- Etz, A., & Vandekerckhove, J. (2017). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin & Review*, 25(1), 5–34. <https://doi.org/10.3758/s13423-017-1262-3>
- European Parliament. (2019). *Multilingualism: The language of the European Union*.
[https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/642207/EPRS_BRI\(2019\)642207_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/642207/EPRS_BRI(2019)642207_EN.pdf)
- Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages after frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery & Psychiatry*, 68(5), 650–652. <https://doi.org/10.1136/jnnp.68.5.650>

- Fabio, R. A., Caprì, T., & Romano, M. (2019). From controlled to automatic processes and back again: The role of contextual features. *Europe's Journal of Psychology*, 15(4), 773–788. <https://doi.org/10.5964/ejop.v15i4.1746>
- Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and brain consequences of conflict. *NeuroImage*, 18(1), 42–57. <https://doi.org/10.1006/nimg.2002.1319>
- Fan, J., Gu, X., Guise, K. G., Liu, X., Fossella, J., Wang, H., & Posner, M. I. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain and Cognition*, 70(2), 209–220. <https://doi.org/10.1016/j.bandc.2009.02.002>
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3), 340–347. <https://doi.org/10.1162/089892902317361886>
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/bf03193146>
- Ferguson, H. J., Victoria, & Elisabeth, B. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports*, 11, e1382. <https://doi.org/10.1038/s41598-020-80866-1>
- 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), 330–336. <https://doi.org/10.1016/j.bandc.2013.09.010>
- Fernandez-Duque, M., Hayakawa, S., & Marian, V. (2023). Speakers of different languages remember visual scenes differently. *Science Advances*, 9(33). <https://doi.org/10.1126/sciadv.adh0064>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (Fifth). SAGE Publications Ltd.
- Filippi, R., Karaminis, T., & Thomas, M. (2014). Language switching in bilingual production: Empirical data and computational modelling. *Bilingualism: Language and Cognition*, 17, 294–315. <https://doi.org/10.1017/s1366728913000485>
- Fitzgerald, J. (2011). Public opinion on bilingual education in Colorado and Massachusetts. *Social Science Journal*, 48(2), 371–396. <https://doi.org/10.1016/j.soscij.2010.07.001>

- Folke, T., Ouzia, J., Bright, P., De Martino, B., & Filippi, R. (2016). A bilingual disadvantage in metacognitive processing. *Cognition*, 150, 119–132.
<https://doi.org/10.1016/j.cognition.2016.02.008>
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1), 152–170.
- Freeman, M. R., Anthony, J. J. D. R., Marian, V., & Blumenfeld, H. K. (2022). Individual and sociolinguistic differences in language background predict Stroop performance. *Frontiers in Communication*, 7, e865965. <https://doi.org/10.3389/fcomm.2022.865965>
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529–534. <https://doi.org/10.1073/pnas.012611199>
- Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, 47(5), 1410–1430.
<https://doi.org/10.1037/a0023750>
- Fyndanis, V., & Lehtonen, M. (2022). Pathological language-switching/mixing and its relationship to domain-general cognitive control. In U. Røyneland & R. Blackwood (Eds.), *Multilingualism across the lifespan* (pp. 209–230). Routledge.
<https://doi.org/10.4324/9781003125815>
- Fyndanis, V., Cameron, S., Hansen, P. B., Norvik, M. I., & Simonsen, H. G. (2023). Multilingualism and verbal short-term/working memory: Evidence from academics. *Bilingualism: Language and Cognition*, 26(3), 490–503.
<https://doi.org/10.1017/s1366728922000621>
- Gallo, F., Myachykov, A., Nelyubina, M., Shtyrov, Y., Kubiak, J., Terekhina, L., & Abutalebi, J. (2023). Linguistic distance dynamically modulates the effects of bilingualism on executive performance in aging. *Bilingualism Language and Cognition*, 1–10.
<https://doi.org/10.1017/s1366728923000743>
- Gallo, F., Novitskiy, N., Myachykov, A., & Shtyrov, Y. (2020). Individual differences in bilingual experience modulate executive control network and performance: behavioral and structural neuroimaging evidence. *Bilingualism Language and Cognition*, 24(2), 293–304. <https://doi.org/10.1017/s1366728920000486>

- Galvao-Carmona, A., Gonzalez-Rosa, J. J., Hidalgo-Muñoz, A. R., Páramo, D., León, R., Izquierdo, G., & Vázquez-Marrufo, M. (2014). Disentangling the attention network test: Behavioral, event related potentials, and neural source analyses. *Frontiers in Human Neuroscience*, 8, e813. <https://doi.org/10.3389/fnhum.2014.00813>
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., Roberts, E. J., Hughes, E. K., & Jones, L. (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, e11. <https://doi.org/10.3389/fpsyg.2014.00011>
- Gayton, A. (2010). Socioeconomic status and language-learning motivation: To what extent does the former influence the latter? *Scottish Languages Review*, 22, 17–28. https://scilt.org.uk/Portals/24/Library/slr/issues/22/3_Gayton_-_Socioeconomic_Status_ML.pdf
- Giavazzi, M., Sambin, S., De Diego-Balaguer, R., Le Stanc, L., Bachoud-Lévi, A. C., & Jacquemot, C. (2018). Structural priming in sentence comprehension: A single prime is enough. *PLOS ONE*, 13(4), e0194959. <https://doi.org/10.1371/journal.pone.0194959>
- Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence*, 52, 71–79. <https://doi.org/10.1016/j.intell.2015.07.006>
- Gill, S. K. (2013). *Language policy challenges in multi-ethnic Malaysia*. Springer.
- Gobet, F., & Sala, G. (2022). Cognitive training: A field in search of a phenomenon. *Perspectives on Psychological Science*, 18(1), 174569162210918. <https://doi.org/10.1177/17456916221091830>
- Gogolin, I. (2021). Multilingualism: A threat to public education or a resource in public education? – European histories and realities. *European Educational Research Journal*, 20(3), 297–310. <https://doi.org/10.1177/1474904120981507>
- Goh, K. S., (1979). Report on the Ministry of Education (Goh Report). Retrieved from <https://www.nas.gov.sg/archivesonline/data/pdfdoc/956--1979-02-10.pdf>
- Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *Journal of Neuroscience*, 33(2), 387–396. <https://doi.org/10.1523/jneurosci.3837-12.2013>

- Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 640–665.
<https://doi.org/10.1037/a0014981>
- Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). What’s easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. *Journal of Experimental Psychology: General*, 143(6), 2167–2195.
<https://doi.org/10.1037/a0038006>
- Gonda, S., Tarrasch, R., & Ben Shalom, D. (2020). The functional significance of the P600. *Medicine*, 99(46), e23116. <https://doi.org/10.1097/md.00000000000023116>
- Goral, M., Campanelli, L., & Spiro, A. (2015). Language dominance and inhibition abilities in bilingual older adults. *Bilingualism: Language and Cognition*, 18(1), 79–89.
<https://doi.org/10.1017/s1366728913000126>
- Gorina, E., Kulikova, A. A., & MacInnes, W. J. (2021). Comparing saccadic and manual responses in the attention network test. *Cortex*, 144, 29–42.
<https://doi.org/10.1016/j.cortex.2021.07.014>
- Gosselin, L., & Sabourin, L. (2023). Language athletes: Dual-language code-switchers exhibit inhibitory control advantages. *Frontiers in Psychology*, 14, e1150159.
<https://doi.org/10.3389/fpsyg.2023.1150159>
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2008). Multivariate latent change modeling of developmental decline in academic intrinsic math motivation and achievement: Childhood through adolescence. *International Journal of Behavioral Development*, 31(4), 317–327. <https://doi.org/10.1177/0165025407077752>
- Green, D. M., & Swets, J. A. (2000). *Signal detection theory and psychophysics*. Los Altos Hills, Calif. Peninsula Publ.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67–81. <https://doi.org/10.1017/s1366728998000133>
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530.
<https://doi.org/10.1080/20445911.2013.796377>

- Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, 29(4), 499–511.
<https://doi.org/10.1080/23273798.2014.882515>
- Greene, D. J., Barnea, A., Herzberg, K., Rassis, A., Neta, M., Raz, A., & Zaidel, E. (2008). Measuring attention in the hemispheres: The lateralized attention network test (LANT). *Brain and Cognition*, 66(1), 21–31. <https://doi.org/10.1016/j.bandc.2007.05.003>
- Grey, S., Sanz, C., Morgan-Short, K., & Ullman, M. T. (2017). Bilingual and monolingual adults learning an additional language: ERPs reveal differences in syntactic processing. *Bilingualism: Language and Cognition*, 21(5), 970–994.
<https://doi.org/10.1017/s1366728917000426>
- Grosjean, F. (1984). *Life with two languages: An introduction to bilingualism* (New edition). Harvard University Press.
- Grosjean, F. (2006). Neurolinguists, beware! The bilingual is not two monolinguals in one person. In K. Bolton and B. B. Kachru (Eds.). *World Englishes: Critical Concepts in Linguistics, Volume IV* (202–213). Abingdon: Routledge.
- Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science*, 4, 177–199.
<https://doi.org/10.1007/s41809-020-00062-5>
- Grundy, J. G., Anderson, J. A. E., & Bialystok, E. (2017). Neural correlates of cognitive processing in monolinguals and bilinguals. *Annals of the New York Academy of Sciences*, 1396(1), 183–201. <https://doi.org/10.1111/nyas.13333>
- Grundy, J. G., Pavlenko, E., & Bialystok, E. (2020). Bilingualism modifies disengagement of attention networks across the scalp: A multivariate ERP investigation of the IOR paradigm. *Journal of Neurolinguistics*, 56, 100933.
<https://doi.org/10.1016/j.jneuroling.2020.100933>
- Grundy, J. G., & Timmer, K. (2016). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, 33, 325–340.
<https://doi.org/10.1177/0267658316678286>
- Guarino, A., Favieri, F., Boncompagni, I., Agostini, F., Cantone, M., & Casagrande, M. (2019). Executive functions in Alzheimer disease: A systematic review. *Frontiers in Aging Neuroscience*, 10, e437. <https://doi.org/10.3389/fnagi.2018.00437>

- Gullifer, J. (2020). jasangullifer/languageEntropy: Release v1.0.1c (v1.0.1c). Zenodo.
<https://doi.org/10.5281/zenodo.3745674>
- Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D. (2018). Bilingual experience and resting-state brain connectivity: Impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia*, 117, 123–134. <https://doi.org/10.1016/j.neuropsychologia.2018.04.037>
- Gullifer, J. W., Kousaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., Klein, D., Baum, S., Phillips, N., & Titone, D. (2020). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, 42(2), 245–278.
<https://doi.org/10.1017/s0142716420000521>
- Gullifer, J. W., & Titone, D. (2019). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23, 283–294.
<https://doi.org/10.1017/s1366728919000026>
- Gullifer, J. W., & Titone, D. (2021). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General*, 150(3), 414–430. <https://doi.org/10.1037/xge0000933>
- Haith, A. M., Pakpoor, J., & Krakauer, J. W. (2016). Independence of movement preparation and movement initiation. *Journal of Neuroscience*, 36(10), 3007–3015.
<https://doi.org/10.1523/jneurosci.3245-15.2016>
- Hällsten, M., & Thaning, M. (2021). Wealth as one of the “Big four” SES dimensions in intergenerational transmissions. *Social Forces*, 100(4), 1533–1560.
<https://doi.org/10.1093/sf/soab080>
- Halsey, L. G. (2019). The reign of the p-value is over: What alternative analyses could we employ to fill the power vacuum? *Biology Letters*, 15(5), 20190174.
<https://doi.org/10.1098/rsbl.2019.0174>
- Han, X., Wei, L., & Filippi, R. (2022). Modulating effects of interactional contexts on bilinguals’ cognitive control: Evidence for the Adaptive Control Hypothesis. *International Journal of Bilingualism*, 27(5), 548–568. <https://doi.org/10.1177/13670069221102011>

- Hannaway, N., Opitz, B., & Sauseng, P. (2017). Exploring the bilingual advantage: Manipulations of similarity and second language immersion in a Stroop task. *Cognitive Neuroscience, 10*(1), 1–12. <https://doi.org/10.1080/17588928.2017.1389874>
- Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition, 150*, 10–19. <https://doi.org/10.1016/j.cognition.2016.01.016>
- Hartanto, A., & Yang, H. (2019). The role of bilingual interactional contexts in predicting interindividual variability in executive functions: A latent variable analysis. *Journal of Experimental Psychology: General, 149*, 609–633. <https://doi.org/10.1037/xge0000672>
- Hart-Gonzalez, L., & Lindemann, S. (1993). Expected achievement in speaking proficiency. *School of Language Studies, Foreign Service Institute, Department of State*.
- Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one- and two-factor fixed effects ANOVA cases. *Journal of Educational Statistics, 17*(4), 315. <https://doi.org/10.2307/1165127>
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation, 22*, 193–225. [https://doi.org/10.1016/s0079-7421\(08\)60041-9](https://doi.org/10.1016/s0079-7421(08)60041-9)
- Haugen, E. (1972). Dialect, Language, Nation. In C. B. Paulston and G. R. Tucke (Eds.), *Sociolinguistics: The Essential Readings* (pp. 411–422). Wiley-Blackwell.
- Haukås, Å., Storto, A., & Tiurikova, I. (2022). School students' beliefs about the benefits of multilingualism. *Journal of Multilingual and Multicultural Development, 45*(7), 2817–2830. <https://doi.org/10.1080/01434632.2022.2075001>
- Hayakawa, S., & Marian, V. (2019). Consequences of multilingualism for neural architecture. *Behavioral and Brain Functions, 15*(1). <https://doi.org/10.1186/s12993-019-0157-z>
- Hayakawa, S., & Marian, V. (2020). Studying Bilingualism through Eye-Tracking and Brain Imaging. In *Cambridge University Press eBooks* (pp. 273–299). <https://doi.org/10.1017/9781316535967.012>
- Hebb, D. O. (1949). *The organization of behavior*. Wiley.
- 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(3), 315–325.
<https://doi.org/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, e234.
<https://doi.org/10.3389/fpsyg.2011.00234>
- Hessels, R. S., Holleman, G. A., Kingstone, A., Hooge, I. T. C., & Kemner, C. (2019). Gaze allocation in face-to-face communication is affected primarily by task structure and social context, not stimulus-driven factors. *Cognition*, 184, 28–43.
<https://doi.org/10.1016/j.cognition.2018.12.005>
- Hessels, R. S., & Hooge, I. T. C. (2019). Eye tracking in developmental cognitive neuroscience – The good, the bad and the ugly. *Developmental Cognitive Neuroscience*, 40, 100710.
<https://doi.org/10.1016/j.dcn.2019.100710>
- Hilchey, M. D., Hashish, M., MacLean, G. H., Satel, J., Ivanoff, J., & Klein, R. M. (2014). On the role of eye movement monitoring and discouragement on inhibition of return in a go/no-go task. *Vision Research*, 96, 133–139.
<https://doi.org/10.1016/j.visres.2013.11.008>
- Hilchey, M., & Klein, R. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, 18, 625–658. <https://doi.org/10.3758/s13423-011-0116-7>
- Hilchey, M. O., Saint-Aubin, J., & Klein, R. M. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. W. Schwieter, (Ed.) *The Cambridge Handbook of Bilingual Processing* (pp. 586–613). Cambridge University Press. <https://doi.org/10.1017/CBO9781107447257.026>
- Hillyard, S. (2009). Event-Related Potentials (ERPs) and cognitive processing. In *Elsevier eBooks* (pp. 13–18). <https://doi.org/10.1016/b978-008045046-9.00311-9>
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, 49(1), 4–14. <https://doi.org/10.1037/a0027238>
- Hoff, E. (2021). Why bilingual development is not easy. *Advances in Child Development and Behavior*, 129–167. <https://doi.org/10.1016/bs.acdb.2021.03.002>

- Hoff, E., & Ribot, K. M. (2015). Language development: influence of Socio-Economic status. In *Elsevier eBooks* (pp. 324–328). <https://doi.org/10.1016/b978-0-08-097086-8.23132-2>
- Hoff, E., Trecca, F., Højen, A., Laursen, B., & Bleses, D. (2024). Context and education affect the quality of parents' speech to children. *Journal of Applied Developmental Psychology*, *91*, e101632. <https://doi.org/10.1016/j.appdev.2024.101632>
- Hofweber, J., Marinis, T., & Treffers-Daller, J. (2016). Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism*, *6*(5), 648–668. <https://doi.org/10.1075/lab.15052.hof>
- Hofweber, J., Marinis, T., & Treffers-Daller, J. (2018). Predicting executive functions in bilinguals using ecologically valid measures of code-switching behavior. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 181–205). John Benjamins. <https://doi.org/10.1075/sibil.54>
- Hofweber, J., Marinis, T., & Treffers-Daller, J. (2020a). Experimentally induced language modes and regular code-switching habits boost bilinguals' executive performance: Evidence from a within-subject paradigm. *Frontiers in Psychology*, *11*, e542326. <https://doi.org/10.3389/fpsyg.2020.542326>
- Hofweber, J., Marinis, T., & Treffers-Daller, J. (2020b). How different code-switching types modulate bilinguals' executive functions: A dual control mode perspective. *Bilingualism: Language and Cognition*, *23*, 909–925. <https://doi.org/10.1017/s1366728919000804>
- Hommel, B., & Dreisbach, G. (2022). Editorial: The editor's challenge: Cognitive resources. *Frontiers in Psychology*, *13*, e994801. <https://doi.org/10.3389/fpsyg.2022.994801>
- Howe, A. S. (2014). Meta-analysis of the endogenous N200 latency event-related potential subcomponent in patients with Alzheimer's disease and mild cognitive impairment. *Clinical Neurophysiology*, *125*(6), 1145–1151. <https://doi.org/10.1016/j.clinph.2013.10.019>
- Hsu, H. (2014). Effects of bilingualism and trilingualism in L2 production: Evidence from errors and self-repairs in early balanced bilingual and trilingual adults. *Journal of Psycholinguistic Research*, *43*(4), 357–379. <https://doi.org/10.1007/s10936-013-9257-3>
- Hu, X., Sidhu, G. K., & Lu, X. (2022). Relationship between growth mindset and English language performance among Chinese EFL university students: The mediating roles of

- grit and foreign language enjoyment. *Frontiers in Psychology*, 13, e935506.
<https://doi.org/10.3389/fpsyg.2022.935506>
- Hübner, R., & Töbel, L. (2019). Conflict resolution in the Eriksen flanker task: Similarities and differences to the Simon task. *PLOS ONE*, 14(3), e0214203.
<https://doi.org/10.1371/journal.pone.0214203>
- Hudon, M.-E. (2023). *The Official Languages Act: Understanding Its Principles and Implementation*. Library of Parliament.
https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/201155E#:~:text=2a%20and%202b-
- Huster, R. J., Messel, M. S., Thunberg, C., & Raud, L. (2020). The P300 as marker of inhibitory control: Fact or fiction? *Cortex*, 132, 334–348.
<https://doi.org/10.1016/j.cortex.2020.05.021>
- Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain and Cognition*, 68(3), 327–340. <https://doi.org/10.1016/j.bandc.2008.08.021>
- Incera, S. (2018). Measuring the timing of the bilingual advantage. *Frontiers in Psychology*, 9, e1983. <https://doi.org/10.3389/fpsyg.2018.01983>
- Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (2007). Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology*, 30(1), 10–22. <https://doi.org/10.1111/j.1469-8986.1993.tb03200.x>
- Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching*, 10, 89–109.
<https://doi.org/10.14746/ssllt.2020.10.1.5>
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18(4), 394–412.
<https://doi.org/10.1080/09658211003702171>
- Jaekel, N., Jaekel, J., Willard, J., & Leyendecker, B. (2019). No evidence for effects of Turkish immigrant children's bilingualism on executive functions. *PLOS ONE*, 14(1), e0209981.
<https://doi.org/10.1371/journal.pone.0209981>
- Jahrani, A., & Listia, R. (2023). The impact of exposure on second language acquisition. *Journal of Language Teaching and Learning, Linguistics and Literature*, 11(1), 416–424.
<https://doi.org/10.24256/ideas.v11i1.3581>

- Janic, A., Cavanagh, P., & Rivest, J. (2020). Effect of bilingualism on visual tracking attention and resistance to distraction. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-71185-6>
- Jared, D. (2015). What is the theory? *Cortex*, 73, 361–363.
<https://doi.org/10.1016/j.cortex.2015.07.009>
- Javan, S. S., & Ghonsooly, B. (2017). Learning a foreign language: A new path to enhancement of cognitive functions. *Journal of Psycholinguistic Research*, 47(1), 125–138.
<https://doi.org/10.1007/s10936-017-9518-7>
- Jeon, H., & Friederici, A. D. (2013). Two principles of organization in the prefrontal cortex are cognitive hierarchy and degree of automaticity. *Nature Communications*, 4(1).
<https://doi.org/10.1038/ncomms3041>
- Jiao, L., Grundy, J. G., Liu, C., & Chen, B. (2020). Language context modulates executive control in bilinguals: Evidence from language production. *Neuropsychologia*, 142, 107441. <https://doi.org/10.1016/j.neuropsychologia.2020.107441>
- Jiao, L., Liu, C., Liang, L., Plummer, P., Perfetti, C. A., & Chen, B. (2019). The contributions of language control to executive functions: From the perspective of bilingual comprehension. *Quarterly Journal of Experimental Psychology*, 72, 1984–1997.
<https://doi.org/10.1177/1747021818821601>
- Jiao, L., Liu, C., Schwieter, J. W., & Chen, B. (2021). Switching between newly learned languages impacts executive control. *Psychophysiology*, 58(10).
<https://doi.org/10.1111/psyp.13888>
- Jiao, L., Liu, C., Wang, R., & Chen, B. (2017). Working memory demand of a task modulates bilingual advantage in executive functions. *International Journal of Bilingualism*, 23(1), 102–117. <https://doi.org/10.1177/1367006917709097>
- Jovanovic, D., & Matejevic, M. (2014). Relationship between rewards and intrinsic motivation for learning: Researches review. *Procedia - Social and Behavioral Sciences*, 149(149), 456–460. <https://doi.org/10.1016/j.sbspro.2014.08.287>
- Judd, N., Sauce, B., Wiedenhoeft, J., Tromp, J., Chaarani, B., Schliep, A., Noort, van, Penttilä, J., Grimmer, Y., Insensee, C., Becker, A., Banaschewski, T., Arun, B., Quinlan, E. B., Desrivières, S., Flor, H., Grigis, A., Gowland, P., Heinz, A., & Ittermann, B. (2020). Cognitive and brain development is independently influenced by socioeconomic status

- and polygenic scores for educational attainment. *Proceedings of the National Academy of Sciences*, 117, 12411–12418. <https://doi.org/10.1073/pnas.2001228117>
- Kałamala, P., Drożdżowicz, A., Szewczyk, J., Marzecová, A., & Wodniecka, Z. (2018). Task strategy may contribute to performance differences between monolinguals and bilinguals in cognitive control tasks: ERP evidence. *Journal of Neurolinguistics*, 46, 78–92. <https://doi.org/10.1016/j.jneuroling.2017.12.013>
- Kałamala, P., Senderecka, M., & Wodniecka, Z. (2021). On the multidimensionality of bilingualism and the unique role of language use. *Bilingualism: Language and Cognition*, 25(3), 471–483. <https://doi.org/10.1017/s1366728921001073>
- Kałamala, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition*, 204, 104373. <https://doi.org/10.1016/j.cognition.2020.104373>
- Kałamala, P., Walther, J., Zhang, H., Diaz, M., Senderecka, M., & Wodniecka, Z. (2021). The use of a second language enhances the neural efficiency of inhibitory control: An ERP study. *Bilingualism Language and Cognition*, 25(1), 163–180. <https://doi.org/10.1017/s1366728921000389>
- Kamal, S., Kamal, A., Bayoumy, H., Abdel Mawla, Shereen, Roshdy, R., & Garas, O. (2016). Executive dysfunction in patients with schizophrenia and bipolar disorders. *Middle East Current Psychiatry*, 23, 79–84. <https://doi.org/10.1097/01.xme.0000475262.10853.09>
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 615–622. <https://doi.org/10.1037/0278-7393.33.3.615>
- Kang, C., & Lust, B. (2018). Code-switching does not predict Executive Function performance in proficient bilingual children: Bilingualism does. *Bilingualism: Language and Cognition*, 22, 366–382. <https://doi.org/10.1017/s1366728918000299>
- Kang, K., Xiao, Y., Yu, H., Diaz, M. T., & Zhang, H. (2023). Multilingual language diversity protects native language production under different control demands. *Brain Sciences*, 13(11), 1587–1587. <https://doi.org/10.3390/brainsci13111587>

- Kapong, J. M. (2022). Othering: When Malaysians are mocked for “speaking London.” *ANP Journal of Social Science and Humanities*, 3(Special1), 45–51.
<https://doi.org/10.53797/anp.jssh.v3sp.6.2021>
- Karimi, M. N., & Rad, Z. H. (2021). Bilingual advantage hypothesis: Testing the fit among L1-dominant child and adolescent bilinguals. *Journal of Neurolinguistics*, 60, 101017.
<https://doi.org/10.1016/j.jneuroling.2021.101017>
- Katz, B., & Shah, P. (2017). The role of child socioeconomic status in cognitive training outcomes. *Journal of Applied Developmental Psychology*, 53, 139–150.
<https://doi.org/10.1016/j.appdev.2017.10.003>
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The language experience and proficiency questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 23, 945–950. <https://doi.org/10.1017/s1366728919000038>
- Keith, T. Z. (2019). *Multiple Regression and Beyond* (3rd ed.). Routledge.
- Keye, D., Wilhelm, O., Oberauer, K., & Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: Testing means and covariance hypothesis about the Simon and the Eriksen Flanker task. *Psychological Research*, 73(6), 762–776.
<https://doi.org/10.1007/s00426-008-0188-9>
- Kheder, S., & Kaan, E. (2021). Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition*, 209, 104575. <https://doi.org/10.1016/j.cognition.2020.104575>
- Kim, S., Jeon, S. G., Nam, Y., Kim, H. soo, Yoo, D.-H., & Moon, M. (2019). Bilingualism for dementia: Neurological mechanisms associated with functional and structural changes in the brain. *Frontiers in Neuroscience*, 13, e1224. <https://doi.org/10.3389/fnins.2019.01224>
- Kircher, R., & Zipp, L. (2022). *Research methods in language attitudes*. Cambridge University Press.
- Kleinman, D., & Gollan, T. H. (2016). Speaking two languages for the price of one. *Psychological Science*, 27(5), 700–714. <https://doi.org/10.1177/0956797616634633>
- Knight, D., Morris, S., Tovey, B., Fitzpatrick, T., & Anthony, L. (2020). *Yr Amliadur: Frequency lists for contemporary Welsh (Version 1.0.0)*.
<https://doi.org/10.17035/d.2020.0120164107>
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577. <https://doi.org/10.1017/s0048577201990559>

- Kong, A. P.-H., Abutalebi, J., Lam, K. S.-Y., & Weekes, B. (2014). Executive and language control in the multilingual brain. *Behavioural Neurology*, 2014, 1–7.
<https://doi.org/10.1155/2014/527951>
- Kopp, B., Steinke, A., & Visalli, A. (2020). Cognitive flexibility and N2/P3 event-related brain potentials. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-66781-5>
- Kousaie, S., Chai, X. J., Sander, K. M., & Klein, D. (2017). Simultaneous learning of two languages from birth positively impacts intrinsic functional connectivity and cognitive control. *Brain and Cognition*, 117, 49–56. <https://doi.org/10.1016/j.bandc.2017.06.003>
- Kousaie, S., & Taler, V. (2015). The bilingual advantage: Elusive but worth the effort? *Cortex*, 73, 338–339. <https://doi.org/10.1016/j.cortex.2015.07.019>
- Kovacs, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, 106(16), 6556–6560.
<https://doi.org/10.1073/pnas.0811323106>
- Kremin, L. V., & Byers-Heinlein, K. (2021). Why not both? Rethinking categorical and continuous approaches to bilingualism. *International Journal of Bilingualism*, 25(6), 1560–1575. <https://doi.org/10.1177/13670069211031986>
- Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language*, 128(1), 34–40. <https://doi.org/10.1016/j.bandl.2013.11.006>
- Kroll, J. F. (2008). Juggling two languages in one mind. *Psychological Science Agenda*, 22, 1–7.
<https://www.apa.org/science/about/psa/2008/01>
- Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25(5), 497–514.
<https://doi.org/10.1080/20445911.2013.799170>
- Kroll, J. F., & Dussias, P. E. (2017). The benefits of multilingualism to the personal and professional development of residents of the US. *Foreign Language Annals*, 50(2), 248–259. <https://doi.org/10.1111/flan.12271>
- Kuipers, J.-R., & Thierry, G. (2013). ERP-pupil size correlations reveal how bilingualism enhances cognitive flexibility. *Cortex*, 49(10), 2853–2860.
<https://doi.org/10.1016/j.cortex.2013.01.012>

- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62(1), 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Kuzyk, O., Friend, M., Severdija, V., Zesiger, P., & Poulin-Dubois, D. (2019). Are there cognitive benefits of code-switching in bilingual children? A longitudinal study. *Bilingualism: Language and Cognition*, 23(3), 1–12. <https://doi.org/10.1017/s1366728918001207>
- Kyllonen, P., Carrasco, C. A., & Kell, H. (2017). Fluid Ability (GF) and complex problem solving (CPS). *Journal of Intelligence*, 5(3), 28. <https://doi.org/10.3390/jintelligence5030028>
- Kyriazos, T. A., & Stalikas, A. (2018). Applied psychometrics: The steps of scale development and standardization process. *Psychology*, 09(11), 2531–2560. <https://doi.org/10.4236/psych.2018.911145>
- Lai, G., & O'Brien, B. A. (2020). Examining language switching and cognitive control through the adaptive control hypothesis. *Frontiers in Psychology*, 11, e1171. <https://doi.org/10.3389/fpsyg.2020.01171>
- Laketa, A., Studenica, A., Chrysochoou, E., Blakey, E., & Vivas, A. B. (2021). Biculturalism, linguistic distance, and bilingual profile effects on the bilingual influence on cognition: A comprehensive multipopulation approach. *Journal of Experimental Psychology General*, 150(11), 2273–2292. <https://doi.org/10.1037/xge0000794>
- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling*. Cambridge University Press.
- Lee, S. K., Lee, Y. S., Wong, F. F., & Ya'acob, A. (2010). The English language and its impact on identities of multilingual Malaysian undergraduates. *GEMA Online Journal of Language Studies*, 10(1), 87–101.
- Lee, S. T., Van Heuven, W. J. B., Price, J. M., & Leong, C. X. R. (2022). Translation norms for Malay and English words: The effects of word class, semantic variability, lexical characteristics, and language proficiency on translation. *Behavior Research Methods*, 55(7), 3585–3601. <https://doi.org/10.3758/s13428-022-01977-3>

- Lee, S. T., Van Heuven, W. J. B., Price, J. M., & Leong, C. X. R. (2023). LexMAL: A quick and reliable lexical test for Malay speakers. *Behavior Research Methods*, 56(5), 4563–4581. <https://doi.org/10.3758/s13428-023-02202-5>
- Lee, Y. J., Jeong, M. Y., Park, S., Kim, J. H., Kim, J.-S., & Woo, S.-I. (2020). Age-related changes in auditory Nogo-N200 latency in medication-naïve children and adolescents with Attention-Deficit/Hyperactivity Disorder. *Psychiatry Investigation*, 17(7), 702–709. <https://doi.org/10.30773/pi.2020.0083>
- Lehle, C., & Hübner, R. (2008). On-the-fly adaptation of selectivity in the flanker task. *Psychonomic Bulletin & Review*, 15(4), 814–818. <https://doi.org/10.3758/pbr.15.4.814>
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, 144, 394–425. <https://doi.org/10.1037/bul0000142>
- Leigh, R. J., & Kennard, C. (2004). Using saccades as a research tool in the clinical neurosciences. *Brain*, 127(3), 460–477. <https://doi.org/10.1093/brain/awh035>
- Leivada, E. (2023). A classification bias and an exclusion bias jointly overinflated the estimation of publication biases in bilingualism research. *Behavioral Sciences*, 13(10), 812. <https://doi.org/10.3390/bs13100812>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(01). <https://doi.org/10.1017/s0140525x99001776>
- Lewendon, J., Mortimore, L., & Egan, C. (2020). The phonological mapping Negativity: history, inconsistency, and future direction. *Frontiers in Psychology*, 11, e1967. <https://doi.org/10.3389/fpsyg.2020.01967>
- Lewis, K., & Davies, I. (2018). Understanding media opinion on bilingual education in the United States. *Journal of Social Science Education*, 17(4), 40–67. <https://doi.org/10.4119/UNIBI/jsse-v17-i3-1096>

- Li, H., Qu, J., Chen, C., Chen, Y., Xue, G., Zhang, L., Lu, C., & Mei, L. (2018). Lexical learning in a new language leads to neural pattern similarity with word reading in native language. *Human Brain Mapping, 40*(1), 98–109. <https://doi.org/10.1002/hbm.24357>
- Li, L., & Smith, D. M. (2021). Neural efficiency in athletes: A systematic review. *Frontiers in Behavioral Neuroscience, 15*, e698555. <https://doi.org/10.3389/fnbeh.2021.698555>
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods, 38*(2), 202–210. <https://doi.org/10.3758/bf03192770>
- Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition, 23*, 938–944. <https://doi.org/10.1017/s1366728918001153>
- Li, T., Wang, L., Huang, W.-Y., Zhen, Y., Zhong, C., Qu, Z., & Ding, Y. (2018). Onset time of inhibition of return is a promising index for assessing cognitive functions in older adults. *Journals of Gerontology, 75*(4). <https://doi.org/10.1093/geronb/gby070>
- Li, W. (2011). Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Journal of Pragmatics, 43*(5), 1222–1235. <https://doi.org/10.1016/j.pragma.2010.07.035>
- Li, X., Ng, K. K., Ju, J., Lee, J. W., Zhou, J. H., & Quin, Y. W. (2021). Bilingual language entropy influences executive functions through functional connectivity and signal variability. *Brain and Language, 222*, 105026. <https://doi.org/10.1016/j.bandl.2021.105026>
- Lijffijt, M., Lane, S. D., Meier, S. L., Boutros, N. N., Burroughs, S., Steinberg, J. L., Gerard Moeller, F., & Swann, A. C. (2009). P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory. *Psychophysiology, 46*(5), 1059–1068. <https://doi.org/10.1111/j.1469-8986.2009.00845.x>
- Lin, Z., & Lei, L. (2020). The research trends of multilingualism in applied linguistics and education (2000–2019): A bibliometric analysis. *Sustainability, 12*, 6058. <https://doi.org/10.3390/su12156058>
- Linck, J. A., Hoshino, N., & Kroll, J. F. (2009). Cross-language lexical processes and inhibitory control. *The Mental Lexicon, 3*, 349–374. <https://doi.org/10.1075/ml.3.3.06lin>

- Ling, D. S., Wong, C. D., & Diamond, A. (2016). Do children need reminders on the day–night task, or simply some way to prevent them from responding too quickly? *Cognitive Development*, 37, 67–72. <https://doi.org/10.1016/j.cogdev.2015.10.003>
- Liu, A. H., & Ricks, J. I. (2012). Coalitions and language politics: Policy shifts in Southeast Asia. *World Politics*, 64(3), 476–506. <https://www.jstor.org/stable/41682887>
- Liu, C., Yang, C.-L., Jiao, L., Schwieter, J. W., Sun, X., & Wang, R. (2019). Training in language switching facilitates bilinguals' monitoring and inhibitory control. *Frontiers in Psychology*, 10, e1839. <https://doi.org/10.3389/fpsyg.2019.01839>
- Liu, J., Li, K., Yeh, S., & Chien, S. (2022). Assessing perceptual load and cognitive load by fixation-related information of eye movements. *Sensors*, 22(3), 1187. <https://doi.org/10.3390/s22031187>
- Liu, J., & Wei, C. (2023). An review of Taiwanese educational leadership. *Chinese Studies*, 12(02), 113–125. <https://doi.org/10.4236/chnstd.2023.122011>
- Liu, Q., & Liu, Z. (2021). A review on the relationship between bilingualism and working memory. *Open Journal of Modern Linguistics*, 11(02), 121–134. <https://doi.org/10.4236/ojml.2021.112010>
- Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance “F” Test. *Review of Educational Research*, 66(4), 579–619. <https://doi.org/10.2307/1170654>
- Ljungberg, J. K., Elbe, P., & Sörman, D. E. (2019). The bilingual effects of linguistic distances on episodic memory and verbal fluency. *Scandinavian Journal of Psychology*, 61, 195–203. <https://doi.org/10.1111/sjop.12609>
- Ljungberg, J. K., Hansson, P., Andrés, P., Josefsson, M., & Nilsson, L.-G. (2013). A longitudinal study of memory advantages in bilinguals. *PLoS ONE*, 8(9), e73029. <https://doi.org/10.1371/journal.pone.0073029>
- Loh, S. Y. (2021). *Auditory Priming and the Adaptive Control Hypothesis* [Unpublished undergraduate dissertation]. University of Nottingham, Malaysia.
- Long, K., & Renbarger, R. (2023). Persistence of poverty: How measures of socioeconomic status have changed over time. *Educational Researcher*, 52(3), 144–154. <https://doi.org/10.3102/0013189x221141409>

- López Zunini, R. A., Knoefel, F., Lord, C., Dzuali, F., Breau, M., Sweet, L., Goubran, R., & Taler, V. (2016). Event-related potentials elicited during working memory are altered in mild cognitive impairment. *International Journal of Psychophysiology*, 109, 1–8.
<https://doi.org/10.1016/j.ijpsycho.2016.09.012>
- López-Penadés, R., Sanchez-Azanza, V. A., Aguilar-Mediavilla, E., Buil-Legaz, L., & Adrover-Roig, D. (2020). Associations of natural language switching with executive control in early bilinguals. *Bilingualism: Language and Cognition*, 23, 1008–1019.
<https://doi.org/10.1017/s1366728920000012>
- López-Rojas, C., Rossi, E., Marful, A., & Bajo, M. T. (2022). Prospective memory in bilinguals and monolinguals: ERP and behavioural correlates of prospective processing in bilinguals. *Brain and Language*, 225, 105059.
<https://doi.org/10.1016/j.bandl.2021.105059>
- López-Soto, T. (2010). Chapter 7. A cascade model of lexical access to explain the phonological activation of recently practiced lexical items. *Multilingual Matters EBooks*, 99–111.
<https://doi.org/10.21832/9781847692900-008>
- Lövdén, M., Garzón, B., & Lindenberger, U. (2020). Human skill learning: Expansion, exploration, selection, and refinement. *Current Opinion in Behavioral Sciences*, 36, 163–168. <https://doi.org/10.1016/j.cobeha.2020.11.002>
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.
- Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance*, 20(4), 887–904. <https://doi.org/10.1037//0096-1523.20.4.887>
- Luerssen, A., & Ayduk, O. (2017). Executive Functions Promote Well-Being: Outcomes and mediators. In *Springer* (pp. 59–75). https://doi.org/10.1007/978-3-319-58763-9_4
- Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition*, 74(3), 347–357.
<https://doi.org/10.1016/j.bandc.2010.09.004>

- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25, 605–621.
<https://doi.org/10.1080/20445911.2013.795574>
- Luk, G., Bialystok, E., Craik, F. I. M., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience*, 31(46), 16808–16813.
<https://doi.org/10.1523/jneurosci.4563-11.2011>
- Lukasik, K. M., Lehtonen, M., Soveri, A., Waris, O., Jylkkä, J., & Laine, M. (2018). Bilingualism and working memory performance: Evidence from a large-scale online study. *PLOS ONE*, 13(11), e0205916. <https://doi.org/10.1371/journal.pone.0205916>
- Lyle, K. B., & Martin, J. M. (2010). Bilateral saccades increase intrahemispheric processing but not interhemispheric interaction: Implications for saccade-induced retrieval enhancement. *Brain and Cognition*, 73(2), 128–134. <https://doi.org/10.1016/j.bandc.2010.04.004>
- Lyu, D.-C., Tan, T.-P., Chng, E. S., & Li, H. (2010). SEAME: A Mandarin-English code-switching speech corpus in South-East Asia. *Linguistic Data Consortium*.
<https://doi.org/10.21437/interspeech.2010-563>
- Ma, D., Wang, X., & Gao, X. (2021). Effects of third-language learning on inhibitory control with consideration of task complexity and language proficiency. *International Journal of Bilingualism*, 26(3), 286–307. <https://doi.org/10.1177/13670069211054599>
- Mackay, J. (2019). An ideal second language self intervention: Development of possible selves in an English as a foreign language classroom context. *System*, 81, 50–62.
<https://doi.org/10.1016/j.system.2019.01.003>
- MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, 24(5), 637–651.
<https://doi.org/10.1037/a0019803>
- Macnamara, B. N., & Conway, A. R. A. (2014). Novel evidence in support of the bilingual advantage: Influences of task demands and experience on cognitive control and working memory. *Psychonomic Bulletin & Review*, 21(2), 520–525.
<https://doi.org/10.3758/s13423-013-0524-y>
- Madjarova, S. J., Williams, R. J., Nwachukwu, B. U., Martin, R. K., Karlsson, J., Ollivier, M., & Pareek, A. (2022). Picking apart p values: common problems and points of confusion.

- Knee Surgery, Sports Traumatology, Arthroscopy*, 30(10), 3245–3248.
<https://doi.org/10.1007/s00167-022-07083-3>
- Mäki-Marttunen, V., Hagen, T., & Espeseth, T. (2019). Task context load induces reactive cognitive control: An fMRI study on cortical and brain stem activity. *Cognitive, Affective, & Behavioral Neuroscience*, 19(4), 945–965. <https://doi.org/10.3758/s13415-019-00691-6>
- Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940–967.
[https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067)
- Marian, V., & Shook, A. (2012). The cognitive benefits of being bilingual. *Cerebrum: The Dana Forum on Brain Science*, 2012(13), 13.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3583091/>
- Marsman, A., Mandl, R. C. W., Klomp, D. W. J., Cahn, W., Kahn, R. S., Luijten, P. R., & Pol, H. E. H. (2017). Intelligence and brain efficiency: Investigating the association between working memory performance, glutamate, and GABA. *Frontiers in Psychiatry*, 8, e154.
<https://doi.org/10.3389/fpsyt.2017.00154>
- Marszalek, J. M., Balagna, D., Kim, A. K., & Patel, S. A. (2022). Self-concept and intrinsic motivation in foreign language learning: The connection between flow and the L2 self. *Frontiers in Education*, 7, e975163. <https://doi.org/10.3389/educ.2022.975163>
- Martin, C. D., & Nozari, N. (2020). Language control in bilingual production: Insights from error rate and error type in sentence production. *Bilingualism: Language and Cognition*, 24(2), 374–388. <https://doi.org/10.1017/s1366728920000590>
- Martinez, M. E. (2013). *Future bright: A transforming vision of human intelligence*. New York, NY: Oxford University Press.
- Martin, N. A., & Brownell, R. (2011). *Receptive One-Word Picture Vocabulary Test*. Pro-Ed
- Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 11(1), 81–93. <https://doi.org/10.1017/s1366728907003227>

- Marzecová, A., Asanowicz, D., Krivá, L., & Wodniecka, Z. (2012). The effects of bilingualism on efficiency and lateralization of attentional networks. *Bilingualism: Language and Cognition*, 16(3), 608–623. <https://doi.org/10.1017/s1366728912000569>
- Mathy, F., Chekaf, M., & Cowan, N. (2018). Simple and complex working memory tasks allow similar benefits of information compression. *Journal of Cognition*, 1(1). <https://doi.org/10.5334/joc.31>
- McCarthy, P., Benuskova, L., & Franz, E. A. (2014). The age-related posterior-anterior shift as revealed by voxelwise analysis of functional brain networks. *Frontiers in Aging Neuroscience*, 6, e301. <https://doi.org/10.3389/fnagi.2014.00301>
- McLean, S., Stewart, J., & Batty, A. O. (2020). Predicting L2 reading proficiency with modalities of vocabulary knowledge: A bootstrapping approach. *Language Testing*, 37(3), 389–411. <https://doi.org/10.1177/0265532219898380>
- McMillan, B. A., & Rivers, D. J. (2011). The practice of policy: Teacher attitudes toward “English only.” *System*, 39(2), 251–263. <https://doi.org/10.1016/j.system.2011.04.011>
- Mecklinger, A., & Pfeifer, E. (1996). Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory. *Cognitive Brain Research*, 4(3), 211–224. [https://doi.org/10.1016/s0926-6410\(96\)00034-1](https://doi.org/10.1016/s0926-6410(96)00034-1)
- Medeiros, M., Von Schoultz, Å., & Wass, H. (2019). Language matters? Antecedents and political consequences of support for bilingualism in Canada and Finland. *Comparative European Politics*, 18(4), 532–559. <https://doi.org/10.1057/s41295-019-00198-x>
- Meisel, J. M. (2007). The weaker language in early child bilingualism: Acquiring a first language as a second language? *Applied Psycholinguistics*, 28(3), 495–514. <https://doi.org/10.1017/s0142716407070270>
- 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. <https://doi.org/10.1017/s1366728913000084>
- Meuter, R., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40. <https://doi.org/10.1006/jmla.1998.2602>

- Mikhailishchina, V., & Utochkin, I. (2022). Many clicks, one d-prime: A simple signal detection account of visual recognition in a multi-try memory test. *Journal of Vision*, 22(14), 3688. <https://doi.org/10.1167/jov.22.14.3688>
- Miller, M., Nevado-Montenegro, A. J., & Hinshaw, S. P. (2012). Childhood executive function continues to predict outcomes in young adult females with and without childhood-diagnosed ADHD. *Journal of Abnormal Child Psychology*, 40, 657–668. <https://doi.org/10.1007/s10802-011-9599-y>
- Ministry of Education. (2013). *Malaysia education blueprint, 2013-2025*. Kementrian Pendidikan Malaysia.
- Ministry of Education. (2018, January). *Kenyataan Media: Pelaksanaan Dual Language Programme (DLP) Tahun 2018 (2)*. <https://www.moe.gov.my/en/pemberitahuan/media-statement/kenyataan-media-kpm-pelaksanaan-dual-language-programme-dlp-tahun-2019>
- Ministry of Justice, Finland (1996). *Language Act - Finlex*. Finlex. https://www.finlex.fi/en/laki/kaannokset/1922/en19220148_19960201.pdf
- Mirsky, J. B., Heuer, H. W., Jafari, A., Kramer, J. H., Schenk, A. K., Viskontas, I. V., Miller, B. L., & Boxer, A. L. (2011). Anti-saccade performance predicts executive function and brain structure in normal elders. *Cognitive and Behavioral Neurology*, 24(2), 50–58. <https://doi.org/10.1097/wnn.0b013e318223f6c6>
- Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: Higher proficiency in a second language is associated with more rapid endogenous disengagement. *Quarterly Journal of Experimental Psychology*, 65(8), 1502–1510. <https://doi.org/10.1080/17470218.2012.657656>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Monnier, C., Boiché, J., Armandon, P., Baudoin, S., & Bellocchi, S. (2021). Is bilingualism associated with better working memory capacity? A meta-analysis. *International Journal of Bilingual Education and Bilingualism*, 25(6), 1–27. <https://doi.org/10.1080/13670050.2021.1908220>

- Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 114(2), 187–202. <https://doi.org/10.1016/j.jecp.2012.09.002>
- 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. <https://doi.org/10.1371/journal.pone.0094169>
- Moriguchi, Y., & Shinohara, I. (2019). Socioeconomic disparity in prefrontal development during early childhood. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-39255-6>
- Morrison, C., Kamal, F., Le, K., & Taler, V. (2019). Monolinguals and bilinguals respond differently to a delayed matching-to-sample task: An ERP study. *Bilingualism Language and Cognition*, 23(4), 858–868. <https://doi.org/10.1017/s1366728919000816>
- Morrison, C., Kamal, F., & Taler, V. (2018). The influence of bilingualism on working memory event-related potentials. *Bilingualism: Language and Cognition*, 22(1), 191–199. <https://doi.org/10.1017/s1366728918000391>
- Morton, J., & Harper, S. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science*, 10, 719–726. <https://doi.org/10.1111/j.1467-7687.2007.00623.x>
- Mostafa, N. A. (2016). Bilingualism And Education: The Malaysian Experience. *English Education International Conference*. English Education International Conference, Banda Aceh, Indonesia. <https://jurnal.usk.ac.id/EEIC/article/viewFile/15855/11677>
- Mueller, J. L. (2009). The influence of lexical familiarity on ERP responses during sentence comprehension in language learners. *Second Language Research*, 25(1), 43–76. <https://doi.org/10.1177/0267658308098996>
- Muftah, M., & Rafic-Galea, S. (2013). Language learning motivation among Malaysian pre-university students. *English Language Teaching*, 6(3), 1-12. <https://doi.org/10.5539/elt.v6n3p92>
- Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, 5(3), 218–228. <https://doi.org/10.1038/nrn1345>
- Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

- Naeem, K., Filippi, R., Periche-Tomas, E., Papageorgiou, A., & Bright, P. (2018). The importance of socioeconomic status as a modulator of the bilingual advantage in cognitive ability. *Frontiers in Psychology, 9*, e1818. <https://doi.org/10.3389/fpsyg.2018.01818>
- Namazi, M., & Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism, 13*(5), 597–616. <https://doi.org/10.1080/13670050.2010.488288>
- Navarro-Torres, C. A., Beatty-Martínez, A. L., Kroll, J. F., & Green, D. W. (2021). Research on bilingualism as discovery science. *Brain and Language, 222*, 105014. <https://doi.org/10.1016/j.bandl.2021.105014>
- Nayak, S., Salem, H., & Tarullo, A. (2020). Neural mechanisms of response-preparation and inhibition in bilingual and monolingual children: Lateralized Readiness Potentials (LRPs) during a nonverbal Stroop task. *Developmental Cognitive Neuroscience, 41*, 100740. <https://doi.org/10.1016/j.dcn.2019.100740>
- Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews, 33*(7), 1004–1023. <https://doi.org/10.1016/j.neubiorev.2009.04.001>
- Neuhaus, A., Urbanek, C., Opgen-Rhein, C., Hahn, E., Ta, T., Koehler, S., Gross, M., & Dettling, M. (2010). Event-related potentials associated with attention network test. *International Journal of Psychophysiology, 76*, 72–79. <https://doi.org/10.1016/j.ijpsycho.2010.02.005>
- Ng, G., & Yang, H. (2021). Code-Switching patterns differentially shape cognitive control: Testing the predictions of the adaptive control hypothesis. *Bilingualism Language and Cognition, 25*(3), 521–535. <https://doi.org/10.1017/s1366728921000754>
- Nguyen, T. T. T., & Hamid, M. O. (2017). Bilingualism as a resource: Language attitudes of Vietnamese ethnic minority students. *Current Issues in Language Planning, 19*(4), 343–362. <https://doi.org/10.1080/14664208.2017.1337922>
- Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods, 5*(2), 241–301. <https://doi.org/10.1037/1082-989x.5.2.241>

- Nielsen, T. R., Antelius, E., & Waldemar, G. (2019). Cognitive advantages in adult Turkish bilingual immigrants: A question of the chicken or the egg. *Journal of Cross-Cultural Gerontology*, 34(2), 115–129. <https://doi.org/10.1007/s10823-019-09375-7>
- Novak, I., & Morvai, L. (2019). Foreign language learning attitude of Hungarian higher education students. *Hungarian Educational Research Journal*, 9(2), 303–317. <https://doi.org/10.1556/063.9.2019.1.26>
- Nuño, L., Gómez-Benito, J., Carmona, V. R., & Pino, O. (2021). A systematic review of executive function and information processing speed in major depression disorder. *Brain Sciences*, 11, 147. <https://doi.org/10.3390/brainsci11020147>
- O’Connell, R. G., Balsters, J. H., Kilcullen, S. M., Campbell, W., Bokde, A. W., Lai, R., Upton, N., & Robertson, I. H. (2011). A simultaneous ERP/fMRI investigation of the P300 aging effect. *Neurobiology of Aging*, 33(10), 2448–2461. <https://doi.org/10.1016/j.neurobiolaging.2011.12.021>
- Oberauer, K. (2019). Working memory and attention: A conceptual analysis and review. *Journal of Cognition*, 2(1), 36. <https://doi.org/10.5334/joc.58>
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, 19(5), 779–819. <https://doi.org/10.3758/s13423-012-0272-4>
- Olguin, A., Cekic, M., Bekinschtein, T., Katsos, N., & Bozic, M. (2019). Bilingualism and language similarity modify the neural mechanisms of selective attention. *Scientific Reports*, 9, 8204. <https://doi.org/10.1038/s41598-019-44782-3>
- Olson, D. J. (2022). The Bilingual Code-Switching Profile (BCSP). *Linguistic Approaches to Bilingualism*, 14(3), 400–433. <https://doi.org/10.1075/lab.21039.ols>
- Olson, D. J. (2023). Measuring bilingual language dominance: An examination of the reliability of the Bilingual Language Profile. *Language Testing*, 40(3), 521–547. <https://doi.org/10.1177/02655322221139162>
- Olszewska, Alicja M, Gaca, M., Herman, A. M., Jednoróg, K., & Marchewka, A. (2021). How musical training shapes the adult brain: Predispositions and neuroplasticity. *Frontiers in Neuroscience*, 15, e630829. <https://doi.org/10.3389/fnins.2021.630829>
- Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From bilingualism to bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently.

- Bilingualism: Language and Cognition*, 21(4), 867–879.
<https://doi.org/10.1017/s1366728918000020>
- Operario, D., Adler, N. E., & Williams, D. R. (2004). Subjective social status: Reliability and predictive utility for global health. *Psychology & Health*, 19, 237–246.
<https://doi.org/10.1080/08870440310001638098>
- Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1), 46–59. <https://doi.org/10.1002/hbm.20131>
- Paap, K. (2019). The bilingual advantage debate: Quantity and quality of evidence. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley Blackwell. <https://doi.org/10.1002/9781119387725.ch34>
- Paap, K. (2023). *The bilingual advantage in executive functioning hypothesis*. Taylor & Francis.
- Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., & Mikulinsky, R. (2020). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cognitive Research: Principles and Implications*, 5, e7. <https://doi.org/10.1186/s41235-020-0207-y>
- 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. <https://doi.org/10.1016/j.cortex.2015.04.014>
- Paap, K. R., Majoubi, J., Balakrishnan, N., & Anders-Jefferson, R. T. (2024). Bilingualism, like other types of brain training, does not produce far transfer: It all fits together. *International Journal of Bilingualism*. <https://doi.org/10.1177/13670069231214599>
- Paap, K. R., Mason, L., & Anders-Jefferson, R. (2021). Predictions about the cognitive consequences of language switching on executive functioning inspired by the adaptive control hypothesis fail more often than not. *Brain Sciences*, 11, 1217.
<https://doi.org/10.3390/brainsci11091217>
- 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, e962. <https://doi.org/10.3389/fpsyg.2014.00962>

- Paap, K., Darrow, J., Dalibar, C., & Johnson, H. (2015). Effects of script similarity on bilingual advantages in executive control are likely to be negligible or null. *Frontiers in Psychology*, 5, e1539. <https://doi.org/10.3389/fpsyg.2014.01539>
- Paap, K., & Greenberg, Z. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66, 232–258. <https://doi.org/10.1016/j.cogpsych.2012.12.002>
- Paplikar, A., Mekala, S., Bak, T. H., Dharamkar, S., Alladi, S., & Kaul, S. (2018). Bilingualism and the severity of poststroke aphasia. *Aphasiology*, 33, 58–72. <https://doi.org/10.1080/02687038.2017.1423272>
- Park, J., Weismer, S. E., & Kaushanskaya, M. (2018). Changes in executive function over time in bilingual and monolingual school-aged children. *Developmental Psychology*, 54(10), 1842–1853. <https://doi.org/10.1037/dev0000562>
- Pascual, A., Muñoz, N., & Robres, A. (2019). The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01582>
- Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: selected studies of the Event-Related Potential. *International Journal of Medical Sciences*, 2(4), 147–154. <https://doi.org/10.7150/ijms.2.147>
- Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, 76, 1–23. <https://doi.org/10.1037/h0093840>
- Peng, P., & Kievit, R. A. (2020). The development of academic achievement and cognitive abilities: A bidirectional perspective. *Child Development Perspectives*, 14, 15–20. <https://doi.org/10.1111/cdep.12352>
- Perfetti, C. A., & Liu, Y. (2005). Orthography to phonology and meaning: Comparisons across and within writing systems. *Reading and Writing*, 18(3), 193–210. <https://doi.org/10.1007/s11145-004-2344-y>
- Perfetti, C. A., Liu, Y., Fiez, J., Nelson, J., Bolger, D. J., & Tan, L. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition*, 10(2), 131–146. <https://doi.org/10.1017/s1366728907002891>

- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184–187. [https://doi.org/10.1016/0013-4694\(89\)90180-6](https://doi.org/10.1016/0013-4694(89)90180-6)
- Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35(1), 73–89. <https://doi.org/10.1146/annurev-neuro-062111-150525>
- Phelps, J., Attaheri, A., & Bozic, M. (2022). How bilingualism modulates selective attention in children. *Scientific Reports*, 12(1), e6381. <https://doi.org/10.1038/s41598-022-09989-x>
- Phyak, P. (2023). Translanguaging as a space of simultaneity: Theorizing translanguaging pedagogies in English medium schools from a spatial perspective. *Modern Language Journal*, 107(1), 289–307. <https://doi.org/10.1111/modl.12830>
- Pietrzyk-Kowalec, P. (2023). Multilingualism and the job market in the European Union: A comprehensive analysis. *Proceedings of the World Conference on Education and Teaching*, 2(1), 1–8. <https://doi.org/10.33422/etconf.v2i1.71>
- Pillai, S., & Ong, L. T. (2018). English(es) in Malaysia. *Asian Englishes*, 20, 147–157. <https://doi.org/10.1080/13488678.2018.1459073>
- Pintner, R. (1932). The influence of language background on intelligence tests. *The Journal of Social Psychology*, 3, 235–240. <https://doi.org/10.1080/00224545.1932.9919147>
- Pliatsikas, C. (2019). Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition*, 23, 459–471. <https://doi.org/10.1017/s1366728919000130>
- Pliatsikas, C., Ansaldo, A. I., & Voits, T. (2020). Bilingualism and the declining brain. *Linguistic Approaches to Bilingualism*, 11(4), 453–458. <https://doi.org/10.1075/lab.00031.int>
- Pliatsikas, C., DeLuca, V., Moschopoulou, E., & Saddy, J. D. (2016). Immersive bilingualism reshapes the core of the brain. *Brain Structure and Function*, 222(4), 1785–1795. <https://doi.org/10.1007/s00429-016-1307-9>
- Polich, J. (1996). Meta-analysis of P300 normative aging studies. *Psychophysiology*, 33(4), 334–353. <https://doi.org/10.1111/j.1469-8986.1996.tb01058.x>
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>

- Polich, J. (2012). Neuropsychology of P300. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components*. Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780195374148.013.0089>
- Pollack, A. E. (2022). Deep versus Shallow Processing: A Learning and Memory Experiment for Asynchronous and Synchronous Online Platforms. *PubMed*, 20(2), 146–149.
<https://doi.org/10.59390/feap4736>
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Posner, M. I. (2008). Measuring alertness. *Annals of the New York Academy of Sciences*, 1129(1), 193–199. <https://doi.org/10.1196/annals.1417.011>
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13(1), 25–42. <https://doi.org/10.1146/annurev.ne.13.030190.000325>
- 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), 1–23.
<https://doi.org/10.1146/annurev.psych.58.110405.085516>
- Pradhan, G. N., Bogle, J., Kleindienst, S., Cevette, M. J., & Stepanek, J. (2018). Correlating multi-dimensional oculometrics with cognitive performance in healthy youth athletes. *Journal of Healthcare Informatics Research*, 2(1-2), 132–151.
<https://doi.org/10.1007/s41666-017-0011-8>
- Preece, S., & Marshall, S. (2020). Plurilingualism, teaching and learning, and Anglophone higher education: An introduction Anglophone universities and linguistic diversity. *Language, Culture and Curriculum*, 33(2), 117–125.
<https://doi.org/10.1080/07908318.2020.1723931>
- Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, 17(4), 682–691.
<https://doi.org/10.1017/s1355617711000580>
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13, 253–262. <https://doi.org/10.1017/s1366728909990526>
- Proverbio, A.M. (2023). EEG and ERPs in the Study of Language and Social Knowledge. In: Boggio, P.S., Wingenbach, T.S.H., da Silveira Coêlho, M.L., Comfort, W.E., Murrins

- Marques, L., Alves, M.V.C. (eds) *Social and Affective Neuroscience of Everyday Human Interaction*. Springer, Cham. https://doi.org/10.1007/978-3-031-08651-9_12
- Pulinx, R., Van Avermaet, P., & Agirdag, O. (2017). Silencing linguistic diversity: The extent, the determinants and consequences of the monolingual beliefs of Flemish teachers. *International Journal of Bilingual Education and Bilingualism*, 20(5), 542–556. <https://doi.org/10.1080/13670050.2015.1102860>
- Purdy, M. (2002). Executive function ability in persons with aphasia. *Aphasiology*, 16(4-6), 549–557. <https://doi.org/10.1080/02687030244000176>
- Purić, D., Vuksanović, J., & Chondrogianni, V. (2017). Cognitive advantages of immersion education after 1 year: Effects of amount of exposure. *Journal of Experimental Child Psychology*, 159, 296–309. <https://doi.org/10.1016/j.jecp.2017.02.011>
- Qian, D. D., & Lin, L. H. F. (2019). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (1st ed., p. 15). Routledge. <https://doi.org/10.4324/9780429291586>
- Qu, L., Low, J. J. W., Zhang, T., Li, H., & Zelazo, P. D. (2015). Bilingual advantage in executive control when task demands are considered. *Bilingualism: Language and Cognition*, 19(2), 277–293. <https://doi.org/10.1017/s1366728914000376>
- Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: a non-technical guide using JASP. *BMC Psychiatry*, 18(1), 178. <https://doi.org/10.1186/s12888-018-1761-4>
- Raisman-Carlovich, A. J., Arias-Trejo, N., & Carrasco-Ortiz, E. H. (2024). Lexical production and cognitive control in sequential bilinguals immersed in two different contexts of language use. *Second Language Research*. <https://doi.org/10.1177/02676583241244606>
- R Core Team. (2021). *R: A language and environment for statistical computing*. <https://www.r-project.org/>
- Radman, N., Jost, L., Dorood, S., Mancini, C., & Annoni, J.-M. (2021). Language distance modulates cognitive control in bilinguals. *Scientific Reports*, 11(1), 24131. <https://doi.org/10.1038/s41598-021-02973-x>
- Rafique, S., Waqas, A., & Shahid, C. (2023). The correlation between vocabulary knowledge and English language proficiency at undergraduate level. *Pakistan Journal of Humanities and Social Sciences*, 11(2), 1162–1171. <https://doi.org/10.52131/pjhss.2023.1102.0422>

- Rahman, M. M., & Singh, M. K. M. (2021). English and Malay language policy and planning in Malaysia. *Training, Language and Culture*, 5(4), 36–46. <https://doi.org/10.22363/2521-442x-2021-5-4-36-46>
- Rahman, M. M., & Singh, M. K. M. (2022). The ideology towards English as a medium of instruction (EMI) adoption in higher education in Malaysia: A case study. *3L the Southeast Asian Journal of English Language Studies*, 28(2), 109–121. <https://doi.org/10.17576/31-2022-2802-08>
- Ramanujan, K. (2019). The impact of relative language distance on bilingual language control: A functional imaging study. *BioRxiv*. <https://doi.org/10.1101/771212>
- Ramírez-Esparza, N., García-Sierra, A., & Jiang, S. (2020). The current standing of bilingualism in today's globalized world: A socio-ecological perspective. *Current Opinion in Psychology*, 32, 124–128. <https://doi.org/10.1016/j.copsyc.2019.06.038>
- Ratiu, I., & Azuma, T. (2014). Working memory capacity: Is there a bilingual advantage? *Journal of Cognitive Psychology*, 27(1), 1–11. <https://doi.org/10.1080/20445911.2014.976226>
- Ratiu, I., Hout, M., Walenchok, S., Azuma, T., & Goldinger, S. (2017). Comparing visual search and eye movements in bilinguals and monolinguals. *Attention, Perception, & Psychophysics*, 79, 1695–1725. <https://doi.org/10.3758/s13414-017-1328-3>
- Raven, J. C. (1941). Standardization of progressive matrices, 1938. *British Journal of Medical Psychology*, 19(1), 137–150. <https://doi.org/10.1111/j.2044-8341.1941.tb00316.x>
- Rentería, M. A., Casalletto, K., Tom, S., Pa, J., Harrati, A., Armstrong, N., Rajan, K. B., Manly, J., Mungas, D., & Zahodne, L. (2019). The Contributions of Active Spanish-English Bilingualism to Cognitive Reserve among Older Hispanic Adults Living in California. *Archives of Clinical Neuropsychology*, 34(7). <https://doi.org/10.1093/arclin/acz029.02>
- Resnik, P. (2018). Multilinguals' use of L1 and L2 inner speech. *International Journal of Bilingual Education and Bilingualism*, 24(1), 72–90. <https://doi.org/10.1080/13670050.2018.1445195>
- Reverberi, C., Kuhlen, A. K., Seyed-Allaei, S., Greulich, R. Stefan., Costa, A., Abutalebi, J., & Haynes, J.-D. (2018). The neural basis of free language choice in bilingual speakers: Disentangling language choice and language execution. *NeuroImage*, 177, 108–116. <https://doi.org/10.1016/j.neuroimage.2018.05.025>

- Riek, H. C., Brien, D. C., Coe, B. C., Huang, J., Perkins, J. E., Yep, R., McLaughlin, P., Orange, J. B., Peltsch, A., Roberts, A., Binns, M. A., Lou, W., Agessandro Abrahão, Arnott, S. R., Beaton, D., Black, S. E., Dowlatshahi, D., Finger, E., Fischer, C. E., & Frank, A. (2023). Cognitive correlates of antisaccade behaviour across multiple neurodegenerative diseases. *Brain Communications*, 5(2), fcad049. <https://doi.org/10.1093/braincomms/fcad049>
- Ries, A. J., Slayback, D., Fulbright, E., Sligh, M., Gorman, K., & Touryan, J. (2019). Congruency effects in the attention network task: The influence of stimulus onset asynchrony and eye movements. *Journal of Vision*, 19(10), 127. <https://doi.org/10.1167/19.10.127>
- Rodriguez-Fornells, A., Krämer, U. M., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiers in Psychology*, 2, e388. <https://doi.org/10.3389/fpsyg.2011.00388>
- Roselló, B., Berenguer, C., Baixauli, I., Mira, Á., Martinez-Raga, J., & Miranda, A. (2020). Empirical examination of executive functioning, ADHD associated behaviors, and functional impairments in adults with persistent ADHD, remittent ADHD, and without ADHD. *BMC Psychiatry*, 20, e134. <https://doi.org/10.1186/s12888-020-02542-y>
- Ross, H. A., Russell, P. N., & Helton, W. S. (2014). Effects of breaks and goal switches on the vigilance decrement. *Experimental Brain Research*, 232(6), 1729–1737. <https://doi.org/10.1007/s00221-014-3865-5>
- Ross, J., & Melinger, A. (2017). Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Developmental Science*, 20(4), e12405. <https://doi.org/10.1111/desc.12405>
- Rothman, J., Bayram, F., DeLuca, V., Di Pisa, G., Duñabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., Kupisch, T., Laméris, T., Luque, A., Van Osch, B., Soares, S. M. P., Prystauka, Y., Tat, D., Tomić, A., Voits, T., & Wulff, S. (2022). Monolingual comparative normativity in bilingualism research is out of “control”: Arguments and alternatives. *Applied Psycholinguistics*, 44(3), 316–329. <https://doi.org/10.1017/s0142716422000315>

- Ryan, R. M., & Deci, E. L. (2000). Self-determination Theory and the Facilitation of Intrinsic motivation, Social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sabourin, L., & Vinerte, S. (2014). The bilingual advantage in the Stroop task: Simultaneous vs. early bilinguals. *Bilingualism: Language and Cognition*, 18(2), 350–355. <https://doi.org/10.1017/s1366728914000704>
- Sadat, J., Martin, C. D., Alario, F. X., & Costa, A. (2011). Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguistic Research*, 41(3), 159–179. <https://doi.org/10.1007/s10936-011-9183-1>
- Saer, D. J. (1923). The effect of bilingualism on intelligence. *British Journal of Psychology: General Section*, 14, 25–38. <https://doi.org/10.1111/j.2044-8295.1923.tb00110.x>
- Salvatierra, J., & Rosselli, M. (2010). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism*, 15, 26–37. <https://doi.org/10.1177/1367006910371021>
- Samuel, S., Roehr-Brackin, K., Pak, H., & Kim, H. (2018). Cultural effects rather than a bilingual advantage in cognition: A review and an empirical study. *Cognitive Science*, 42(7), 2313–2341. <https://doi.org/10.1111/cogs.12672>
- Sánchez-Pérez, M. del M., & Manzano-Agugliaro, F. (2021). Worldwide trends in bilingual education research: A half-century overview. *Education Sciences*, 11(11), 730. <https://doi.org/10.3390/educsci11110730>
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13(2), 231–252. <https://doi.org/10.1017/s1366728909990514>
- Satel, J., Wilson, N. R., & Klein, R. M. (2019). What neuroscientific studies tell us about inhibition of return. *Vision*, 3(4), 58. <https://doi.org/10.3390/vision3040058>
- Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2017). Comparison of the working memory load in n-back and working memory span tasks by means of EEG frequency band power and P300 amplitude. *Frontiers in Human Neuroscience*, 11, e6. <https://doi.org/10.3389/fnhum.2017.00006>

- Schmid, M. S., & Yilmaz, G. (2018). Predictors of language dominance: An integrated analysis of first language attrition and second language acquisition in late bilinguals. *Frontiers in Psychology, 9*, e1306. <https://doi.org/10.3389/fpsyg.2018.01306>
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1–66. <https://doi.org/10.1037/0033-295x.84.1.1>
- Schug, D. & Sarnou, H.(2024). The Reshaping of Language Learning Motivation during COVID-19-imposed Distance Learning. *Journal of Studies in Language, Culture, and Society*, In press.
- Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *Wiley Interdisciplinary Reviews: Cognitive Science, 7*(4), 264–275. <https://doi.org/10.1002/wcs.1393>
- Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty, & M. Long (Eds.), *The handbook of second language acquisition* (pp. 383–406). London: Blackwell.
- Selamat, A., Esa, A., Saad, S. S., & Atim, A. (2010). Teaching and learning mathematics and science in English in primary schools in the State of Johor, Malaysia. *Journal of Education, 16* (2011), 61–73.
- Selvadurai, S., Liu, O. P., Radzi, M. M., Hoon, O. P., Tee, O. P., & Saibeh, B. (2017). Debating education for nation building in Malaysia: National school persistence or vernacular school resistance? *Geografia: Malaysian Journal of Society and Space, 11*(13), 14-23.
- Serva, M., & Petroni, F. (2008). Indo-European languages tree by Levenshtein distance. *EPL (Europhysics Letters), 81*(6), 68005. <https://doi.org/10.1209/0295-5075/81/68005>
- Shalchy, M. A., Pergher, V., Pahor, A., Van Hulle, M. M., & Seitz, A. R. (2020). N-back related ERPs depend on stimulus type, task structure, pre-processing, and lab factors. *Frontiers in Human Neuroscience, 14*, e549966. <https://doi.org/10.3389/fnhum.2020.549966>
- Shen, W., & Park, H. (2020). Working memory and second language learning: A review of the past twenty years' research in China. *Journal of Pan-Pacific Association of Applied Linguistics, 24*(1), 85–106. <https://doi.org/10.25256/paal.24.1.5>
- Shtyrov, Y., Goryainova, G., Tugin, S., Ossadtchi, A., & Shestakova, A. (2013). Automatic processing of unattended lexical information in visual oddball presentation:

- Neurophysiological evidence. *Frontiers in Human Neuroscience*, 7, e421.
<https://doi.org/10.3389/fnhum.2013.00421>
- Sicam, F. P. M., & Lucas, R. I. G. (2016). Language attitudes of adolescent Filipino bilingual learners towards English and Filipino. *Asian Englishes*, 18(2), 109–128.
<https://doi.org/10.1080/13488678.2016.1179474>
- Silva-Corvalán, C., & Treffers-Daller, J. (2016). *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge University Press.
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do “brain-training” programs work? *Psychological Science in the Public Interest*, 17(3), 103–186.
<https://doi.org/10.1177/1529100616661983>
- Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D., & Diamond, A. (2012). Refining the understanding of inhibitory processes: How response prepotency is created and overcome. *Developmental Science*, 15(1), 62–73.
<https://doi.org/10.1111/j.1467-7687.2011.01105.x>
- Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching*, 7(1), 89–103.
<https://doi.org/10.14746/ssllt.2017.7.1.5>
- Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N., Tripoliti, E. E., Marias, K., Fotiadis, D. I., & Tsiknakis, M. (2021). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering*, 99, 260-277. <https://doi.org/10.1109/rbme.2021.3066072>
- Smith, D. K. (1999). Kaufman Brief Intelligence Test (K-BIT). *Diagnostique*, 24(1-4), 125–134.
<https://doi.org/10.1177/153450849902401-412>
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General* 117(1), 34–50. <https://doi.org/10.1037//0096-3445.117.1.34>
- Soon, T. W. (1988). *Singapore's new education system: Education reform for national development*. Singapore: Institute of Southeast Asian Studies. Call no.: RSING 370.95957 SOO.

- Sorge, G. B., Toplak, M. E., & Bialystok, E. (2017). Interactions between levels of attention ability and levels of bilingualism in children's executive functioning. *Developmental Science*, 20(1), e12408. <https://doi.org/10.1111/desc.12408>
- Sörman, D. E., Hansson, P., & Ljungberg, J. K. (2019). Different features of bilingualism in relation to executive functioning. *Frontiers in Psychology*, 10, e269. <https://doi.org/10.3389/fpsyg.2019.00269>
- Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–293. <https://doi.org/10.2307/1412107>
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137–149. <https://doi.org/10.3758/bf03207704>
- Stasenko, A., Matt, G. E., & Gollan, T. H. (2017). A relative bilingual advantage in switching with preparation: Nuanced explorations of the proposed association between bilingualism and task switching. *Journal of Experimental Psychology: General*, 146(11), 1527–1550. <https://doi.org/10.1037/xge0000340>
- Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E.-J. (2019). A tutorial on Bayes factor design analysis using an informed prior. *Behavior Research Methods*, 51(3), 1042–1058. <https://doi.org/10.3758/s13428-018-01189-8>
- Stephen, J. (2013). English in Malaysia: A case of the past that never really went away? *English Today*, 29(2), 3–8. <https://doi.org/10.1017/s0266078413000084>
- Strani, K. (2020). Multilingualism and politics revisited: the state of the art. In *Springer eBooks* (pp. 17–45). https://doi.org/10.1007/978-3-030-40701-8_2
- Studenica, A., Laketa, A., Chrysochoou, E., Blakey, E., & Vivas, A. B. (2022). The influence of bilingualism on adolescent cognition: The roles of biculturalism, the bilingual profile, and linguistic similarity. *Cognitive Development*, 63, 101203. <https://doi.org/10.1016/j.cogdev.2022.101203>
- 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. <https://doi.org/10.1016/j.bandl.2014.10.004>

- Sulpizio, S., Maschio, D., Mauro, D., Fedeli, D., & Abutalebi, J. (2020). Bilingualism as a gradient measure modulates functional connectivity of language and control networks. *NeuroImage*, 205, 116306. <https://doi.org/10.1016/j.neuroimage.2019.116306>
- Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 22, 401–415. <https://doi.org/10.1017/s1366728917000682>
- Surrain, S., & Luk, G. (2019). The perceptions of bilingualism scales. *PsyArXiv*. <https://doi.org/10.31234/osf.io/s32zb>
- Suslak, D. F. (2009). The sociolinguistic problem of generations. *Language & Communication*, 29(3), 199–209. <https://doi.org/10.1016/j.langcom.2009.02.003>
- Swadesh, M. (1952). Lexico-statistic dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society*, 96(4), 452–463.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson Education Limited.
- Tan, S. I. (2009). Lexical borrowing in Malaysian English: Influences of Malay. *Lexis*, 3. <https://doi.org/10.4000/lexis.629>
- 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, e123. <https://doi.org/10.3389/fpsyg.2011.00123>
- Tavakoli, P. (2019). Chapter 2. Automaticity, fluency and second language task performance. In *Task-based language teaching* (pp. 39–52). <https://doi.org/10.1075/tblt.13.03tav>
- Taylor, Z. W. (2021). Breaking the English routine: An argument for polylingual discourse in U.S. higher education. *Journal of Praxis in Higher Education*, 3(1), 5–25. <https://doi.org/10.47989/kpdc103>
- Teichler, U. (2017). Internationalisation trends in higher education and the changing role of international student mobility. *Journal of International Mobility*, 5(1), 177. <https://doi.org/10.3917/jim.005.0179>
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, 104(30), 12530–12535. <https://doi.org/10.1073/pnas.0609927104>

- Thomas-Sunesson, D., Hakuta, K., & Bialystok, E. (2018). Degree of bilingualism modifies executive control in Hispanic children in the USA. *International Journal of Bilingual Education and Bilingualism*, 21(2), 197–206.
<https://doi.org/10.1080/13670050.2016.1148114>
- Timmer, K., Christoffels, I., & Costa, A. (2018). On the flexibility of bilingual language control: The effect of language context. *Bilingualism: Language and Cognition*, 22, 555–568.
<https://doi.org/10.1017/s1366728918000329>
- Timmer, K., Wodniecka, Z., & Costa, A. (2021). Rapid attentional adaptations due to language (monolingual vs bilingual) context. *Neuropsychologia*, 159, 107946.
<https://doi.org/10.1016/j.neuropsychologia.2021.107946>
- Titone, D. A., & Tiv, M. (2022). Rethinking multilingual experience through a Systems Framework of Bilingualism. *Bilingualism Language and Cognition*, 26(1), 1–16.
<https://doi.org/10.1017/s1366728921001127>
- Tong, Y., Hasim, Z., & Abdul Halim, H. (2023). The relationship between L2 vocabulary knowledge and reading proficiency: The moderating effects of vocabulary fluency. *Humanities and Social Sciences Communications*, 10(1), 1–14.
<https://doi.org/10.1057/s41599-023-02058-2>
- Too, W. K. (2023). “People are already doing it”: Malaysian English language teachers’ perceptions on translanguaging. *Asian Englishes*, 25(2), 1–14.
<https://doi.org/10.1080/13488678.2023.2214769>
- Treffers-Daller, J. (2015). Language dominance: The construct, its measurement, and operationalization. *Language Dominance in Bilinguals*, 235–265.
<https://doi.org/10.1017/cbo9781107375345.012>
- Treffers-Daller, J., Ongun, Z., Hofweber, J., & Korenar, M. (2020). Explaining individual differences in executive functions performance in multilinguals: The impact of code-switching and alternating between multicultural identity styles. *Frontiers in Psychology*, 11, e561088. <https://doi.org/10.3389/fpsyg.2020.561088>
- Trofimovich, P., & Turuševa, L. (2019). Language attitudes and ethnic identity: Examining listener perceptions of Latvian-Russian bilingual speakers. *Journal of Language, Identity & Education*, 19(1), 9–24. <https://doi.org/10.1080/15348458.2019.1696682>

- UNESCO (2003). The mother-tongue dilemma. *Education Today*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000130800>
- Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, 18(1), 3–24. <https://doi.org/10.1017/s1366728914000522>
- van den Berg, F., Brouwer, J., Tienkamp, T. B., Verhagen, J., & Keijzer, M. (2022). Language entropy relates to behavioral and pupil indices of executive control in young adult bilinguals. *Frontiers in Psychology*, 13, e864763. <https://doi.org/10.3389/fpsyg.2022.864763>
- van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S.-H., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, 9(3), 27. <https://doi.org/10.3390/bs9030027>
- van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-uk: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, 67, 1176–1190. <https://doi.org/10.1080/17470218.2013.850521>
- van Heuven, W., & Coderre, E. (2015). A call for sophisticated statistical approaches and neuroimaging techniques to study the bilingual advantage. *Cortex*, 73, 330–331. <https://doi.org/10.1016/j.cortex.2015.08.006>
- Vaughn, K. A., Greene, M. R., Ramos Nuñez, A. I., & Hernandez, A. E. (2015). The importance of neuroscience in understanding bilingual cognitive control. *Cortex*, 73, 373–374. <https://doi.org/10.1016/j.cortex.2015.06.010>
- Veen, V. van, & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, 14(4), 593–602. <https://doi.org/10.1162/08989290260045837>
- Verhagen, J., Mulder, H., & Leseman, P. P. M. (2017). Effects of home language environment on inhibitory control in bilingual three-year-old children. *Bilingualism: Language and Cognition*, 20(1), 114–127. <https://doi.org/10.1017/s1366728915000590>
- Verme, P. (2009). Happiness, freedom and control. *Journal of Economic Behavior & Organization*, 71(2), 146–161. <https://doi.org/10.1016/j.jebo.2009.04.008>

- Vivas, A. B., Chrysochoou, E., Ladas, A. I., & Salvati, V. (2020). The moderating effect of bilingualism on lifespan cognitive development. *Cognitive Development*, 55, 100890. <https://doi.org/10.1016/j.cogdev.2020.100890>
- Vlaspolder, T. T., Duits, A. A., Dijkstra, H. T., van Laar, T., & Spikman, J. M. (2018). Effectiveness of ReSET; A strategic executive treatment for executive dysfunctioning in patients with Parkinson's disease. *Neuropsychological Rehabilitation*, 30, 67–84. <https://doi.org/10.1080/09602011.2018.1452761>
- Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438(7067), 500–503. <https://doi.org/10.1038/nature04171>
- von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, 145(2), 246–258. <https://doi.org/10.1037/xge0000120>
- Wagner, D., Bekas, K., & Bialystok, E. (2023). Does language entropy shape cognitive performance? A tale of two cities. *Bilingualism: Language and Cognition*, 1–11. <https://doi.org/10.1017/s1366728923000202>
- Walters, S. O., & Weaver, K. A. (2003). Relationships between the Kaufman Brief Intelligence Test and the Wechsler Adult Intelligence Scale-Third Edition. *Psychological Reports*, 92(3_suppl), 1111–1115. <https://doi.org/10.2466/pr0.2003.92.3c.1111>
- Wang, A., Yue, Z., Zhang, M., & Chen, Q. (2015). Interaction between spatial inhibition of return (IOR) and executive control in three-dimensional space. *Experimental Brain Research*, 233(11), 3059–3071. <https://doi.org/10.1007/s00221-015-4374-x>
- Wang, S., Wu, X., & Xiong, Z. (2021). The Effect of Teachers' Reinforcement on English Learning Engagement of EAL learners. *Advances in Social Science, Education and Humanities Research/Advances in Social Science, Education and Humanities Research*. <https://doi.org/10.2991/assehr.k.210617.126>
- Wang-Taylor, Y., & Clenton, J. (2022). Exploring the relationship between L2 vocabulary size and academic speaking. *System*, 107, 102822. <https://doi.org/10.1016/j.system.2022.102822>
- Warchoła, Ł., & Zająć-Lamparska, L. (2023a). The relationship of N200 and P300 amplitudes with intelligence, working memory, and attentional control behavioral measures in young

- healthy individuals. *Advances in Cognitive Psychology*, 19(4), 63–75.
<https://doi.org/10.5709/acp-0404-2>
- Ware, C., Dautricourt, S., Gonneaud, J., & Chételat, G. (2020). Does second language learning promote neuroplasticity in aging? A systematic review of cognitive and neuroimaging studies. *Frontiers in Aging Neuroscience*, 13, e706672.
<https://doi.org/10.3389/fnagi.2021.706672>
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45, 1191–1207.
<https://doi.org/10.3758/s13428-012-0314-x>
- Wass, S. V., Scerif, G., & Johnson, M. H. (2012). Training attentional control and working memory: Is younger, better? *Developmental Review*, 32(4), 360–387.
<https://doi.org/10.1016/j.dr.2012.07.001>
- Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, 44(3), 219–246. <https://doi.org/10.1177/0095798418771807>
- Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, 38(6), 998–1003.
<https://doi.org/10.1111/1469-8986.3860998>
- Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence (4th ed.)*. Williams & Wilkins Co. <https://doi.org/10.1037/11167-000>
- Wen, Y., Qiu, Y., Leong, C. X. R., & Van Heuven, W. J. B. (2023). LexCHI: A quick lexical test for estimating language proficiency in Chinese. *Behavior Research Methods*, 56(3), 2333–2352. <https://doi.org/10.3758/s13428-023-02151-z>
- Wickens, C. D. (1980). The structure of attentional resources In R. S. Nickerson (Ed.), *Attention & Performance, VIII*, (pp. 239–257). Hillsdale, N.J.: Erlbaum.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. <https://ggplot2.tidyverse.org>
- Von Grebmer Zu Wolfsturn, S., Gupta, A., Pablos, L., & Schiller, N. O. (2022). When left is right: The role of typological similarity in multilinguals' inhibitory control performance. *Bilingualism Language and Cognition*, 26(1), 165–178.
<https://doi.org/10.1017/s1366728922000426>

- Wong, A. L., Goldsmith, J., Forrence, A. D., Haith, A. M., & Krakauer, J. W. (2017). Reaction times can reflect habits rather than computations. *eLife*, 6, e28075. <https://doi.org/10.7554/elife.28075>
- Wong, F. F., Lee, K. S., Lee, S. K., & Ya'acob, A. (2012). English use as an identity marker among Malaysian undergraduates. *3L: Language, Linguistics, and Literature, The Southeast Asian Journal of English Language Studies*, 18(1), 145-155.
- Wu, Y., & Thierry, G. (2013). Fast modulation of executive function by language context in bilinguals. *Journal of Neuroscience*, 33, 13533–13537. <https://doi.org/10.1523/jneurosci.4760-12.2013>
- Xie, W., Altarriba, J., & Ng, B. C. (2022). Bilingualism, culture, and executive functions: Is there a relationship? *Languages*, 7(4), 247. <https://doi.org/10.3390/languages7040247>
- Xie, W., & Cavallaro, F. (2016). Attitudes towards Mandarin–English bilingualism: A study of Chinese youths in Singapore. *Journal of Multilingual and Multicultural Development*, 37(6), 628–641. <https://doi.org/10.1080/01434632.2015.1122603>
- Yang, E. (2017). Bilinguals' working memory (WM) advantage and their dual language practices. *Brain Sciences*, 7(12), 86. <https://doi.org/10.3390/brainsci7070086>
- Yang, H., Tng, G. Y. Q., Ng, G. R., & Ng, W. Q. (2022). Bilingual interactional contexts predict executive functions in older adults. *Bilingualism Language and Cognition*, 26(1), 36–47. <https://doi.org/10.1017/s1366728922000190>
- Yang, L. F., & Ishak, Md. S. A. (2012). Framing controversy over language policy in Malaysia: The coverage of PPSMI reversal (teaching of mathematics and science in English) by Malaysian newspapers. *Asian Journal of Communication*, 22(5), 449–473. <https://doi.org/10.1080/01292986.2012.701312>
- Yang, M., Cooc, N., & Sheng, L. (2017). An investigation of cross-linguistic transfer between Chinese and English: A meta-analysis. *Asian-Pacific Journal of Second and Foreign Language Education*, 2(1), 15. <https://doi.org/10.1186/s40862-017-0036-9>
- Yang, S., & Yang, H. (2016). Bilingual effects on deployment of the attention system in linguistically and culturally homogeneous children and adults. *Journal of Experimental Child Psychology*, 146, 121–136. <https://doi.org/10.1016/j.jecp.2016.01.011>

- Yao, Z., Wu, J., Zhang, Y., & Wang, Z. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods*, 49, 1374–1385. <https://doi.org/10.3758/s13428-016-0793-2>
- Yap, M., Liow, S., Jalil, S., & Faizal, S. (2010). The Malay Lexicon Project: A database of lexical statistics for 9,592 words. *Behavior Research Methods*, 42, 992–1003. <https://doi.org/10.3758/brm.42.4.992>
- Ye, Y., Mo, L., & Wu, Q. (2017). Mixed cultural context brings out bilingual advantage on executive function. *Bilingualism: Language and Cognition*, 20, 844–852. <https://doi.org/10.1017/s1366728916000481>
- Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: why variations in bilingual experiences matter. *Frontiers in Psychology*, 6(6), e164. <https://doi.org/10.3389/fpsyg.2015.00164>
- Yu, L., Lee, L., Hao, S., Wang, J., He, Y., Hu, J., Robert, L. K., & Zhang, X. (2016). Building Chinese affective resources in valence-arousal dimensions. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 540–545. <https://doi.org/10.18653/v1/n16-1066>
- Yun, J.-Y., Jang, J. H., Jung, W. H., Shin, N. Y., Kim, S. N., Hwang, J. Y., & Kwon, J. S. (2017). Executive dysfunction in obsessive-compulsive disorder and anterior cingulate-based resting state functional connectivity. *Psychiatry Investigation*, 14, 333. <https://doi.org/10.4306/pi.2017.14.3.333>
- Yunus, M. M., & Sukri, S. I. A. (2017). The use of English in teaching mathematics and science: The PPSMI Policy vis-à-vis the DLP. *Advances in Language and Literary Studies*, 8(1), 133. <https://doi.org/10.7575/aiac.all.v.8n.1p.133>
- Zhang, L., & Tsung, L. (2019). Language attitudes and ethnic identity of the present YI. In *Multilingual education* (pp. 85–111). https://doi.org/10.1007/978-3-030-03454-2_5
- Zhang, Y., Wang, T., Huang, P., Li, D., Qiu, J., Shen, T., & Xie, P. (2015). Free Language selection in the bilingual brain: An event-related fMRI study. *Scientific Reports*, 5(1), 11704. <https://doi.org/10.1038/srep11704>

- Zhao, J., Jiang, W., Wang, X., Cai, Z., Liu, Z., & Liu, G. (2020). Exercise, brain plasticity, and depression. *CNS Neuroscience & Therapeutics*, 26, 885–895.
<https://doi.org/10.1111/cns.13385>
- Zheng, X., & Zhou, Y. (2021). The cohort effect of political change on language speaking: Evidence from Hong Kong. *Economics of Transition and Institutional Change*, 29(4), 575–596. <https://doi.org/10.1111/ecot.12286>
- Zhou, B., & Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychonomic Bulletin & Review*, 23(4), 1221–1230.
<https://doi.org/10.3758/s13423-015-0981-6>
- Zhou, B., & Krott, A. (2018). Bilingualism enhances attentional control in non-verbal conflict tasks: Evidence from ex-Gaussian analyses. *Bilingualism: Language and Cognition*, 21(1), 162–180. <https://doi.org/10.1017/s1366728916000869>
- Zhou, W. (2020). Structural priming: A new perspective of language learning. *American Journal of Applied Psychology*, 9(4), 124. <https://doi.org/10.11648/j.ajap.20200904.16>
- Zhu, J. D., Blanco-Elorrieta, E., Sun, Y., Szakay, A., & Sowman, P. F. (2022). Natural vs forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage*, 247, 118797.
<https://doi.org/10.1016/j.neuroimage.2021.118797>
- Zirnstien, M., Bice, K., & Kroll, J. F. (2019). Chapter 3. Variation in language experience shapes the consequences of bilingualism. In *Studies in bilingualism* (pp. 35–47).
<https://doi.org/10.1075/sibil.57.03zir>

Appendix 1: Scales and Questionnaires

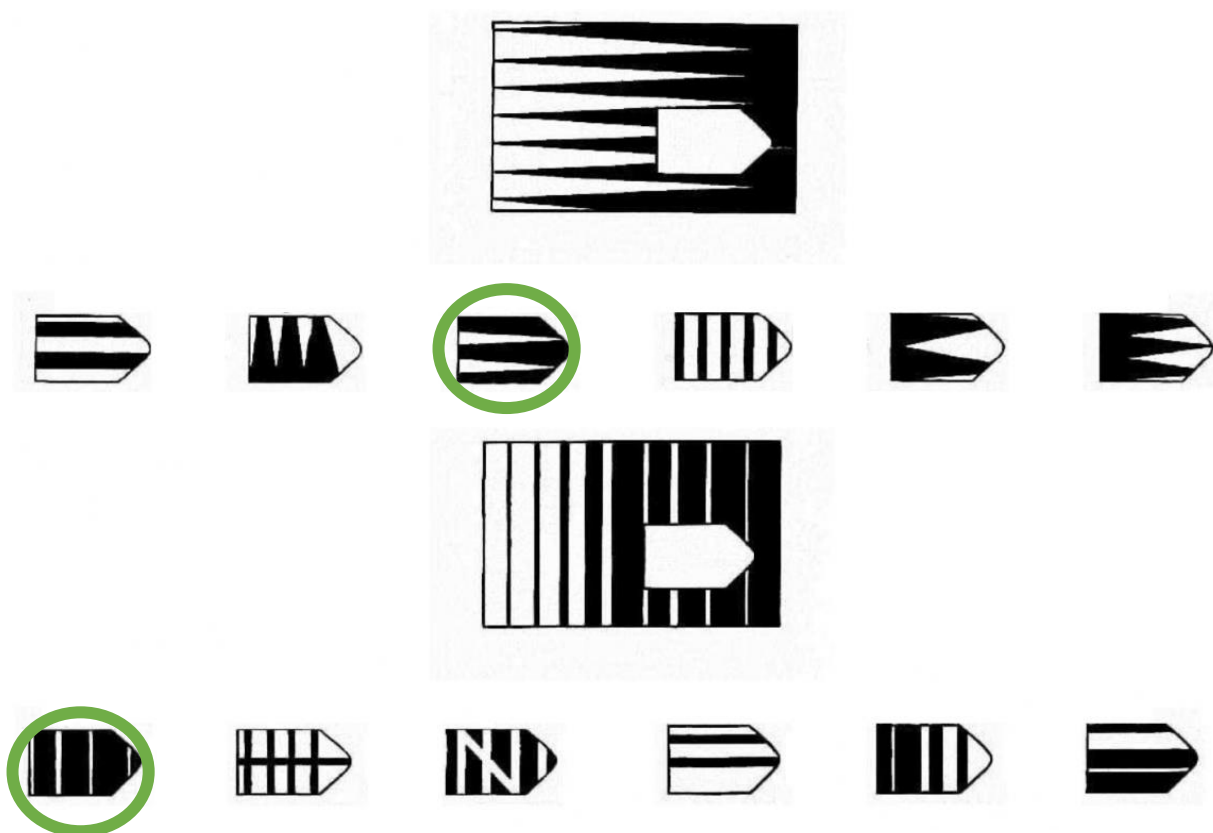
MacArthur's SSS Scale

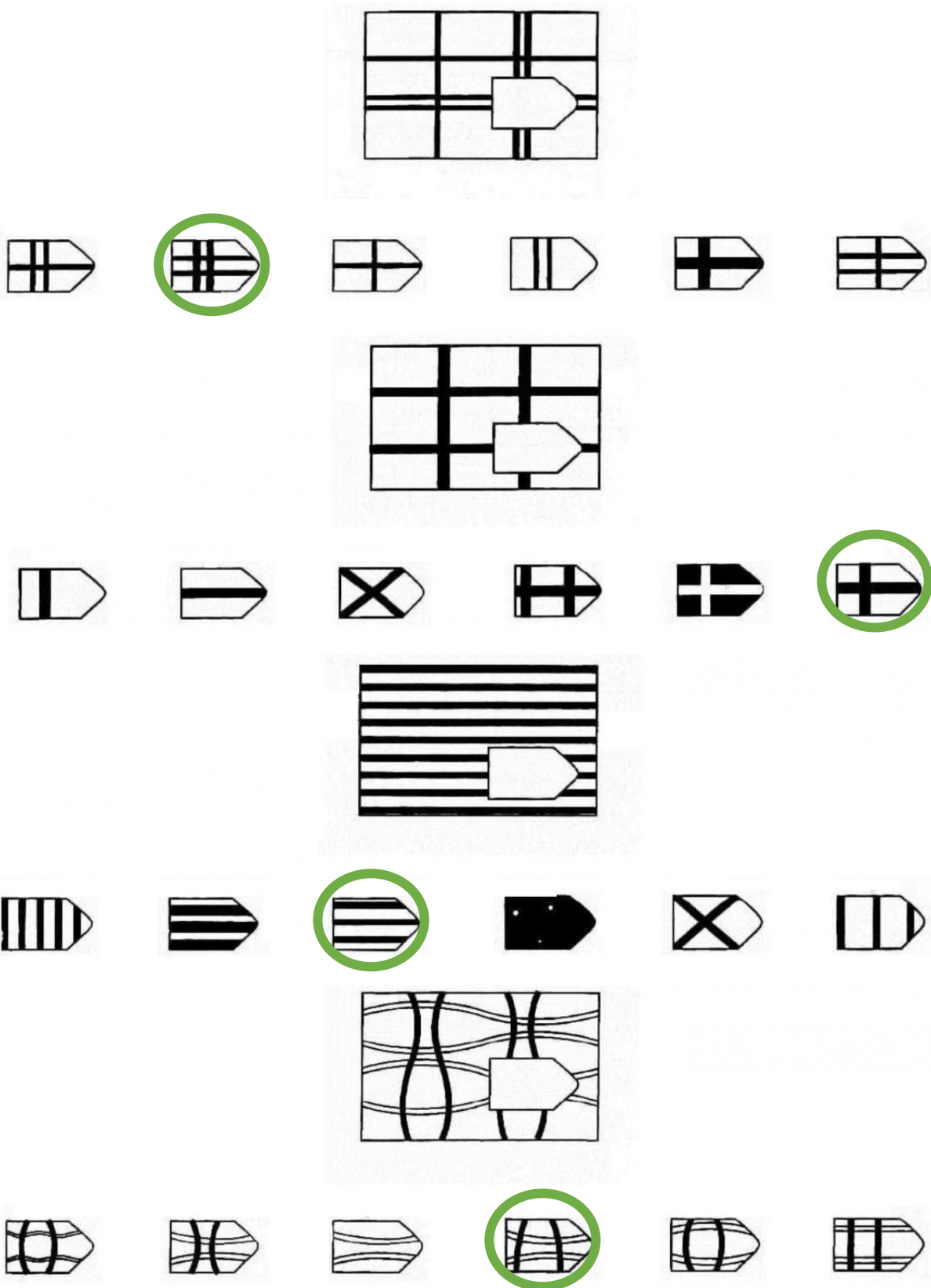
Imagine at the top of the ladder are the people in Malaysia who are the best off, those who have the most money, most education and best jobs. At the bottom are the people in Malaysia who are the worst off, those who have the least money, least education and worst jobs, or no job. Please select the rung number that best represents where you think you stand on the ladder

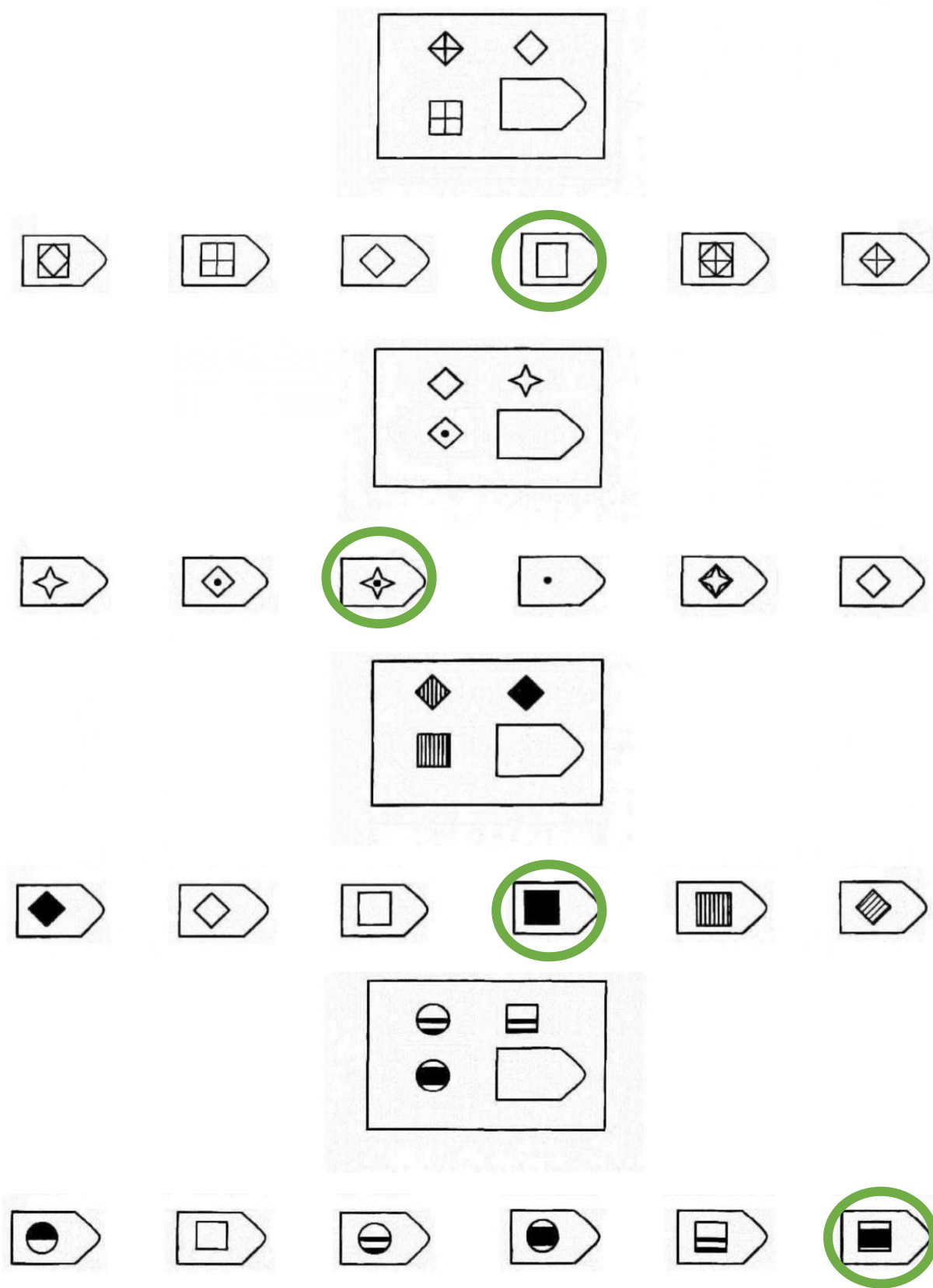


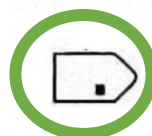
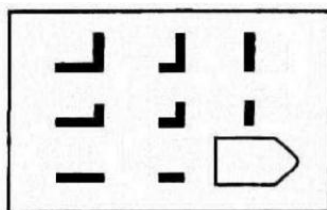
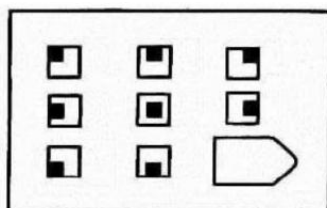
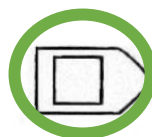
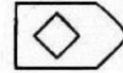
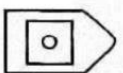
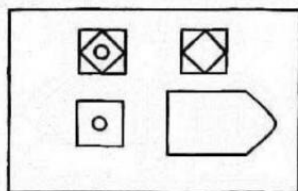
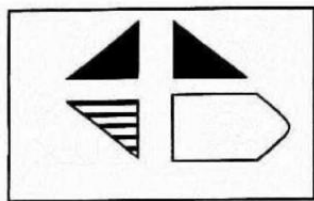
Raven's Standard Progressive Matrices (Adapted)

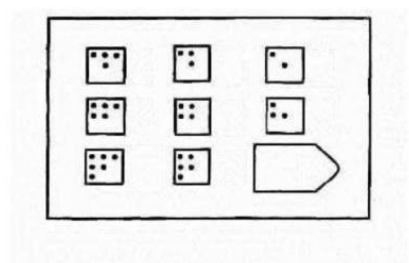
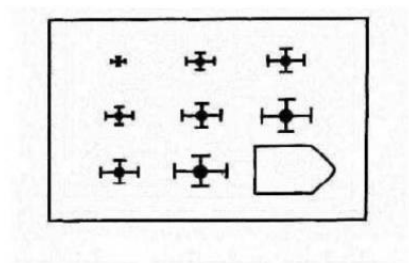
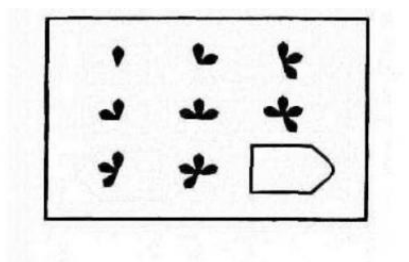
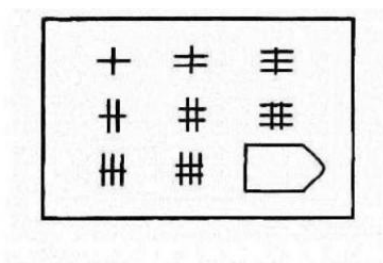
Below, you can see patterns with a piece cut out of it. Each of the pieces below the pattern is the right shape to fill the space, but only one of them is the right pattern, there are 30 patterns with a piece missing. You have to choose which of the pieces below is the right one to complete the pattern. When you think you have found the right piece, click the corresponding number below to register your answer. Continue with the rest of the questions by yourself until you reach the end of the test. The problems are simple at the beginning and get harder as you go on. There is no catch, if you pay attention to the way the answers to the easy problems are found, you will find the later ones less difficult. Work at your own pace Do not miss any out. If you are not sure, guess, as guesses are sometimes right. If you get stuck, move on to the next problems and then come back to the one you had difficulty with. You can have as much time as you like

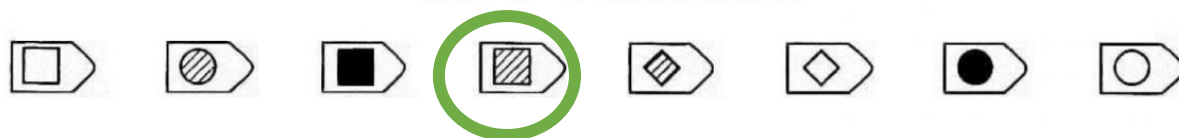
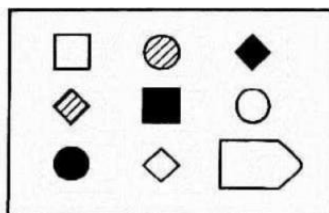
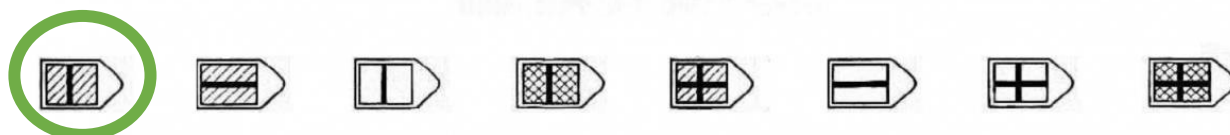
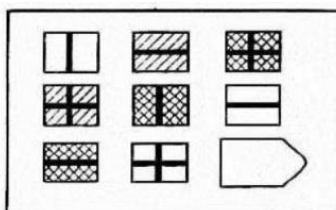
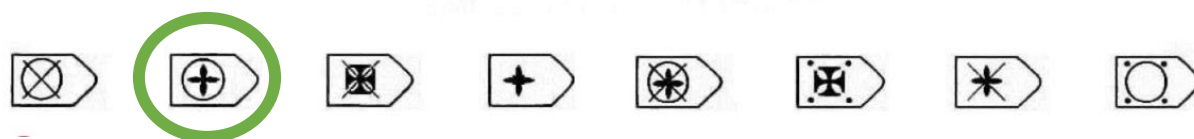
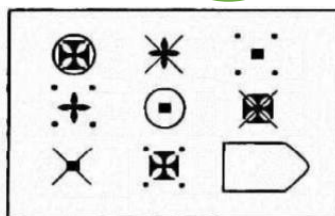
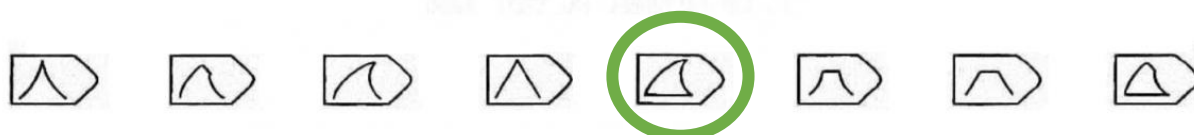
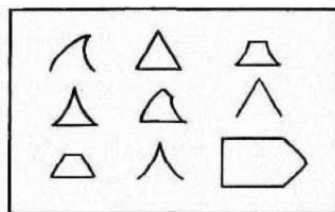


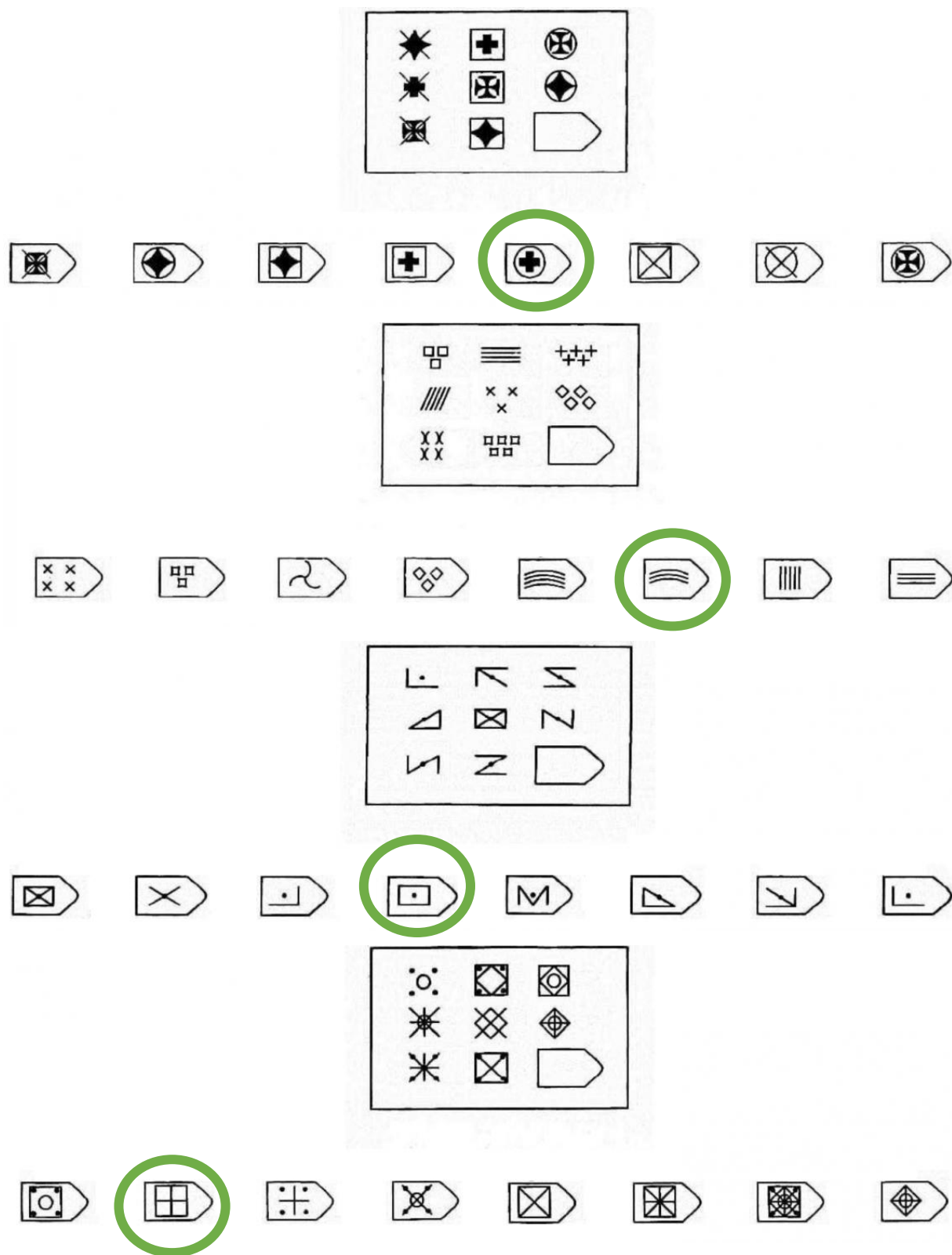












LAQ-LL

For each statement below, on the scale of 1 – Strongly Disagree to 7 Strongly Agree, please rate how much you agree or disagree with the statements. Answer as honestly as you can

LLE_1: Language classes are always among my favorite classes

LLE_2: Language classes are more entertaining than other classes

LLE_3: Studying languages is fun

LLE_4: I like learning languages

LLE_5_R: Language classes are boring^a

LLE_6: I like communicating with multilingual speakers

LLE_7: Studying languages is interesting

LLE_8: Learning languages is worth the effort that it takes

LLE_9: I would take language classes even if they were not required

VoM_1_R: I have no interest in learning languages

VoM_2_R: There is no need to learn other languages beyond my mother tongue^a

VoM_3: I like the idea of being/becoming multilingual

VoM_4: Speaking more than one language is a valuable skill

VoM_5: Language study is a valuable pursuit

VoM_6: I admire people who speak multiple languages

VoM_7_R: Speaking just my mother tongue is enough for me^a

VoM_8_R: Learning a second language is a waste of time^a

LLE = Language Learning Experience; VoM = Value of Multilingualism

LexTALE English Word List

Word	Answer
mensible	FAKE
scornful	REAL
stoutly	REAL
ablaze	REAL
kermshaw	FAKE
moonlit	REAL
lofty	REAL
hurricane	REAL

flaw	REAL
alberation	FAKE
unkempt	REAL
breeding	REAL
festivity	REAL
screech	REAL
savoury	REAL
plaudate	FAKE
shin	REAL
fluid	REAL
spaunch	FAKE
allied	REAL
slain	REAL
recipient	REAL
exprate	FAKE
eloquence	REAL
cleanliness	REAL
dispatch	REAL
rebondicate	FAKE
ingenious	REAL
bewitch	REAL
skave	FAKE
plaintively	REAL
kilp	FAKE
interfate	FAKE
hasty	REAL
lengthy	REAL
fray	REAL
crumper	FAKE
upkeep	REAL
majestic	REAL

magrity	FAKE
nourishment	REAL
abergy	FAKE
proom	FAKE
turmoil	REAL
carbohydrate	REAL
scholar	REAL
turtle	REAL
fellick	FAKE
destraption	FAKE
cylinder	REAL
censorship	REAL
celestial	REAL
rascal	REAL
purrage	FAKE
pulsh	FAKE
muddy	REAL
quirty	FAKE
pudour	FAKE
listless	REAL
wrought	REAL

LexTALE Chinese Characters List

FAKE Characters (30 Characters)

刷

闷

歛

駛

斫

𢇛

𢇛

羽

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

𢇛

吸 碎 财 耘 汲

斂 殍 裕 裨 泡

REAL Characters (60 Characters)

恨 奉 掠 昏

坪 桨 稻 阕

乖 滋 鸣 欢

痹 涕 辜 秃

聿 虜 坤 隅

劈 隧 朔 匈

悼 烹 倘 峒 瞭

冗 凇 蟑 倆

鑱 踊 篋 夜

腋 嗫 椰 跚

丘

顎

觔

堦

淬

蘊

褻

逖

殛

級

髀

鼈

麗

鄣

筵

呷

簌

樵

齡

Malay Cloze Test

Question	Choices	Answer	Source
Suasana _____ di bandar telah menyebabkan Hafiz mengambil cuti beberapa hari dan pulang ke kampungnya bagi menenangkan fikiran.	A huru-hara	C	Oxford Fajar
	B kacau-bilau		
	C hingar-bingar		
	D porak-peranda		
Hazlin berasa _____ untuk memandang ke bawah ketika berada di atas Menara Taming Sari.	A risau	B	Sasbadi
	B gayat		
	C rimas		
	D lemah		
Kereta yang dipandu oleh pemuda itu telah hilang kawalan lalu _____ ke dalam gaung.	A terjerlus	C	Oxford Fajar
	B tersadung		
	C terhumban		
	D terperosok		
Perselisihan yang timbul baru-baru ini telah _____ hubungan dua hala antara kedua-dua negara itu.	A mengeruhkan	A	Sasbadi
	B mengotorkan		
	C melunturkan		
	D melucutkan		
Majlis makan malam diraja diadakan di Istana Negara bersempena dengan hari _____ Sultan Perak.	A persemayaman	B	Pep
	B keputeraan		
	C pemasyhuran		
	D penganugerahan		
Adik lelaki saya yang baru berusia empat tahun itu _____ bercakap apabila orang lain bertanya soalan.	A fasih	D	Cemerlang
	B pandai		
	C mahir		
	D petah		
	A gugup	A	Cemerlang
	B kekok		

Dia masih kelihatan _____ dan teragak-agak ketika berdiri di atas pentas walaupun telah banyak kali menjalankan sesi latihan.	C gerun		
	D pelik		
Kerongsang mutiara itu amat sesuai jika _____ pada tudung bercorak keluaran syarikat ternama tersebut.	A disemat	A	Vision
	B disemad		
	C dicucuk		
	D disangkut		
Pihak Lebu Raya Utara Selatan atau PLUS menasihati pengguna jalan raya agar sentiasa berwaspada semasa memandu ketika jerebu kerana jarak penglihatan agak _____.	A balam-balam	B	Soalan Sebenar
	B samar-samar		
	C sayup-sayup		
	D dekat-dekat		
_____ darah kelihatan di sana sini sejurus kemalangan yang melibatkan dua buah kenderaan itu berlaku.	A Renjisan	D	Vision
	B Tumpahan		
	C Palitan		
	D Tompokan		
Tilam kekabu itu berbau _____ kerana terkena air kencing.	A hamis	D	Oxford Fajar
	B hapak		
	C hanyir		
	D hancing		
Sejak pagi tadi anak lelaki saya asyik _____ kerana kurang sihat.	A mengamuk	B	Cemerlang
	B merengek		
	C merajuk		
	D merintih		
Tok Mat _____ setandan pisang nangka yang diambil dari kebunnya.	A memikul	A	Oxford Fajar
	B mengandar		
	C menggalas		
	D menjunjung		
Pasukan bola baling sekolah kami menang _____ dalam perlawanan akhir peringkat daerah.	A tipis	A	Soalan Sebenar
	B nipis		
	C kecil		

	D sedikit		
Filem terbaharu arahan pengarah muda itu pasti akan mendapat sambutan _____ daripada para penggemar filem komedi.	A luas	D	Sasbadi
	B tinggi		
	C pantas		
	D hangat		
Bunyi gesekan _____ itu sungguh merdu kerana dimainkan oleh pemuzik tersohor tanah air.	A seruling	C	Soalan Sebenar
	B piano		
	C biola		
	D gitar		
Penari dalam persembahan kuda kepang itu semakin _____ pergerakannya apabila rentak muzik semakin rancak.	A tangkas	A	Sasbadi
	B bingkas		
	C cekap		
	D deras		
Wajah gadis yang berjilbab labuh itu _____ semula kenangan pemuda itu bersama bekas kekasihnya.	A mengundang	C	Soalan Sebenar
	B mengimpi		
	C mengimbau		
	D menghimpit		
_____ dalam syarikat itu menjadi isu yang hangat diperkatakan kini.	A Keretakan	D	Oxford Fajar
	B Keburukan		
	C Kesalahan		
	D Kepincangan		
Tiga orang kanak-kanak yang sedang bermain di tepi padang itu lari _____ apabila ternampak orang gila mengamuk.	A bersorakan	C	Vision
	B berkejaran		
	C bertempiran		
	D berpelesen		

LHQ

The LHQ is designed to be a short and simple measure of your language background. Here, you will be tasked to make estimations on certain aspects of your language experience (e.g. Proficiency, Total Years of Use). We understand that these estimations may sometimes be quite difficult to make, however, we ask that you attempt to provide the most accurate estimates while completing the LHQ.

Age of Acquisition

Indicate the language(s) and the **age at which you started** using each language in terms of reading, listening, writing, speaking, and the total number of years you have spent using each language. Only list the languages that you are able to communicate in. For example, if you are just learning a new language, you may not be fluent enough to communicate in it. If this is the case, you may leave this language out of the report.

Note:

*In some cases, you may know how to speak but have never acquired the skill to read in a language. If this applies to you, **please leave the answer box blank for the column that does not apply to you**. For example, if you know how to Listen and Speak in Mandarin, but have never acquired the Reading and Writing skill, leave the columns of Reading and Writing blank.*

	Reading	Listening	Writing	Speaking	Years of Use

	The age you started reading in this language	The age you started listening in this language	The age you started writing in this language	The age you started speaking in this language	You may have learned a language, stopped using it, and then started using it again. Please give the total number of years
First Language					
Second Language					
Third Language					
Fourth Language					

Proficiency

For each of your languages, rate your current ability in terms of reading, listening, writing and speaking

Note: If you have a hard time rating your language skills, consider the following:

"Excellent" is a person who can fluently describe what s/he did in one day without fumbling over vocabulary or grammar

"Very Poor" is someone who is unable to describe what tasks they performed during his/her day

	Reading	Listening	Writing	Speaking
First Language	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)
Second Language	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)
Third Language	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)
Fourth Language	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)	▼ Very Poor (1 ... Excellent (7)

Reading, Writing, Listening Dominance

Estimate how many hours per day you spend engaged in the following activities in each of the languages you have studied or learned

Note: Put the number "0" in the answer box if you do not engage in an activity using a particular language (i.e. you do not watch television in your second language)

	Reading for Fun	Reading for School/Work	Watching Television	Listening to Music/Radio/Podcasts	Using Social Media and Internet	Writing for School/Work
	Estimated hours spent reading for fun in this language	Estimated hours spent reading for school/work in this language	Estimated hours spent watching television in this language	Estimated hours spent listening to Music/Radio/Podcasts in this language	Estimated hours spent using social media and internet in this language	Estimated hours spent writing for school/work in this language

First Language						
Second Language						
Third Language						
Fourth Language						

Speech Dominance

Estimate how many hours per day you spend speaking with the following groups of people in each of the languages you have studied or learned

Note: Put the number "0" in the answer box if you do not speak with a particular group of people using a particular language (i.e. you do not speak to your family in your second language)

	Family Members	Friends*	Classmates	Others (Co-Workers**, Roommates, etc.)
	This category can include extended families as well (e.g. relatives, grandparents)	* Include significant others in this category if you did not include them as family members (e.g. married partners)	If you are no longer studying, please leave this column blank	** Include anyone in the work environment in this category (e.g. if you are a teacher, include students as co-workers)

First Language				
Second Language				
Third Language				
Fourth Language				

Multilingual Switching Questionnaire (MSWQ)

Please, try to answer to what degree the following questions are representative of the manner you use to talk or speak in the language you know (e.g., English-Mandarin, English-Mandarin-Malay). Many of these questions ask you to report your tendency to switch or mix languages during a conversation. Switching and mixing languages is a characteristic of some multilingual contexts or environments (i.e. Malaysia). The present questionnaire aims to identify the language switching patterns that exist in these languages. If you have doubts about how to rate yourself in the following questions, please try to compare your manner of speaking and talking with that of most people, or those who you know very well.

	Never (1)	Very Infrequently (2)	Occasionally (3)	Frequently (4)	Always (5)
I do not remember, or I cannot recall, some words in my other languages when I am speaking in L1	☐	☐	☐	☐	☐
I do not remember, or I cannot recall, some words in my other languages when I am speaking in L2	☐	☐	☐	☐	☐
I do not remember, or I cannot recall, some words in my other languages when I am speaking in L3	☐	☐	☐	☐	☐
I tend to switch languages during a conversation (for example, I switch from L1 to L2 or any other combinations of your languages)	☐	☐	☐	☐	☐
When I cannot recall a word in my other languages, I tend to immediately produce it in L1	☐	☐	☐	☐	☐
When I cannot recall a word in my other languages, I tend to immediately produce it in L2	☐	☐	☐	☐	☐
When I cannot recall a word in my other languages, I tend to immediately produce it in L3	☐	☐	☐	☐	☐
I do not realize when I switch the language during a conversation (e.g., from L2 to L1) or when I mix two or more languages; I often realize it only if I am informed of the switch by another person. (8)	☐	☐	☐	☐	☐
When I switch languages, I do it consciously. (9)	☐	☐	☐	☐	☐

It is difficult for me to control the language switches I introduce during a conversation (e.g., from L2 / L3 to L1 (10)	☐	☐	☐	☐	☐
Without intending to, I sometimes produce the L1 word faster when I am speaking in any of my other languages. (11)	☐	☐	☐	☐	☐
Without intending to, I sometimes produce the L2 word faster when I am speaking in any of my other languages. (12)	☐	☐	☐	☐	☐
Without intending to, I sometimes produce the L3 word faster when I am speaking in any of my other languages. (17)	☐	☐	☐	☐	☐
There are situations in which I always switch between my languages. (13)	☐	☐	☐	☐	☐
There are certain topics or issues for which I normally switch between my languages. (14)	☐	☐	☐	☐	☐

Appendix 2: Priming Stimuli List

English Word Priming List (Experiment 2A, 2B, 2C and 3)

Experiment	Word	Word Length	Valence	Arousal	Zipf
2A & 3	little	6	5.91	4.2	6.03
2A & 3	demand	6	4.1	5.85	4.56
2A & 3	god	3	5.9	5.56	5.38
2A & 3	banking	7	5.47	4.24	4.45
2A & 3	final	5	4.56	4.4	5.46
2A & 3	tiny	4	5.1	4.6	4.87
2A & 3	physical	8	5.9	5.5	4.45
2A & 3	drunk	5	4.06	5.05	4.46
2A & 3	approach	8	5.45	4.48	4.71
2A & 3	weird	5	5.09	4.29	4.75
2A & 3	garlic	6	5.67	4.12	4.45
2A & 3	forces	6	5.58	4.33	4.81
2A & 3	exist	5	5.9	4	4.41
2A & 3	hot	3	5.73	4.76	5.14
2A & 3	serious	7	5.88	4.05	5.16
2A & 3	screen	6	5.55	4.09	4.61
2A & 3	matter	6	5.82	4.05	5.24
2A & 3	mouth	5	5.59	4.14	4.78
2A & 3	auction	7	5.43	4.5	5.17
2A & 3	crucial	7	5.16	4.14	4.6
2A & 3	spread	6	5.57	4.15	4.62
2A & 3	boom	4	4.74	5.57	4.5
2A & 3	training	8	5.19	4.1	4.99
2A & 3	lion	4	5.84	5.29	4.45
2A & 3	break	5	5.1	4.71	5.23
2A & 3	king	4	6	4.45	5.09
2A & 3	beat	4	4.38	5.52	5.07

2A & 3	league	6	5.35	4.25	4.97
2A & 3	odd	3	4.56	4.26	4.7
2A & 3	remind	6	4.81	4.2	4.56
2A & 3	large	5	5.77	5.35	5.1
2A & 3	trade	5	5.91	4.45	4.83
2A & 3	chamber	7	4.73	4.58	4.39
2A & 3	finger	6	5.8	4.15	4.53
2A & 3	main	4	5.27	4.1	5.14
2A & 3	effect	6	5.5	4.08	4.84
2A & 3	news	4	4.96	4.61	5.41
2A & 3	show	4	5.91	4.1	5.77
2A & 3	beer	4	5.67	4.3	4.47
2A & 3	daily	5	5.74	4.2	4.63
2A & 3	want	4	6	5.29	6.22
2A & 3	big	3	5.64	4.33	5.94
2A & 3	army	4	4.65	4.49	4.91
2A & 3	bring	5	5.68	4.29	5.4
2A & 3	dare	4	4.67	5.6	4.4
2A & 3	prank	5	4.79	5.36	4.39
2A & 3	steam	5	5.4	4	4.39
2A & 3	theory	6	5.65	4.41	4.44
2A & 3	spend	5	5.91	4.48	5.13
2A & 3	bone	4	5.24	4.75	4.51
2A & 3	enormous	8	5.68	5.05	4.62
2A & 3	object	6	5.52	4.08	4.47
2A & 3	plot	4	4.9	4.52	4.38
2A & 3	medal	5	5.2	5.3	4.64
2A & 3	incident	8	4.2	5.65	4.39
2A & 3	wet	3	5.64	5	4.76
2A & 3	pack	4	5.55	4.1	4.62

2A & 3	move	4	4.95	4.24	5.49
2A & 3	force	5	4.17	5.35	4.96
2A & 3	bloke	5	5	4.24	4.44
2A & 3	race	4	5.44	5.2	5.11
2A & 3	massive	7	5.18	5.79	5.04
2A & 3	unusual	7	5.81	4.15	4.81
2A & 3	military	8	4.9	5.8	4.82
2A & 3	border	6	4.68	4.3	4.61
2A & 3	bell	4	5.67	4.7	4.54
2A & 3	strategy	8	5.95	4.15	4.6
2A & 3	sector	6	4.67	4	4.95
2A & 3	machine	7	5	4.39	4.83
2A & 3	impact	6	5.76	5.85	4.85
2A & 3	religion	8	5.42	5	4.39
2A & 3	network	7	5.05	4.43	4.44
2A & 3	staff	5	5.18	4.3	4.95
2A & 3	majority	8	4.95	4.3	4.71
2A & 3	grab	4	5.44	4.5	4.55
2A & 3	growth	6	6	4.95	4.91
2A & 3	pass	4	5.73	4.38	5.14
2A & 3	squad	5	5.35	4.33	4.39
2A & 3	aim	3	5.95	4	4.49
2A & 3	prepare	7	5.32	4.15	4.48
2A & 3	quiz	4	4.53	4.95	4.57
2A & 3	man	3	5.42	4.36	5.86
2A & 3	debate	6	4.95	4.86	4.92
2A & 3	empire	6	5.36	4.59	4.43
2A & 3	vote	4	5.24	5.05	5.16
2A & 3	tour	4	6	4.8	4.76
2A & 3	wheel	5	5.9	4	4.51

2A & 3	crazy	5	5.14	5.9	4.79
2A & 3	block	5	4.48	4.25	4.57
2A & 3	concern	7	5.05	4.7	4.7
2A & 3	heat	4	4.77	5.75	4.76
2A & 3	throw	5	5.55	4.52	4.88
2A & 3	case	4	5.18	4.14	5.43
2A & 3	drop	4	4.23	4.67	4.9
2A & 3	editor	6	5.43	4.09	4.39
2A & 3	try	3	5.64	4	5.65
2A & 3	leave	5	4.68	4.48	5.49
2A & 3	source	6	5.48	4.5	4.44
2A & 3	premier	7	5.37	4.62	4.39
2A & 3	test	4	4.44	4.3	5.06
2A & 3	role	4	5.1	4.4	4.94
2A & 3	escape	6	5.5	4.55	4.66
2A & 3	vast	4	5.68	4.71	4.53
2A & 3	welsh	5	4.53	4	5
2A & 3	bar	3	5	4.53	4.75
2A & 3	cue	3	5.84	4.35	4.39
2A & 3	busy	4	5.41	4.52	5
2A & 3	pot	3	5.81	4	4.76
2A & 3	boy	3	5.84	4.11	5.28
2A & 3	official	8	5.75	4.04	4.58
2A & 3	secret	6	5.33	4.14	4.93
2A & 3	industry	8	5.07	4.25	4.95
2A & 3	job	3	5.64	4	5.6
2A & 3	deliver	7	6	4.75	4.69
2A & 3	project	7	5	4.55	4.94
2A & 3	rent	4	4.41	4.25	4.53
2A & 3	chest	5	5.18	4.95	4.48

2A & 3	dark	4	5.08	4.09	4.94
2A & 3	attitude	8	4.72	4.76	4.44
2A & 3	wine	4	5.85	4.24	4.84
2A & 3	prince	6	5.44	5.15	4.74
2A & 3	fox	3	5.52	4.36	4.46
2A & 3	stop	4	4.73	4.14	5.55
2A & 3	famous	6	5.89	4.2	5.03
2A & 3	airport	7	6	5.5	4.54
2A & 3	garage	6	5.5	4.1	4.46
2A & 3	adult	5	5.9	4.36	4.4
2A & 3	prefer	6	5.62	4.65	4.48
2A & 3	firm	4	5.36	4.3	4.54
2A & 3	past	4	5.09	4.71	5.3
2A & 3	loose	5	4.52	4.3	4.41
2A & 3	tumble	6	5.65	4.9	4.5
2A & 3	power	5	5.97	5.41	5.29
2A & 3	player	6	5.55	5.95	4.99
2A & 3	use	3	5.18	4.19	5.57
2B	time	4	5.31	4.64	6.35
2B	say	3	5.91	4.43	6.13
2B	take	4	4.82	4.52	6.08
2B	little	6	5.91	4.2	6.03
2B	world	5	6.5	5.32	5.88
2B	man	3	6.73	5.24	5.86
2B	thought	7	6.39	4.83	5.84
2B	nice	4	6.55	4.38	5.83
2B	part	4	5.11	3.82	5.67
2B	sorry	5	4.81	4.62	5.65
2B	try	3	5.64	4	5.65
2B	top	3	6	4.33	5.64

2B	job	3	5.64	4	5.6
2B	moment	6	5.76	3.83	5.57
2B	idea	4	7	5.86	5.56
2B	answer	6	6.63	5.41	5.56
2B	game	4	6.98	5.89	5.54
2B	water	5	6.61	4.97	5.53
2B	high	4	5.77	4.1	5.53
2B	young	5	6.89	5.64	5.51
2B	police	6	4.59	5.95	5.5
2B	mind	4	6.68	5	5.47
2B	final	5	4.56	4.4	5.46
2B	chance	6	6.02	5.38	5.46
2B	school	6	5.41	4.57	5.45
2B	face	4	6.39	5.04	5.44
2B	hand	4	5.95	4.4	5.44
2B	case	4	5.18	4.14	5.43
2B	news	4	5.3	5.17	5.41
2B	red	3	6.41	5.29	5.41
2B	saw	3	4.16	4.18	5.4
2B	city	4	6.03	5.24	5.4
2B	cut	3	3.64	5	5.35
2B	person	6	6.32	4.19	5.34
2B	huge	4	5.72	4.24	5.32
2B	girl	4	6.87	4.29	5.29
2B	space	5	6.78	5.14	5.29
2B	garden	6	6.71	4.39	5.28
2B	boy	3	6.32	4.58	5.28
2B	black	5	5.39	4.61	5.26
2B	door	4	5.13	3.8	5.26
2B	white	5	6.47	4.37	5.25

2B	history	7	5.24	3.93	5.25
2B	matter	6	5.82	4.05	5.24
2B	air	3	6.34	4.12	5.22
2B	green	5	6.18	4.28	5.22
2B	prime	5	6	4.71	5.2
2B	street	6	5.22	3.39	5.2
2B	fish	4	6.04	4	5.19
2B	kids	4	6.91	5.27	5.19
2B	cost	4	4.11	4.9	5.17
2B	auction	7	5.43	4.5	5.17
2B	vote	4	5.24	5.05	5.16
2B	blue	4	6.76	4.31	5.16
2B	serious	7	5.08	4	5.16
2B	art	3	6.68	4.86	5.15
2B	wife	4	6.33	4.93	5.15
2B	hot	3	5.73	4.76	5.14
2B	bid	3	5.16	4	5.14
2B	health	6	6.81	5.13	5.14
2B	control	7	4.43	4.76	5.14
2B	pass	4	5.73	4.38	5.14
2B	spend	5	5.91	4.48	5.13
2B	office	6	5.24	4.08	5.13
2B	tough	5	4.45	4.24	5.12
2B	race	4	5.44	5.2	5.11
2B	large	5	5.77	5.35	5.1
2B	board	5	4.82	3.36	5.1
2B	expect	6	5.95	4.86	5.09
2B	royal	5	5.25	4.23	5.09
2B	queen	5	6.44	4.76	5.08
2B	rain	4	5.08	3.65	5.08

2B	trouble	7	3.03	6.85	5.08
2B	month	5	5.15	4.03	5.08
2B	key	3	5.68	3.7	5.07
2B	beat	4	4.38	5.52	5.07
2B	test	4	4.44	4.3	5.06
2B	shop	4	5.89	4.76	5.05
2B	major	5	5.52	4.24	5.05
2B	catch	5	5.64	4	5.04
2B	massive	7	5.18	5.79	5.04
2B	famous	6	6.98	5.73	5.03
2B	doctor	6	5.2	5.86	5.02
2B	church	6	6.28	4.34	5.02
2B	class	5	5.73	4.43	5.01
2B	welsh	5	4.53	4	5
2B	busy	4	5.41	4.52	5
2B	player	6	5.55	5.95	4.99
2B	society	7	5.24	4.55	4.98
2B	total	5	5.32	4.25	4.97
2B	league	6	5.35	4.25	4.97
2B	social	6	6.88	4.98	4.97
2B	corner	6	4.36	3.91	4.96
2B	staff	5	5.18	4.3	4.95
2B	taste	5	6.66	5.22	4.95
2B	sector	6	4.67	4	4.95
2B	role	4	5.1	4.4	4.94
2B	charge	6	4.61	4.25	4.94
2B	river	5	6.85	4.51	4.93
2B	debate	6	4.95	4.86	4.92
2B	foot	4	5.02	3.27	4.92
2B	glass	5	4.75	4.27	4.92

2B	push	4	5.39	4.4	4.92
2B	popular	7	5.71	4.74	4.91
2B	growth	6	6	4.95	4.91
2B	army	4	4.72	5.03	4.91
2B	manager	7	4.82	4.15	4.9
2B	present	7	6.95	5.12	4.9
2B	drop	4	4.23	4.67	4.9
2B	clock	5	5.14	4.02	4.89
2B	natural	7	6.59	4.09	4.89
2B	bear	4	5.33	4.5	4.88
2B	union	5	5.85	4.05	4.88
2B	nervous	7	3.29	6.59	4.87
2B	tiny	4	5.1	4.6	4.87
2B	coast	5	5.98	4.59	4.86
2B	excuse	6	4.05	4.48	4.86
2B	broken	6	3.05	5.43	4.86
2B	impact	6	5.76	5.85	4.85
2B	shut	4	4.5	4.05	4.85
2B	letter	6	6.61	4.9	4.85
2B	dress	5	6.41	4.05	4.85
2B	rate	4	5.68	4.17	4.85
2B	wine	4	5.95	4.78	4.84
2B	effect	6	5.5	4.08	4.84
2B	window	6	5.91	3.97	4.84
2B	trade	5	5.91	4.45	4.83
2B	machine	7	5.09	3.82	4.83
2B	cat	3	5.72	4.38	4.83
2B	explain	7	5.41	4	4.83
2B	radio	5	6.73	4.78	4.82
2B	forces	6	5.58	4.33	4.81

2B	damage	6	3.05	5.57	4.81
2B	farm	4	5.53	3.9	4.81
2B	unusual	7	5.81	4.15	4.81
2C	sir	3	5.45	4.33	5.3
2C	thin	4	5.47	4.5	4.48
2C	credit	6	5	4.38	4.76
2C	bother	6	4.44	4.2	4.45
2C	shop	4	5.89	4.76	5.05
2C	red	3	5.67	5.02	5.41
2C	royal	5	5.25	4.23	5.09
2C	hammer	6	5.17	4.56	4.66
2C	knife	5	4.33	4.86	4.49
2C	degree	6	5.63	4.1	4.49
2C	speak	5	5.91	4.19	5.09
2C	prime	5	6	4.71	5.2
2C	push	4	5.39	4.4	4.92
2C	reveal	6	5.86	4.14	4.77
2C	must	4	5.45	4.1	5.67
2C	chain	5	4.79	4.05	4.49
2C	total	5	5.32	4.25	4.97
2C	bet	3	4.95	4.76	4.8
2C	alcohol	7	4.11	5.17	4.52
2C	edge	4	4.43	4.65	4.8
2C	limit	5	4.53	4.29	4.44
2C	direct	6	5.75	4.29	4.52
2C	school	6	5.41	4.57	5.45
2C	due	3	4.94	4.38	4.79
2C	charge	6	4.61	4.25	4.94
2C	chat	4	5.75	4.27	4.6
2C	mayor	5	4.59	4.15	4.39

2C	subject	7	5.33	4.2	4.74
2C	sell	4	5.33	4.19	5.22
2C	split	5	4.5	5.55	4.63
2C	work	4	5.05	4.33	5.99
2C	huge	4	5.72	4.24	5.32
2C	hurry	5	4.32	5.62	4.53
2C	stadium	7	5.53	5.25	4.47
2C	kick	4	4.55	5.1	4.91
2C	top	3	6	4.33	5.64
2C	complex	7	4.21	4.6	4.44
2C	speaker	7	5.32	4.35	4.53
2C	officer	7	4.91	4.38	4.74
2C	rise	4	5.73	5.3	4.78
2C	bid	3	5.16	4	5.14
2C	button	6	5.48	4.1	4.46
2C	target	6	5.24	5.35	4.81
2C	police	6	4.59	5.95	5.5
2C	head	4	5.86	4.45	5.61
2C	measure	7	5.14	4.36	4.48
2C	boss	4	4.76	4.43	4.68
2C	treasury	8	5.8	4.29	4.41
2C	cope	4	5.1	4.1	4.51
2C	nurse	5	5.41	4.64	4.5
2C	hard	4	4.35	4.5	5.58
2C	frame	5	5.32	4.04	4.76
2C	audience	8	5.89	4.91	4.86
2C	change	6	5.09	4.05	5.49
2C	storm	5	5.74	5.86	4.46
2C	coast	5	5.9	4.17	4.86
2C	shut	4	4.5	4.05	4.85

2C	opening	7	5.95	4.6	4.79
2C	sorry	5	4.81	4.62	5.65
2C	western	7	5.35	4.48	4.62
2C	crown	5	6	4.52	4.53
2C	class	5	5.73	4.43	5.01
2C	pound	5	4.62	4.26	4.51
2C	plane	5	5.72	4.91	4.58
2C	mood	4	5.29	4.5	4.5
2C	load	4	4.43	4.55	4.49
2C	wild	4	5.94	5.9	4.96
2C	cheap	5	5.24	4.47	4.65
2C	trick	5	4.89	5.57	4.62
2C	pub	3	5.45	4.36	4.71
2C	deputy	6	5.3	4.05	4.47
2C	vehicle	7	6	4.68	4.52
2C	bank	4	6	4.19	5.28
2C	nation	6	6	4.78	4.76
2C	anti	4	4.47	4.23	4.54
2C	rate	4	5.68	4.17	4.85
2C	say	3	5.91	4.43	6.13
2C	manager	7	4.82	4.15	4.9
2C	shake	5	4.59	5.2	4.58
2C	take	4	4.82	4.52	6.08
2C	speech	6	5.62	4.1	4.72
2C	fashion	7	5.26	5.25	4.63
2C	increase	8	5.9	5.06	4.8
2C	tricky	6	4.78	4.86	4.51
2C	location	8	5.77	4.58	4.64
2C	lift	4	5.32	4.9	4.75
2C	reform	6	5.05	4.78	4.63

2C	expect	6	5.95	4.86	5.09
2C	scratch	7	4.95	4.35	4.39
2C	paint	5	5.74	4.5	4.7
2C	high	4	5.77	4.1	5.53
2C	society	7	5.24	4.55	4.98
2C	body	4	5.95	4.62	5.18
2C	medical	7	5.22	4.05	4.63
2C	election	8	4.5	4.23	5.07
2C	word	4	5.77	4	5.29
2C	cost	4	4.11	4.9	5.17
2C	property	8	6	4.75	5.37
2C	vital	5	5.84	5.15	4.59
2C	bear	4	5.33	4.5	4.88
2C	popular	7	5.71	4.74	4.91
2C	van	3	4.43	4.35	4.69
2C	maximum	7	4.7	5.05	4.48
2C	explain	7	5.41	4	4.83
2C	reality	7	5.73	4.1	4.7
2C	crowd	5	4.48	5.15	4.76
2C	foreign	7	4.86	5.43	4.76
2C	pocket	6	5.67	4.5	4.63
2C	spare	5	5.57	4.45	4.44
2C	carbon	6	4.5	4.05	4.39
2C	saw	3	4.16	4.18	5.4
2C	act	3	5.64	4.19	4.99
2C	maker	5	5.74	4.16	4.47
2C	union	5	5.85	4.05	4.88
2C	mirror	6	5.9	4.55	4.45
2C	tackle	6	5.38	4.43	4.59
2C	drama	5	4.37	5.14	4.52

2C	mystery	7	5.29	5.04	4.7
2C	sharp	5	5.24	6	4.55
2C	stretch	7	5	4.73	4.38
2C	control	7	4.43	4.76	5.14
2C	twist	5	4.61	4.27	4.4
2C	doctor	6	5.93	4.05	5.02
2C	giant	5	5.41	5.1	4.64
2C	finance	7	4.81	5.6	4.41
2C	catch	5	5.64	4	5.04
2C	major	5	5.52	4.24	5.05
2C	attempt	7	6	4.94	4.6
2C	master	6	5.72	4.4	4.76
2C	drug	4	4.11	4.48	4.5
2C	pull	4	5.59	4.1	4.94
2C	gas	3	4.06	4.4	4.77
2C	tough	5	4.45	4.24	5.12
2C	roman	5	5.55	4.36	4.52
2C	rugby	5	5.94	4.1	4.63

Pinyin Word Priming List (Experiment 2A & 3)

Chinese Characters (Experiment 3)	Pinyin (Experiment 2)	English translations	Valence	Arousal	ZipF
问题	wèn tí	problem	4.55	4.77	6
时间	shí jiān	time	4.05	4.77	5.96
不要	bú yào	Don't	4	4.20	5.79
生活	shēng huó	life	5.21	4.5	5.67
很多	hěn duō	a lot of	5.6	4.20	5.65
世界	shì jiè	world	4.82	4.45	5.54
重要	zhòng yào	important	6	5.60	5.54
得到	dé dào	get	5.6	4.00	5.52
不能	bù néng	Can't	4	4.20	5.43
注意	zhù yì	note	5.6	4.60	5.33
小心	xiǎo xīn	Be careful	4.6	6.00	5.32
奇怪	qí guài	strange	4.2	5.20	5.24
控制	kòng zhì	control	4.2	5.00	5.22
容易	róng yì	easy	6	6	5.18
理解	lǐ jiě	understanding	5.8	4.20	5.18
夫人	fū ren	lady	4.58	4.19	5.18
事实	shì shí	fact	4.77	4.23	5.17
玩笑	wán xiào	joke	5.6	6.00	5.17
任务	rèn wu	task	4.38	4.29	5.17
抓住	zhuā zhù	Catch	4.4	4.80	5.12
突然	tū rán	suddenly	4.6	5.20	5.11

介意	jiè yì	Mind	4.1	4.50	5.08
人类	rén lèi	Humanity	4.63	5.07	5.06
必要	bì yào	necessary	5	5.40	5.04
撒谎	sā huǎng	lie	5.07	4.9	5.04
遇到	yù dào	Encounter	5	4.20	5.02
消失	xiāo shī	disappear	4	4.00	4.99
命令	mìng lìng	command	4.2	5.40	4.97
居然	jū rán	Unexpectedly	4.6	4.60	4.93
人生	rén shēng	life	5.2	4.83	4.93
感情	gǎn qíng	feeling	5.4	5.20	4.86
公平	gōng píng	fair	5.8	4.00	4.84
习惯	xí guàn	habit	5.05	5.17	4.82
公园	gōng yuán	park	4.13	5.17	4.82
健康	jiàn kāng	health	6	4.00	4.81
逮捕	dài bǔ	arrest	4	5.40	4.8
取消	qǔ xiāo	cancel	4	4.40	4.75
在乎	zài hu	care	5.6	4.40	4.74
竟然	jìng rán	actually	4.8	4.80	4.72
记者	jì zhě	reporter	4.89	4.43	4.7
过程	guò chéng	process	4.78	4.61	4.7
事业	shì yè	cause	5.64	5.55	4.69
正式	zhèng shì	formal	5.4	4.20	4.68
士兵	shì bīng	soldier	4.39	4.61	4.68
规则	guī zé	rule	4.67	4.63	4.67
付出	fù chū	Give	5.8	5.60	4.67
想念	xiǎng niàn	miss	4.8	5.40	4.64

情绪	qíng xù	mood	4.74	5	4.63
速度	sù dù	speed	4.02	4.44	4.63
道理	dào lǐ	reason	4.47	5.17	4.62
保安	bǎo ān	Security guard	4.06	4.51	4.61
巨大	jù dà	huge	5.2	4.80	4.6
退出	tuì chū	drop out	4.2	4.00	4.59
伤口	shāng kǒu	wound	5.11	5.96	4.59
掌握	zhǎng wò	grasp	5.2	4.20	4.58
思考	sī kǎo	Thinking	5.36	5.04	4.56
地图	dì tú	map	5.14	5.27	4.55
幻想	huàn xiǎng	fantasy	4.6	4.20	4.55
清理	qīng lǐ	Clean up	4.8	4.20	4.55
不可	bù kě	No	4	4.20	4.54
坦白	tǎn bái	frank	6	5.60	4.54
好奇	hào qí	curious	5	4.20	4.52
自豪	zì háo	pride	4.98	4.49	4.5
热情	rè qíng	enthusiasm	4.64	4.11	4.5
投入	tóu rù	Invest in	6	5.20	4.5
海军	hǎi jūn	navy	4.11	4.42	4.49
微笑	wēi xiào	smile	4.47	4.97	4.49
搜索	sōu suǒ	search for	4.4	4.80	4.47
海盗	hǎi dào	pirate	5.36	5.5	4.45
增加	zēng jiā	increase	5.8	4.00	4.45
渴望	kě wàng	desire	5.6	5.80	4.44
全力	quán lì	Full force	5.8	5.40	4.44
古怪	gǔ guài	eccentric	4.2	4.80	4.44

期望	qī wàng	expect	6	5.20	4.43
精力	jīng lì	energy	5.4	4.00	4.42
深刻	shēn kè	profound	5.4	4.60	4.4
豁免	huò miǎn	Exempt	5.2	4.60	4.39
罪名	zuì míng	Charge	5.54	5.98	4.38
经典	jīng diǎn	classic	5.72	4.86	4.36
死刑	sǐ xíng	death penalty	4.64	4.29	4.36
苹果	píng guǒ	apple	4.47	4.73	4.36
顾问	gù wèn	consultant	5.06	5.53	4.35
敏感	mǐn gǎn	sensitive	4.6	5.00	4.35
文化	wén huà	culture	5.31	4.74	4.35
流行	liú xíng	popular	5.6	5.20	4.34
封面	fēng miàn	Cover	4.94	5.19	4.32
奴隶	nú lì	slave	5.51	5.03	4.31
食品	shí pǐn	food	5.23	5.45	4.31
提高	tí gāo	improve	5.8	5.00	4.31
支援	zhī yuán	support	5.6	4.20	4.29
限制	xiàn zhì	limit	4.2	4.20	4.29
肮脏	āng zāng	dirty	5.11	5.64	4.29
排除	pái chú	exclude	4.2	4.40	4.28
迷路	mí lù	get lost	4.2	5.20	4.27
措施	cuò shī	Measures	5.46	5.91	4.26
包裹	bāo guǒ	package	5.72	5.08	4.25
伪装	wěi zhuāng	camouflage	5.5	4.86	4.25
蜜蜂	mì fēng	bee	4.07	5.39	4.24
护照	hù zhào	passport	5.07	4.36	4.23

知识	zhī shì	know how	4.32	4	4.23
收养	shōu yǎng	adoption	5.42	4.5	4.23
硬币	yìng bì	coin	5.11	4.26	4.23
轻易	qīng yì	easily	5.09	5	4.22
困惑	kùn huò	confused	4.2	4.80	4.22
有意	yǒu yì	Intentionally	4.6	4.40	4.21
使命	shǐ mìng	mission	4.54	5.86	4.21
害羞	hài xiū	shy	4.2	4.60	4.21
呵呵	hē hē	Ha ha	6	4.60	4.2
掩盖	yǎn gài	cover	5.97	5.57	4.2
除掉	chú diào	Get rid of	4	4.20	4.2
遭遇	zāo yù	encounter	4	4.80	4.19
矛盾	máo dùn	contradiction	4.2	4.60	4.19
清单	qīng dān	Checklist	4.59	4.61	4.19
充分	chōng fèn	full	5.8	4.40	4.19
迅速	xùn sù	rapid	5.8	4.80	4.18
领带	lǐng dài	tie	4.68	4.61	4.18
保镖	bǎo biāo	bodyguard	4.78	4.71	4.15
旅游	lǚ yóu	tourism	5.14	4.54	4.15
哭泣	kū qì	cry	5.18	5.39	4.14
清除	qīng chú	Clear	4	4.40	4.14
手铐	shǒu kào	Handcuffs	4	4.27	4.14
名声	míng shēng	reputation	5.15	5.27	4.13
征服	zhēng fú	conquer	5.2	5.60	4.13
香烟	xiāng yān	cigarette	5.2	4.88	4.12
克服	kè fú	get over	5.2	5.40	4.12

奇妙	qí miào	wonderful	5.8	4.20	4.12
消除	xiāo chú	eliminate	4	4.40	4.12
防止	fáng zhǐ	prevent	4.8	5.00	4.11
贿赂	huì lù	bribe	4.63	4.2	4.1
贵族	guì zú	noble	6	5.32	4.09
不足	bù zú	insufficient	4.2	4.20	4.09
火炬	huǒ jù	torch	5.51	5.57	4.09
滑稽	huá jī	funny	4.2	5.00	4.08
有名	yǒu míng	famous	6	4.20	4.08
列车	liè chē	train	5	4.93	4.08

Malay Words Priming List (Experiment 2B)

Malay	English	Valence	Arousal	Number of Letters	ZipF
kata	word	5.77	4	4	6.26
acara	event	6.21	5.1	5	6.06
utama	main	5.27	4.1	5	6.03
besar	big	5.64	4.33	5	6.02
nama	name	5.55	4.25	4	5.8
harus	must	5.45	4.1	5	5.74
tuan	sir	5.45	4.33	4	5.6
ingin	want	6	5.29	5	5.55
kerja	work	5.05	4.33	5	5.55
kuning	yellow	5.61	4.43	6	5.33
bangkit	rise	5.73	5.3	7	5.3
kuasa	power	6.54	6.67	5	5.25
amanah	trust	6.68	5.3	6	5.23
minat	interest	6.97	5.66	5	5.22
mutu	quality	6.25	4.48	4	5.18
rasmi	official	5.75	4.04	5	5.17
sumber	source	5.48	4.5	6	5.17
kampung	village	5.92	4.08	7	5.12
segera	quick	6.64	6.57	6	5.08
hendak	want	6	5.29	6	5.07
alih	move	4.95	4.24	4	5.05
wujud	exist	5.9	4	5	5.04
tinggal	leave	4.68	4.48	7	5.04
kena	hit	4.33	5.73	4	5.03

pecut	sprint	5.67	5.9	5	5.01
murah	cheap	5.24	4.47	5	4.98
kayuh	paddle	5.38	4.17	5	4.95
batu	rock	5.56	4.52	4	4.9
henti	stop	4.73	4.14	5	4.88
lancar	smooth	6.58	4.91	6	4.87
gerak	move	4.95	4.24	5	4.85
pasar	market	5.66	4.12	5	4.81
sasar	target	5.24	5.35	5	4.8
guna	use	5.18	4.19	4	4.78
hambar	bland	4.1	3.29	6	4.78
kasar	rough	4.74	5.33	5	4.76
jarum	needle	3.82	5.36	5	4.76
tunjuk	show	5.91	4.1	6	4.76
kepala	head	5.86	4.45	6	4.75
buku	book	5.72	4.17	4	4.74
rahsia	secret	5.33	4.14	6	4.73
mimpi	dream	6.73	4.53	5	4.71
naluri	instinct	5.67	4.44	6	4.71
anggap	assume	4.41	4.2	6	4.69
kuda	horse	5.89	3.89	4	4.69
lutut	knee	4.7	4.25	5	4.67
kapal	ship	5.55	4.38	5	4.66
ladang	field	6.2	4.08	6	4.62
wayang	movie	6.86	4.93	6	4.62
palang	bar	6.42	5	6	4.61
putera	prince	5.44	5.15	6	4.6

tonton	watch	5.78	4.1	6	4.59
jual	sell	5.33	4.19	4	4.58
putus	break	5.1	4.71	5	4.58
bawa	bring	5.68	4.29	4	4.56
dewasa	adult	6.49	4.76	6	4.54
pokok	tree	6.32	3.42	5	4.53
guru	teacher	5.68	4.05	4	4.53
dada	chest	5.18	4.95	4	4.51
pakai	use	5.18	4.19	5	4.51
permata	jewel	7	5.38	7	4.51
desa	village	5.92	4.08	4	4.47
gula	sugar	6.74	5.64	4	4.45
bahang	heat	4.77	5.75	6	4.42
ambak	bank	6	4.19	5	4.41
sewa	rent	4.41	4.25	4	4.4
amar	warn	4.9	4.65	4	4.39
basah	wet	5.64	5	5	4.38
sejuk	cold	4.02	5.19	5	4.38
laku	act	5.64	4.19	4	4.34
sempit	narrow	4.95	4.53	6	4.33
tubuh	body	5.55	5.52	5	4.32
aneh	weird	5.09	4.29	4	4.3
payah	hard	5.22	5.12	5	4.26
gelap	dark	4.71	4.28	5	4.25
belah	split	4.5	5.55	5	4.24
licin	smooth	6.58	4.91	5	4.21
masak	cook	6.16	4.44	5	4.21

cermin	mirror	5.9	4.55	6	4.2
keliru	confused	3.21	6.03	6	4.19
liar	wild	5.94	5.9	4	4.17
anjur	project	5	4.55	5	4.15
gugur	fall	4.09	4.7	5	4.08
tukar	change	5.09	4.05	5	4.08
jahat	evil	3.23	6.39	5	4.07
pindah	move	4.95	4.24	6	4.05
singkap	reveal	5.86	4.14	7	4.05
lampau	past	5.09	4.71	6	4.03
papar	show	5.91	4.1	5	4.03
bolos	escape	5.5	4.55	5	3.99
masam	sour	3.93	5.1	5	3.97
pohon	tree	6.32	3.42	5	3.97
pecah	break	5.1	4.71	5	3.95
mabuk	drunk	4.06	5.05	5	3.92
taja	sponsor	5.3	4.05	4	3.92
gerbang	gate	5.32	4.5	7	3.9
peluh	sweat	4.38	5.1	5	3.9
raga	basket	5.45	3.63	4	3.9
renung	stare	4.45	4.68	6	3.9
buta	blind	3.05	4.39	4	3.87
kacang	nut	5.44	4.24	6	3.87
angan	dream	6.73	4.53	5	3.85
heboh	noisy	5.02	6.38	5	3.82
terik	scorching	3.76	5	5	3.82
goreng	fry	5.95	5	6	3.78

duri	thorn	3.64	5.14	4	3.75
helah	trick	4.89	5.57	5	3.75
paksa	force	4.17	5.35	5	3.75
ajak	invite	5.68	5.1	4	3.71
labur	invest	5.93	5.12	5	3.71
madah	word	5.77	4	5	3.71
rujuk	refer	5.18	4.09	5	3.71
tabal	crown	6.58	4.28	5	3.71
tikus	rat	3.02	4.95	5	3.71
lalang	weed	4.65	4.78	6	3.67
riuh	noisy	5.02	6.38	4	3.67
jimat	save	6.45	4.95	5	3.62
tendang	kick	4.31	4.9	7	3.62
lejang	loose	4.52	4.3	6	3.57
lempar	throw	5.55	4.52	6	3.57
ribut	storm	4.95	5.71	5	3.57
bongkak	arrogant	3.69	5.65	7	3.51
curiga	suspicious	3.76	6.25	6	3.51
dedah	expose	4.14	5.89	5	3.51
londeh	slip	4.33	4.7	6	3.51
sentap	pull	5.59	4.1	6	3.51
ucap	speak	5.91	4.19	4	3.51
banjir	flood	3.19	6	6	3.45
danau	lake	6.82	3.95	5	3.45
tabik	salute	5.92	5.31	5	3.45
tekak	throat	4.76	4.2	5	3.45
aci	shaft	4.86	4.65	3	3.37

cegah	prevent	5.42	4.18	5	3.37
endah	concern	5.05	4.7	5	3.27
ponteng	skip	4.76	4.25	7	3.27

Welsh Words Priming List (Experiment 2C)

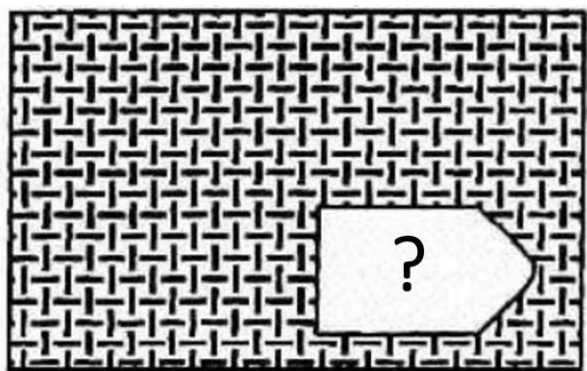
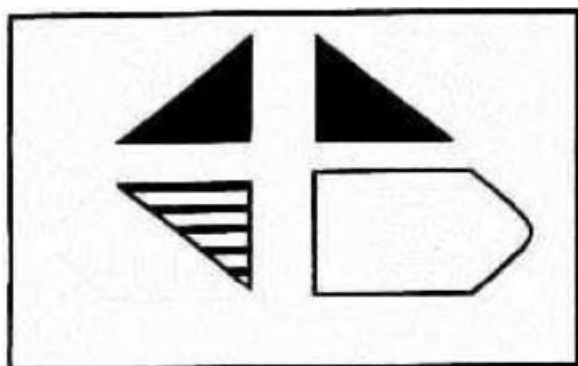
Welsh	Character Count
cymraeg	7
peth	4
pobl	4
rhan	4
gwaith	6
iaith	5
cyfer	5
plentyn	7
ysgol	5
byd	3
dydd	4
rhaid	5
angen	5
amser	5
ffordd	6
enw	3
lle	3
plaid	5
mis	3
mwyn	4
tro	3
llyfr	5
cynnwys	7
cwrs	4
rhaglen	7
diolch	6
aelod	5

gair	4
oed	3
cyfle	5
hanes	5
bodd	4
hyd	3
llun	4
diwrnod	7
rhyw	4
gwlad	5
saesneg	7
cyfnod	6
bywyd	5
nifer	5
cal	3
gwefan	6
ystod	5
wythnos	7
cyd	3
teulu	5
cwmni	5
meddwl	6
pryd	4
arian	5
diwedd	6
dechrau	7
ardal	5
cilydd	6
stori	5
gallu	5

syniad	6
taith	5
pobol	5
merch	5
cyngor	6
blaen	5
tân	3
nos	3
bwyd	4
ffrind	6
llaw	4
dyn	3
mudiad	6
tîm	3
eglwys	6
awr	3
sŵn	3
sylw	4
maes	4
cyfrwng	7
oen	3
eisiau	6
addysg	6
hanner	6
hawl	4
cyfres	6
rhywun	6
gwybod	6
cynnig	6
iechyd	6

cofnod	6
digon	5
duw	3
adroddiad	9
gofal	5
ymchwil	7
cefnidir	7
pwyllgor	8
ymateb	6
popeth	6
dyfodol	7
sgwrs	5
gêm	3
grŵp	4
rhyfel	6
ochr	4
hwyl	4
cymuned	7
tywydd	6
rhifyn	6
clwb	4
cwbl	4
papur	5
swydd	5
profiad	7
drws	4
cefn	4
sul	3
canrif	6
gŵyl	4

dewis	5
rheswm	6
munud	5
gweinidog	9
arfer	5
testun	6
rhestr	6
canol	5
bwrdd	5
cartref	7
gogledd	7
llawr	5
tir	3
cymorth	7
tipyn	5
siop	4
llais	5
ateb	4

NoLanguage Stimuli (Experiment 4)**Wrong Answer****Correct Answer****Wrong Answer****Correct Answer**

English Dialogues (Experiment 4)

Priming #1

Amy: So which country will you be going to next year?

George: Actually, I think I will be completing my degree here, the 3+ 0 program, I don't think my family can afford to send me overseas to study.

Amy: Oh, have you tried to apply for financial aid? Our university will subsidise the cost of your travel and living expenses there. For me, I think spending one year overseas is quite valuable, it will teach you a lot of things, more than what you can learn when you just stay in your comfort zone.

George: I actually have not inquired about the financial aid, maybe after lunch I will drop by the admin office and ask about it. Going overseas sound nice, but if my family cannot afford, and got no financial aid, I think it is best if I completed my degree locally.

<George will be going to the admin office after lunch>

True/False?

(Correct Ans = True)

Priming #2

Jared: What are you going to do this coming holiday?

Ashley: Eh, coming got holiday ah? When is it eh?

Jared: Next week! I'm planning to maybe go to Langkawi or Tioman and go snorkelling, or maybe go Ipoh and Penang for a food trip, not sure yet, you want to come along? I am going with my family since we don't really have much time together once everyone starts working again

Ashley: Oh! Sure sure! I miss going to beaches, especially in Tioman, the water there very clear, can see all the way down when going for snorkelling.

Jared: Exactly! We are planning to leave KL sometime around 7am, go out have early breakfast, then make our way to Tioman, we want to escape the jam since it will be holiday and everyone will also be going vacation

Ashley: Sounds good, I will come over to your place around 6:30 am, then we go together, I will be bringing my boyfriend also, he really needs a break from work

<Jared plans to leave after 7am to Tioman>

True/False?

(Correct Ans = True)

Mandarin Dialogues (Experiment 4)

Priming #1

欣怡：我有一段时间没有运动了，但不是因为我懒，是因为我找不到时间运动。

伟祺：那到时候你变胖了能怪谁？你自己，对吧？

欣怡：是这样没错，但是有时候吃东西能让我们即时满足欲望，体会一种虽然内疚但是满足的快感。

伟祺：其实我也这样认为。我觉得考试或者压力大的时候，变胖也是不能避免的，不过当然考完试后我们就应该多注意自己的饮食和运动习惯，这样才可以把考试期间吃进去的毒素清除干净。

欣怡：对，其实我打算下个星期开始跑步了，但也不是跑马拉松，就是绕着公园慢跑而已。

Xin Yi: I haven't been exercising for a while, not because I'm lazy, but because I can't find time to do it.

Wei Qi: Then when you gain weight later, who's to blame? Yourself, right?

Xin Yi: That's correct. However, sometimes eating food gives us an immediate sense of satisfaction, even though there's some guilt involved.

Wei Qi: I also think that's true. I believe that during exams or times of stress, gaining weight is unavoidable. After all, we should pay attention to our diet and exercise habits after the exams to eliminate the toxins we consumed during that time.

Xin Yi: Right, I plan to start running next week, but it won't be a marathon; it'll just be a slow run around the park.

<欣怡最近不常运动>

对/错?

(Correct Ans = 对)

Xinyi hasn't been exercising much recently

True/False?

(Correct Answer = True)

Priming #2

晓婷：因为我男朋友的妈妈很会烤芝士蛋糕，我吃过觉得很好吃，所以就问她怎样做。她就把步骤写下来给我，我只是跟着她去那件冷藏公司超市买一样的材料而已。

李冰：那是什么芝士蛋糕来的？我也想试做看看。

晓婷：我觉得应该是费城芝士蛋糕，但是我觉得你应该先试试提拉米苏或者意大利芝士蛋糕，这些比较容易做。如果你能去冷藏公司超市买更省事，因为在那里你可以买到那种免考的芝士蛋糕混合粉。

李冰：喔，它的盒子是蓝色和银色的是吗？

晓婷：对对对，很容易认的。它们应该是二十五令吉，一盒是250克。你跟着盒子上的指示准备一个消化饼干做的底，然后把蛋糕糊倒进去，冷藏过夜就可以了。

Xiaoting: My boyfriend's mom is really good at baking cheesecake. I tried it and thought it was delicious, so I asked her how to make it. She wrote down the steps for me, and I just followed them and went to that frozen goods supermarket to buy the exact same ingredients.

Libing: What kind of cheesecake is it? I'd like to try making it too.

Xiaoting: I think it' s Philadelphia cheesecake, but I suggest you try making tiramisu or Italian cheesecake first—they' re easier. If you can go to the frozen goods supermarket, it' ll be even more convenient because you can find that no-bake cheesecake mix there.

Libing: Oh, is the box blue and silver?

Xiaoting: Yes, yes! It' s easy to recognize. It should be 25 ringgit per box, and each box is 250 grams. Just follow the instructions on the box, prepare a crust using digestive biscuits, pour in the cheesecake mixture, and refrigerate it overnight—that' s it!

<晓婷觉得她男朋友的妈妈不会烤芝士蛋糕>

对/错?

(Correct Ans = 错)

Xiaoting thinks her boyfriend's mom can't bake cheesecake.

True/False?

(Correct Answer = False)

Mixed Dialogues (Experiment 4)

Priming #1

Alex: 嗯，那你比较喜欢讲英语还是华语？

Jane: 我两个语言 also can 啦，因为我读华中。

Alex: 喔，原来如此。

Jane: 我觉得我心里是 feel that my both languages 都是好的，可是实际上两个都很烂，真的。

Alex: 为什么你这样觉得呢？

Jane: 其实我很难判断, in my heart I think my Chinese is better but people tell me that my English 比较好。At home I speak in English, 因为我爸爸是 English educated 的, but 在我婆婆家我是讲 Chinese 的, 因为她只会讲 Chinese。跟工人的话也是讲 Chinese, 因为她听不懂英语。

Alex: Hmm, do you prefer speaking English or Chinese?

Jane: Both languages are fine for me because I studied in a Chinese-medium school.

Alex: Oh, I see.

Jane: I feel like both my languages are okay in my heart, but in reality, both are actually quite bad, really.

Alex: Why do you think so?

Jane: It' s actually hard for me to tell. In my heart, I think my Chinese is better, but people say my English is better. At home, I speak English

because my dad is English-educated, but at my grandma' s house, I speak Chinese because she only understands Chinese. When talking to the domestic worker, I also speak Chinese because she doesn' t understand English.

<Jane 只会说英语>

对/错?

(Correct Ans = 错)

Jane can only speak English.

True/False?

(Correct Answer = False)

Priming #2

Sarah: But 香港的local food 会贵吗?

June: I mean 如果你跟朋友split the bill 的话就不会贵。就, 你pay 一点, 他们 pay 一点, 然后一起 share, 这样就 quite affordable。

Sarah: 喔, 这样activities呢, 除了eating 还可以做什么?

June: 去香港都是去吃的, 如果要的话你可以去那边的 Disneyland, but most people 都是去那边吃的, 因为那边真的很多good food!

Sarah: 喔, 那你觉得那边什么好吃?

June: I think 有一个你一定要try 的是老婆饼, 那个很好吃。Actually 我有想过要再去香港玩, 因为last time 我都是去sightseeing 而已。Next time 我去我invite你, then 我们一起去night market找good food!

Sarah: But is local food in Hong Kong expensive?

June: I mean, if you split the bill with friends, then it' s not expensive. Like, you pay a bit, they pay a bit, and you share, so it' s quite affordable.

Sarah: Oh, what about activities? Besides eating, what else can you do?

June: People mainly go to Hong Kong for the food. If you want, you can visit Disneyland, but most people go there just to eat because there' s really a lot of good food!

Sarah: Oh, what do you think is good to eat there?

June: I think one thing you definitely have to try is wife cakes (lou po beng)—they' re really delicious. Actually, I' ve been thinking about going to Hong Kong again because last time, I only did sightseeing. Next time I go, I' ll invite you, then we can go to the night markets together and look for good food!

<June 打算跟Sarah一起去香港>

对/错?

(Correct Ans = 对)

June plans to go to Hong Kong with Sarah.

True/False?

(Correct Answer = True)

Appendix 3: Error Rate Analysis Output for Experiment 2A, 2B, and 2C

Experiment 2A - ANOVA (Err)

Cases	df _{effect} , df _{res}	F	p	$\eta^2 p$
Congruency	1, 74	75.916	<.001	0.506
Language	3, 222	1.151	0.329	0.015
Congruency * Language	2.59, 191.82	0.751	0.505	0.01

Descriptives

Congruency	Language	N	Mean	SD	SE	Coefficient of Variation
Congruent	NoLanguage	75	0.804	1.616	0.187	2.01
	English	75	0.508	1.547	0.179	3.046
	Mandarin	75	0.656	1.17	0.135	1.783
	Mixed	75	1.481	5.109	0.59	3.449
Incongruent	NoLanguage	75	8.395	11.364	1.312	1.354
	English	75	8	8.033	0.928	1.004
	Mandarin	75	7.062	6.92	0.799	0.98
	Mixed	75	8.247	9.286	1.072	1.126

Experiment 2B - ANOVA (Err)

Cases	df _{effect} , df _{res}	F	p	$\eta^2 p$
Congruency	1, 71	66.464	<.001	0.484

Language	2.81, 199.62	0.705	0.542	0.010
Congruency * Language	3, 213	1.419	0.238	0.020

Descriptives

Congruency	Language	N	Mean	SD	SE	Coefficient of Variation
Congruent	NoLanguage	72	1.213	1.836	0.216	1.514
	English	72	0.816	1.685	0.199	2.065
	Malay	72	1.389	4.575	0.539	3.294
	Mixed	72	1.323	2.832	0.334	2.141
Incongruent	NoLanguage	72	9.722	12.563	1.481	1.292
	English	72	11.265	11.653	1.373	1.034
	Malay	72	10.957	12.486	1.471	1.140
	Mixed	72	11.471	11.895	1.402	1.037

Experiment 2C - ANOVA (Err)

Cases	df _{effect} , df _{res}	F	p	$\eta^2 p$
Congruency	1, 74	129.346	< .001	0.636
Language	3, 222	1.304	0.274	0.017
Congruency * Language	3, 222	0.979	0.403	0.013

Descriptives

Congruency	Language	N	Mean	SD	SE	Coefficient of Variation
Congruent	NoLanguage	75	1.122	1.687	0.195	1.504
	English	75	0.910	1.938	0.224	2.130
	Welsh	75	1.545	4.359	0.503	2.821
	Mixed	75	1.206	2.287	0.264	1.896
Incongruent	NoLanguage	75	8.444	6.045	0.698	0.716
	English	75	9.975	9.452	1.091	0.948
	Welsh	75	10.321	10.121	1.169	0.981
	Mixed	72	11.471	11.895	1.402	1.037

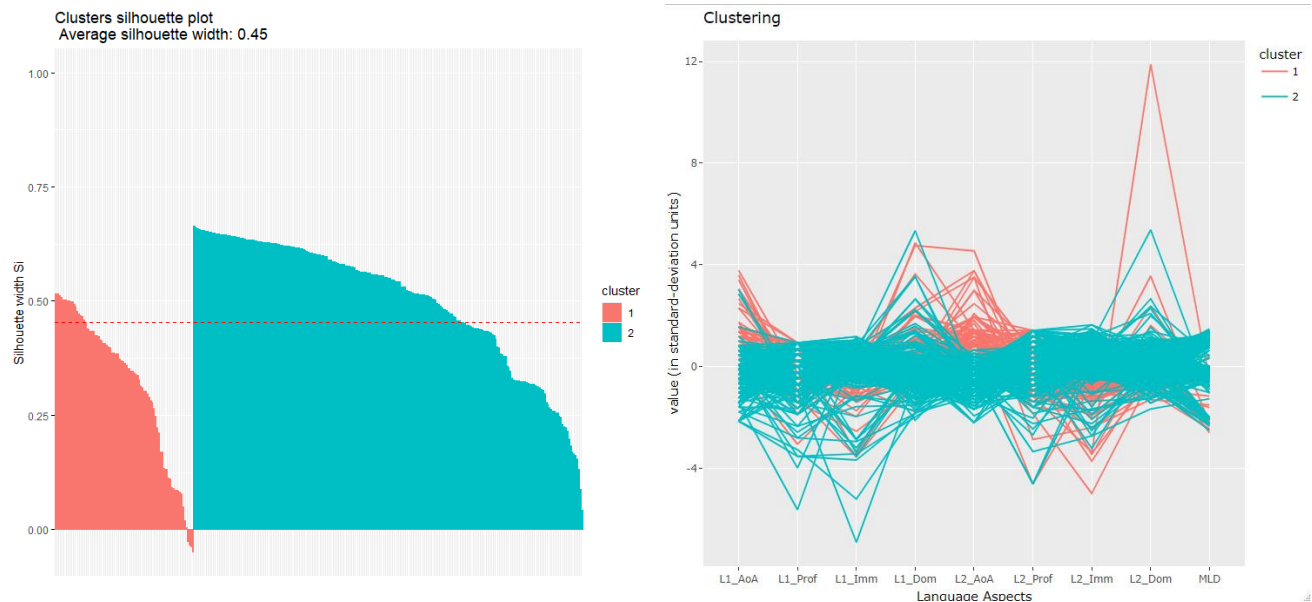
Appendix 4: IES Calculation

The following calculations summarize the findings from Chapter 6, Comishen and Bialystok (2021) and Morrison et al. (2018) studies. Specifically, only results from the bilinguals performing the 2-Back experiments were included to ensure cross-study comparability. Important to note is that the calculations for Comishen and Morrison's studies are an approximation given the data that was reported in the studies. The actual calculations using the raw data may be different.

Study		RT (ms)	Accuracy to Targets (1 – ERR)	IES
Chapter 6	NoLanguage	569.24	0.80	711.55
	English	563.30	0.78	722.18
	Mandarin	552.67	0.80	690.84
	Mixed	562.04	0.81	693.88
Morrison et al. (2018)		471.15	0.75	628.20
Comishen & Bialystok (2021)		653	0.72	906.94
Note: A lower IES value indicates better efficiency				

Appendix 5: K-Means Cluster Analysis

The language background of the first and second language (AoA, Proficiency, Immersion, and Dominance) data of all 324 participants across Experiments 2A, 2B, 2C, 3 and 4 were collated and a K-means clustering analysis was conducted using the *kmeans* function from the *stats* package in R. The appropriate *k* value was determined using the silhouette coefficient. The model with the largest silhouette coefficient was with *k* = 2 (SI = 0.45). The outputs are visualized in the following figures



To further test whether these two clusters are significantly different from each other, a three-way mixed ANOVA was conducted with Language (L1, L2) and Experience (AoA, Proficiency, Immersion, Dominance) as the within-subjects variables, and the cluster (Cluster 1, Cluster 2) as the between-subjects variables. The analysis reveals that there was a significant main effect of language and experience. A significant two-way interaction between Language and Experience was also found, with significant differences between acquisition of L1 and L2. However, this does not imply that the sample is sequential, as the difference between L1 AoA and L2 AoA is, on average, 1.3 years apart. Importantly, there were no main effects, nor any interaction effects, of cluster. While experiences were significantly different between languages, cluster 1 and cluster 2 do not significantly differ from each other on either across languages or experiences.

Within-Subjects ANOVA

Cases	df	F	p
Language	1	136.185	< .001
Language × cluster	1	0.982	0.323
Experience	3	1132.906	< .001
Experience × cluster	3	0.932	0.424
Language × Experience	3	199.187	< .001
Language × Experience × cluster	3	1.043	0.373

Between-Subjects ANOVA

Effects	df	F	p
cluster	1	1.16	0.282

Post Hoc Comparisons - Language * Experience

Comparison	Mean Difference	SE	t	Cohen's d	p
L1, AoA - L2, AoA	-1.365	0.050	-27.053	-1.612	< .001***
L1, Prof - L2, Prof	0.096	0.050	1.910	0.114	1.000
L1, Imm - L2, Imm	0.058	0.050	1.143	0.068	1.000
L1, Dom - L2, Dom	0.104	0.050	2.055	0.122	1.000