Cronfa
Swansea University's Research Repository

Swansea University
2024

Does Welsh-English bilingualism affect cognition across the lifespan?

Submitted to Swansea University in fulfilment of

the requirements for the Degree of Doctor of Philosophy

Hanna Elan Thomas MSc, BSc (Hons)

Copyright: The author, Hanna E. Thomas, 2024. Thesis is released under the terms of a Creative Commons Attribution-No-Derivatives (CC-BY-ND) license. Third party content is excluded for use under the license terms.

Abstract

It is currently unclear whether Welsh-English bilingualism affects cognition across the lifespan, particularly cognitive reserve. Therefore, this thesis considered the effect of this unique type of bilingualism on cognition in young and older adults, and across general and specific cognitive processes, in an attempt to locate a possible bilingual advantage that may be a more suitable proxy of CR. Study One found that Welsh-English bilingual older adults performed comparably overall across Welsh and English versions of two major cognitive assessment batteries, and there was also little evidence that Welsh-English bilingualism had either a detrimental or advantageous impact on performance compared to normative data. The second study then moved away from using general cognitive processes to locate a possible bilingual advantage, by considering the impact of bilingualism on semantic cognition. Age and language status were not found to have distinct effects on semantic control, but Welsh-English bilinguals exhibited faster reaction times and increased accuracy when semantic control demands were high. The final study then investigated the impact of age and bilingualism on response consistency in uncertain situations involving general knowledge and semantic categories. Overall, participants were more consistent in their responses to general knowledge than category statements, but neither age nor bilingualism status impacted consistency of responding. In conclusion, the findings of this thesis suggest that Welsh-English bilingualism may impact on, and lead to an advantage in particular aspects of cognition in young and older adults, like semantic control processes, as opposed to an uplift of general cognitive domains. Consequently, a more effective proxy of cognitive reserve might be processes of semantic control. Overall, and even though further research is needed, this thesis highlights the importance of exploring the complex relationship between bilingualism and cognitive functioning to facilitate understanding of how bilingualism may contribute to cognitive reserve.

Declarations

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed [redacted]

Date .....20.05.24.....

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed [redacted]

Date .....20.05.24.....

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access approved by the Swansea University.**

Signed [redacted]

Date .....20.05.24.....

The University's ethical procedures have been followed and, where appropriate, that ethical approval has been granted.

Signed

Date .....20.05.24.....

Contents

Abstract I
Declarations II
Contents IV
Acknowledgements VIII
List of Tables X
List of Figures XI
List of Abbreviations XIII
Dissemination of Research XIV
Chapter 1 – The Impact of Bilingualism on Cognition Across the Lifespan 1
1.0 Thesis Overview 1
1.1 Chapter Overview 1
1.2 Defining Bilingualism 1
1.3 Classifications of Bilingualism 3
1.4 Welsh-English Bilingualism 7
1.5 Bilingualism and Cognitive Function 12
1.6 Bilingual Language Activation 15
1.7 Bilingualism, Cognitive Function and Cognitive Reserve 17
1.8 Bilingualism, Ageing and CR 25
1.9 Bilingualism and CR - Possible Moderating Factors 35
1.10 Conclusion 39
1.11 Thesis Aims 40
Chapter 2 – Study 1 – Collecting Normative Data for Clinical Neuropsychological Assessments for the Purposes of Testing Welsh Speaking (Bilingual) Individuals. 41
2.0 Chapter Overview 41
2.1 Background 41
2.1.1 The Challenge 43
2.1.2 Study Aims 53
2.2 Methods 54
2.2.1 Ethical Approval Procedure 54
2.2.2 Participants 54
2.2.3 Data Exclusions 55
2.2.4 Materials 58
2.2.4.1 Demographic and Background Details 59

2.2.4.2 Language Proficiency Questionnaire 59
2.2.4.3 Lexical Test for Advanced Learners of English (LexTALE) 60
2.2.4.4 Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) – Form A 61
2.2.4.5 Addenbrooke's Cognitive Examination-III (ACE-III) 64
2.2.4.6 Additional Verbal Fluency Task 67
2.2.5 Design 68
2.2.6 Procedure 70
2.2.6.1 Pre-COVID-19 Procedure 74
2.2.6.2 Post-COVID-19 Procedure 75
2.3 Results 76
2.3.1 Global Level Analysis – is there a Difference Between Overall Scores? 76
2.3.1.1 RBANS Findings 76
2.3.1.2 ACE-III Findings 79
2.3.2 Interim Summary: The Removal of One Counterbalancing Order 84
2.3.3 Individual Level Analysis: Does Language Format Predict Participant Scores? 87
2.3.4 Performance on the RBANS and ACE-III by Index/Domain 89
2.3.5 Comparison of Welsh-English Bilingual Performance with Published Norms 91
2.4 Discussion 94
2.4.1 Limitations, Strengths, and Unanswered Questions 102
2.4.2 Conclusion 104
Chapter 3 – Study 2 – Semantic Cognition in Ageing and Bilingualism 106
3.0 Chapter Overview 106
3.1 Introduction 106
3.1.1 Models of Semantic Organisation 111
3.1.2 Semantic Processing in Bilinguals 115
3.1.3 Semantic Cognition (SC) 118
3.1.4 Semantic Cognition – The Mechanisms of Control that Operate on Semantic Memory 119
3.1.5 Hoffman's (2018) Aims 123
3.1.6 The Semantic Control Paradigm – Probing Different Semantic Control Mechanisms 124
3.1.7 Findings from Hoffman (2018) 126
3.1.8 Potential Effects of Bilingualism on Hoffman's Paradigm and Findings 129
3.1.9 Study Aims 132
3.2 Methods 134
3.2.1 Ethical Approval Procedure 134
3.2.2 Participants 134
3.2.3 Data Exclusions 135

3.2.4 Materials 136
3.2.4.1 General Cognitive Tasks 136
3.2.4.1.1 Verbal Fluency Tasks 137
3.2.4.1.2 Eriksen Flanker Task – Executive Function (Eriksen, 1995) 137
3.2.4.1.3 Adapted version of the Wisconsin Card-Sorting Task (WCST) – Complex Executive Function 138
3.2.4.2 Breadth of Semantic Knowledge 139
3.2.4.2.1 Spot-the-Word Task (Lexical Decision) 140
3.2.4.2.2 Mill Hill Vocabulary Test - modified version from Hoffman (2018). (Synonym Judgement) 141
3.2.4.3 Test of Semantic Control 142
3.2.5 Design 144
3.2.6 Procedure 145
3.3 Results 145
3.3.1 General Cognitive Assessments (Effects of age/bilingualism) 145
3.3.2 Differences in Performance on General Cognitive Tasks and Breadth of Semantic Knowledge 146
3.3.3 Semantic Control Task: Effects of Age and Language Status on Semantic Control via Analysis of Accuracy Data 148
3.3.4 Reaction Time Data Analysis 152
3.3.5 Relationships Between the Semantic Control Task and General Cognitive Tasks 155
3.3.6 Principal Component Analysis (PCA) – Older Participants 157
3.3.7 PCA – Younger Participants 158
3.4 Discussion 159
3.4.1 General Cognitive Tasks 160
3.4.2 Semantic Control Task Findings and Interpretations – Accuracy 162
3.4.3 Semantic Control Task Findings and Interpretations – Reaction Time 166
3.4.4 Relationship Between the Semantic Control Task and General Cognitive Tasks 168
3.4.5 The Effects of Bilingualism on SC in the Wider Context of CR 171
3.4.6 Strengths and Limitations 172
3.4.7 Conclusion 173
Chapter 4 – Study 3 – Bilingualism & The Rumsfeld Effect 176
4.0 Chapter Overview 176
4.1 Background 177
4.1.1 Thresholds – How are Decisions Made? 180
4.1.2 Domains of Knowledge 184
4.1.3 Findings from Hampton et al (2012) – The 'Rumsfeld Effect' 186
4.1.4 Present Study 191

4.2 Methods 194
4.2.1 Participants 194
4.2.2 Data Exclusions 195
4.2.3 Materials 196
4.2.3.1 LexTALE 196
4.2.3.2 Knowledge Judgement Task 196
4.2.4 Design 198
4.2.5 Procedure 200
4.3 Results 201
4.3.1 Descriptive Statistics 201
4.3.2 Inferential Statistical Analyses 205
4.3.2.1 Within Participants IV Analyses – Statement Type, Threshold of Response Type, and Re-Test Interval 205
4.3.2.2 Between Participants IV Analyses – Age of Participant and Language Status 207
4.3.3 Overall Summary 208
4.4 Discussion 208
4.4.1 Strengths and Limitations 214
4.4.2 Conclusion 215
Chapter 5 – General Discussion 217
5.0 Chapter Overview 217
5.1 Summary of Thesis Aims 217
5.2 Summary of Each Study: Their Contribution to Knowledge and Practical and Theoretical Implications 217
5.3 Strengths, Limitations, and Future Directions 230
5.4 Key Standards for Researching the Bilingual Advantage 235
5.5 Thesis Conclusion 237
References 238
Appendices 294

Acknowledgements

I would like to take this opportunity to express my deepest appreciation to my supervisors Professor Jeremy Tree, Dr David Playfoot and Dr Claire Williams for their expertise, dedication, and for guiding and supporting me throughout this doctoral journey. They were always available to answer any questions and gave **very** generously of their time and knowledge. Their enthusiasm and encouragement were extremely helpful, and I am very grateful to the three of them.

I further acknowledge Professor Jeremy Tree for his ability to use accessible analogies to break down complex ideas, and Dr David Playfoot for always helping me to see 'the wood for the trees', and for his superb statistical knowledge. Additionally, I further acknowledge Dr Claire Williams for her meticulous scrutiny and critical insight, invaluable advice, thoughtfulness, for being a constant source of strength, and for giving the confidence boost and mentorship that I needed on many occasions over the course of this journey. You have been instrumental.

I am also grateful to Maria Davis and Katie Retallick from the Faculty's Postgraduate Research Office who directly supported me during challenging periods, and when extensions were required because of the COVID-19 pandemic.

I would also like to thank Swansea University and the Coleg Cymraeg Cenedlaethol for funding my research. This work would not have been possible without them – diolch yn fawr iawn!

A debt of gratitude is also owed to all participants for their time and efforts in taking part in each study, and also to those who advertised the study to support recruitment. In particular, a special thank you is owed to Rhian who went on numerous 'recruitment drives' for the first study of this thesis. Diolch yn fawr iawn Rhian!

I also thank my parents, Gwil and Yvonne, and my sister, Nerys, for their patience, constant encouragement, and daily FaceTime calls. I would also like to thank my godmother, Nia, for her endless support and for sending me motivational cards in the post to keep me going!

Last but not least, I thank my partner Aaron, for being so understanding, for keeping on top of household chores, for being the voice of reason on bad days, and for his unwavering faith in my ability when I doubted myself over the last four and a half years. He was always there to listen, as a shoulder to cry on, and to always remind me *"one step at a time."* I **absolutely** could not have done this without him.

Diolch yn fawr iawn i bawb!

Thank you very much everyone!

List of Tables

Table 1 *Different Types and Classifications of Bilinguals*…… 6
Table 2 *Summary of Demographic Characteristics*……... 57
Table 3 *Performance Comparison on the Welsh LexTALE across Current and Normative Data*…….. 58
Table 4 *Examples of Welsh and English Words and Non-Words Used in the LexTALE Task*……. 61
Table 5 *Description of the RBANS Indexes*……... 62
Table 6 *A Description of each Subtest Within the ACE-III Domains*…… 65
Table 7 *RBANS Descriptive Statistics for each Order*……... 77
Table 8 *RBANS Descriptive Statistics Regardless of Order* ……. 77
Table 9 *ACE-III Descriptive Statistics Regardless of Order*……..... 80
Table 10 *ACE-III Descriptive Statistics for each Order*…… 80
Table 11 *Number of Participants in each Order and Mode*…….. 83
Table 12 *Welsh and English LexTALE Performance of the Final Sample and a Comparison on the Welsh LexTALE across Current and Normative Data*…… 85
Table 13 *Summary of Demographic Characteristics*..... 86
Table 14 *Descriptive Statistics of Final Sample (n = 70)* 87
Table 15 *ACE-III Domains Means per Language of Assessment*..... 90
Table 16 *RBANS Index Scores Means per Language of Assessment*……. 91
Table 17 *Descriptive Statistics for Comparison with Published Normative Data* 92
Table 18 *Test-retest Correlation*…… 93
Table 19 *Summary of Demographic Characteristics*……. 135
Table 20 *Performance Comparison on the Welsh LexTALE across Current and Normative Data*…….. 136
Table 21 *Descriptive Statistics for the 'Domain-general' and* Breadth of *Knowledge Tasks according to Age (young versus old) and Language Status (bilingual versus monolingual)*……... 147
Table 22 *Principal Component Analysis - Older Participants*…….. 157
Table 23 *Principal Component Analysis - Younger Participants*…….. 159
Table 24 *Summary of Demographic Characteristics of each Group*…… 196

List of Figures

Figure 1 *Biomarkers of AD and the Cognitive Response Curves* 26
Figure 2 *First Trial of the RBANS Line Orientation*……. 72
Figure 3 *Example of the Dot Counting Task in the ACE-III Visuospatial Domain*..... 72
Figure 4 *Bar graph of RBANS Mean Scores in each Counterbalancing Order*…….. 78
Figure 5 *Bar graph of ACE-III Mean scores in each Counterbalancing Order*…….. 82
Figure 6 *Diagram of the Semantic Network Model Related to Animals originally proposed by Collins & Quillian, 1969*……... 112
Figure 7 *Diagram of a Spreading Activation Model of Semantic Information*…. 113
Figure 8 *Processing Assumptions of the Feature-Comparison Model taken from Smith & Estes (1978)*…..... 115
Figure 9 *Key Elements of the Semantic Control Task (Hoffman, 2018)*…..... 125
Figure 10 *Congruent and Incongruent Trials of the Eriksen Flanker Task*…….. 138
Figure 11 *WCST Stimuli and Feedback*……. 139
Figure 12 *Spot-the-Word Stimuli Example*…… 140
Figure 13 *Mill Hill Stimuli Example*…….. 141
Figure 14 *Examples of Global Semantic Association Trials with Four and Two Response Options*……... 142
Figure 15 *Examples of Colour Trials with Four and Two Response Options*…... 144
Figure 16 *Examples of Size Trials with Four and Two Response Options*…… 144
Figure 17 *Average Accuracy across the Four Experimental Conditions and Age* 150
Figure 18 *Average Accuracy for Younger and Older Monolinguals*……. 151
Figure 19 *Average Accuracy for Younger and Older Bilinguals*…….. 152
Figure 20 *Average RTs across the Four Experimental Conditions and Age*……. 155
Figure 21 *Thresholds Proposed by Hampton et al. (2012)*……... 181
Figure 22 *Examples of Questions/Statements and Response Options from the Knowledge Judgement Task*……... 197

Figure 23 *Analysis of Consistency across Instances when Initial 100% True Responses Remain 100% True and when Initial Responses Outside of the 100% True Box Remain Outside of this Box* …… 199
Figure 24 *Analysis of Consistency when Initial 100% False Responses Remain 100% False and when Initial Responses Outside of the 100% False Box Remain Outside of this Box* ……. 199
Figure 25 *Analysis of Consistency when 'All True' and 'All False' Responses are Collapsed* ……..... 199
Figure 26 *Mean Consistency of Response Selection across Question/Statement Type and Response Thresholds* …….. 202
Figure 27 *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Monolinguals only* ……. 203
Figure 28 *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Bilinguals only*…… 203
Figure 29 *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Young Participants only* …… 203
Figure 30 *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Old Participants only* …… 203
Figure 31 *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – One Hour Re-test Interval*…….. 204
Figure 32 *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – One Week Re-test Interval*…….. 204
Figure 33 *Mean Consistency of Response Selection across Re-test Intervals*….. 207

List of Abbreviations

ACE-III	Addenbrooke's Cognitive Examination – Third Edition
AD	Alzheimer's Disease
ANOVA	Analysis of Variance
AoA	Age of Acquisition
BAASN	Batri Amlroddadwy ar gyfer Asesiad o Statws Niwroseicolegol (Welsh version of the RBANS)
CALD-EE	Culturally and Linguistically Diverse Background Educated in English
CALD-NE	Culturally and Linguistically Diverse Background with Non-English and Mostly Non-Western Education
CR	Cognitive Reserve
DV	Dependent Variable
EFs	Executive Functions
ESB	English speaking background
fMRI	Functional magnetic resonance imaging
GK	General Knowledge
IV	Independent Variable
LexTALE	Lexical Test for Advanced Learners of English
MCI	Mild Cognitive Impairment
PCA	Principal Component Analysis
PFC	Prefrontal Cortex
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
SC	Semantic Cognition
SES	Socioeconomic Status
WAIS-III	Wechsler Adult Intelligence Scale - Third Edition
WCST	Wisconsin Card Sorting Task

Dissemination of Research

Presentations

Thomas, H., Tree, J., Playfoot, D., Williams, C. (October 2021). Validating clinical neuropsychological assessments for Welsh speakers. *BPS Division of Neuropsychology*, Virtual Conference (oral presentation).

Thomas, H. (October 2021). Dementia services in Welsh – A clinical need. *Cardiff & Vale Regional Partnership Board* (oral presentation).

Thomas, H. (December 2021). Dementia in Wales (oral presentation to MSc Dementia Studies students at Bangor University).

Thomas, H., Tree, J., Playfoot, D., Williams, C. (January 2022). Dilysu asesiadau clinigol niwroseicolegol ar gyfer siaradwyr Cymraeg. *Centre for Ageing & Dementia Research: Dementia and the Welsh language* (webinar – oral presentation).

Thomas, H. (December 2019). *Llandysul Lunch Club for Older Adult Males* (oral presentation).

Media Interviews

Thomas, H. (February 2019). Prynhawn Da S4C (TV).

Thomas, H. (August 2021). BBC Radio Cymru (Radio).

Thomas, H. (January 2022). BBC Radio Cymru & Newyddion S4C (TV).

Thomas, H. (October 2022). BBC Radio Cymru (Radio).

Published Outputs

Thomas, H. (2021). Clefyd Alzheimer: Yr angen am asesiadau Cymraeg. *Gwerddon Fach.* https://golwg.360.cymru/gwerddon/2064860-clefyd-alzheimer-angen-asesiadau-cymraeg

Chapter 1 – The Impact of Bilingualism on Cognition Across the Lifespan

1.0 Thesis Overview

The impact of bilingualism, particularly Welsh-English bilingualism, on cognition across the lifespan remains uncertain. Little is also known about how bilingualism influences ageing and cognitive reserve (CR). Therefore, this thesis will consider the effect of Welsh-English bilingualism on cognition in young and older adults, and across general and specific cognitive processes, in an effort to identify a possible bilingual advantage that may be a more suitable proxy of CR. Understanding what characteristics of bilingualism, or specific types of bilingualism, may contribute to a bilingual advantage is crucial for advancing knowledge in this field.

1.1 Chapter Overview

This chapter will begin by providing an overview of bilingualism, including different classification systems and a discussion of the distinctive nature of Welsh-English bilingualism. It will then explore the impact of bilingualism on general cognitive functions (e.g., executive functions) across the lifespan, and the evidence for a bilingual advantage in the context of cognitive reserve (CR). Furthermore, it will consider bilingualism and CR in the context of ageing and dementia, exploring how various factors associated with knowing two languages may influence the relationship between bilingualism and CR. In summary, the aim of this chapter is to determine what is already known about the impact of bilingualism on cognition across the lifespan and to evaluate the evidence for a bilingual advantage in ageing.

1.2 Defining Bilingualism

To capture the complexity that arises from its various types and characteristics, this section explores multiple definitions of bilingualism, including those that are not as central to this

thesis. Bilingualism is the norm in most countries with its prevalence extending across all age groups and various segments of society (Grosjean, 2013). Indeed, it is estimated that half of the world's population is bilingual (Grosjean, 2008). However, whilst bilingualism is a common phenomenon, no precise definition has gained acceptance (Bhatia, 2017; Butler, 2012; Wald, 1974). For example, Bloomfield (1933) defined bilingualism as *native-like control over two languages*. However, this definition may be too narrow and unrealistic because it limits the number of people who can be considered bilingual (Butler & Hakuta, 2006). By emphasising the need for native-like control over both languages, the definition excludes a significant portion of individuals who use more than one language but are not fluent in each language. As Grosjean (1994) points out, if native-like control is required to meet the criteria for bilingualism, no label would exist for the majority of people. In contrast, others have defined bilinguals as individuals who are fluent in one language but who can produce *complete and meaningful utterances* in the other language (Haugen, 1953). However, whilst such a definition can include people who possess various degrees of language proficiency in a second language, it gives little weight to those who are fluent in both languages.

Additionally, such definitions also fail to specify whether bilingualism refers to knowing and/or speaking two languages to some extent, and in turn, what 'knowing' and 'speaking' mean (Dewaele, 2015; Thiery, 1978). It is also unclear whether the ability to read, write and speak an additional language to some extent is necessary to be classified as being bilingual. Thus, it has been argued that 'passive knowledge' of a second language must be taken into account when defining bilingualism (Jacobson & Schwartz, 2002; Wald, 1974). After all, it is possible for someone to understand a second language without being able to produce complete and meaningful speech in that language.

Consequent to such issues, many researchers have argued that bilingualism should be defined more broadly and encompass things such as the ability to produce meaningful utterances in more than one language, the control of at least one language skill (reading, writing, speaking and listening) in another language, as well as the ability to switch between languages (Grosjean, 1994; Harris & Nelson, 1992; Haugen, 1953). As argued by Grosjean (1994), there may be large differences between individuals from each group, but they all live life with more than one language. Therefore, it seems logical to consider bilinguals as individuals who have attained a certain level of proficiency that is equivalent, yet distinct in its nature, to that of monolinguals (Grosjean, 1994).

In line with this, bilingualism can be viewed as occurring on a continuum (Beardsmore, 1986), with the monolingual speaker at one end, and individuals who have learned two languages in childhood and can speak both languages with equal and native-like fluency, at the other. However, even though the latter ambi-bilingualism is considered the ideal form of bilingualism, it is ultimately rare in reality (Lyons, 1981). Thus, a continuum approach allows for individuals who have different degrees of bilingualism, such as those who speak two languages fluently but use one more than the other, or individuals who learned a second language in adulthood but with varying degrees of fluency, to be captured (Moradi, 2014).

1.3 Classifications of Bilingualism

In addition to the definitions discussed above, it is also possible to classify bilinguals in relation to the level of fluency and proficiency in their spoken languages and by age, context, and method of acquisition (Moradi, 2014). In relation to age of acquisition (AoA), early bilingualism is defined as the acquisition of both languages in the pre-adolescent phase of life (Beardsmore, 1986). In contrast, late bilingualism is defined as the acquisition of one language before, and a second language after, eight years of age (Beardsmore, 1986). Early

bilinguals are thought to have native-like abilities in both languages, whereas most late bilinguals are considered non-native speakers of the second language (Beardsmore, 1986). Further, early bilingualism can also be classified into two further sub-types: (a) simultaneous early bilingualism, and (b) successive early bilingualism. The first refers to children who learn two languages at the same time from birth, and typically leads to strong bilingualism (Moradi, 2014), whereas the second occurs when a child has already acquired one language to some extent but learns a second language in early childhood. The latter situation can also lead to strong bilingualism, but the individual typically needs more time to acquire the second language (Moradi, 2014).

With regards to the level of fluency and proficiency in each language, it is possible to distinguish between balanced and dominant bilinguals (Peal & Lambert, 1962). Individuals who acquire a similar degree of proficiency in both languages are referred to as balanced bilinguals, whereas dominant bilinguals have a higher proficiency in one language (Moradi, 2014). Therefore, it is important to note that balanced bilinguals may be different in nature due to practical (e.g., their day-to-day experience), historical and cultural reasons, but most research focused on the impact of language on cognition tends to focus on monolinguals and other types of bilinguals.

In terms of the organisation of linguistic codes in bilinguals, Weinreich (1953) argued that it may be possible to distinguish between compound, coordinate and subordinate bilinguals. According to Ervin & Osgood (1954), compound bilingualism represents the product of foreign language learning in school, but also a child who grows up in a home where both languages are used concurrently or interchangeably. In contrast, coordinate bilingualism refers to individuals who have learned two languages in two different situations and contexts (e.g., second language learning by moving to a different country) (Ervin & Osgood, 1954),

and subordinate bilingualism refers to when an individual is proficient in one of the two languages (Pitta et al., 1978).

Further, such distinctions between different bilingual individuals also relate to how a bilingual organises and stores linguistic codes in both languages. Compound bilinguals are thought to develop two linguistic codes (e.g., word for 'dog' in both languages) that are stored in one meaning component. That is, there may be one unit that stores the meaning of words for both languages (Moradi, 2014). In contrast, coordinate bilinguals may have two separate systems for the meaning of words in each language - one system for the meaning of words in one language, and another system for the meaning of words in the other language (Moradi, 2014). Furthermore, in subordinate bilinguals, the linguistic codes of their second language are thought to be understood through the linguistic codes of their first language. Put simply, they are considered to possess two sets of linguistic codes (i.e., one for each language) but only one meaning component, which is only accessible through their first language (Moradi, 2014). Even so, and despite many suggestions in the literature regarding compound, coordinate and subordinate bilinguals, no experimentation has led to conclusive evidence to suggest that bilinguals should be classified in such a way (Grosjean, 1994).

In addition to the possible classifications already mentioned, there are many other potential bilingual classifications which deal with social factors, cultural identity and language usage (Moradi, 2014). For example, Fishman (1977) argued that it is possible to differentiate between folk and elite bilinguals. Folk bilinguals often speak a minority language and live in a community in which the language does not have a high status compared to the community's dominant language (Fishman, 1977). On the contrary, elite bilinguals are those who can speak the dominant language of a society but may also benefit from having the ability to speak another language in the community (Fishman, 1977). See Table 1 for a summary of the different types and classifications of bilingualism.

Table 1

Different Types and Classifications of Bilingualism

Term	Definition
Early Bilingualism	The acquisition of both languages in the pre-adolescent phase of life (Beardsmore, 1986).
Late Bilingualism	The acquisition of one language before and the second language after the age of eight years (Beardsmore, 1986).
Simultaneous Early Bilingualism	Children who learn two languages at the same time, from birth (Moradi, 2014).
Successive Early Bilingualism	Occurs when a child has already acquired one language to some extent and learns a second language in early childhood (Moradi, 2014).
Balanced Bilinguals	Individuals who acquire a similar degree of proficiency in both languages (Moradi, 2014).
Dominant Bilinguals	Individuals who have a higher proficiency in one language (Moradi, 2014).
Compound Bilingualism	The product of foreign language learning in school but also a child who grows up in a home where both languages are used concurrently or interchangeably (Ervin & Osgood, 1954).
Coordinate Bilingualism	Refers to individuals who have learned two languages in two different situations (e.g., second language learning by moving to a different country) (Ervin & Osgood, 1954)
Subordinate Bilingualism	Occurs when an individual is proficient in one of the two languages (Pitta et al., 1978).
Folk Bilinguals	Often speak a minority language and live in a community in which the language does not have a high status compared to the community's dominant language (Fishman, 1977)
Elite Bilinguals	Those who can speak the dominant language of a society but also benefit from having the ability to speak another language in the community (Fishman, 1977).

To summarise, bilingualism is a complex skill which has multi-dimensional aspects. Arguably, bilinguals are best perceived as an integrated whole which has led researchers to adopt a wholistic view of bilingualism and accept that bilinguals are unique communicators (Grosjean, 2008). Even so, definitions of bilingualism may be inadequate when applied to specific populations of bilinguals, the nature of language use in different communities, as well as how certain bilinguals operate within their social and cultural context. Consequently, definitions of bilingualism need to be considered along with a myriad of other factors that contribute to different types of bilingualism (e.g., cultural identity, situational use of language, language switching) (Marian & Hayakawa, 2021; Ramírez-Esparza & García-Sierra, 2014).

1.4 Welsh-English Bilingualism

This thesis is interested in a very specific and largely ignored population of bilinguals – lifelong Welsh speakers who are fully immersed in a bilingual society. Welsh-English bilinguals often show a balanced dominance to both languages due to the bilingual environment in which they live (Antón et al., 2019). Consequently, there may be important differences in how such a unique population of bilinguals behave with respect to both ageing, and the organisation of cognitive functions. This will be discussed in more detail later in the thesis. Hereafter, the thesis will be referring to balanced bilingualism as individuals who attain a comparable degree of proficiency in both languages. That is, individuals who consider themselves as native Welsh speakers but demonstrate balanced dominance and proficiency in both Welsh and English.

Welsh (Cymraeg) is a Brythonic Celtic language spoken mainly in Wales (Cymru) (Morris, 2014). The language has the oldest roots compared to other languages spoken in

Britain, such as English and Gaelic. Its origin dates to approximately 4,000 years ago, compared to 1,500 in the case of both English and Gaelic (Davies, 2014).

However, and despite its lineage, there have been many attempts to suppress the language and diminish its cultural significance (Griffiths, 1994). For example, the Act(s) of 'Union' of 1536-1543 marked the desire of the English state to eliminate the Welsh language. English became the sole official language in Wales and no official positions were given to Welsh speakers (Penhallurick, 1993). The Elementary Education Act of 1870, along with subsequent related legislation in the 1880s, also resulted in similar suppression of the Welsh language. That is, elementary education became compulsory and free in England and Wales, but children would only be taught through the medium of English (Penhallurick, 1993). If a child was caught speaking Welsh at school, a wooden board with the words 'Welsh Not' (also known as 'Welsh Knot' or 'Welsh Note') was hung around their necks and was only removed if the child caught somebody else speaking Welsh (Khleif, 1979). At the end of the school day, the child wearing the board would receive a hit to the hand with a cane (Penhallurick, 1993). However, the creation of Cymdeithas yr Iaith Gymraeg (The Welsh Language Society) in 1962 marked the beginning of a counter campaign which compelled the English state to accept and subsidize Welsh identity (Jones & Merriman, 2009; Penhallurick, 1993). This led to the creation of a Welsh television channel (S4C) and the availability of education through the medium of Welsh (Penhallurick, 1993). Subsequently in 2011, The Welsh Language Measure led to the creation of the Welsh language standards. This legislation declares that the Welsh language has official status and ensures that all public services treat both languages equally. This also led to the creation of the post of Welsh Language Commissioner.

Considering the above, the fact that the Welsh language continues to survive as a mother tongue and a medium of communication in everyday life in Wales, is therefore an extraordinary cultural achievement (Aitchison & Carter, 2003). Indeed, Wales was one of the

first areas of the United Kingdom to be merged into the English state that was developing in 1284 and consequently, many assumed that Welsh would naturally cease to exist, and that the language spoken would be English (Aitchison & Carter, 2003). Nevertheless, the Welsh language persisted, defying English expectations of its survival (Aitchison & Carter, 2003).

According to the last census in 2021, Welsh is currently spoken by approximately 17.8% of the population of Wales (538,300 speakers) (Welsh Government, 2022). Consequently, Welsh is still a minority language in Wales and all Welsh speakers are bilingual with English (Morris, 2014). The language is more widely spoken in counties such as Anglesey, Gwynedd, Ceredigion and Carmarthenshire (Jones, 2012). In such areas, the Welsh language is the heart of Welsh identity, but for individuals in other areas of Wales, the language has very little influence on their lives (Davies, 2014). Although children and adults can learn Welsh as a second language either at school or at Welsh courses for adults, it is possible to argue that the language will only have limited influence over their daily lives due to less cultural understanding or rarely engaging with the Welsh language in other contexts.

With regards to the notion that all Welsh speakers are bilingual with English, it is important to note the significant influence that English has had on the Welsh language. There is a disparity between the written standard of the Welsh language and the language or dialect spoken by Welsh people in day-to-day life (Thomas, 1992). Indeed, the vernacular language has been subject to a process of significant relexification from English and is now a fact of life for the Welsh language and people (Thomas, 1992). Due to this influence from the dominant local language, it is therefore possible that Welsh-English bilinguals may differ in terms of language use from other kinds of bilinguals who are not influenced in such a way.

The coexistence of Welsh and English (for those who associate with both languages) in Wales has profound implications for the cultural and linguistic identities of its inhabitants. Research indicates that individuals who feel less confident in their Welsh language skills also

tend to hold less favourable attitudes towards the Welsh language (Morris, 2014). Conversely, those who receive their education primarily through the medium of Welsh and use Welsh at home tend to rate their proficiency in speaking, reading, and writing more highly (Morris, 2014). Interestingly though, this high self-rating of proficiency appears to be unrelated to the level of Welsh spoken in their area (Morris, 2014), suggesting that immersion in Welsh culture, such as attending the Eisteddfod at school, watching Welsh television, and listening to Welsh radio at home, may play a significant role in shaping perceptions of language proficiency. This is further supported by Murphy (2017), who explored children's perceptions of Welsh identity. Despite children aged 9-10 not prioritising language when discussing their own identity, they placed particular importance on the Welsh Language when asked to indicate their perceptions of a Welsh individual. Such findings may therefore reflect the influence of cultural immersion and geographic factors on identity formation. Specifically, it is possible that the young age of participants contributed to their emphasis on language in defining Welsh identity, while the geographic context of south-east Wales, where the Welsh language is less prevalent compared to other areas of Wales, might have also influenced their perceptions.

Another language in a similarly uncommon position as Welsh is Basque. Basque is a language spoken in the Basque Country, a name that refers to the seven territories in which the Basque language is spoken (Hualde et al., 1996). The Basque Country spans across a small area on both sides of the Pyrenees (mountains in southwest Europe that form a border between Spain and France), and on the banks of the Bay of Biscay. The seven provinces are Biscay, Gipuzkoa, Alava, Navarre, Labourd, Low Navarre and Soule (Zuazo, 1995). The Basque Country has a population of approximately 2,900,000, but the number of people who can clearly express themselves in Basque is only around 660,000 (Hualde et al., 1996). Thus, the dominant languages of the Basque Country are Spanish and French (Lasagabaster, 2007).

Similar to Welsh, the Basque language was also influenced by historical events (e.g., mass immigration) and in turn, dilution of the Basque language (Hualde et al., 1996). However, it was also heavily influenced by repressive government policies (Barros et al., 2009). Indeed, legal restrictions were in place so that the language was not used in education, official discourse, or in publishing, and it also became illegal to register new-born babies under Basque names (Cenoz & Gorter, 2017). It is possible therefore to draw similarities between the minority languages of Welsh and Basque in terms of the small number of speakers, limited territory and the suppressive attitudes and actions they have faced (Hualde et al., 1996).

In addition, other similarities can also be drawn in terms of the language shift in both Welsh and Basque. In both countries, a language shift led to a more common language dominating society, creating an environment in which both minority and dominant languages are heard (Urla, 2012). This means that in such unique societies, people are more likely to hear two languages spoken in everyday life compared to bilinguals who live in a community in which only one of those languages is normally used. Bilinguals in such contexts may therefore use words for specific items in one language but not in the other, and vice versa (Mueller Gathercole et al., 2008). Additionally, individuals who grow up in this kind of bilingual environment tend to be exposed to some of their two languages in overlapping contexts (e.g., they may hear both languages when with friends), and some of the two languages in non-overlapping contexts (e.g., they may hear more of one language in the grocery store and hear more of the other language in the waiting room of a medical centre) (Mueller Gathercole et al., 2008). This is important to note, as the way individuals are exposed to each language will not only affect which words are known in each language, but also word usage in specific contexts (Mueller Gathercole et al., 2008). Indeed, it has been suggested that the environment in which Welsh individuals speak Welsh is often situational

(Henley & Schott, 1999). For instance, it is common for minority languages to be used in certain situations such as with family and friends as opposed to a more formal setting (Henley & Schott, 1999). This situational impact on language use is also seen in the Basque community. For example, some school children tend to speak Basque with teachers, but then converse in Spanish with fellow school children (Cenoz, 2008). In Wales, people may use Welsh in familiar situations (Henley & Schott, 1999). Situations that are considered 'familiar' vary from person to person, with some using Welsh only with friends and family, and others using Welsh at home but also at work and in formal settings. This may depend on the area in which the individual lives and how much of their lives include the Welsh language (e.g., occupation). It may be common for people to speak with healthcare professionals in English (Henley & Schott, 1999), but this may be due to a lack of Welsh speaking clinicians.

Consequently, it seems logical to conclude that the environments in which two languages are used may in turn influence the way in which bilinguals use each of their languages. In particular, if it is assumed that switching between languages leads to a switching cost, defined as reduced task performance due to the cognitive effort and fatigue induced by task switching (Macizo et al., 2012; Meuter & Allport, 1999; Thomas & Allport, 2000), and that this may then subsequently contribute to a form of reserve (reserve will be defined later and discussed in detail in the thesis in section 1.7) (Bialystok et al., 2007; Schweizer et al., 2012), then it seems necessary to explore whether this applies to all types of bilinguals. This also means that this Welsh-English bilingualism may be different from the bilingual populations often studied in the literature (e.g., French, German).

1.5 Bilingualism and Cognitive Function

The notion that bilingualism may impact upon developing minds has long been suggested (Bialystok et al., 2012). Early research argued that bilingualism could lead to negative

consequences on cognition, with some evidence showing that bilingual children performed poorly at school, scored lower on tests of intelligence, and presented with more emotional problems also (e.g., Darcy, 1953). However, such early studies were not without methodological limitations. For instance, many failed to consider the impact of confounding variables such as socioeconomic status (SES), where middle-class monolingual children were often compared with bilingual children from a lower-class background (Cummins, 1984).

Even so, the negative consequences of bilingualism were emphasised for many years until a study by Peal and Lambert (1962) found that bilingual French children outperformed English monolinguals on a variety of cognitive tasks. Further, these findings were found to remain even after controlling for group differences in terms of age, gender and SES (Hakuta & Diaz, 1985). However, the cognitive advantages found in Peal and Lambert's (1962) bilingual group may have been somewhat exaggerated by the way in which participants were selected. Firstly, when it was unclear whether a child was bilingual, scores on an English vocabulary test were used to determine which participants would be included. The bilingual sample only included children who scored over a certain level on the English vocabulary test. Therefore, the concern is that performance on the English vocabulary test might have been influenced by their overall intelligence rather than solely reflecting their proficiency in English as a second language (Hakuta & Diaz, 1985). Therefore, it is possible that the cognitive advantages observed in the bilingual group might reflect higher intelligence levels rather than their bilingualism per se. Secondly, although all participants were ten years old, the bilingual group, on average, were from a higher school grade than the monolinguals. Thus, their longer exposure to education may have also contributed, at least in part, to the study's findings (Hakuta & Diaz, 1985).

Nevertheless, subsequent empirical work has also reported that bilinguals have an advantage over monolinguals in many cognitive domains (e.g., Christoffels et al., 2015;

Hakuta & Diaz, 1985). As requisite examples, positive associations have been found between bilingualism and cognitive flexibility (Christoffels et al., 2015), concept formation (Liedtke & Nelson, 1968) and intellectual development (Peal & Lambert, 1967).

However, studies exploring the impact of bilingualism on cognitive function in adulthood have reported contrasting findings (e.g., Gollan et al., 2007; Portocarrero et al., 2007). For instance, some evidence now suggests that bilinguals have weaker verbal skills in each language compared to monolinguals (Bialystok et al., 2012). That is, bilingual individuals may control a smaller vocabulary in the societal language compared to monolinguals who speak the same societal language (Bialystok & Luk, 2012). Research also suggests that adult bilinguals are slower (Gollan et al., 2005) and less accurate (Gollan et al., 2007) on picture-naming tasks (tasks that measure verbal functions and language skills) compared to monolinguals. In addition, the performance of adult bilingual individuals on verbal fluency tasks tends to reveal deficits, particularly in semantic fluency conditions (Portocarrero et al., 2007), and even when they are permitted to respond in any language (Gollan & Ferreira, 2009). This suggests that it may be more effortful for bilinguals to retrieve words compared to monolinguals. Thus, there is a body of evidence to suggest that bilingualism may have negative consequences on various aspects of cognitive processing. Even so, it is crucial to remember that different types of bilinguals differ from each other in terms of their daily experience of using both languages, as well as the situations in which each language is used. In other words, the impact of bilingualism on various aspects of cognition may not be uniform or consistent.

However, it is possible that bilingualism may also result in cognitive advantages because of a potential enhancement of executive functions (EFs) (Bialystok et al., 2009). For example, research suggests that bilinguals outperform monolinguals on tasks that involve conflict or a need to inhibit a learned response (Bialystok, 2011). EFs are a set of general-

purpose control processes considered to play a central role in the self-regulation of thoughts and behaviours that are vital for goal attainment (Paap et al., 2014). Three principle EFs are also thought to exist: inhibition, working memory and cognitive flexibility (Diamond, 2013). Higher-order EFs such as problem solving, reasoning, and planning are developed from the three key functions (Collins & Koechlin, 2012). In terms of the neuropsychology of EFs, they are viewed as a group of processes involving the prefrontal cortex (PFC), with each individual element engaging with other areas of cortical function (Paap et al., 2014). The PFC is also a region involved in the controlled use of semantic knowledge (therefore overlapping with general-purpose EFs) (Hoffman, 2018; Noonan et al., 2013) and may be a process that is different in the bilingual brain.

1.6 Bilingual Language Activation

In reference to why bilingualism may support executive type function, the underlying premise is that bilingual individuals may have more cognitive control compared to monolinguals because of their ability to control and coordinate two languages (Paap et al., 2014). Consistent with this, evidence from research focusing on bilingual language activation suggests that both languages of a bilingual are constantly active to some extent, meaning that language processing is more effortful (e.g., Van Heuven et al., 2008). This means that words in both languages may be activated even when the individual is conversing in one language, meaning that the individual must effectively control this dual activation (Blanco-Elorrieta & Caramazza, 2021). For example, a bilingual speaker is required to monitor patterns of conversation and switch between languages depending on situational changes, such as someone entering or leaving a conversation (Paap et al., 2014). This behaviour therefore requires the individual to inhibit one language whilst activating the other. In addition, more cognitive control is necessary when a bilingual needs to manage the competition between two

languages during speech production (Paap et al., 2014). Consequently, it is also possible that the cognitive demands of bilingualism may impact upon the structure of the brain, where it has long been established that the brain is responsive, rather than rigid, to the environment (Guzmán-Vélez & Tranel, 2015). In this respect, the lifelong experience of using and regulating more than one language, in addition to switching between languages, may contribute to a form of reserve known as cognitive reserve (CR) (Guzmán-Vélez & Tranel, 2015). However, it is unclear whether bilingual cognitive processes and bilingual language activation is similar across different types of bilinguals or whether these processes and their consequences are different in certain bilinguals. For instance, in a study exploring the neural signatures of language co-activation and control during spoken word comprehension in Korean-English bilinguals and English monolinguals, Chen et al. (2017) found that the extent to which both languages are activated may by influenced by the level of regular exposure to each language (Chen et al., 2017). They also found evidence that bilinguals who are better able to manage the dual-language activation (i.e., less affected by the activation of both languages) may also be better and more efficient at managing interference in non-linguistic tasks. Such evidence suggests that the amount of dual-language activation may not be the same across all types of bilingualism, and that it may also lead to differential impacts across certain classifications of bilingualism (Chen et al., 2017). To reinforce this, Marian & Spivey (2003) found parallel activation for words within and between languages. However, they also noted that the competition between the two languages of a bilingual may be mediated by language background and the stimuli presented. Building on this section, the bilingual cognitive processes thought to contribute to CR will be discussed further in section 1.8.

1.7 Bilingualism, Cognitive Function and Cognitive Reserve

Neuroplasticity and CR are two related concepts that attempt to explain how cognitive processes deteriorate or are maintained or improved (Vance et al., 2010). Neuroplasticity is the ability of the brain's neural networks to adapt through growth and reorganisation. That is, the brain is sensitive to life experiences, and therefore its structure and organisation can alter because of experiences throughout life (Guzmán-Vélez & Tranel, 2015). Consequently, the cognitive status of an individual can change across the lifespan due to factors that lead to either positive or negative neuroplasticity (Vance et al., 2010). Positive neuroplasticity is the brain's physiological ability to create and strengthen dendritic connections, produce structural changes, and increase CR (Vance et al., 2010). Factors such as social interaction, education, physical activity (Vance et al., 2010), and bilingualism are thought to stimulate positive neuroplasticity (Baum & Titone, 2014). While some studies have suggested potential negative effects of bilingualism on certain cognitive functions, considering the broader perspective reveals that bilingualism may also contribute to positive neuroplasticity. Thus, avoiding factors that contribute to negative neuroplasticity throughout life (e.g., a decrease in dendritic connections between neurones, a reduction in brain size, and a decrease in neurotrophic factors) (Vance & Crowe, 2006) is arguably crucial to ensure the prolongation of good cognitive health. These factors include poor health, poor sleep, poor nutrition, substance abuse, as well as depression and anxiety (Vance et al., 2010).

In contrast, CR is a concept that refers to those brain changes that support cognitive function (Vance et al., 2010). Fundamentally, because these changes to the structure of the brain can support and improve cognition, the more experiences that an individual encounters throughout life that contribute to positive neuroplasticity, the greater the level of CR (Vance et al., 2010). This theory describes an active form of reserve in which a range of mechanisms, such as an increase of synapses and dendritic endings in addition to other neural elements,

lead to a greater efficiency of the brain's neural networks and/or the adoption of secondary strategies (Stern, 2002). Put simply, the concept argues that the brain copes with brain damage by using cognitive processing methods that already exist or by using compensatory approaches (Stern, 2012). Therefore, individuals with a higher CR may have developed a larger array of other networks to choose from to complete the task at hand and reach the goal, even if the main solution strategy was affected by brain damage or disease (Stern, 2002). Consequently, an individual with a higher CR would have a better ability to cope post brain damage and maintain cognitive health than an individual with a lower CR (Stern, 2012; Vance et al., 2010).

However, determining the level of CR possessed by an individual is challenging because it cannot be measured directly (Guzmán-Vélez & Tranel, 2015). Consequently, it is often estimated by measuring factors such as education, occupation, premorbid intelligence and SES (Guzmán-Vélez & Tranel, 2015; Nucci et al., 2012), where individuals with a higher SES, education level or premorbid intelligence, are thought to have a higher level of CR. However, as these factors may be confounded with each other, it is not possible to establish how each factor may influence cognition (Jones et al., 2011). To illustrate, formal education is the most frequently used measure of CR, but determining the effect of this single factor on CR is challenging because other variables such as SES and premorbid intelligence also impact upon educational achievement (Jones et al., 2011). Thus, and to obtain a more accurate estimate, all factors that may contribute to CR need to be considered.

Further, as CR cannot be directly measured, this also means that establishing a cause-and-effect relationship between measures of CR and the consequences of this reserve, such an advantage on cognitive tasks, has not been possible (Guzmán-Vélez & Tranel, 2015). Therefore, correlational relationships are the only confirmed relationships. Even so, and despite this difficulty, a wealth of evidence supports the existence of CR and that factors,

such as SES and education achievement, contribute to it (e.g., Costa et al., 2008; Guzmán-Vélez & Tranel, 2015).

As mentioned previously, if more cognitive control is required to manage two languages which may contribute to CR throughout life, then it seems reasonable to suggest that bilinguals may show an advantage on tasks that involve switching or conflict resolution (Green, 1998). Costa et al. (2008) compared the performance of young bilinguals and monolinguals (mean age of 22 years in both groups) on an attentional network task developed by Fan et al. (2002). This task assesses three attentional networks: alerting, orienting and executive control, through the presentation of a sequence of visual stimuli. The monolingual group (n = 100) were Spanish speakers, whereas the bilingual group (n=100) had learned two languages (Catalan-Spanish) almost at the same time, received education in both their languages, and lived in a bilingual environment. Costa et al. (2008) found that bilinguals were faster in completing the task, better at conflict resolution, and had a reduced switching cost compared to the monolinguals. Therefore, these findings suggest that bilinguals – and arguably a population with similarity to Welsh-English bilinguals - have more efficient executive control processes, enabling them to complete tasks more effectively and to manage additional cognitive load (Costa et al., 2008). Importantly, it should be noted though that all participants were at an age where their attentional system would have been performing at its best, meaning that there should not have been a difference in the performance of the two groups if bilingualism itself does not provide a cognitive advantage. Therefore, this study supports the suggestion that the increased cognitive control in the daily lives of bilinguals may lead to the ability to complete tasks faster, switch between tasks more efficiently, and resolve conflict more efficiently between congruent and incongruent information.

Consistent with this, Prior and Gollan (2011) also provided evidence of a switching advantage. They investigated performance on a colour-shape switching task in English

monolinguals (n = 47), Mandarin-English bilinguals (n = 43) and Spanish-English bilinguals (n =41). However, after controlling for differences across the groups in response speed and parent's educational level, a bilingual advantage in switching costs was only observed for ly the Spanish-English bilinguals. The Spanish-English bilinguals also reported that they also used both languages every day and switched languages more frequently than the Mandarin-English bilinguals, and they also had significantly smaller switching costs in a language-switching task. Consequently, Prior and Gollan argued that certain types of bilingual language use, which may not be universal to all types of bilinguals, may therefore introduced the advantage. Specifically, that the magnitude of a switching advantage may be determined by the frequency of language switching in day-to-day life. In contrast though, Tare and Linck (2011) found no evidence of bilingual switching costs in a sample of 35 bilinguals who were individually matched to monolingual controls on a range of demographic factors, including age, pay grade, and education, as well as measures of general intelligence and verbal ability. Therefore, while language switching frequency may contribute to a potential switching advantage, it does not appear to be the sole determinant.

In line with this, Bialystok et al. (2004) compared the performance of monolingual and bilingual middle-aged and older adults on the Simon-task, a measure of interference or conflict resolution. There were 40 participants in total - twenty middle-aged adults *(M* = 43.0 years) and twenty older adults (*M* = 71.9 years). In each age group, half of the participants were monolingual English speakers living in Canada, and the other half were bilingual Tamil-English speakers living in India. The bilingual group learnt Tamil as their first language and were educated in both languages from the age of 6 years. The average daily usage of English was 56% and 44% for Tamil, meaning that the bilingual group were considered balanced bilinguals. Bialystok et al. (2004) found that both bilingual groups had smaller Simon effect costs and that they were also quicker to respond to conditions that placed more demands on

working memory. Therefore, incongruent items and the additional load attached, did not lessen the efficiency of the bilingual participants. However, whilst participants did not differ in terms of background experiences and performance on baseline cognitive measures (e.g., abstract non-verbal reasoning, receptive vocabulary), the bilingual advantage was found to be more pronounced in older adults. Thus, while ageing is often considered to have negative impacts on cognitive processes, the study showed that the bilingual advantage seemed particularly enhanced in older adults. One potential explanation for this is that the benefits conferred by bilingualism may only manifest or become more evident during the ageing process. Consequently, this study provides additional evidence that bilinguals outperform monolinguals in tasks involving conflict resolution, but also suggests that the experience of controlling two languages may help to maintain certain cognitive functions in older adulthood (Bialystok et al., 2004).

In contrast to such findings, Paap and Greenberg (2013) found no bilingual advantage on the Simon task in a group of bilingual and English monolingual university students. However, the way in which participants were allocated to groups may explain the findings. Participants self-rated their level of bilingual proficiency and only those who described their level of proficiency in both languages as an *"Advanced Intermediate - Can converse with little difficulty with a native speaker on most everyday topics, but with less fluency than a native speaker"* (Paap & Greenberg, 2013, p.237) or above (Advanced Intermediate – Super Fluency) were included in the bilingual sample. For those who rated their proficiency in English as an "Advanced Intermediate" or above but rated their proficiency in all other languages as *"Intermediate - Can converse with a native speaker on most everyday topics, but with some difficulty"* (Paap & Greenberg, 2013, p.237) or less (Intermediate – Beginner), were classified as monolinguals. This is an issue because it means that the monolingual group had some knowledge of other languages, which may help explain why performance

differences were not found between the bilingual and monolingual groups on the Simon task. This is especially important considering that it is generally accepted that broader definitions of bilingualism are more realistic (Butler & Hakuta, 2006; Grosjean, 1994) and that as argued previously, passive knowledge of a second language must be considered when describing what it means to be bilingual (Wald, 1974). Thus, despite difficulty in identifying participants without any skills in another language, particularly due to foreign language learning in school, future research needs to ensure that the way in which participants are allocated to groups truly reflects their language abilities. It is crucial not to categorise individuals as monolinguals in bilingual research if they possess "advanced intermediate" abilities in a second language (as above).

Even so, other studies have also provided evidence against a bilingual advantage. For instance, Duñabeitia et al. (2014) compared Spanish monolinguals (n=252) to Basque-Spanish bilingual (n=252) children at six successive grades using verbal and non-verbal versions of the Stroop task, which measures the ability to selectively attend to specific information while ignoring distracting information. Both groups performed equivalently, with performance not modulated by other investigated factors, including amount of time exposed to Spanish, the AoA of both languages, and parents' scores of their child's language proficiency (Duñabeitia et al., 2014). However, and consistent with several other studies researching the effects of bilingualism on cognitive function, this study did not fully consider the type of bilingualism involved. As previously explained, Basque-Spanish bilinguals live in an environment in which two languages are heard simultaneously. Therefore, they may be more likely to show balanced dominance to both languages (Antón et al., 2019). This may then impact upon the way in which they use their languages, and in turn, have consequences for the organisation and governance of their cognitive functions. In turn, this could potentially confer a greater degree of cognitive advantage than experienced by bilinguals who

are not fully immersed in a bilingual society. Consequently, it is vital that researchers recognise the importance and potential influence of the type of bilinguals included in their samples when interpreting key findings.

As mentioned previously, it is possible to draw similarities between the Welsh and Basque languages, and consistent with the findings of Duñabeitia et al. (2014), Gathercole et al. (2014) also found no evidence for a bilingual advantage in a group of Welsh-English bilinguals. A sample of 650 children and adults completed a card sorting task, where they were required to sort cards according to a given criterion. Furthermore, 557 also completed the Simon task which assesses interference and conflict resolution, and 354 participated in a metalinguistic task, where they judged the grammatical correctness of sentences and corrected them if necessary. The bilingual participants were recruited from North Wales and were classified as either Welsh-dominant, balanced, or English-dominant, with only Welsh, Welsh and English, or only English at home. The monolinguals were English and were recruited from Chester. The children were from five different age groups (3, 4, 5, 8 [primary schoolers] and 15 years [teenagers]), and there were two groups of adults: younger [18-39 years] and older [57-90 years]. All age groups completed the card sorting task and Simon task, and participants of primary school age and above also completed the metalinguistic task (Gathercole et al., 2014). Overall, Gathercole et al. (2014) did not find strong evidence of a bilingual advantage, even though there was occasionally better performance by monolinguals or persons dominant in the language being tested. For instance, they found that the language spoken at home provided an advantage on the metalinguistic tasks that were conducted in that language. Thus, for the English task, the bilinguals who only spoke English at home outperformed the other bilinguals. Also, in the primary age group, on a task condition that required greater levels of sentence analysis as opposed to attentional control, the bilinguals who only spoke English at home performed better than those who only spoke Welsh at home.

In Welsh tasks, those who only spoke Welsh at home and those who spoke both Welsh and English at home outperformed participants who only spoke English at home, and on the task condition that required greater levels of sentence analysis as opposed to attentional control, those who only spoke Welsh at home, performed better than those who only spoke English at home. Consequently, the findings of Gathercole et al. (2014) support the argument that not all bilinguals are the same and that this should be considered when interpreting research findings. It may be that language switching - the act which is considered the principal contributor to a bilingual advantage - for such uncommon types of balanced and simultaneous bilinguals is more automatic and thus, the cognitive advantages may be less relevant to such populations. On the other hand, it may be that the tasks used in most research studies to explore the potential bilingual advantage are not measuring the relevant cognitive processes. That is, the advantage may be more to do with the activation of links or differences in the search of particular memory systems (e.g., processes operating on semantic memory) as opposed to regular switching, as both languages are constantly 'on-line' (Gathercole et al., 2014; Ordin et al., 2020).

The fact that most of the literature that has failed to find a bilingual advantage has been conducted using participants from fully fluent, bilingual environments, such as those conducted in North Wales and the Basque Country (e.g., Duñabeitia et al., 2014; Gathercole et al., 2014) is significant, particularly as such bilinguals represent the best groups controlling for a number of other confounding variables (e.g., SES, relative levels of proficiency). Future research needs to ensure that participants and cognitive tasks are carefully selected so that it is possible to determine which bilingual individuals show an advantage, as well as what conditions are necessary for the advantage to present itself.

In sum, this section has introduced the concept of CR, reviewed the mixed evidence for cognitive performance differences in bilingual populations, and highlighted the importance of

considering the type of bilingualism in research designs and when interpreting key findings. Building on this, the next section will further explore CR but in the context of ageing and dementia specifically.

1.8 Bilingualism, Ageing and CR

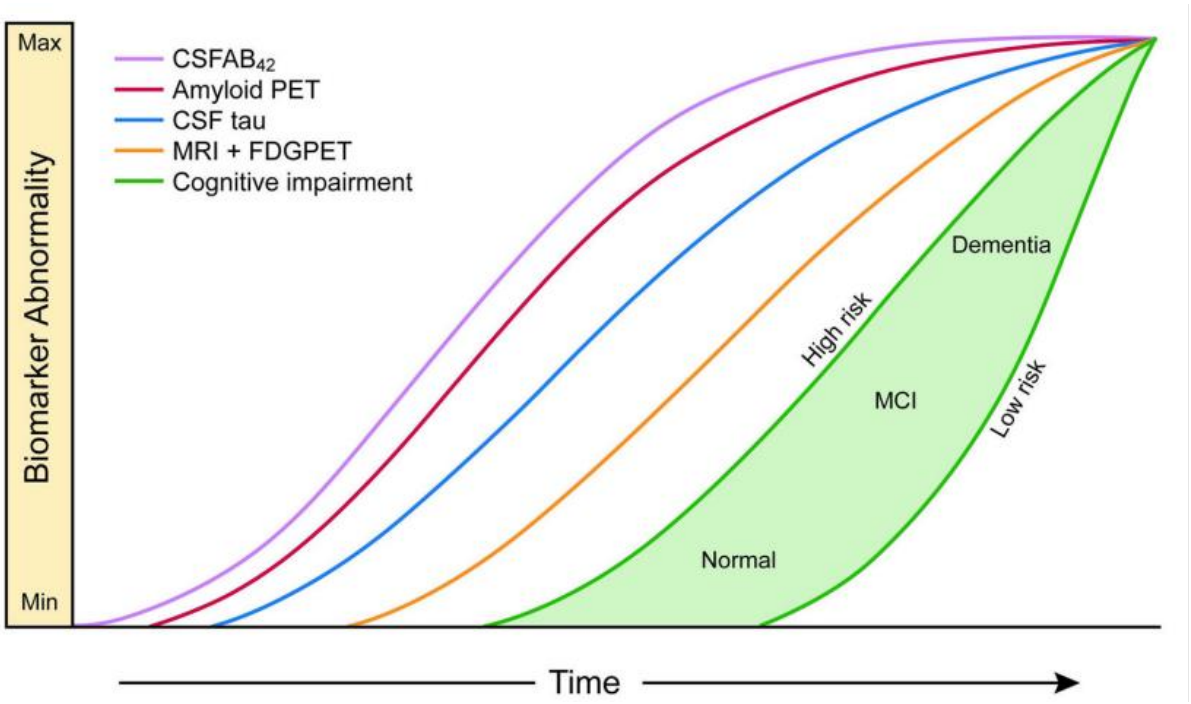
The world's population is ageing rapidly due to lower birth rates and people living longer because of better health care systems (Musselwhite & Scott, 2019). This means that there are increased numbers of older adults living within society. Approximately 11.7 percent of the world's population is over 60 years of age, equating to approximately 900 million people. This is expected to increase to 22 percent by 2050 (United Nations, 2015). In terms of the United Kingdom, it is thought that one in four people (25%) will be over 60 years of age by 2030 (Office for National Statistics, 2020). As a result, the prevalence of age-related conditions such as dementia is also likely to increase (Prince et al., 2013; Wittenberg et al., 2020).

More than half of the world's population is also bilingual or multilingual (Bialystok, 2017; Digard et al., 2022; Grosjean, 1992), with the Ethnologue (2020) estimating that 7,000 languages are spoken in the 195 countries of the world. Therefore, the increased prevalence of dementia is likely to affect many bilingual individuals and is especially important to consider because bilingualism is poorly understood generally, but also because evidence suggests that bilingualism may impact upon the ageing process (Bialystok et al., 2007).

Cognitive health, broadly defined as the improvement or maintenance of cognitive function as well as the absence or delay in the onset or progression of dementia (Jedrziewski et al., 2007; Laditka et al., 2009), is known to contribute to quality of life and is central to whether older adults maintain a level of independence (Clare et al., 2017). That is, cognitive health allows older individuals to continue to use learned skills in the appropriate context,

remain socially connected, and maintain a sense of purpose (Hendrie et al., 2006). However, some loss of cognitive health in older adulthood is inevitable, meaning that some decline in cognitive health can be anticipated as adults age. However, the extent or severity of the decline is dependent on a myriad of factors, including such things as genetics and gender (Polidori et al., 2010). Nevertheless, evidence also suggests that certain lifestyle factors may protect against higher levels of cognitive decline and the development of Alzheimer's disease (AD) (Claire et al., 2017). Indeed, the key construct of CR has been used to account for the differential outcomes seen for individuals across the ageing process (Stern et al., 2019; Whalley et al., 2004; Yegla & Foster, 2022). A classic example of this situation is captured by individual disease progression in patients with AD (see Figure 1). Over time with the accumulation of certain biomarkers, cognitive decline will manifest (Jack et al., 2010; Jack et al., 2013), but the onset and acceleration of this decline seems to be much earlier or later for some individuals (Snowdon, 2003; Stern, 2002).

Figure 1. *Biomarkers of AD and the Cognitive Response Curves (Jack et al. 2013 - published article is under the CC-BY-NC-ND licence)*



It seems that some individuals have a greater ability to cope with the onset of dementia compared to others, presenting with cognitive deficits much later than expected (Jack et al., 2013; Stern, 2002). Indeed, CR may help to explain the individual differences observed during the course of cognitive decline, which to reiterate, is a concept that argues that the brain may become more efficient as a result of life experiences increasing synapses, dendritic endings and other neural elements, thus allowing the brain to cope with brain damage by using cognitive processing methods that already exist or by adopting compensatory approaches (Guzmán-Vélez & Tranel, 2015; Stern, 2002; Stern, 2012). Therefore, individuals with a higher CR may have developed a larger array of other networks to choose from to complete the task at hand and subsequently, would be able to reach the end goal even if the main solution strategy was affected by brain damage or disease (Stern, 2002).

In relation to the life experiences that are thought to be implicated in CR, research has found a plethora of factors that may play a role (Clare et al., 2017). These factors include an active lifestyle, level of intelligence, educational and occupational level, a balanced diet (Whalley et al., 2004), and engagement in social and intellectual leisure activities (Anstey et al., 2013; Scarmeas & Stern, 2003; Stern, 2002). For example, Zhou et al. (2018) found that a group of elderly Chinese participants with high levels of social engagement had a lower risk of dementia compared to those with lower levels of social engagement. Similarly, other studies have also reported a protective effect of social engagement against the onset of dementia (e.g., Gureje et al., 2011; Penninkilampi et al., 2018). Additionally, findings from the famous Nun Study also demonstrated that despite having a genetic predisposition to AD as well as many AD lesions in the neocortex, Sister Bernadette scored very high on various cognitive tests (Snowdon, 2003). Post-mortem magnetic resonance imaging (MRI) also showed that although she was in the 90th percentile for neurofibrillary tangles in her neocortex, she scored in the 90th percentile for the volume of gray matter in her brain, which

is the principal material of the neocortex. Subsequently, Mortimer et al. (2003) argued that a higher level of education and a larger head size may have played a role in preventing the clinical symptoms of dementia from manifesting. Consistent with this, Fratiglioni and Wang (2007) concluded from their systematic review of the literature, that education, occupational complexity, and a socially engaged lifestyle in older adulthood may contribute to delayed onset of AD. That is, mentally stimulating activities throughout the lifespan may preserve late-life cognitive health, and in turn, protect against dementia risk. In addition to these factors and as mentioned previously, it is thought that the lifelong experience of using and regulating more than one language, as well as switching between languages, may also contribute to increased CR (Guzmán-Vélez & Tranel, 2015).

Even so, as the evidence supporting the notion that these factors impact upon CR is merely correlational, it remains unclear whether maintaining an active lifestyle properly influences cognitive function, or whether individuals with higher functioning brains, due to genetics, have a higher likelihood of performing well intellectually and are also naturally protected against a form of dementia (Craik et al., 2010). However, this is not the case for bilingualism. Importantly, bilingualism differs from the other lifestyle factors mentioned above because many bilingual individuals have been born into an environment in which two languages are spoken, as opposed to having chosen to learn another language (Craik et al., 2010). This is important because CR is often thought to be associated with individuals who simply have higher socioeconomic status, or perhaps a high level of education or intelligence. However, bilingualism crosses the typical boundaries of such socioeconomic factors. Consistent with this, research suggests that bilingualism may contribute to the normal ageing process and lead to cognitive advantages in healthy ageing but may also play a protective role when a neurodegenerative condition is present (Bak et al., 2014).

Several factors for how bilingualism might enhance CR have been proposed in the literature. For example, some studies propose that inhibition and/or monitoring may be the aspects of bilingualism that contribute to CR and provide an advantage (Bialystok, 2015; Costa et al., 2009; Hilchey & Klein, 2011). Building on this notion, the Bilingual Interactive Activation model (Grainger and Dijkstra, 1992; van Heuven et al., 1998) considers how features, letters, and words interact with each other, suggesting that when presented with a word, features and letters initiate word processing by mutually activating each other. The activated letters then activate words in the target language, but for bilinguals, words from both languages may become active and interact with each other, suggesting an integrated lexicon. This interaction therefore requires the bilingual to adopt an inhibitory control process to manage cross-language activation (Grant et al., 2019; Thomas & Van Heuven, 2005). For example, when presented with the word 'dog', the corresponding node for that word may be activated in both languages (Dijkstra et al., 2019). However, due to the focus on orthographic representations within this model, the model was updated to account for phonological and semantic representations (i.e., the Brain Activation Plus model). This means that the recognition of words is not only impacted by interacting orthography across languages, but also by interacting semantic and phonological information (Dijkstra & van Heuven, 2002). The suggestion, therefore, is that selectively attending to the required language and inhibiting information from the other language may enhance CR. An alternative model with a similar emphasis on inhibition, is the Inhibition Control model (Green, 1998). This model suggests that bilinguals inhibit competing information from both languages via a supervisory attentional system depending on the context or goal. For example, during a conversation in Welsh, words from the non-target language (i.e., English) are inhibited by the individual. This form of inhibition is suggested to react in the moment using an "as and when needed" approach (Linck et al., 2008). Reflecting the diversity among bilinguals and consistent with

the notion of inhibition, the Adaptive Control Model (Bialystok & Craik, 2022; Green & Abutalebi, 2013) proposes that the need for inhibition depends on the individual's linguistic environment, such as the frequency and nature of exposure to both languages. These differences between bilingual linguistic environments may place differing levels of demand on cognitive control processes (Bialystok & Craik, 2022).

Considering the above evidence and models, it could be that bilingualism is therefore an important proxy for CR, and for potentially explaining variability in outcomes for individuals with dementia. In fact, the notion that bilingualism may enhance CR and lead to a delayed onset of dementia has been suggested by several studies (e.g., Alladi et al., 2013; Berkes & Bialystok, 2022; Bialystok et al., 2007; Woumans et al., 2015). Such studies demonstrated that the constant use of two or more languages may lead to better cognitive function in ageing (e.g., Bialystok et al., 2004). Using the hospital records of 184 bilingual and monolingual patients diagnosed with a form of dementia, Bialystok et al. (2007) further investigated this possibility. At the beginning of the study, the groups (bilingual versus monolingual) did not perform differently on the Mini-Mental State Examination (a screening tool for cognitive impairment), and they did not differ in terms of occupational status. However, the monolingual group had been educated for more years than the bilingual group, which should have been advantageous to these individuals. The findings demonstrated however that the age at onset of dementia was 71.4 years for the monolingual participants compared to 75.5 years for the bilingual individuals. Additionally, the monolingual patients were first seen at clinic at 75.4 years compared to 78.6 years for the bilinguals. In essence, the bilingual group first presented with symptoms of dementia 3-4 years later than the monolingual group, suggesting that a psychological construct may have impacted upon the development of a biological disease (Bialystok et al., 2007).

Moreover, factors such as immigration, cultural differences and employment status could not account for the above findings (Bialystok et al., 2007), which is important to note because these results would only be of importance if the participants in both language groups only differed in terms of their language abilities. Even so, it must be noted that the estimate of the age at onset of dementia is somewhat subjective. Subjectivity is also an issue which affects the time point at which patients and their families decide to seek support, and therefore the age at which the diagnosis is made is delayed (Bialystok et al., 2007). Thus, the findings are suggestive rather than definitive, but it is nevertheless possible to conclude that bilingualism may be a factor involved in delaying the symptoms of dementia.

Even so, Craik et al. (2010) replicated the above findings in a sample of individuals with AD, as opposed to any form of dementia. Overall, bilingual patients were found to have been diagnosed 4.3 years later and reported symptoms 5.1 years later than the monolingual group. The groups did not differ in terms of cognitive abilities, occupational level, and there was also no effect of immigration status. The monolingual group had also received more education. Therefore, the findings of Craik et al. (2010) support the notion that bilingualism may increase CR and delay the onset of AD, with the results not accounted for by the presence of various confounding variables. However, it is not possible to determine whether bilingualism is truly the factor influencing CR in the bilingual group, or if other unexplored factors associated with this group are potentially responsible (Craik et al., 2010).

Another important study and one that explores the Welsh population specifically, is that of Clare et al. (2016a). They found that Welsh-English bilinguals with AD were more impaired at the time of diagnosis compared to a monolingual sample, suggesting that they were also diagnosed at a later stage in the disease progression. However, these findings could be explained by the fact that individuals living in rural communities in Wales may have more

informal support networks and consequently, do not seek formal support until later in their cognitive decline (Shannon et al., 2019).

Additionally, studies using neuroimaging techniques have also investigated the relationship between bilingualism and CR to evaluate the differences in brain structure and cognitive performance between lifelong bilingual and monolingual groups with AD (e.g., Gallo & Abutalebi, 2023; Perani et al., 2017; Schweizer et al., 2012). For example, Schweizer et al. (2012) found no significant difference between a monolingual versus bilingual group in terms of cognitive function, education level, daily activities, or age at diagnosis. However, a head computed tomography (CT) scan showed evidence of significantly more medial temporal lobe atrophy in the bilingual compared to monolingual group. Therefore, although the bilingual group had more atrophy, their daily activities and cognitive function did not differ from that of the monolinguals. Further, the monolingual group also had higher occupational status and had received more formal education, meaning that the findings could not be explained by such factors (Schweizer et al., 2012). Even so, and whilst these findings are encouraging, more research is required to establish whether bilingualism contributes to CR and leads to a delayed onset of the manifestations of dementia.

In keeping with the discussion on brain structure in bilinguals, the left middle and inferior frontal gyrus have been found to play a central role in bilingual language control, and particularly, for inhibiting automatic and irrelevant information (Abutalebi & Green, 2016). Research has also demonstrated that the inferior frontal gyrus is also a key element in the processing of semantic decisions that are executively demanding (Whitney et al., 2011). Again, it may be that the bilingual advantage lies in processes related to semantic memory.

Evidence also suggests that there is a positive association between the number of languages spoken by an individual and the age of onset and diagnosis of AD (Chertkow et al., 2010). This is consistent with other findings that suggest that multilingualism is associated

with better cognitive health (Kavé et al., 2008). Indeed, it seems that the number of languages spoken may predict scores obtained on cognitive assessments more so than the impact of other variables such as age, gender and education (Kavé et al., 2008). This is supported by the findings of Perquin et al. (2013) who found that the likelihood of *not* developing cognitive impairment increased up to four times with each spoken language, and that results remained even after accounting for factors that are known to contribute to CR. Please note that having briefly mentioned multilingualism in the context of CR, the impact of specific factors on CR will be discussed further in the 'Bilingualism and CR – Possible Moderating Factors' section of the thesis.

However, and even though a body of evidence suggests that bilingualism contributes to CR which may in turn lead to a delay in the onset of AD symptoms, some studies have found contradictory results (e.g., Calvo et al., 2016). For example, Crane et al. (2010) assessed healthy second-generation Japanese-Americans for dementia on three occasions over a six-month period, finding that the rates of cognitive decline were not associated with the use of spoken or written Japanese. However, participants' use of the Japanese language was only assessed through self-report measures, meaning that their language use may not have been correctly estimated (Crane et al., 2010). Indeed, if the participants rarely used both languages, it is perhaps unsurprising that no significant relationship was found between the rate of cognitive decline and the use of Japanese.

Zahodne et al. (2014) assessed 1,067 bilingual participants (Spanish-English) at 18–24-month intervals for up to 23 years. During the study, 300 participants developed dementia, but bilingualism was only associated with better memory and executive function at baseline and was not found to be independently related to the rate of cognitive decline or the development of dementia (Zahodne et al., 2014). However, as different types of bilingualism may impact upon CR differently, it is important to acknowledge that participants were native

Spanish speaking immigrants to the United States who became bilingual by learning English as adults. Therefore, it may be that late bilingualism does not offer protection against cognitive decline and dementia, where individuals who acquire a second language later in life may not have the same degree of proficiency or frequency of language use as those who are bilingual from childhood. Indeed, if cognitive advantages associated with bilingualism stem from the continuous and active use of two languages, which then stimulates cognitive processes (e.g., cognitive flexibility, executive control), individuals who acquire a second language later in life may not experience the same level of cognitive advantage to offer protection against later cognitive decline. Alternatively, it may be that the executive challenge of switching between English and Spanish is not enough to impart changes significant enough to protect against cognitive decline and dementia, where it has been suggested that different language pairs may confer different levels of protection against cognitive decline. Namely, that the cognitive burden of maintaining two languages that are more phonemically and orthographically similar (e.g., Spanish and English), may confer weaker protection from cognitive decline than less structurally similar languages (e.g., Korean and English).

Moreover, it is also possible that participants level of engagement in their second language, as well as the contexts in which they use the language, may also offer some explanation for the findings observed by Zahodne et al. If the act of constantly inhibiting one language is the main factor that leads to a bilingual advantage as suggested by Emmorey et al. (2008), then the protective effects of bilingualism may only be present in those who use both languages in day-to-day life and who are required to search through large (or multiple) knowledge stores every day for required information whilst inhibiting irrelevant information (Bartolotti & Marian, 2012). Consequently, future research also needs to assess how often and in what contexts participants use each of their languages. Factors such as language use,

context of language use and proficiency, will be discussed in greater detail in the 'Bilingualism and CR - Possible Moderating Factors' section below.

That said, findings from Lawton et al. (2015) also suggest that bilingualism may not impact upon cognitive decline. In a large sample of participants over 60 years of age who were assessed every 12-15 months for ten years, Lawton et al. (2015) investigated whether the age at diagnosis of AD and vascular dementia were different between monolinguals and bilinguals in immigrant and US born Hispanic Americans. Overall, they found that age at diagnosis was not significantly different between the monolingual and the bilingual participants. However, and representing a common limitation with such research, bilingualism status was only assessed using self-report measures. Consequently, the language proficiency of bilingual participants may have been somewhat exaggerated, which may help to explain the observed null effects.

In sum, there are mixed findings surrounding whether bilingualism contributes to CR and delays the onset of dementia. Findings vary from study to study, with some showing that bilinguals were diagnosed with AD later than monolinguals, and other studies finding no difference between monolinguals and bilinguals in terms of the rate of cognitive decline and age at AD diagnosis (e.g., Calvo et al., 2016). It may be that bilingualism contributes to CR, but only under certain conditions and dependent on such factors as the type of bilingualism, age of acquisition (AoA), language proficiency and language use (e.g., Guzmán-Vélez & Tranel, 2015).

1.9 Bilingualism and CR - Possible Moderating Factors

To explore the above possibilities further, this next section will consider the potential factors that may moderate the relationship between bilingualism and CR. Although it is recognised that a plethora of factors may modulate this relationship (see Van Den Noort et al.,

2019), the moderating factors mentioned in this section will focus on those related to bilingualism (e.g., language proficiency, type of bilingualism). Indeed, and as already indicated previously, there appear to be many factors in relation to language/bilingualism that may influence the relationship between bilingualism and CR, including education, language proficiency, the amount of daily usage of the two language, the type of bilingualism, and the language itself (Bartolotti & Marian, 2012; Guzmán-Vélez & Tranel, 2015; Nour et al., 2019; Perani et al., 2017). However, and even though the literature studying the impact of these factors is somewhat disorganised and unclear, some exploratory studies have attempted to decipher what role they might play in shaping the complex relationship between bilingualism and CR in the context of ageing.

Many studies (e.g., Crane et al., 2010; Lawton et al., 2015; Zahodne et al., 2014) investigating the effects of bilingualism on CR and ageing do not account for language proficiency and only focus on whether the participants have used their languages on a regular basis before adulthood. However, language proficiency must be taken into consideration because proficiency and AoA of the second language have been found to cause functional and structural brain changes (Guzmán-Vélez & Tranel, 2015), and have also been associated with the onset of dementia (Gollan et al., 2011).

For example, Mechelli et al. (2004) found that native English bilinguals who learnt their second language before five years of age and had regularly used both languages since, had greater grey-matter density in the left and right inferior parietal cortex compared to bilinguals who learnt their second language between the ages of ten and fifteen. However, whilst similar findings have been reported elsewhere (e.g., Abutalebi et al., 2014 and 2015), both Olsen et al. (2015) and Gold et al. (2013) found no differences in grey or white matter volume between language groups, and Anderson et al. (2019) reported poorer volume in bilinguals than monolinguals. One possibility for such inconsistent findings relates to the different ages

of participants used across the studies. For example, in studies finding better brain structure, participants tended to be around 60 years of age – essentially being on the border of being older adults. In contrast, participants tended to be older (65-75 years) in those finding comparable grey matter volume (e.g., Olsen et al., 2015), and older still (75 years) when poor white matter structure has been found in bilinguals (e.g., Anderson et al., 2019). In response, Bialystok (2021) therefore raised the possibility that with increasing age, bilingual and monolingual participants may maintain parity on cognitive measures, but that bilinguals may show more brain deterioration than monolinguals in older age (i.e., a dissociation between brain level and cognitive level indicative of cognitive reserve). Additionally, other evidence suggests that less proficient bilinguals may also use more brain resources compared to highly proficient bilinguals, suggesting that more proficient bilinguals may use the brain's available resources more efficiently (Leonard et al., 2011). Therefore, there is at least some evidence to suggest that either or both language proficiency and AoA may modulate the impact of bilingualism on CR.

With regards to language proficiency and the delay in dementia manifestation, Jafari et al. (2015) found that Persian-English individuals with a higher degree of bilingualism, indicated by higher proficiency, achieved better lexical memory scores, where participants listen to a list of words and are required to recall as many as possible. Additionally, Gollan et al. (2011) found that individuals who are more proficient in their second language presented with symptoms of dementia later than less proficient individuals. However, whilst higher degrees of bilingualism (Spanish-English) were related to a delay in the age at diagnosis and onset of symptoms, this was only true for those with low levels of education. Therefore, it may be that education does not provide additional benefits because bilinguals with a high education level reach a threshold at which the contribution of bilingualism to CR decreases (Gollan et al., 2011). Consistent with this line of thinking, it could be that bilingualism and formal

education may use the same or similar mechanisms to mitigate the impact of brain damage or disease. This may help explain why education does not lead to additional significant benefits for bilinguals (Gollan et al., 2011). The findings of Gollan et al. (2011) also demonstrated that objective measures of bilingualism predicted age at diagnosis, whereas self-reported measures did not. Consequently, this study supports the existence of a relationship between bilingualism and a delay of AD onset, but also highlights how the effects of bilingualism are not always obvious due to an interaction between education and bilingualism and due to the lack of objective measurements of bilingualism. More broadly, these findings also provide evidence that bilingualism may contribute to CR and that this reserve may delay cognitive decline. However, there may be an upper limit on the extent to which this reserve can protect against such decline and the development of AD (Gollan et al., 2011).

As previously mentioned, there is also evidence suggesting that the protection against the onset of cognitive impairment increases with the number of languages acquired by an individual (e.g., Bialystok et al., 2016; Perquin et al., 2013). For example, and as covered in greater detail previously, Perquin et al. (2013) found an increase in this protection of up to four times with each spoken language. They also found that the age at which an individual learns more than two languages plays a role in determining whether multilingualism contributes to the delay of AD. Those who learnt more than two languages in early childhood were protected against cognitive decline to a higher degree than those who acquired the languages at an older age. Therefore, protection against cognitive decline appears to increase with the number of languages spoken and that early multilingualism may also contribute to CR to a greater extent than late multilingualism.

Interestingly, the language in which the bilingual is most proficient also seems to modify the effect of bilingualism on CR. Kavé et al. (2008) examined a group of bilinguals (*M* age = 83) who spoke Hebrew fluently and spoke various combinations of other languages. Those

who were less proficient in their native language performed better on a cognitive-screening test than bilinguals who were less proficient in their second language. This may be because learning a second language is both effortful and time consuming and may also require more cognitive resources, which may then provide a higher level of CR (Kavé et al., 2008). Despite this, studies have observed a delay in the onset of dementia regardless of the languages spoken by participants (e.g., Perquin et al., 2013; Zahodne et al., 2014).

In sum, although little research has focused on the potential factors that may influence the relationship between bilingualism and CR, the available evidence suggests that bilingualism could affect CR and may also delay the onset of AD. However, more research is needed to determine the conditions in which bilingualism may impact upon CR, as well the factors that may moderate the complex relationship. Future studies also need to improve their definition of bilingualism so that factors such as language proficiency and language use are incorporated, as the diverse ways in which bilingualism is defined and measured appears to impact on study findings (Ljungberg et al., 2016). This may explain why the current literature has not yet ascertained the precise nature of the relationship between bilingualism and CR and its impact on the development of dementia.

1.10 Conclusion

This chapter presented the various definitions and classifications of bilingualism, and additionally emphasised how not all bilingual language use is the same which could lead to differential impacts on cognition. However, the latter remains largely unexplored. This chapter also discussed the key construct of cognitive reserve (CR) and how it may be linked to bilingualism, noting how there is mixed evidence of cognitive performance differences across bilingual and monolingual groups. Additionally, this chapter explored whether a bilingual advantage exists in older adults and discussed how CR may account for large inter-

individual variability in dementia outcomes. In addition, research investigating the role of bilingualism in CR was discussed, noting that similar elements of bilingualism as proxies of CR (e.g., inhibition, attention switching) tend to be examined, and that the evidence for the impact of bilingualism in this context is also mixed (e.g., Calvo et al., 2016; Bialystok, 2021). Consequently, it may be that the bilingual advantage lies in other cognitive processes (e.g., processes that operate on language or processes involved in making semantic decisions).

1.11 Thesis Aims

The preceding literature review indicates mixed evidence of a bilingual advantage in the context of cognitive reserve (CR). Moreover, studies examining this phenomenon are plagued by methodological limitations, and research involving Welsh-English bilinguals is notably lacking. Additionally, most studies investigating a bilingual advantage tend to focus on common cognitive processes, yet there remains inconclusive evidence supporting or refuting its presence. Therefore, the aim of this thesis is to explore the potential bilingual advantage in Welsh-English bilinguals. Namely, to explore the impact of this type of bilingualism on cognition in young and older adults, and across general and specific cognitive processes, in an effort to identify a possible bilingual advantage that may serve as a more suitable proxy for CR. To achieve this goal, three areas will be explored: (1) performance on Welsh/English language versions of cognitive assessments in a sample of Welsh-English bilinguals, with comparison against monolingual normative data; (2) performance across monolingual and bilingual populations on semantic tasks when semantic control demands are manipulated, and (3) performance across monolingual and bilingual populations when making decisions about semantic and general knowledge under uncertainty.

Chapter 2 – Study 1 – Collecting Normative Data for Clinical Neuropsychological Assessments for the Purposes of Testing Welsh Speaking (Bilingual) Individuals.

2.0 Chapter Overview

As discussed in Chapter 1, there is mixed evidence and a lack of clarity around the impact of bilingualism, let alone Welsh-English bilingualism, on cognitive functions and its potential role in CR. Consequently, the aim of this chapter was to explore gross performance variations in Welsh and English versions of two cognitive assessment batteries. The goal was to determine whether (a) performance differs across Welsh and English language assessments within a group of Welsh-English bilingual participants, and (b) the performance of bilinguals differs from that of monolingual normative data, investigating the potential influence of a bilingual advantage in CR. In addition, this chapter highlights the broader importance of ensuring that individuals have access to cognitive assessments in their language of preference. It explores the way in which cognitive assessments are currently administered in Wales, the legal requirements surrounding the provision of health care services in Welsh, and the growing desire amongst Welsh people for cognitive assessments to be conducted through the medium of Welsh.

2.1 Background

As mentioned in the Chapter Overview, the aims of this study were two-fold. The study aimed to: (a) investigate performance across Welsh and English language assessments to determine whether the language format of cognitive assessment yields meaningful performance differences, and (b) explore if performance of Welsh-English bilinguals is different to that of normative data to determine a possible bilingual advantage. The following sections will therefore address each aim in order.

Language impairment plays a crucial role in the development of Alzheimer's disease (AD) and the way in which professionals monitor its progression (Obler & Albert, 1984). More than half of the world's population is bilingual or multilingual (Grosjean, 1992), with the Ethnologue (2020) estimating that 7,000 languages are spoken across the 195 countries of the world. Due to the ageing population, the number of bilingual individuals suffering from AD is likely to increase, along with the increased prevalence of various dementias in the general population (Grosjean, 2001; Ponjoan et al., 2019). Indeed, whilst some evidence suggests that bilingualism may potentially serve as a protective factor in the development and progression of AD (see Chapter 1), it does not mean that the prevalence rate of AD will not increase. Nonetheless, little is known about the way in which AD influences the two languages of a bilingual (Stilwell et al., 2016). In the specific context of Wales, the last census in the United Kingdom (UK) in 2021 showed that approximately 17.8% of the population (538,300 speakers) speak Welsh (Welsh Government, 2022). However, this number is likely to increase by 78% (437,984 speakers) given the Welsh Government's aim to reach one million Welsh speakers by 2050 (Williams, 2017). Therefore, the prevalence of AD may increase in the complex context of Welsh-English bilingualism. This is important because the broader impacts of increasing rates of bilingualism and AD is unknown.

Consequently, due to the ageing population and likely associated increased prevalence of dementia (Ponjoan et al., 2019), along with a potential increase in the number of bilingual individuals in Wales also (Stilwell et al., 2016), the ability to detect the condition at the earliest opportunity is essential for clinicians to ensure that the required interventions and support are established and implemented in a timely and efficacious manner (Yoshida et al., 2011).

2.1.1 The Challenge

The early signs of dementia are often misdiagnosed or unrecognised (Bluvshtein, 2004), and this is further confounded when individuals do not have the opportunity to communicate their symptoms or express their feelings in their native language (Braun et al., 1996). Perhaps more concerning though, is that such individuals do not routinely have access to formal language-sensitive and culturally appropriate cognitive assessments either (Chou et al., 2002), despite the fact that a significant portion of the general population in Europe is composed of non-Western migrant populations (Nielsen et al., 2011). Further, populations in most European countries are reaching an age where dementia is becoming more prevalent. That is, low birth rates and higher life expectancy are transforming the shape of the world's age pyramid, meaning that the population will be much older and as a result, dementia will be more prevalent. Therefore, and despite the fact that the symptoms of dementia are consistent across various ethnicities and that assessment instruments have also been found to affect diagnosis rates in people who do not speak the majority language (e.g., Daker-White et al., 2002; Hassett et al., 1999; Nielsen et al., 2011, it is perhaps even more surprising that there is generally a lack of culturally validated, and language appropriate, cognitive assessments available for these individuals.

Such issues are particularly relevant to Wales, where there is a growing desire amongst Welsh people for cognitive testing to be conducted through the medium of Welsh (Welsh Language Commissioner, 2012; Older People's Commissioner for Wales, 2016). For example, the Welsh Language Commissioner (2012) report describes the experience of a female carer who took her Welsh speaking mother to a GP surgery to undertake a dementia test in English:

> *"Carys feels that if the test had been in Welsh her mother would have responded better, and would have felt more at ease. In one instance, her mother could only remember the*

Welsh word for 'harp' when asked to name a number of images, and was told that she was wrong, which she found upsetting" (Welsh Language Commissioner, 2012, p.55).

This type of experience continues to occur despite the fact that 'Standard 78' of The Welsh Language Standards (No.7) Regulations of 2018 states the following:

You must publish a policy on providing a primary care service which requires you to take the following into account when you make decisions in relation to providing a primary care service:

(a) what effects, if any (and whether positive or negative), the decision would have on -

(i) opportunities for persons to use the Welsh language, and

(ii) treating the Welsh language no less favourably than the English language;

(b) how that decision could be taken or implemented so that it would have positive effects, or increased positive effects, on:

(i) opportunities for persons to use the Welsh language, and

(ii) treating the Welsh language no less favourably than the English language; and

(c) how the decision could be taken or implemented so that it would not have adverse effects, or so that it would have decreased adverse effects on—

(i) opportunities for persons to use the Welsh language, and

(ii) treating the Welsh language no less favourably than the English language

("The Welsh Language Standards (No. 7) Regulations 2018", 2018, schedule 2, part 1, standard 78).

However, despite such reports from patients and the Welsh Government's regulatory standards (Welsh Government, 2015), as well as the knowledge that culture and language play an important role in the lives of older adults (Braun et al., 1995; Mueller et al., 2016), language appropriate and sensitive assessments are rarely available. Further, even when they may be available, clinicians may not have the necessary language skills to utilise the

assessments and to communicate with patients in their first or preferred language, bilingual normative data is rarely available (Mindt et al., 2008), and translated tests do not always have established psychometric qualities either (Mindt et al., 2008), making interpretation of scores impossible. Evidence also suggests that language plays an important role in the progression of AD, meaning that a lack of language appropriate assessments may result in differential impacts depending on the stage of the disease. For example, patients may find it difficult to find the correct words during a spontaneous conversation (i.e., anomia) or in a word-list generation task in the early stages of cognitive change (Nicholas et al., 1985). In contrast, and by the middle stages, the production and comprehension aspects of language become more challenging (Dassa & Amir, 2014). Then at later stages of decline, speech production may become uninterpretable, and patients may not engage in any social interaction by communicating (Kempler, 1995). Thus, the key message here is that impeding communication by not allowing individuals to 'self-select' their language of preference (e.g., in example above) will only add to the developing difficulties (and distress) of patients during this pattern of cognitive decline.

To further reinforce the importance of appropriate assessments, performance on cognitive tests may also be hampered if the test is not administered in a bilingual's native language – or perhaps performance may not reflect the individual's true ability if it is completed in their second as opposed to first/native language (Escobar et al., 1986). This is relevant to the current study, as research suggests that it is best practice to conduct assessments in both languages of a bilingual, irrespective of whether they are balanced or unbalanced in their proficiency. This approach allows for a more comprehensive evaluation of an individual's cognitive profile, capturing information that may come more readily to mind in one language than the other (Paradis, 2008; Rivera Mindt et al., 2008). Thus, the current study also has relevance to all types of bilinguals – not just Welsh-English bilingualism.

Cognitive assessments in Wales are usually conducted through the medium of English, potentially disadvantaging native Welsh speakers. The potential underpinning reasons for this not only include a lack of appropriate Welsh medium assessments as mentioned previously, but also that most Welsh speakers are fluent in English, and so clinicians may assume that it is appropriate to assess a Welsh individual's ability in English only. However, this assumption not only fails to comply with 'Standard 78' of The Welsh Language Standards (No.7) Regulations of 2018, but it is also an inaccurate assumption because conducting an assessment in English with balanced Welsh-English bilinguals may still place them at a disadvantage. For example, a balanced Welsh-English bilingual may still lack confidence in English comparatively speaking with them needing to translate the instructions (in their mind) to fully understand the task and before providing the answer in English (Morgan & Crowder, 2003). This is important to consider for two reasons. Firstly, such individuals may take longer to process and understand task instructions, and in turn, provide an answer (due to translating and processing instructions). Secondly, the latter point is particularly important to consider in the context of timed tasks, or those that use reaction time as an index of performance and cognitive ability (Morgan & Crowder, 2003). The result of these issues means that there is a risk that the participant's scores do not provide a reliable indication of their performance and cognitive status. In extreme cases, this may mean the difference between a normal and abnormal score, or the possible presence or absence of a cognitive impairment, with serious ramifications for subsequent healthcare, interventions and support. For example, an individual who may underperform because of the language of test administration may be incorrectly diagnosed with cognitive decline.

Additionally, being unable to communicate in the medium of preference may have broader impacts beyond an individual's test score and cognitive profile. Undergoing a neuropsychological assessment can already invoke stress/anxiety in patients, with test anxiety

implicated in poor test performance (Dorenkamp & Vik, 2018). For example, Dorenkamp and Vik (2018) conduced a systematic review of 78 articles and found that two thirds of the papers reported a relationship between test performance and anxiety. Specifically, test, social, state and math anxiety were regularly associated with poor test performance, whereas anxiety was related to poor performance on tasks of verbal memory, inhibition, attention and working memory (Dorenkamp & Vik, 2018). Therefore, conducting an assessment in a language that is not the individual's first choice could add further to test anxiety/stress, leading to additional negative impacts on performance and increased risk of misinterpretations and incorrect diagnoses as a result (Morgan & Crowder, 2003). However, to the best of the author's knowledge, no evidence currently exists concerning the language of an assessment and misdiagnosis rates, including whether bilinguals not tested in their preferred language are more likely to be misdiagnosed.

Even so, a growing body of evidence has begun to demonstrate how bilingualism may impact on cognitive test performance depending on the language of the assessment. However, it is important to highlight that the findings that follow encompasses various types of bilingualism, rather than exclusively focussing on balanced bilingualism. For example, Walker et al. (2010) found that individuals with a brain injury who have a non-English educational background do not perform as well on English administered neuropsychological assessments compared to those educated in English. Indeed, they compared the performance of three groups of individuals with moderate-severe traumatic brain injury on the Wechsler Adult Intelligence Scale–Third edition (WAIS-III) and Wechsler Memory Scale–Third Edition (WMS-III). The three groups were: (a) people of English-speaking background (ESB), (b) people of a culturally and linguistically diverse background educated in English (CALD-EE), and (c) people of a culturally and linguistically diverse background with non-English and mostly non-Western education (CALD-NE). Overall, neuropsychological

performance following brain injury was found to be significantly related to cultural background. Specifically, the CALD-EE group scored lower (small to moderate effect sizes) than the ESB group on the WAIS-III Verbal Intelligence Quotient (VIQ), as well as the WAIS-III Vocabulary and Similarities subtests. In terms of the CALD-NE group, lower scores were obtained on almost half of the cognitive measures used compared to the ESB group with large effect sizes. The CALD-NE group also obtained significantly lower scores than the CALD-EE group on many of the WAIS-III subtests (Walker et al., 2010). These findings suggest that non-native English individuals, or those who have not received English or Western education, may be at a disadvantage when completing English cognitive assessments. It is important to highlight, however, that while language use (English) may play a significant role in the observed differences, cultural factors may also contribute to these disparities. Nevertheless, this underscores the importance of considering both language and cultural factors when evaluating cognitive function in bilinguals.

Kisser et al. (2012) also reported performance differences on English-administered neuropsychological tests between native and non-native English-speaking undergraduate students (*M* age = 21.35 years). English speakers were found to perform significantly better on many English language mediated tasks compared to non-native English speakers. Specifically, non-native English speakers performed significantly worse on various language-mediated tasks, such as letter fluency, category fluency, the cognitive estimation test and Trails A. However, the two groups performed similarly on tests of executive functions, verbal memory and visuospatial abilities. Whilst consistent with prior findings by Boone et al. (2007) who also found that non-native English speakers performed worse on tests of language and attention, but not tests of memory and executive functioning, the pattern of results therefore partly contradict those of Walker at al. (2010) above. However, as Walker et al. (2010) focused on a clinical sample, it is possible that other factors, alongside non-native

language status, may have influenced their results. Further, and despite some differences in specific findings across studies, the overall findings of Kisser et al. (2012) nevertheless highlight how non-native language status may impact neuropsychological performance even in individuals who would be expected to perform at or near ceiling.

Additionally, Kowoll et al. (2015) compared the performance of bilingual patients (14 different native languages and 11 different second languages) with AD or Mild Cognitive Impairment (MCI) and assessed them in both their dominant and non-dominant languages on the verbal fluency task and the Boston Naming Test. The bilingual sample consisted of 41 lifelong bilinguals and 33 were immigrants to Germany. Bilingual patients with MCI performed worse in their dominant language than bilingual controls, but bilingual AD patients exhibited poorer performance when assessed in their non-dominant language when compared to both bilingual MCI patients and bilingual controls (Kowoll et al., 2015). Therefore, without paying much attention to the type of bilingualism, there seems to be evidence for differential performance on cognitive assessments depending on the language of the assessment, as well as an individual's diagnosis and stage of disease. Overall, therefore, such findings further highlight the importance of considering bilingualism and language status when conducting neuropsychological assessments, thus supporting the development of better care for bilingual patients (Kowoll et al., 2015).

Additionally, bilinguals have also been found to make more cross-language intrusion errors (e.g., accidentally using a Welsh word during an English conversation/task) on verbal fluency tasks when speaking the non-dominant language, but almost no intrusion errors when speaking the dominant language (Sandoval et al., 2010). Given that verbal fluency tasks are timed, bilingual individuals may produce less words when asked to do so in their non-dominant language due to intrusion errors from the dominant language. However, this may

differ depending on the type of bilingualism. As a result, the ability of the individual may be underestimated, and performance may not be representative of their ability.

Investigating this further, Roselli et al. (2000) explored the impact of bilingualism on verbal fluency and repetition tests in a sample of older Hispanic bilinguals. All English monolingual participants were born in the United States and only spoke English, and all Spanish monolingual participants were Latin American immigrants living in Florida. With regards to the bilinguals, one participant was born in the US, six in Puerto Rico, seven in Central America or the Caribbean, and five in South America. All bilingual participants reported speaking Spanish as their first language. Equivalent performance was found between bilingual and monolingual individuals on all tasks except for semantic verbal fluency. However, bilingual individuals who learnt English before 12 years of age performed significantly better on the English repetition test and produced significantly more words to describe a picture compared to the bilinguals who learnt English after the age of 12 (Rosselli et al., 2000). This suggests that early exposure to a language may enhance proficiency, and then potentially confer additional advantages. In contrast though, Soltani et al. (2019) found that a group of 12 Persian monolinguals outperformed a group of 12 Arabic-Persian bilinguals on a verbal fluency task when the bilinguals were assessed in their native language. However, there was no significant difference between the performance of the groups when the bilinguals were tested in their second and most frequently used language (Soltani et al., 2019). These findings suggest that bilinguals may be disadvantaged but this may, at least in part, be dependent on factors such as AoA and whether they use the language on a regular basis.

In addition to verbal fluency and intrusion errors, research has also shown that in an unexpected test of free recall, bilingual participants remembered more emotional than neutral words but only for words presented in their native language (Anooshian & Hertel, 1994;

Ferré et al., 2019). Gollan et al. (2007) also demonstrated that older Spanish-English bilinguals had higher naming scores on the Boston Naming Task in their dominant language compared to their less dominant language. Therefore, such evidence suggests that bilingual individuals may recall fewer words if the assessment is conducted in their non-native language.

In sum, the literature so far indicates that the impact of bilingualism on cognitive test performance in non-native language appears to be inconsistent across cognitive domains. However, as the type of bilingualism examined differed across the studies discussed above (e.g., unbalanced, non-native bilinguals), it is important to recognise that the impact of bilingualism may be different again in the context of Wales, given that Welsh speakers are often 'balanced bilinguals.' For example, Clare et al. (2016b) found no significant differences in performance on the attentionally demanding Stroop Task depending on whether the task was administered in Welsh or English in a group of healthy bilinguals. In effect, such findings may be consistent with the fact that the nature of Welsh-English bilingualism is different to that of other bilingual groups, as mentioned previously, due to the impact of the dominant local language creating an environment in which both minority and dominant languages are heard simultaneously and consistently. Although this situation may not seem different to that of some other bilingual communities, such as the Pakistani community in Leicester (UK) for example, the difference may lie in the scale of the environment in which both languages are heard. However, a study by Yeung et al. (2014) found that older adults who spoke English as a second language performed worse on an English test of global cognitive function (modified Mini-Mental State Examination) compared to participants who spoke English as a first language (Yeung et al., 2014). It is worth noting though that drawing definitive conclusions in relation to the reasons behind such findings is challenging, because the impact of other contributory factors (e.g., mood) on cognitive test performance is often overlooked. As

previously mentioned, undergoing a neuropsychological assessment can already invoke stress/anxiety in patients (Dorenkamp & Vik, 2018) but this is likely to increase if the assessment is not conducted in their first language or language of preference, and moreover such mood related effects may even become amplified under the conditions of early onset dementia.

In sum, although the studies discussed above do not exclusively focus on balanced bilinguals, when considered collectively, they suggest that bilingualism may have varying effects on cognitive performance depending on the language of the assessment. Considering that language impairment plays a major role in the progression of AD and that bilingualism is so complex in nature, there is an urgent need for cognitive assessments that are sensitive to the cognitive experiences of bilinguals (Bluvshtein, 2004). It is crucial that individuals have access to cognitive assessments in their native language so that their cognitive abilities are not misjudged, meaning that the necessary interventions are put in place early and that any unnecessary further investigation (due to misleading test scores) does not occur. That is, appropriate cognitive assessments would ensure that appropriate treatment and care is provided based on actual need. In the context of Wales, this is even more pressing given that accessing health services in Welsh is a legal requirement that is not currently being met.

Having discussed the need for cognitive assessments in the individual's language of preference and the possible differential performance across languages, this next section will focus on the second aim of this study. As covered in Chapter 1, research suggests that bilingualism may contribute to CR and delay the onset of pathological cognitive decline (Bialystok, 2021; Gold et al., 2013; Schweizer et al., 2012). As covered in detail previously, some studies argue that this may be because bilinguals are required to manage two languages, meaning that general cognitive skills (e.g., attentional control, conflict resolution, memory, inhibition) may be more sophisticated or efficient in the bilingual brain (Luk, 2015;

Macnamara & Conway, 2014). Consequently, and as presented in Chapter 1, studies have compared bilinguals' and monolinguals' performance on cognitive tasks that tap widespread cognitive processes, finding mixed evidence for a bilingual advantage (Bialystok et al., 2004; Costa et al., 2008; Duñabeitia et al., 2014; Gathercole et al., 2014; Paap and Greenberg, 2013). However, few studies have considered Welsh-English bilinguals in this context.

2.1.2 Study Aims

Considering the above, the first aim of the study was to collect normative data for two major cognitive assessment batteries for the purposes of testing Welsh speaking (bilingual) participants. Thus, the study investigated the impact of Welsh-English bilingualism on cognitive test performance, and whether native Welsh speakers perform differently when assessed in English and in Welsh on the Addenbrooke's Cognitive Examination (Third Edition) (ACE-III) and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) (Form A). Specifically, it sought to determine whether those who speak Welsh as a first language would perform better on these assessments when administered in their native language (Welsh). The study also sought to explore whether there would be a difference in performance across Indexes/Domains in the Welsh and English versions. That is, the study aimed to determine whether there are any differences at a global level but also at specific Index/Domain levels. The second aim of the study was to compare the performance of the Welsh-English bilinguals on the English cognitive assessment batteries with that of normative data, to provide insight into whether this type of bilingualism may lead to an advantage across general cognitive domains (e.g., language, attention, memory, fluency, visuospatial abilities) and potentially in CR.

To the author's knowledge, this is the first study that has: (a) collected normative data for routinely used cognitive assessment batteries for Welsh speakers, therefore responding to the

desires of service users, (b) studied the impact of a uncommon form of bilingualism (where both languages [Welsh/English] are used daily) across a range of broad cognitive domains (e.g., Memory, Visuospatial, Attention), and (c) considered whether performance differs on such cognitive assessment batteries across bilingual versus monolingual populations.

2.2 Methods

2.2.1 Ethical Approval Procedure

The original aim of the study was to collect data from a sample of healthy older adults (control participants), as well as a sample of individuals living with dementia (clinical sample). Therefore, favourable ethical approval was obtained from both the Integrated Research Application System (NHS – clinical sample) and the Department of Psychology at Swansea University (control participants). However, due to the COVID-19 pandemic, amendments were necessary as in-person data collection became unfeasible, and obtaining a clinical sample was no longer viable. Consequently, an ethics amendment reflecting a revised protocol for the control sample was approved by the Departmental Ethics Board (see Appendix AA). Further information concerning pre- and post-COVID-19 protocols are covered under the 'Procedure' section below, including coverage of how participants were supplied with a detailed information sheet and provided written informed consent before taking part in the research.

2.2.2 Participants

The study was advertised as 'Validating clinical neuropsychological assessments for Welsh speakers.' Participants were drawn from the general adult population and were recruited from social media adverts, by contacting various organisations/groups (e.g., Lunch Club for older adults; Merched y Wawr), via the researcher's own social and personal networks, by word of mouth (e.g., via participants), and via media presentations and interviews (e.g., radio, TV)

about the project. Participation was on an entirely voluntary basis, with no reimbursement or compensation arrangements in place.

Individuals were eligible to participate if they were bilingual (Welsh – English) speakers, with Welsh being their *first language*. No monolingual participants were recruited nor were they eligible to participate in this study. Participants also had to be between 55-89 years of age, and report no history of brain injury (e.g., traumatic brain injury, stroke), mood disorders (e.g., depression, anxiety) or substance abuse.

2.2.3 Data Exclusions

The following a priori post-data collection exclusion criteria were applied before conducting any data analysis: Participants were excluded if they:

- Scored within the 'abnormal' range on either or both cognitive assessments. As the intention was to collect normative data from a healthy general population sample (i.e., without cognitive decline or impairment), participants who scored in the 'abnormal' range on either or both cognitive assessments were excluded. For the RBANS, a neuropsychological screening battery (see 'Materials'), an abnormal score was defined as the overall index score falling ≥2 standard deviations below the mean (Duff et al., 2010). For the ACE cognitive assessment battery (see 'Materials), a total raw score <82 was used as the threshold to indicate an abnormal score (Giebel & Challis, 2017). When scores were observed in the 'abnormal' range, the participants GP was notified accordingly. Participants consented to this GP contact as part of the information and consent procedures for the research.
- Chose to withdraw part way through the study, leading to partial/incomplete data.
- Could not complete the study in accordance with the established protocol (e.g., completing the second session 2-3 months after the first; being unable to complete the second session under the same conditions or in the same 'mode' as the first) owing to

the COVID-19 pandemic. Please note that this will be explained further under 'Procedure'.

- Obtained an RBANS effort scale index score (a measure used to evaluate the degree of effort or motivation exerted by an individual) in Session 1 or 2 indicative of suboptimal effort. Such participants were excluded because interpretation of neuropsychological test performance is only valid when appropriate effort has been made (McGuire et al., 2019). An Effort Index score of >3 is indicative of suboptimal effort on the RBANS (Silverberg et al., 2007). More information about the RBANS Effort Index is described in the 'Materials' section.

In total, 109 participants took part in the study. Using the above data exclusions, 16 participants were subsequently excluded for the following reasons: 'abnormal' RBANS and/or ACE-III score (n = 2); withdrew from the study (n = 6); were unable to complete the study in line with the protocol (n = 6), and an RBANS effort scale index score indicating suboptimal effort (n = 2). After these exclusions, the sample consisted of 93 participants. Participants' ages ranged from 55 to 87 (M = 67.10, SD = 7.00), and 49 participants (52.7%) were female. Further demographic information is presented in Table 2.

Table 2

Summary of Demographic Characteristics

Demographic Characteristics	**Sample (n = 93)**	
	n	**%**
Gender		
Females	49	52.7
Males	44	47.3
Employment Status		
Retired	64	68.9
Semi-retired	3	3.2
Employed	19	20.4
Self-employed	7	7.5
Education Level		
Primary School	2	2.2
Secondary School	15	16.1
Undergraduate degree	44	47.3
Postgraduate degree	31	33.3
Unknown	1	1.1
Level of education in Welsh		
Did not attend formal education in Welsh	2	2.2
Primary school	12	12.9
Secondary school	37	39.8
Further education (college, diploma, cert)	7	7.5
Undergraduate degree	26	27.9
Postgraduate degree	9	9.7
Language currently used with family		
Welsh	87	93.5
English	1	1.1
Both	5	5.4
Language of majority of TV programmes watched		
Welsh	24	25.8
English	54	58.1
50/50	14	15.0
No TV	1	1.1
Language most confident in using		
Welsh	57	61.3
English	8	8.6
No preference	27	29.0
Unknown	1	1.1

Additionally, and confirming their bilingual status, participants scored higher on the Welsh version of the Lexical Test for Advanced Learners of English **(**LexTALE) compared to normative samples from another study (see Table 3). The LexTALE and the normative sample is described further under 'Materials'.

Table 3

Performance Comparison on the Welsh LexTALE across Current and Normative Data

	Current Study	**Welsh Normative Data**	
N	93	44	22
Age (years)	55-87	40-69	70-87
M Raw Score	48.30	-	-
SD Raw Score	6.08	-	-
M Accuracy (%)	81	65.95	69.32
SD Accuracy (%)	12.58	12.21	9.93

Note: Normative data provided by researcher (Kyle Jones) at Swansea University from unpublished thesis. Please note that the raw scores were not provided for the normative data set.

Note: That the English LexTALE, M Accuracy for the current study was 96.43% and the SD Accuracy was 4.37%.

2.2.4 Materials

All participants were asked to complete the below materials. Please note that the interval between tasks as well as the format of testing is explained in the design section and pre- and post-COVID-19 procedure sections below.

1) Demographic and Background Details Questionnaire
2) Language Proficiency Questionnaire
3) Lexical Test for Advanced Learners of English (LexTALE). This was completed in both Welsh and English.
4) Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) – Form A. Both Welsh and English versions were completed.

5) Addenbrooke's Cognitive Examination-III (ACE-III). Both Welsh and English versions were completed.
6) An additional Verbal Fluency Task, completed in both Welsh and English.

Further details about each are presented below.

2.2.4.1 Demographic and Background Details

Participants were asked to supply standard demographic details (e.g., age, gender, employment status). Participants were also asked about handedness, social/leisure activities, and a range of questions about their parents'/guardians', such as their level of education and occupation. These latter questions were included to enable further characterisation of the sample, as well as giving scope to potentially form the basis of an estimate of SES. However, whilst such variables can be used as single or combined proxy indicators of SES (Sasaki et al., 2021; Chen et al., 2018), their use varies substantially across research and is not without challenges. Further, the response categories used in the current study for the applicable proxy variables, were subsequently judged to lack specificity and range. As a result, these variables were not used as a proxy measure of SES and where applicable, were instead considered as individual items or were not considered further in the analysis (Appendix A).

2.2.4.2 Language Proficiency Questionnaire

Participants were asked to complete a questionnaire about their Welsh/English language proficiency and were also probed about any additional language ability. To begin, participants were asked about their use of languages, such as the language used with parents/guardians/siblings and the language in which they received primary and secondary school education. They then indicated their AoA of both Welsh and English and rated their level of proficiency (e.g., speech, listening, reading, writing) on a 6-point scale (1 = not at all

to 6 = fluent). They then proceeded to answer a set of questions concerning their use of the Welsh language specifically, such as how frequently they used the language, the level of education they received in Welsh, and in which language they felt most confident. The final set of questions focused on any additional languages used beyond Welsh and English. For each additional language declared, participants were asked about their AoA, frequency of usage, and perceived proficiency (e.g., speaking, listening, reading and writing). See Appendix B for a copy of the questionnaire.

2.2.4.3 Lexical Test for Advanced Learners of English (LexTALE)

The LexTALE is a test of vocabulary knowledge for medium to highly proficient speakers of English as a second language (Lemhöfer & Broersma, 2012). Consisting of 60 words/trials (40 real words, 20 non-words – each between 4 and 12 letters long), participants are asked to make a lexical decision on each item (word or non-word?) by pressing specific keys on the keyboard ('S' if the word is a real word and 'K' if the word is a non-word) or by pressing certain buttons if using an iPad/tablet. A fixation cross appears for 1000ms following each word/trial, and no feedback is provided to participants between trials. See Table 4 for examples of real and non-words. The task takes approximately 5 minutes to complete, and participants are informed that speed and accuracy will be measured. The LexTALE is scored by calculating the mean accuracy of the scores across all participants. Previous research has found the English version of the LexTALE to be a good predictor of English vocabulary knowledge and is also considered superior to self-rated measures of proficiency (Lemhöfer & Broersma, 2012). The LexTALE has also been used in many studies as an objective measure of proficiency and translated into other languages (e.g., Spanish, Dutch, German, and French) (Brysbaert, 2013; Izura et al., 2014; Lemhöfer & Broersma, 2012).

Table 4

Examples of Welsh and English Words and Non-Words Used in the LexTALE Task

Real Welsh words	Welsh non-words	Real English words	English non-words
Anferthedd	Lleliad	Moonlit	Proom
Penrhydd	Elithio	Breeding	Interfate
Mewnblyg	Blysog	Fluid	Plaudate
Gwrthymosod	Llungapwr	Cleanliness	Mensible

In addition, participants completed a Welsh version of the same task which was developed by a researcher at Swansea University (Kyle Jones', unpublished thesis). The structure of the Welsh version was entirely consistent with that of the English version. During its development, consideration was given to orthographic similarity and frequency of words across both languages to replicate the complexity of the English version.

2.2.4.4 Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) – Form A

The English version of the RBANS was developed to identify and characterise cognitive decline in older adults and as a neuropsychological screening battery for younger individuals (Randolph et al., 1998). The battery takes approximately 30 minutes to administer and contains 12 subtests which are used to create five index scores: immediate memory, visuospatial/construction, language, attention, and delayed memory (Randolph et al., 1998; Carlozzi et al., 2008). Each of the five domains include at least two subtests (Beatty et al., 2003) and a further description is provided in Table 5 (Loughan et al., 2019). The five domains contribute to a total score of cognitive functioning, with higher scores indicating superior cognitive abilities (Houston et al., 2019). Parallel versions of the assessment have been developed for the purposes of repeat testing, with Forms A and B adapted to be

appropriate for use in the United Kingdom (Brown et al., 2019). There were no practice items included in the RBANS and the assessment was administered according to the administration guidelines.

Table 5

Description of the RBANS Indexes

RBANS Index	Description	Subtests
Immediate Memory	The participant's ability to recall information immediately after it is provided to them.	List Learning Story Memory
Visuospatial/Construction	The participant's ability to recognise spatial relations.	Figure Copy Line Orientation
Language	The participant's ability to respond verbally by naming stimuli or recalling pre-existing information.	Picture Naming Semantic Fluency
Attention	The participant's ability to remember and use both visual and auditory information stored in short-term memory.	Digit Span Coding
Delayed Memory	The participant's ability to recall information acquired earlier.	List Recall List Recognition Story Recall Figure Recall
Total Scale	A score made up of the 5 index scores.	Five indexes combined

The RBANS manual provides normative data for ten-year age bands from 20-89 (Duff et al., 2003), and Index scores have a mean of 100 and standard deviation of 15 (Brown et al., 2019). A score that is two or more standard deviations below the mean is thought to indicate impairment or an 'abnormal' performance (Duff et al., 2010). In addition, the RBANS also includes an embedded measure of effort, allowing evaluation of performance validity by

using scores from two of the RBANS' subtests: Digit Span and List Recognition. This measure is called the Effort Index (EI) which appears to be effective at detecting insufficient effort, with an EI of >3 indicative of suboptimal effort (Silverberg et al., 2007; Moore et al., 2013).

The RBANS has been used with a variety of patient populations due to its ability to reliably and rapidly assess a patient's neurocognitive functions (McKay et al., 2007), and has also been translated and/or validated in other languages, such as Greek (Tsatali et al., 2019), Russian (Bluvshtein et al., 2004), Spanish (Encinas et al., 2012), and Sinhala (Suraweera et al., 2016). Further, the RBANS has also been translated into Welsh (BAASN: Batri Amlroddadwy ar gyfer Asesiad o Statws Niwroseicolegol) by Siwan Roberts, a trainee Clinical Psychologist at Bangor - with this translated version used in the current study. At the time of translation (approx. 2010), it was recognised that simply translating the English version was insufficient, as word frequency and usage across languages needed to be considered. Consequently, some items were adapted in the Welsh version. For example, the word 'saddle' ('cyfrwy' in Welsh) in the list learning subtest was changed to 'feather' ('pluen' in Welsh), which was a word already included in Form B of the RBANS. The word 'saddle' ('cyfrwy) is rarely used in the Welsh language but 'pluen' ('feather') is heard and used more frequently. Additionally, the place named in the story memory subtest was translated to 'Porthcawl in Glamorgan' to be a more appropriate Welsh alternative. For the word recognition subtest, the word 'grassland' was also translated to 'glaswellt' (grass) rather than 'glaswelltir' ('grassland') because the latter is used less frequently. Following the adaptation of the BAASN, it was back-translated to English by a professional translator employed by the NHS. The final version of the BAASN was also reviewed by Professor Bob Woods, Dr. Eleri Jones and Siwan Roberts (researchers and/or clinicians) who deemed it appropriate for clinical use. Given the nature of Welsh-English bilingualism, they also recommended that

both Welsh and English responses are scored as correct - with this guidance adopted for the semantic fluency subtest (i.e., one minute to name as many fruits and vegetables) in the current study.

On creation, and as explained on the BAASN form (the Welsh version of the RBANS), the BAASN was administered to 100 people who attended a memory clinic for a dementia assessment and concurrent validity was assessed. Scores obtained in the BAASN were compared to scores obtained in the Addenbrooke's Cognitive Examination Revised (ACE-R), and no significant difference was found across the assessments. However, the BAASN was not subjected to comprehensive investigation in relation to validity and reliability, and no normative data was collected. Nevertheless, the initial efforts made to collect clinical data provided fairly good evidence of concurrent validity.

2.2.4.5 Addenbrooke's Cognitive Examination-III (ACE-III)

The original ACE was developed in the late 1990s (Noone, 2015) as a simple bedside cognitive assessment battery to detect mild dementia and to differentiate between AD and frontotemporal dementia (Pigliautile et al., 2011). Due to the diagnostic accuracy of the ACE, it is widely used (Noone, 2015), including the most recent version utilised here (the ACE-III) which replaced many items on the previous versions of the assessment (Noone, 2015).

In its most recent version, the ACE-III includes 5 cognitive domains (attention, memory, fluency, language and visuospatial) and is scored out of 100 (Velayudhan et al., 2014). It takes 15-20 minutes to administer and a score below 82/100 is typically used to indicate dementia (Beishon et al., 2019). There were no practice items contained within the ACE-III and it was administered in line with administration guidelines. See Table 6 for a description of each subtest within each domain (Bruno & Vignaga, 2019).

Table 6

A Description of each Subtest Within the ACE-III Domains

Cognitive Domain	Description	Subtests
Attention	The ability to select and focus on relevant information.	Date, seasons, current location, repetition of three words, serial subtraction.
Memory	The ability to remember information immediately and after a period of time.	Recall three words repeated previously, memorise and recall an address, and to recall general current and historical facts.
Fluency	Measures information retrieval from memory that requires self-monitoring, selective attention, and selective inhibition.	Name as many words beginning with a certain letter in one minute, and to name as many animals as possible in one minute.
Language	Abilities in reading, writing and speaking. This includes responding verbally to pictures and following verbal instructions.	Complete a set of physical commands using a pencil and a piece of paper, write two sentences that are grammatically correct, repeat several polysyllabic words and two proverbs, a picture naming task, answer questions on the objects in the picture naming task, and read aloud five words that are often mispronounced.
Visuospatial	Ability to recognise spatial relations. This includes being able to see a diagram as a set of parts and being able to re-create the diagram.	Copy two diagrams, draw a clock face with the hands set at a particular time, count sets of dots without pointing, and identify four fragmented letters.

The ACE-III was translated to Welsh by ‘Language Awareness Infrastructure Support’, which is embedded in the North Wales Organisation for Randomised Trials in Health, part of the Bangor Institute of Health and Medical Research (Welsh Language Health Measures, n.d.). Please note that this translation was completed independently and not specifically for this study and a comprehensive approach to development was adopted and involved the following steps: preparation, translation, back translation, review processes, cognitive testing, final review and report.

Whilst every effort was made by those who translated the assessment to keep the Welsh version as close as possible to the English original, several small adaptations were deemed necessary because of things like variations in letter category size across languages. For example, the three-word repetition and recall (‘lemon’, ‘key’, ‘ball)’ was adapted, with ‘key’ being replaced by ‘cloch’ (‘bell’). There are many different words for ‘key’ in Welsh, whereas there is only one direct translation for ‘bell’. Similarly, the Welsh version asks participants to produce as many words as possible beginning with the letter ‘B’ rather than ‘P’, where there are more everyday words beginning with ‘B’ in the Welsh language. The name and the address details included in the memory subtest were also changed in the Welsh version. Namely, the name ‘Harry Barnes’ was replaced with ‘Harri Puw’, and the house number and street name was directly translated to be in keeping with the English version. However, the name of the town and county were changed from ‘Kingsbridge, Devon’ to ‘Bontgoch, Gwynedd’. Two (‘caterpillar, ‘eccentricity’) of the 4 words requiring repetition in the language subtest were also changed to ‘doves’ and ‘individualistic’ (‘colomennod’ and ‘unigolyddol’, respectively). In the language subtest, the proverb ‘a stitch in time saves nine’ was also changed to ‘i’r pant y rhed y dŵr’. This directly translates to ‘the water runs into the hollow’ (meaning if someone has something, they are likely to get more of it). The list of 5 words in the language subtest were also changed, where ‘sew’, ‘pint’, ‘soot’, ‘dough’,

‘height’, were replaced with ‘cwyn’, ‘ynghyd’, ‘englyn’, ‘hwyr’, ‘pryfyn’ (‘complaint’, ‘together’, ‘englyn’, ‘late’, ‘insect’). Note that there is no English translation for the word ‘englyn’. It describes a short traditional Welsh poem with strict rules around rhyming and number of syllables.

Finally, it should be noted that the letter ‘K’ remains in the Welsh version of the letter identification task (visuospatial subtest) even though the letter ‘K’ is not included in the Welsh alphabet. No explicit explanation was provided by LLAIS to explain this decision. However, the letter ‘K’ may have been retained in the Welsh version because no other letter with similar physical properties exists in the Welsh alphabet.

2.2.4.6 Additional Verbal Fluency Task

Research suggests that bilinguals’ performance can vary across categories (Rosselli et al., 2002). For example, Acevedo et al. (2000) found that English participants scored higher than Spanish participants on the category ‘vegetable’, and others have additionally found that more exemplars are provided for the category ‘animals’ compared to ‘fruit and vegetables’ (Ardila et al., 1994). Therefore, performance on verbal fluency tasks in a particular language may depend on the category size available in that same language, thus potentially leading to interference from the other language (Rosselli et al., 2002). Owing to this, an additional category-based fluency task was included to explore if any subsequent language-based differences (if observed) across the verbal fluency tasks included in the cognitive assessment batteries (i.e., RBANS, ACE-III) reflected a real difference in performance, or simply limitations of the categories used in the assessments. Thus, the intention was to analyse this additional category-based verbal fluency task only *if* subsequent language differences were observed on the verbal fluency related Domains/Indexes from the RBANS and ACE-III. For the task, participants were asked to name as many household items as possible in one minute

in either Welsh or English, depending on the language of the session. There was a two-to-three-month interval between sessions. Further, and consistent with the administration and scoring guidance adopted for the BAASN (RBANS), Welsh responses in the English verbal fluency tasks were scored as correct. This was also considered appropriate due to the regularity and normality of language switching found in Welsh-English bilingualism (e.g., balanced dominance).

2.2.5 Design

This study used a within-subjects design. The independent variable (IV) was language (Welsh or English), and the dependent variable (DV) was the participants' performance on the RBANS and the ACE-III. Four counterbalancing orders were used in the study:

- Session 1 - English ACE-III / RBANS -> Session 2 - Welsh RBANS / ACE-III
- Session 1 - English RBANS / ACE-III -> Session 2 – Welsh ACE-III / RBANS
- Session 1 - Welsh ACE-III / RBANS -> Session 2 – English RBANS / ACE-III
- Session 1 - Welsh RBANS / ACE-III -> Session 2 – English ACE-III / RBANS

Participants completed the Welsh and English versions across two sessions, with each session corresponding to a single language. The duration of each session typically lasted between 60-70 minutes. If a participant was assigned to counterbalancing order 1, they completed the English ACE-III, followed by the English RBANS in Session 1. Then in session 2, they completed the Welsh RBANS followed by the Welsh ACE-III. Participants were not randomly allocated to one of the four counterbalancing orders or placed in one of the four counterbalancing conditions via sequential rotation (i.e., order 1, order 2, order 3, order 4). Rather, participants were allocated in blocks consisting of two counterbalancing orders which alternated. To explain, participants 1 to 50 (prior to initial exclusion criteria), were allocated to order 1 and 4 (alternated between each participant – i.e., 1, 4, 1, 4, 1, 4).

The next 41 participants were allocated to orders 2 and 3 (alternated between each participant – i.e., 2, 3, 2, 3, 2, 3). Then, a block of 9 participants were allocated to orders 4 and 1 (alternated between each participant – i.e., 4, 1, 4, 1, 4, 1), and the final 9 participants recruited were allocated to orders 2 and 3 (alternated between each participant – i.e., 2, 3, 2, 3, 2, 3).

To ensure an equal distribution of participants across each counterbalancing order, the original intention was to rotate around blocks of participants. However, as this approach is not resilient to changes due to unforeseen circumstances and as it also assumes that the recruitment target will be achievable, it subsequently transpired that an even distribution of participants across orders was not possible. Largely, the impact of COVID-19 (explained further below), coupled with the need to reduce the total sample size due to the time lost because of the pandemic, meant that comparably lower numbers of participants were allocated in blocks 3 and 4. After the initial exclusion criteria had been applied, the final sample size was spread across the four counterbalancing orders as follows: 36 participants were in block 1, 40 in block 2, 9 in block 3, and 8 in block 4. Please note that the participants did not significantly differ in terms of age, $F(3, 95) = 0.350$, $p = 0.789$, or gender, $F(3, 95) = 1.607$, $p = 0.193$, across the four counterbalancing orders.

Finally, note that repeated administration of test materials enables statistical analyses of: (a) comparative performance between the two major cognitive assessment batteries (and their Domain/Index scores) as a means of determining the impact of language format, both at a general and specific (domain level) and thus determine whether relative performance differs depending on the language medium used for presentation, and (b) comparisons between the Welsh bilingual sample and normative data on the English versions of both tests (normative data has already been obtained for this latter group).

2.2.6 Procedure

Prior to outlining the procedure, it is necessary to note that the COVID-19 pandemic happened approximately halfway through the planned participant recruitment cycle (e.g., February/March 2020 – UK), at which point data collection was taking place at participants' homes. Consequently, the study was temporarily suspended as visiting participants in their homes was no longer possible. At the same time, the longevity and impacts of the pandemic were rapidly unfolding and changing on a frequent basis (e.g., 'Stay at Home' order, two households permitted to meet outdoors, extended households to meet up indoors and stay overnight), meaning that was extremely challenging to establish a date to resume the project.

After approximately 11 months' suspension, the likelihood of the study resuming in its original format, as well as maintaining the overarching timeline of the thesis, was deemed unlikely. At that point, it was also recognised that even if restrictions were lifted, that the longer-term effects of the pandemic would pose further challenges (e.g., people remaining uncomfortable with close contact despite lockdown restrictions being lifted). Therefore, the decision was reached to explore alternative arrangements and procedural processes which would allow the study to continue in some form.

To begin, the possibility of conducting the study via a remote means was considered. At that time, video conferencing (e.g., Facetime, Skype, Zoom, Teams) had become the primary mode of communication in the workplace, as well as the main method of making social contact with loved ones. Therefore, it was apparent that appropriate platforms were available to deliver remote testing sessions with participants. However, participants' ability, familiarity, and willingness to use various video conferencing platforms, as well as their accessibility, also needed to be considered. For example, Zoom was identified as an open-source platform - available to everyone – participants can join a call without needing to subscribe. Additionally, the functions of each platform were taken into consideration (e.g., the ability to share screen

whilst simultaneously viewing the participants' video). Again, Zoom seemed the most appropriate option as it included all the functions required for online testing, but also appeared to be the most secure platform - it had a comprehensive General Data Protection Regulation policy and was an approved platform by the host institution.

After establishing that a suitable platform existed, the feasibility of conducting sessions online was then considered, such as ensuring that remote versions of the RBANS and ACE-III accurately represented the paper-pen versions of the assessments. That is, the stimuli from the assessments needed to be in digital format so that they could be included in a PowerPoint presentation and shared on a screen to participants. Given the COVID restrictions (i.e., lockdown) and lack of access to facilities, this involved having to capture an appropriate image of each stimulus using a tripod and an iPhone (i.e., equipment available to the researcher at the time). Whilst the digitisation of stimuli was straightforward for certain subtests (e.g., ACE-III picture-naming task), there were some occasions where this procedure was more complex. For example, each trial in the RBANS line orientation task appears on different pages of the stimulus booklet (i.e., one trial per page), meaning that an appropriate image of each trial had to be captured separately, and that the perspective, orientation and size of the stimuli had to be consistent across all trials and as close as possible to the pen-paper version of the assessment. Please see Figure 2 and 3 for examples of stimuli presented electronically for display to participants via PowerPoint.

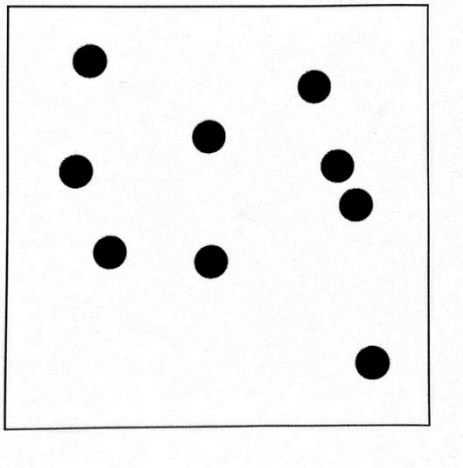
Further, and even after digitising all stimuli so that they could be shared with the participants, there were some subtests that could not be responded to verbally. Indeed, some tasks within the RBANS and ACE-III require participants to record a response on paper (e.g., copying a diagram). Therefore, a method was needed to ensure that it was possible for participants to complete tasks remotely, whilst also being able to provide hard copies of their responses. In addition to the participants' assessment responses, written consent was also

required from all participants. This led to practical considerations such as sending the participants plain paper, consent forms and a pencil via post before each testing session. This then led to the consideration of funding for the costs attached with this method (e.g., the purchasing of stamps, envelopes, paper, pencils etc.), with funding subsequently secured from the Coleg Cymraeg Cenedlaethol.

Figure 2. *First Trial of the RBANS Line Orientation*

[INFORMATION HAS BEEN REMOVED BY THE AUTHOR OF THIS THESIS FOR COPYRIGHT REASONS]

Figure 3. *Example of the Dot Counting Task in the ACE-III Visuospatial Domain. The ACE-III is available for free. The copyright is held by Professor John Hodges who is happy for the test to be used in clinical practice and research projects*



Further, ethics and related conduct codes also needed to be considered, such as consideration of how consent would be obtained and stored whilst awaiting the return of the paper copies in the post. The agreed method was to ask participants to sign consent forms (sent via post) whilst on Zoom and to hold the signed consent form to the camera so that a screenshot of the document could be obtained. Participants were then asked to return the consent form along with their written (paper) responses via post.

Next, and having determined that the adaptations were possible, feasible and practical, the final consideration was whether the revised format was appropriate. To inform such considerations, the guidelines released in May 2020 by the British Psychological Society on '*Psychological Assessment Undertaken Remotely*' were reviewed (The British Psychological Society, 2020). These guidelines recognised that conducting assessments under typical circumstances was not possible due to the pandemic and considered whether remote assessment was an appropriate alternative. The guidelines outlined a number of practical considerations such as privacy, the security of the technology used, the accessibility of the platform used, the response method for participants, and how to maintain confidentiality. Much of these practical considerations had already been considered (see above), but the guidelines additionally highlighted the need to consider the validity of test scores resulting from remote administration. In response, a brief literature search was conducted, and little evidence was found to suggest that scores on a range of assessments, including the RBANS, differed across in-person and remote testing (e.g., Galusha-Glasscock et al., 2016; Marra et al., 2020; Parks et al., 2021; Wadsworth et al., 2018). Thus, the decision was reached to pursue remote administration.

However, prior to resuming data collection under the new format, the researcher undertook the additional step of conducting an informal pilot with Dr Claire Williams (a member of the supervisory team), to test the mechanics of running the session on Zoom and to discuss possible technical difficulties that could arise during a testing session. This proved beneficial because effective ways of managing and resolving such situations with participants were then considered and implemented (e.g., ensuring that all participants were provided with another method of contacting the researcher). Additionally, the pilot led to the realisation that participants may need to be sent a booklet via post containing detailed instructions on how to access, download and use Zoom. It was also realised that the PowerPoint shared on screen

with the participants needed to include a ‘test’ slide to ensure that participants could see all corners of the screen before any stimuli were presented. A ‘sound check’ slide was also additionally added as a prompt to the researcher to check that there were no sound problems, and that the participant could hear the researcher before commencing any verbal instructions or formal testing.

Finally, at that time and before proceeding to formal recruitment, the implications of remote testing and the new procedure were also considered in relation to the study design and methodological arrangements. For instance, the change of delivery format ultimately meant that a new confound of ‘mode’ (online testing versus in-person testing) would be introduced into the design. It was also recognised how ‘mode’ would impact the counterbalancing procedure as well as the recruitment pool (i.e., some participants would be more/less willing/able to take part online compared to in-person). However, the COVID-19 pandemic persisted for longer than originally expected and with the research timeline in mind, the pragmatic decision was ultimately made to pivot the study to an online version.

2.2.6.1 Pre-COVID-19 Procedure

In Session 1, all participants read a participant information sheet and signed a consent form (see Appendix C and D for copies). Participants were then successively given the demographic questionnaire, language proficiency questionnaire, the Welsh OR English version of the LexTALE task, a Welsh OR English version of the RBANS, a Welsh OR English version of the ACE-III and Welsh OR English verbal fluency task.

In Session 2, which took place two to three months later, participants successively completed the Welsh OR English version of the LexTALE, Welsh OR English version of the RBANS (whichever was not completed in the previous session), a Welsh OR English version of the ACE-III and Welsh OR English verbal fluency task (whichever was not completed in

the previous session). A debrief form was given to all participants at the end of Session 2. The two to three months duration between testing sessions was determined based on previous literature. This demonstrated that a range of interval lengths are typically used ranging from a few days to a full year, all of which result in good test-retest reliability (Dong et al., 2013; Zhang et al., 2008; Duff et al., 2005). Both testing sessions took place at participants' homes, and with the exception of the LexTALE which was administered electronically via the researcher's HP Pavilion laptop (using the PsychoPy software), all questionnaires and assessments were completed on paper.

2.2.6.2 Post-COVID-19 Procedure

As mentioned above, due to the impact of the pandemic and the lockdown measures that were put in place in 2020 (mid-way through data collection), data collection methods were adjusted to be conducted remotely online. Participants were sent a copy of the participant information sheet via email, and an information pack via post that included an 'invitation' letter, instructions on how to download/use Zoom, and a further envelope containing the consent forms, plain paper and a pencil. The invitation letter stated that participants should refrain from opening the enclosed envelope until the Zoom session to avoid losing important documentation. Participants were sent a Gorilla link for the LexTALE task via email before each session which they completed online either on a laptop or tablet any time following receiving the link. Participants completed the consent form and answered the questionnaires on Zoom and the assessments' stimuli were digitised and presented to them in a PowerPoint presentation via the screen sharing function on Zoom. Participants recorded their responses on the provided paper for any subtests requiring a drawing or other form of written response. Participants then returned the information packs via post to the Psychology Department, Swansea University. A debrief form was emailed to all participants after Session 2.

2.3 Results

This study aimed to investigate whether there was a difference in performance across the Welsh and English versions of the RBANS and ACE-III.

2.3.1 Global Level Analysis – is there a Difference Between Overall Scores?

2.3.1.1 RBANS Findings

The descriptive statistics indicate that participants generally performed better on the second administration of the RBANS regardless of counterbalancing order. See Table 7 for the descriptive statistics of each order[1] and Table 8 for the descriptive statistics for the RBANS performance in Welsh and English.

A mixed ANOVA was undertaken to determine any performance difference between languages and to confirm whether the counterbalancing was successful (Appendix E). The within subjects factor was the language in which the assessment was conducted (Welsh, English), and the between subjects factor was the counterbalancing order (1-4; see 'Design', Section 2.2.5 for further details). The assumption of equal variance was upheld in this analysis. The effect of mode of presentation (i.e., online vs in-person) was not considered as a factor, as no participants completed the assessments in-person in two counterbalancing orders (see Table 11 for participant numbers in each mode and counterbalancing order).

The ANOVA showed a statistically significant main effect of language, $F(1, 89) = 4.831$, $p = 0.031$, $\eta^2 = 0.008$, such that scores were significantly higher when the RBANS was completed in English. The ANOVA also revealed a non-statistically significant main effect of the between subjects factor of counterbalancing order, $F(3, 89) = 1.496$, $p = 0.22$, $\eta^2 = 0.038$, such that the performance on the RBANS was not affected by counterbalancing order.

[1] Order corresponds to counterbalancing order; Mode corresponds to online or in-person; Language corresponds to Welsh or English

Table 7

RBANS Descriptive Statistics for each Order

Language of first session	**English**		**Welsh**	
First assessment completed	ACE-III	RBANS	ACE-III	RBANS
	Order 1	Order 2	Order 3	Order 4
n	22	24	24	23
RBANS Welsh *M*	112.37	117.27	107.63	104.47
RBANS Welsh *SD*	12.66	8.12	11.03	11.96
RBANS English *M*	108.14	112.95	115.56	113.35
RBANS English *SD*	11.71	8.84	13.16	13.42

Note: Participants completed both the ACE-III and the RBANS in each session, and in different orders and languages e.g., a participant who completed RBANS first in Welsh also completed the ACE-III in Welsh during their first session, but then competed the ACE-III in English followed by the RBANS in English in their second session.
The counterbalancing orders are: 1 = English ACE-III/RBANS; 2 = English RBANS/ACE-III; 3 = Welsh ACE-III/RBANS; 4 = Welsh RBANS/ACE-III. To be clear, participants in order 1 completed the English assessments in the first session. The ACE-III was completed first, followed by the RBANS. In their second session, they completed the assessments in Welsh. They completed the Welsh RBANS first, followed by the Welsh ACE-III.
To facilitate the reader's understanding of the table, a detailed example is included here: - The second column on the left (i.e., ACE-III, Order 1) shows the English and Welsh RBANS M and SD for participants who completed the assessments in the order of English ACE-III-RBANS in their first session (i.e., 108.14). This means that this is the first time they completed the RBANS. Their RBANS score in Welsh is 112.37, as this was their second session, as denoted by their counterbalancing order.

Table 8

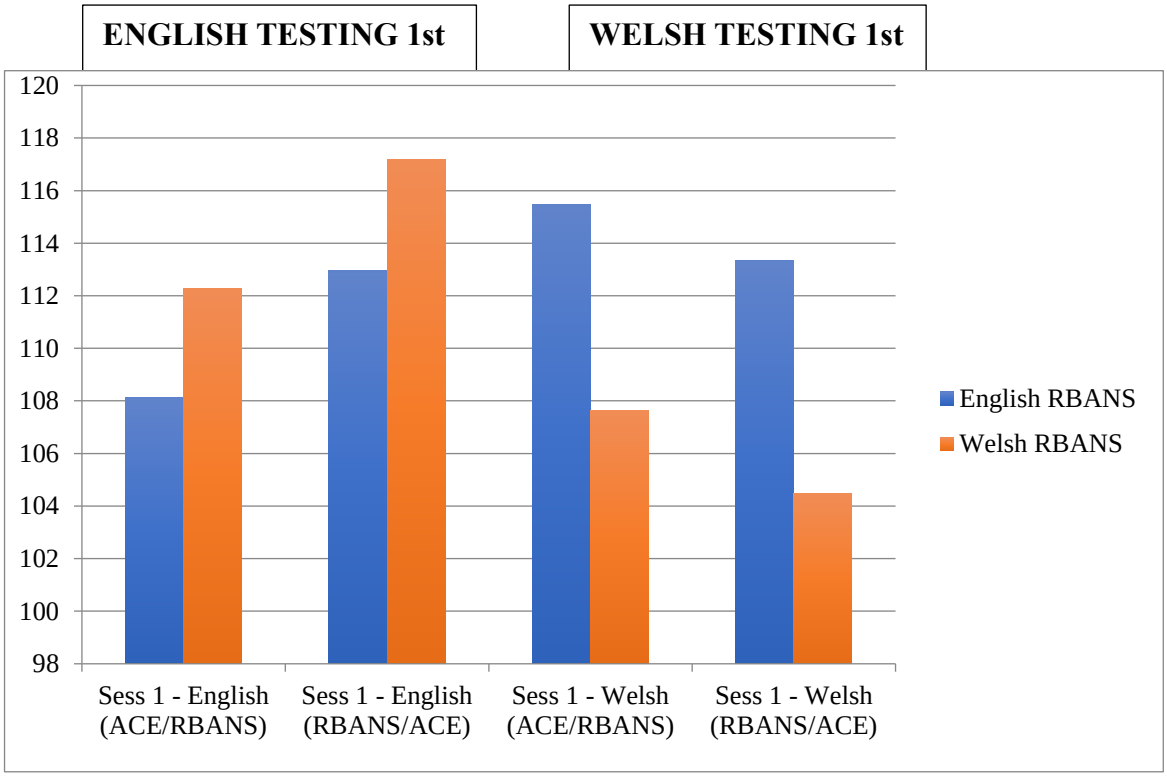
RBANS Descriptive Statistics Regardless of Order

	Welsh RBANS	**English RBANS**
M	110.41	112.56
SD	11.89	11.98

A statistically significant interaction between language and order was also found, $F(3, 89) = 14.508$, $p < .001$, $\eta^2 = 0.069$. This interaction is shown in Figure 4 below. Numerically,

participants scored higher in the second session compared to their first, regardless of language. It is likely that this pattern reflects a natural practice effect, with participants performing better the second time they completed the assessment. This is precisely why a counterbalancing approach was adopted as part of the study design (i.e., that across all participants, equivalent numbers had tests in the first or second language first).

Figure 4. *Bar Graph of RBANS Mean Scores in each Counterbalancing Order*



Post-hoc pairwise comparisons (with Bonferroni correction applied) indicated that participants who completed RBANS in Welsh in their first session (orders 3 [Welsh ACE-III/RBANS] and 4 [Welsh RBANS/ACE-III]) scored significantly higher when doing the assessments in English in the second session (order 3 p = .002; order 4 p <.001), but those who completed the English version in their first session did not improve by a significant

amount in their second session (order 1 [English ACE-III/RBANS] $p = 1.000$; order 2 [English RBANS/ACE-III] $p = 0.753$).

Participants in order 2 (English RBANS/ACE-III) scored significantly higher in Welsh (second session) than order 4 (Welsh RBANS/ACE-III) in Welsh during the first session ($p = .007$). Participants in order 3 (Welsh ACE-III/RBANS) scored higher in English by a significant amount (second session) ($p = 0.002$) compared to those in order 4 (Welsh RBANS/ACE-III) who completed the Welsh RBANS first ($p < .001$). This suggests that participants performed better in their second attempt on the RBANS. No other pairwise comparison reached statistical significance (Appendix F).

In sum, the interaction reflects that in general, participants scored higher in their second attempt at the RBANS irrespective of the language in which it was administered, though the practice effect was greater for those who completed the Welsh assessments first.

The next part of the analysis examined whether there was a similar pattern of practice effects observable in scores from the ACE-III.

2.3.1.2 ACE-III Findings

It is evident from the descriptive statistics that participants generally performed similarly or better on the second administration of the ACE-III regardless of order. See Table 9 for the descriptive statistics for the ACE-III performance in Welsh and English, and Table 10 for the descriptive statistics of each order.

Table 9

ACE-III Descriptive Statistics Regardless of Order

	Welsh ACE-III	**English ACE-III**
M	92.27	93.29
SD	4.33	4.50

Table 10

ACE-III Descriptive Statistics for each Order

Language of first session	**English**		**Welsh**	
First assessment completed	ACE-III Order 1	RBANS Order 2	ACE-III Order 3	RBANS Order 4
n	22	24	24	23
Welsh ACE-III *M*	92.64	94.75	91.95	89.65
Welsh ACE-III *SD*	4.76	2.44	3.84	4.61
English ACE-III *M*	91.27	94.71	93.95	92.91
English ACE-III *SD*	5.30	3.61	3.93	4.62

Note: Participants completed both the ACE-III and the RBANS in each session, and in different orders and languages e.g., a participant who completed ACE-III first in Welsh also completed the RBANS in Welsh during their first session, but then competed the RBANS in English followed by the ACE-III in English in their second session.
The counterbalancing orders are: 1 = English ACE-III/RBANS; 2 = English RBANS/ACE-III; 3 = Welsh ACE-III/RBANS; 4 = Welsh RBANS/ACE-III. To be clear, participants in order 1 completed the English assessments in the first session. The ACE-III was completed first, followed by the RBANS. In their second session, they completed the assessments in Welsh. They completed the Welsh RBANS first, followed by the Welsh ACE-III.
To facilitate the reader's understanding of the table, a detailed example is included here: - The second column on the left (i.e., ACE-III, Order 1) shows the English and Welsh ACE-III M and SD for participants who completed the assessments in the order of English ACE-III-RBANS in their first session (i.e., 91.27). This means that this is the first time they completed the ACE-III. Their ACE-III score in Welsh is 92.64, as this was their second session, as denoted by their counterbalancing order.

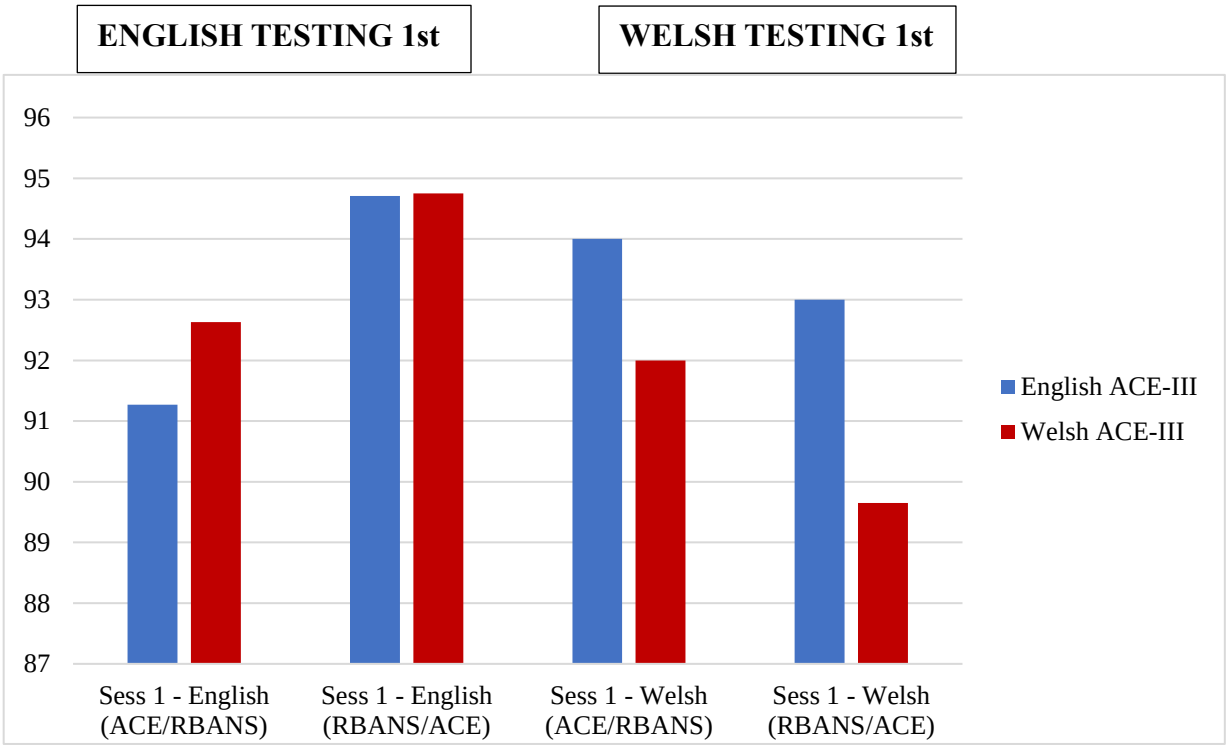
As with the RBANS, a mixed ANOVA was conducted to determine any performance differences between languages and to investigate the effectiveness of counterbalancing order

(Appendix G). The within-subjects factor was the language (Welsh, English) in which the assessment was conducted, and the between-subjects factor was the counterbalancing order. The assumption of equal variance was violated in this analysis for the Welsh ACE-III total scores. This means that *p* values will be increased if large groups are associated with large variances, making it more difficult to find a significant result, and *p* values will be decreased if small groups are associated with large variances, making it easier to find a significant result. However, the violation of the equality of variances assumption only poses a challenge if the sample size in the groups (i.e., counterbalancing orders) are not equal which is not the case in this study. Research suggests that if the largest group size is less than one and a half times the smallest group size, the ANOVA is sufficiently robust for the violation not to have a significant impact on the analysis (Blanca et al., 2018).

The ANOVA showed a statistically significant main effect of language, $F(1, 89) = 7.541$, $p = 0.007$, $\eta^2 = 0.012$, such that scores were significantly higher on the English version of the assessment. A statistically significant main effect of order, $F(3, 89) = 3.585$, $p = 0.017$, $\eta^2 = 0.087$, was also found. Post-hoc pairwise comparisons (with Bonferroni correction applied) showed that participants in order 4 (Welsh RBANS/ACE-III) performed significantly poorer than all other orders *(*$p = 0.017$).

A statistically significant interaction between language and order was also found, $F(3, 89) = 8.462$, $p < .001$, $\eta^2 = 0.040$. This interaction is shown in Figure 5 below.

Figure 5. *Bar Graph of ACE-III Mean Scores in each Counterbalancing Order*



Post-hoc pairwise comparisons (with Bonferroni correction applied) indicated that one of the four counterbalancing conditions (first assessment = RBANS, first language = Welsh) had very low scores on the Welsh ACE-III compared to other orders. Scores on the ACE-III were significantly lower for participants who completed the Welsh RBANS first followed by the Welsh ACE-III (order 4), compared to scores on the English ACE-III for those who completed the English RBANS followed by the English ACE-III ($p = 0.002$). In addition, scores on the ACE-III were significantly lower for the same group (order 4 - participants who completed the Welsh RBANS first followed by the Welsh ACE-III), compared to scores on the English ACE-III for those who completed the Welsh ACE-III first followed by the Welsh RBANS ($p = 0.017$). Scores were also significantly lower for the same group of participants (order 4 - participants who completed the Welsh RBANS first followed by the Welsh ACE-III) on the Welsh ACE-III compared to scores on the English ACE-III ($p < .001$). It is important to note that all other pairwise comparisons did not significantly differ (Appendix

H). In sum, and as shown in Figure 5 above, it seems that when participants completed the ACE-III for the first time in Welsh, immediately after the Welsh RBANS, their scores on the ACE-III were significantly lower.

It was then considered that mode of presentation (in-person vs online) may have been impacting on performance and offer an explanation for that finding. As mentioned in the methods section, due to the COVID-19 pandemic the format of testing changed from in-person to online administration midway through participant recruitment. This impacted on the method of counterbalancing adopted at the beginning of the study (pre-COVID-19), whereby participants who completed the sessions either in-person or online were not evenly distributed across the different counterbalancing conditions. 36 completed the sessions in-person and 57 participated online. To be clear, the six participants that completed the first session in-person and the second online, were removed from the study during pre-screening (see Section 2.2.3). Therefore, no participants carried forward to data analysis completed the first session in-person and the second online (i.e., mixed 'mode'). See Table 11 for number of participants in each order and mode.

Table 11

Number of Participants in each Order and Mode

Mode	**English ACE-III RBANS**	**English RBANS ACE-III**	**Welsh ACE-III RBANS**	**Welsh RBANS ACE-III**
	Order 1	Order 2	Order 3	Order 4
n In-person	18	0	0	18
n Online	4	24	24	5

Note: this table demonstrates that 36 participants completed the study in-person (prior to COVID-19), and 57 completed it online (during/post-COVID-19).

On review of Table 11, it is evident that there was a similar number of participants who completed the assessments online and in-person in two counterbalancing orders (n = 18, 4; n = 18, 5).

2.3.2 Interim Summary: The Removal of One Counterbalancing Order

To ensure clarity of the information presented thus far, participants who completed the Welsh RBANS followed by the Welsh ACE-III in their first session (i.e., Order 4), exhibited poorer performance on the Welsh ACE-III compared to participants in other counterbalancing orders. This difference in performance was consistent across various metrics: they performed worse than those who completed the assessments in the same order (i.e., RBANS followed by ACE-III) but in English, and compared to participants who completed the English ACE-III in their second session (i.e., those who completed [practiced] the assessment in Welsh in their first session). Additionally, participants in Order 4 performed worse on the Welsh ACE-III compared to the English ACE-III. Notably, their poorer performance could not be attributed to mode of presentation, and they also shared similar demographic characteristics with participants in the other three counterbalancing orders. Consequently, a pragmatic decision was reached to subsequently exclude participants from 'Order 4' (n=23) from the individual level analysis below. Further discussion of the rationale behind this decision, including its merits and potential implications, will be provided within the 'Discussion' section.

Thus, the final data set for the analysis below comprises 70 participants. Participants' ages ranged from 55 to 84 (M = 67.00, SD = 6.51) and 38 participants (54%) were female. LexTALE outcomes and further demographic information are captured across Tables 12 and 13.

Table 12

Welsh and English LexTALE Performance of the Final Sample and a Comparison on the Welsh LexTALE across Current and Normative Data

	Current Study	**Welsh Normative Data**	
n	70	44	22
Age (years)	67	40-69	70-87
M Welsh Raw Score	49	-	-
SD Welsh Raw Score	5.29	-	-
M Welsh Accuracy (%)	82	65.95	69.32
SD Welsh Accuracy (%)	10.79	12.21	9.93
M English Raw Score	58.01	-	-
SD English Raw Score	2.46	-	-
M English Accuracy (%)	97	-	-
SD English Accuracy (%)	2.53	-	-

Note: Normative data provided by researcher (Kyle Jones) at Swansea University from unpublished thesis. Please note that the raw scores and the English scores were not provided for the normative data set.

Table 13

Summary of Demographic Characteristics

Demographic Characteristics	**Sample (n = 70)**	
	n	**%**
Gender		
Females	38	54.3
Males	32	45.7
Employment Status		
Retired	52	74.3
Semi-retired	2	2.9
Employed	11	15.7
Self-employed	5	7.1
Education Level		
Secondary School	8	11.4
Undergraduate degree	35	50.0
Postgraduate degree	26	37.1
Unknown	1	1.5
Level of education in Welsh		
Did not attend formal education in Welsh	1	1.4
Primary school	5	7.1
Secondary school	30	43.0
Further education (college, diploma, cert)	5	7.1
Undergraduate degree	22	31.4
Postgraduate degree	7	10.0
Language currently used with family		
Welsh	64	91.5
English	1	1.4
Both	5	7.1
Language of majority of TV programmes watched		
Welsh	17	24.3
English	41	58.6
50/50	11	15.7
No TV	1	1.4
Language most confident in using		
Welsh	42	60.0
English	6	8.6
No preference	21	30.0
Unknown	1	1.4

2.3.3 Individual Level Analysis: Does Language Format Predict Participant Scores?

See Table 14 for the descriptive statistics for the final sample of participants carried forward to the next part of data analysis.

The first part of the data analysis presented above explored performance across the Welsh and English versions of the cognitive assessments at a *group* level, and no evidence in favour of Welsh assessment was found overall. However, it is possible that there may be performance differences across languages at an *individual* level. That is, someone may perform better in one language compared to the other, but this difference may not emerge at group level analysis because some individuals may have performed more poorly in the same language – namely, perhaps some participants show a Welsh assessment preference, whilst others the reverse, and a combined average would cancel out each preference such that no overall effect was determined. Therefore, the aim of the next part of the analysis was to determine what factors are important in predicting participants' performance in each language on an individual level (e.g., what factors predict whether a participant performs better in Welsh or English). A logistic regression analysis was conducted to ascertain this.

Table 14

Descriptive Statistics of Final Sample (n = 70)

	ACE-III Total Welsh	**ACE-III Total English**	**RBANS Total Scale Welsh**	**RBANS Total Scale English**
M	93.13	93.37	112.35	112.30
SD	3.90	4.49	11.27	11.57
Minimum Score	83.00	82.00	84.00	82.00
Maximum Score	99.00	100.00	139.00	141.00

For each participant, their English score was subtracted from their Welsh score to determine which language they performed better in, thus an individual with a positive score would indicate a preference for Welsh format testing and a negative score the reverse. Those who did not show a difference across Welsh and English (e.g., same score on the assessments across languages and thus a difference score of zero) were excluded (RBANS n = 3; ACE-III n = 7) from the regression.

A series of models were carried out for each assessment which included the Welsh and English LexTALE, age and *one* other predictor such as counterbalancing order, education level, level of education in Welsh, language currently used with family, language of majority of TV programmes watched, and language most confident in using. It was necessary to carry out a series of models rather than including all predictors in the same analysis because the predictors were closely related to one another. This means it would not be possible to reliably estimate the regression coefficients of each predictor if all predictors were included in the analysis. Therefore, a series of models were carried out with each subjective measure of language proficiency/usage (e.g., level of education in Welsh, language currently used with family, language of majority of TV programme watched, and language most confident in using), whilst the objective proficiency measures were included in all models. This approach was also adopted to comprehensively determine the potential impact of the additional measures across all kinds of combinations, to fully explore the potential impact (if any) of the predictive measures.

For both the ACE-III and RBANS, only the model including counterbalancing order significantly improved the prediction: ACE-III - χ^2 (57) = 13.72, p = 0.017, McFadden R^2 = 0.157, and RBANS - χ^2 (61) = 17.35, p = 0.004, McFadden R^2 = 0.187. The Wald statistics showed that the order in which the sessions were completed was a significant factor in predicting which language they scored better in. This mirrors the ANOVA findings reported

above and is consistent with an expected natural 'practice effect' across tests administered. For both the ACE-III and the RBANS, none of the other predictive variables significantly improved the prediction; all *p* values were greater than .05. In all models, the Variance Inflation Factor was below 10, indicating the absence of multicollinearity. See Appendix I - N for each ACE-III[2] model and Appendix O - T for each RBANS[3] model.

Importantly therefore, this study found no statistical evidence of an advantage or preference for Welsh format testing using the total scores on the cognitive assessment batteries across both types of analyses. Given that these test batteries are comprised of several Indexes/Domains (e.g., attention, language, visuospatial etc.), it was then important to explore whether there may be any evidence of Welsh preference in any of these specific instances.

2.3.4 Performance on the RBANS and ACE-III by Index/Domain

Literature suggests that bilinguals may perform differently across cognitive domains (e.g., attention, memory, language) (López et al., 2016; Gasquoine et al., 2007). For example, research has demonstrated that bilinguals score higher in letter fluency tasks (Patra et al., 2020) and lower in some memory tasks compared to monolingual participants (Bialystok, 2009). With that in mind and for full examination, mean averages on Welsh/English formats for each of the Index/Domain scores (as applicable) were calculated and paired samples t-tests conducted. To control for familywise error, a Bonferroni correction was applied (0.05/5 = $p < .01$). Significant differences were found on the ACE-III Fluency Domain (Table 15) and

[2] For completeness, the analysis was also conducted with the participants in counterbalancing order 4 (i.e., those that were removed from the analysis) included in the data. For the ACE-III, the same pattern of findings was discovered except that age was a significant predictor in the model that included counterbalancing order. See Appendix N1 for all ACE-III model including group 4.

[3] For the RBANS, the pattern of findings was similar to that found without the inclusion of order 4 in that only the model including counterbalancing order significantly improved the prediction. See Appendix T1 for all RBANS model including group 4.

RBANS Language Index (Table 16), such that performance was significantly higher on the English versions. No other significant pairwise differences were found (Tables 15 and 16).

To further explore differences in the ACE-III verbal fluency domain and as performance on such tasks can differ across languages because of differences in category size (Acevedo et al., 2000; Rosselli et al., 2002) as mentioned previously (see 'Additional Verbal Fluency Task, Section 2.2.4.6), performance on the additional verbal fluency task (household items) included in this study was examined. No statistically significant difference was found across languages: $t(69) = -1.956$, $p = 0.055$.

In summary, this study found no evidence of performance differences across Welsh and English assessments at whole/total level, but there was some indication that performance differed on two Domains/Indexes.

Table 15

ACE-III Domains Means per Language of Assessment

Domains	**Language of Assessment**		***t***	***p***	**Cohen's *d***
	Welsh *M* (*SD*)	**English *M* (*SD*)**			
Attention	17.47 (0.88)	17.47 (0.98)	0.000	1.000	0.173
Memory	23.68 (1.95)	23.44 (2.35)	0.945	0.348	0.118
Fluency	11.20 (1.90)	11.84 (1.87)	-2.908	0.005*	0.121
Language	25.28 (0.98)	24.97 (1.38)	2.456	0.017	0.104
Visuospatial	15.50 (0.95)	15.64 (0.66)	-1.490	0.141	0.135

*Note: df = 69; * denotes significant findings at the Bonferroni adjusted alpha level of p <.01*

Table 16

RBANS Index Scores Means per Language of Assessment

Index Scores	Language of Assessment		*t*	*p*	*Cohen's d*
	Welsh *M* (*SD*)	English *M* (*SD*)			
Immediate Memory	106.02 (12.86)	103.58 (14.68)	1.440	0.154	0.123
Visuospatial	116.84 (12.19)	116.58 (12.20)	0.170	0.865	0.124
Language	98.35 (9.73)	102.35 (11.74)	-3.308	0.001*	0.115
Attention	112.84 (12.22)	114.65 (13.63)	-1.412	0.162	0.099
Delayed Memory	109.57 (10.75)	107.15 (12.52)	1.610	0.112	0.129

*Note: df = 69; * denotes significant findings at the Bonferroni adjusted alpha level of p <.01*

2.3.5 Comparison of Welsh-English Bilingual Performance with Published Norms

The main aims of this research were to generate normative data for the Welsh versions of the RBANS and the ACE-III, and to compare the English performance of the bilinguals in this study on both assessments with existing normative data sourced from the literature. This comparative analysis aimed to investigate a possible bilingual advantage. Please note that such comparisons were not possible for the Welsh data due to a lack of other Welsh studies. Therefore, only the English data was compared.

For the sake of completeness, a number of normative samples were considered which were identified via Google Scholar searches. These are presented in Table 17 below. The majority of the studies (i.e., Green et al., 2008; Hsieh et al., 2013; Olaithe et al., 2019; So et al., 2018) were conducted in Australia, and Duff et al. (2003) was conducted in America. In each study, it is assumed that participants were assessed in their dominant language, and thereby the language native to the country. Unfortunately, though, this is not always explicitly stated in

the published work, and some also neglect to explicitly specify the language status of their participants. Most studies included older adults over the age of 60 (e.g., Hsieh et al., 2013; Olaithe et al., 2019; So et al., 2018), Duff et al. (2003) included participants over the age of 70, and Green et al. (2008) included a sample of participants ranging from 20-89 years with a mean age of 58. To confirm, the normative data presented are from healthy participants (i.e., no history of neurological conditions, substance abuse history, psychiatric illness).

One sample t-tests found multiple statistically significant differences between the mean scores of the two groups of participants (See Table 17). With respect to the RBANS published sample comparisons, the sample in the current study tended to perform better across the board, but this was not the case for the ACE-III comparisons where participants in the current study tended to do perform slightly worse. On balance therefore, the overall picture from the normative data that was available for comparison, suggests that the Welsh participants in the present study are unlikely to be performing at levels that may suggest a general *detrimental effect* of English format testing, and thus provides further evidence (consistent with the above) of no Welsh format testing preference per se.

Table 17

Descriptive Statistics for Comparison with Published Normative Data

Study	**Assessment Used**	**N Controls**	**Number of sessions**	***M* (*SD*)**	**Present Study *M* English**	**t**	***p***	**Cohen's *d***
Hsieh et al. (2013)	ACE-III	25	1	95.4 (3.3)	93.371	-3.778	< .001	-0.452
So et al. (2018)	ACE-III	52	1	95.7 (3.3)	93.371	-4.337	< .001	-0.518
Olaithe et al. (2019)	RBANS	180	1	100.5 (12.5)	112.300	8.127	< .001	0.993
Duff et al. (2003)	RBANS	718	1	97.9 (15.9)	112.300	9.932	< .001	1.213
Green et al. (2008)*	RBANS	172	1	108 (?)*	112.300	2.921	0.005	0.357

*Note: * = no report of SD*

It was also important to examine whether the correlation between the two time points (Session 1 and 2) in the current study was similar to that in other studies. It must be noted that the correlations were conducted using only the RBANS data due to a lack of normative test-retest literature on the ACE-III. See Table 18 for test-retest correlations. It is evident from Table 18 that the correlation between the two time points in the current study is fairly large (indicating high test-retest reliability). A Fisher r-to-z transformation determined that: (a) compared to the work of Dong et al. (2013), the findings of the present study are consistent such that there was no statistical difference, $z = 1.26$, $p = 0.208$, and (b) compared to the work of Duff et al. (2005), their finding was of a 'stronger' correlation, $z = 2.41$, $p = 0.016$. However, it must be noted that the normative studies included in the table below conducted the RBANS in the same language across both time points, with their re-test interval also differing. This was not the case for the current study as participants completed the Welsh or English versions in Session 1 and Session 2, and thus a straightforward interpretation of such a comparison is somewhat complex. Nonetheless, it is possible to interpret the findings in this case as largely consistent with other work.

Table 18

Test-retest Correlation

Citation	Assessment Used	N Controls	Pearson's correlation (*r*)
Dong et al. (2013)[a]	RBANS	30	0.80
Duff et al. (2005)[b]	RBANS	445	0.81
Current Study	RBANS	70	0.57

Note: [a] two-week re-test interval, ≥50 years of age, no objective cognitive impairment; [b] approx. one-year re-test interval, ≥65 years of age, community dwelling adults

2.4 Discussion

This study aimed to collect data from healthy participants in Welsh for two major cognitive assessment batteries to determine whether performance differs across languages within a group of bilingual older adults, as well as between bilinguals and English normative data. Aforementioned literature suggests that performance on cognitive assessments may differ if they are conducted in the individual's second or non-preferred language (Escobar et al., 1986; Walker et al., 2010; Yeung et al., 2014). There is also growing desire in Wales for cognitive assessments to be conducted in Welsh (Welsh Language Commissioner, 2012; Older People's Commissioner for Wales, 2016). However, research investigating the performance of this unique population of bilinguals on both Welsh and English cognitive assessments is limited. The present study investigated whether performance differs on the Welsh and English versions of the ACE-III and RBANS in a group of Welsh-English bilingual older adults. Specifically, it sought to understand whether native Welsh speakers, or those who prefer to speak Welsh, would perform better on the Welsh versions of the assessments.

Consistent with preliminary evidence reported by Clare et al. (2016b), the findings of this study demonstrated that healthy Welsh-English bilinguals performed comparably on the Welsh and English versions of the cognitive assessment batteries at whole/total assessment level. In terms of Domain/Index level performance, performance was found to differ on the ACE-III Fluency Domain and RBANS Language Index, such that performance was superior on the English versions. However, when interpreting the significance and meaning of these findings, several possibilities need to be considered to aid interpretation. First, it is unclear how challenging it may be to complete various categories or letter fluency in the Welsh verbal fluency domain (ACE-III) as well as various subtests in the language index (RBANS) compared to completing them in English. In relation to the letter fluency in the Welsh ACE-

III, participants are asked to name as many words as possible beginning with the letter ‘B’ (compared to the letter ‘P’ in English). Although, even though there are more everyday words beginning with ‘B’ than ‘P’ in the Welsh language, a potential issue with including this letter in verbal fluency tasks is that the letter ‘B’ is often the first letter of a word after a mutation has taken place (e.g., ‘pont’ [bridge] changes to ‘bont’), meaning that participants may have to suppress mutated forms (e.g., ‘bont’) of the radical word (e.g., ‘pont’) (i.e., the ‘basic’ form listed in the dictionary) before identifying correct answers. Thus, and given that this is a timed task (one minute), participants may spend time inhibiting mutated responses and consequently preform poorer. Indeed, a number of participants in the present study produced mutated words beginning with ‘B’ before realising that they had mutated from the radical form. Thus, further research is required to determine the best letters to use in Welsh verbal fluency tasks and whether mutated words should be accepted, even though they may not be radical forms of a word (and therefore not meeting task requirements). Even so, note that the challenges that may arise due to mutations are only relevant to letter fluency tasks (as opposed to category fluency) because the task requires the individual to state as many words *beginning with* a particular letter. In Welsh, the first letter of a word sometimes changes due to a mutation.

In line with the above idea of inhibiting responses, and so that participants do not have to spend time inhibiting words that come to mind in the opposite language, it is also unclear whether participants should be instructed from the start that both Welsh and English responses are appropriate. This may explain the difference in performance on the RBANS language Index and may also serve as an additional explanation for the difference in performance on the ACE-III verbal fluency domain. Semantic fluency is one of only two subtests (along with picture naming) that make up the RBANS Language Index. Even though English responses were scored as correct in the Welsh semantic fluency tasks across

assessments, participants may not have had to contend with the inhibition of Welsh responses in the English versions, especially when language intrusions seem to be more common in the Welsh assessments (that is, more intrusions from the English language when completing a Welsh assessment).

However, it is important to highlight that no difference was found on the additional category fluency measure (household items), which may appear to contradict the suggestion that participants may have been inhibiting responses in the verbal fluency measures in the ACE-III (animals) and RBANS (fruit and vegetables). However, English words are more commonly used for the household items category compared to others. That is, household categories may generate more English responses. Consequently, participants may have experienced less of a need to inhibit responses in the household items category because they often use English words for household items. This may explain the lack of a significant difference between the Welsh and English version of the additional category fluency measure. Again, more research is required to determine the most appropriate categories to use with this population of bilinguals, as well as whether accepting bilingual responses and/or explicitly instructing participants that bilingual responses are accepted is a valid and appropriate measure of verbal fluency.

A second possibility is that observed differences on the ACE-III Fluency Domain and RBANS Language Index may reflect some other possible confounding variable (i.e., counterbalancing order), rather than reflecting meaningful performance differences. However, this seems unlikely as subsequent exploratory analysis showed that counterbalancing order was unable to fully account for the findings. Additionally, and even though the difference in performance across the two Domains/Indexes was statistically significant, it should be noted that the effect sizes were very small. Thus, the findings may have little practical significance in that they are perhaps unlikely to lead to meaningful

differences in clinical outcomes or total scores (i.e., difference between a typical overall score and a score indicative of cognitive decline). In addition, the variation in performance across languages in these two Domains/Indexes may simply reflect the general variation in standard repeat testing. That is, the variation that is inevitably present when re-testing a patient in the same language over time (i.e., when a monolingual undergoes repeated assessments). If variation is inevitable when assessing monolinguals over time, this is likely to also be the case for bilinguals. Overall, therefore, there appears to be little evidence of consistent and pervasive differences at global level and at Domain/Index level.

The study also failed to find consistent evidence of a bilingual advantage or disadvantage across total scores when compared with the normative data, *suggesting* that the mechanisms involved in increasing CR, may be less to do with general cognitive domains. However, it is essential to note that based on the comparisons made with available normative data, it is unclear whether the findings reflect real differences or artefacts of the study designs. Moreover, the number of suitable comparative studies available in the literature is limited, and those that were available are not direct comparisons due to the lack of information provided about the sample of participants used. That is, such studies commonly fail to explicitly describe the language status of the participants (monolingual/bilingual) and/or to specify in which language the assessments were administered. Consequently, the most logical conclusion is that there is no clear evidence of a bilingual advantage or disadvantage given the mixed findings observed across the ACE-II and RBANS normative comparisons, as well as the limitations attached to the comparisons that were made.

The main explanation for these findings is that the nature of Welsh-English bilingualism is different to that of other bilingual populations. As mentioned in the introduction section of this chapter, both the minority and dominant languages are heard simultaneously and consistently in Wales due to the impact of the dominant language (English). This means that

it may be the case that the participants in this study are so familiar with hearing and using both languages on a daily basis, that they do not show a preference for one language over the other and also experience less competition between languages (Clare et al., 2016b; Macnamara & Conway, 2014). Additionally, although there is a limited amount of literature available on the effects of balanced bilingualism on cognitive assessment performance, the focus is usually on countries where both languages are not heard and used consistently and simultaneously in daily life (they are not balanced via their own personal use of both languages), meaning it is different from the unique balanced bilingualism seen in Wales. Therefore, and whilst this study goes some way in addressing an important gap in the knowledge base, further research is still needed to fully elucidate the impact of language testing with such unique balanced bilinguals. Indeed, it is important to recognise the inherent complexity in drawing conclusions when multiple variables are potentially involved and the interrelationships between them unclear. This issue will be discussed further in Chapter 5, as it pertains to the entirety of the thesis.

Further, and despite the lack of difference in performance across languages, it is unclear whether the findings would be different if: (a) the testing sessions were conducted in a clinical setting, and (b) if a clinical population was included in the sample. In relation to the first suggestion, if the testing sessions were conducted in a formal, clinical setting, participants may have experienced increased anxiety because of the clinical/medical environment which may have impacted on their performance, as evidenced in the literature (Dorenkamp & Vik, 2018). To address the second suggestion, it may be the case that people show an advantage or a preference for one language when they are affected by cognitive impairment and when language becomes affected by dementia. Consistent with this suggestion, Morgan & Crowder (2003) found that 81% of Welsh-English bilinguals with a diagnosis of dementia performed differently on the Welsh and English versions of the Mini-

Mental State Examination. 42% scored better in Welsh, 39% scored better in English and 19% obtained consistent scores in both tests. Further, even though the proportion of participants whose scores were higher in one language than another was similar, observed differences were noted to be much larger when in favour of the Welsh language. However, Morgan & Crowder did not examine whether such differences reached statistical significance. Further, the study has several limitations which could have influenced the findings. For example, in addition to a small sample size (n = 31) which affected the generalisability of the results, the minimum and maximum duration between assessments was 48 hours and one week. This timeframe may not have been long enough to reduce or eliminate any practice effects (Morgan & Crowder, 2003). Indeed, 52% of participants who demonstrated a difference in their scores, achieved a higher score in their second assessment.

A further reason why the findings of the current study may have been different if a clinical sample were examined, is that some studies indicate that bilingual individuals can regress to the primary language or show preferential preservation of the first language during the development of cognitive impairment (McMurtray et al., 2009; Mendez et al., 1999). Therefore, conducting an assessment in a language that is no longer used by, familiar to, or preferred by the patient may merely worsen the experience for patients. Consequently, more longitudinal rather than cross-sectional research is needed so that any changes in the characteristics of the Welsh-English bilingual population can be detected.

In addition, future research is needed to determine whether a clinical sample would score differently across the Welsh and English versions of the RBANS and ACE-III. Findings of such research could have major clinical implications, such as if those who prefer to speak Welsh (who are also affected by cognitive impairment) were found to score higher on the Welsh versions of these assessments. For example, older adult services in Wales would need to adapt assessment processes, ensure sufficient numbers of Welsh speaking staff within

services, and that such staff are also appropriately trained to conduct cognitive assessments. On the other hand, if a clinical sample were found to score comparably across languages (in line with the healthy sample in this study) and considering the notion that some may regress to the primary language during the development of cognitive impairment (e.g., McMurtray et al., 2009), then it may be sensible for clinicians to conduct the assessments in both Welsh and English. This may lead to a cognitive profile that more accurately reflects actual cognitive status. However, and as previously emphasised, more research is needed before any definitive conclusions can be made.

Previously, some clinicians have also used the Welsh language translated assessments utilised here, despite a lack of Welsh normative data. Therefore, the practical implications of the current study are notable as it has provided important information, including: (a) that healthy individuals who have already received an assessment in Welsh (non-validated translated versions) are unlikely to have been disadvantaged, (b) that healthy individuals who have received English assessments due to a lack of Welsh versions are unlikely to have been disadvantaged, and (c) as a normative data set has been generated, individuals should and can be offered an assessment in Welsh. Offering assessments in Welsh will also ensure services meet part of the 'Active Offer' mentioned in the Welsh Governments 2016 'More than just words: A strategic framework for promoting Welsh language in health, social services and social care' guidance (Welsh Government, 2016). Put simply, many individuals living with dementia are extremely vulnerable and it should not be their responsibility to ask for an assessment or other aspects of their care in Welsh. Indeed, person-centred approaches, which are considered best-practice for dementia care (e.g., Kong et al., 2022), tailor care to the individual's clinical needs, personal preferences, and life circumstances (Marulappa et al., 2022). Thus, language considerations are integral to such provision. Of course, offering

assessments in Welsh will also mean that professional standards, as well as legal and statutory requirements in Wales, are met (Welsh Government, 2016).

In addition to the main findings of this research, there were also some unexpected outcomes. The study found that performance on the Welsh ACE-III was reduced if it was completed immediately after the Welsh RBANS (i.e., counterbalancing Order 4). This finding was unique to participants who completed the Welsh RBANS followed by Welsh ACE-III in their first session, prompting their subsequent exclusion from the main analysis. Whilst the precise reason for this performance discrepancy remains unclear and warrants further investigation, potential factors such as fatigue or cognitive load may have contributed. For instance, there may be something about doing the Welsh RBANS first that differentially impacts performance on the Welsh ACE-III. For example, it may be that the Welsh RBANS is more challenging (compared to the English version), which in turn, may lead to increased fatigue and impact on subsequent performance in remaining assessments in the same session. Participants performed significantly lower than expected on the first exposure (Welsh RBANS/ACE-III), but then performed at a 'typical' level when they completed the English ACE-III in their second session – with that gain perhaps larger than could be attributable to practice effects alone. However, more research is required to confirm the reason for this unexpected finding.

Nevertheless, while a rationale for the exclusion of this participant group (i.e., counterbalancing order four) was outlined in the results section, it is important to discuss the potential implications of this decision. The main aim of this study was to collect normative data, necessitating the need for high-quality and 'clean' data. Thus, the removal of the data from participants in counterbalancing order four, improved the reliability of the normative dataset. That is, the distinctive underperformance of participants in that counterbalancing order on the ACE-III (Welsh) could have skewed the results, potentially leading to inaccurate

conclusions being drawn and unreliable normative data being produced. However, it is important to acknowledge that removing a set of data can also increase the risk of introducing bias and losing potentially valuable information. Even so, given the unexplained and distinctive underperformance on the ACE-III observed across participants in counterbalancing order 4, coupled with the objective of acquiring normative data, the decision to exclude this data from the analysis was deemed appropriate on balance, after carefully considering potential biases and trade-offs.

2.4.1 Limitations, Strengths, and Unanswered Questions

One of the most significant accomplishments in the current study was adapting the data collection method due to the COVID-19 pandemic to achieve the study's aims. The process of determining the new method was thorough and a plethora of factors were taken into consideration, as outlined in the methods section. Even so, the study is not without limitations. As mentioned previously, prior research investigating the performance of Welsh-English bilinguals on both Welsh and English cognitive assessments is limited, and therefore, it was not possible to compare the performance of the participants in the present study with those with a similar nature of bilingualism. Instead, the English scores were compared with normative scores of other healthy participants from other studies, where the current sample scored higher than some published norms and lower than others. However, due to the limitations attached to the available studies (mentioned previously), it was not possible to conclude whether Welsh-English bilingualism leads to a general improvement in cognitive processes. However, with respect to the key focus of the current work, findings suggest that Welsh participants in the present study are unlikely to be performing at levels that may suggest a general *detrimental effect* of English format testing, and thus provides further evidence (consistent with the above) of no Welsh format testing preference per se.

Nonetheless, and with the sample's level of education in mind, one of the limitations of this study is that the sample's demographics were not representative of the whole population of interest. Previous research has found significant correlations between education level and the RBANS subtests' scores (Duff et al., 2010). In addition, Wang et al. (2009) found evidence that the cut-off score for the Chinese version of the ACE-III is impacted by level of education, and educational bias has also been reported in studies including the Mini Mental State Examination in Asian populations (Rosli et al., 2016). Consequently, further research is required to look at the performance of a more representative group of Welsh-English bilinguals on the Welsh and English versions of the RBANS and the ACE-III.

Another limitation is that the approach taken with counterbalancing meant that mode of testing could not be explored because of the low numbers of participants in each counterbalancing order. That is, the approach taken in the current study was reliant on the full sample being met, but also no changes being made due to unforeseen circumstances like the pandemic. Therefore, and if the study were to be conducted again, a different approach to counterbalancing would be advisable. Namely, rather than alternating two orders with a group of participants before switching to another two orders, participants should be randomly allocated to all counterbalancing orders (e.g., 1, 2, 3, 4).

This study has also raised an important question around how to score bilingual responses during a cognitive assessment, where it remains unclear how best to score bilingual responses on either a Welsh or English assessment (e.g., an English response in a Welsh assessment). Due to the nature of Welsh-English bilingualism, code switching is common in Wales and has been mentioned in publications since the 1980s (Thomas, 1982; Deuchar et al., 2016), with code switching particularly prominent in the case of balanced bilinguals or those who learn both Welsh and English simultaneously (Deuchar et al., 2016). This raises the question of whether bilingual responses should be scored as correct at all, whether they should be scored

as correct for every subtest forming Index/Domain scores, or whether they should be scored as correct for language sensitive subtests only (e.g., verbal/category fluency, picture naming tasks). Further research investigating this issue would inform the development and adoption of clearer administration and scoring guidelines to ensure reliable and valid test scores are produced.

2.4.2 Conclusion

In sum, this study is the first to the researcher's knowledge to collect data from healthy participants for the Welsh versions of the RBANS and ACE-III. The outcomes of this research have provided insight into the performance of Welsh-English bilinguals across Welsh and English versions of two major cognitive assessment batteries. The research also furthers our understanding of the nature of Welsh-English bilingualism and raises many important questions for future research. Even so, whilst examination revealed no consistent performance differences on general cognitive assessments between the bilinguals in this study and the monolingual normative data, much remains unknown about the potential effects of ageing and bilingualism on *particular aspects* of cognition – those aspects that are not widely studied in the literature and those that are not necessarily measured in such general cognitive assessment batteries (e.g., semantic control). With the theme of CR in mind, investigating metacognitive processes like semantic control may also be helpful in detecting the bilingual advantage. Thus, the subsequent experimental chapters will explore this further.

Key Messages

- This study aimed to collect data from healthy Welsh-English bilinguals for two major cognitive assessment batteries to determine whether performance differs across Welsh and English versions, and to generate a normative data set (for the first time) that can be used in future clinical and research work.
- This study also aimed to compare the performance of the Welsh-English bilinguals on the English versions of the assessments with that of normative data to explore evidence for a possible bilingual advantage.
- Previous studies suggest that performance in cognitive assessments is related to language and culture, and that conducting an assessment in the individual's second or non-preferred language may place them at a disadvantage (Kisser et al., 2012; Kowoll et al., 2015; Morgan & Crowder, 2003; Walker et al., 2010).
- This study found that healthy Welsh-English bilingual older adults performed comparably across the Welsh and English versions of the ACE-III and RBANS at total level, with only two small significant differences found at Index/Domain level.
- Based on performance comparisons with available normative data, there was no consistent or strong evidence of a bilingual advantage or disadvantage.
- More research is required to investigate whether clinical populations would also perform comparably across Welsh and English versions, or whether they would perform better in one language.

Chapter 3 – Study 2 – Semantic Cognition in Ageing and Bilingualism

3.0 Chapter Overview

The findings from study 1 provided little evidence that Welsh-English bilingualism leads to differential performance across the Welsh and English versions of the ACE-III and RBANS. Comparisons with normative data also provided little evidence that Welsh-English bilingualism may have either a detrimental or advantageous impact on cognitive performance at global levels. In line with this, existing research has found mixed evidence of a bilingual advantage when using general executive type processes (e.g., inhibition, language switching) to measure CR (Bialystok., 2021; Calvo et al., 2016). This highlights the need to move away from merely using *general* and well-studied cognitive processes, to locate the bilingual advantage, by additionally considering the impact of bilingualism on *specific* cognitive processes. This Chapter aims to explore whether processes of semantic cognition (SC) are a more suitable proxy of CR by investigating the impact of Welsh-English bilingualism and ageing on a semantic control paradigm.

3.1 Introduction

The previous chapters of this thesis introduced the notion that language plays a crucial role in the development of AD and the way in which clinicians monitor its progression (Obler & Albert, 1984). The ability to detect dementia as soon as possible is essential for ensuring that patients receive appropriate and timely support (Yoshida et al., 2011). One of the main steps in determining a dementia diagnosis, is to evaluate cognitive function via validated cognitive assessment batteries, such as the ACE-III and RBANS. Therefore, the first study in this thesis (Chapter 2) considered whether the performance of healthy bilingual older adults differed across Welsh and English cognitive assessments. The previous chapters also examined the

lifelong experience of using, regulating, and switching between more than one language, and how it may contribute to one's level of CR, which is a crucial aspect in discussions around the development and onset of dementia (Guzmán-Vélez & Tranel, 2015). The previous Chapter investigated this in the Welsh context by comparing the performance of Welsh-English bilinguals on two cognitive assessment batteries to normative data.

During ageing, various biomarkers of dementia may develop which help diagnose different types of dementia (e.g., amyloid plaques, tau protein, atrophy in various brain areas) and predict decline (Ahmed et al., 2014). Over time, with the accumulation of certain biomarkers, cognitive decline will manifest (Jack et al., 2010; Jack et al., 2013) and the acceleration of this decline seems to be much earlier or later for some individuals than others (Stern, 2002; Snowdon, 2003). However, the reason for this interindividual variability remains unclear. Some suggest that genetics, lifestyle factors or mood/wellbeing may play a role, whereas others suggest that such interindividual variability may be something to do with CR (Cuyvers et al., 2015; Müllers et al., 2019; Penninkilampi et al., 2018; Stern, 2012). See Chapter 1, Section 1.8 for further details.

However, systematically exploring the latter possibility has proved challenging because CR is a hypothetical construct and thus difficult to measure (Guzmán-Vélez & Tranel, 2015). For instance, previous studies have used proxy factors like education level, occupation, intelligence, and SES (Jones et al., 2011; Sattler et al., 2012; Tucker & Stern, 2011; Wilson et al., 2019) as variables that might predict higher or lower CR. Importantly though, these factors may be confounded with each other which means establishing how each factor may affect cognition is challenging (Jones et al., 2011). Further, such proxy factors do not speak directly to the kinds of cognitive processes that may underpin the CR advantage. In other words, even if education level is an effective proxy of CR, it is unclear whether this reflects a

general uplift of the entire cognitive system or whether it is linked to an uplift of a specific cognitive process.

On the basis of the work presented in the previous chapter, no strong or consistent evidence was found of a widespread uplift (or disadvantage) on general cognitive tasks in the context of Welsh-English bilingualism. Compared to available normative data, the Welsh-English bilingual sample performed more poorly on the RBANS, but better on the ACE-III. Thus, findings were mixed and inconclusive. However, when exploring a novel proxy for CR, and shifting the focus to specific cognitive processes instead of overall system differences, language emerges as the most suitable candidate. This is because the main distinction between bilinguals and monolinguals lies in their differing language abilities. It seems plausible to assume therefore, that specific cognitive system differences across these populations may be found in this domain. Consequently, this will be the focus of the empirical work presented in this chapter.

Complementing the discussion already presented in Chapter 1 (see Sections 1.6 and 1.8), while many studies suggest the existence of a bilingual advantage, the exact process or component of bilingual cognition responsible for this advantage remains unclear (Bialystok, 2015). Some research indicates that inhibition may play a central role in bilingual cognitive processes (Bialystok & Craik, 2022; Grainger and Dijkstra, 1992; Green, 1998; Green & Abutalebi, 2013) and, consequently, in the bilingual advantage. Bilinguals often outperform monolinguals on tasks measuring inhibitory control (e.g., flanker task). However, as they have also been found to outperform monolinguals on tasks that do not require inhibitory skills, it is unclear whether enhanced inhibition abilities are solely responsible for the bilingual advantage (Bialystok, 2010). Additionally, the type of inhibition may be an important consideration, as bilinguals have been found to show an advantage in certain types of inhibition compared to others, such as response inhibition or interference suppression

(Carlson & Meltzoff, 2008). Nonetheless, other research suggests that the advantage may be less attributable to inhibition. For instance, a bilingual advantage has been observed in both the congruent and incongruent trials of a flanker task, even though specific inhibition is not required in the congruent trials where there is no conflicting information to inhibit (Costa et al., 2008).

This has led researchers to consider monitoring, rather than inhibition, as the specific cognitive process that may be the source of the bilingual advantage (Bialystok, 2021; Costa et al., 2009; Hilchey & Klein, 2011). As captured by Blanco-Elorrieta & Caramazza (2021), the argument is that bilinguals may be more efficient in using monitoring mechanisms to adjust their inhibitory control to effectively cope with interference. Thus, it is the combination of both mechanisms that confers an advantage to bilinguals where they can more effectively cope with overall interference. In line with this, some studies have found a bilingual advantage in tasks requiring participants to hold a rule in mind and to monitor changes to the rule (e.g., dimension change card sorting task) (Bialystok, 1999; Okanda et al., 2010; Zelazo et al., 1996) - suggesting bilingual advantages may extend to multiple components of executive function.

Further, Bialystok & Craik (2022) highlight how monitoring also involves working memory, as successful monitoring requires holding a rule in mind. However, studies examining working memory differences across monolingual and bilingual children have typically produced mixed findings (e.g., Blom et al., 2014; de Abreu, 2011). In adult populations, findings are also generally mixed, albeit differences across monolingual and bilingual groups have more consistently been reported when working memory tasks involve non-verbal than verbal materials (e.g., Bialystok et al., 2014). Thus, there is at least some evidence that the bilingual advantage may stem from cognitive processes such as inhibition,

monitoring, and working memory, with all three accounts presented above assuming that language control relies on domain-general executive control.

However, the tasks used to identify the key processes around inhibition, monitoring and working memory that may reveal a bilingual advantage, tend to be relatively simple, involving visual stimuli that change (e.g., flanker task) across different trials. Therefore, the cognitive processes required to complete the task (e.g., inhibition, monitoring, control) operate at a single level of knowledge/information (i.e., whether the stimuli are the same or different to a target item such as arrows pointing in certain directions). Therefore, they are only positioned to identify single components of executive function that may be responsible for the bilingual advantage. Consistent with this line of thinking, de Bot (2017) also highlights how the evidence for a bilingual advantage is mixed, and that most associated literature, as mentioned previously (Bialystok, 1999; Costa et al., 2008; Okanda et al., 2010), is focused on general areas of cognitive processing. These include switching, inhibition, and updating/monitoring (de Bot, 2017). De Bot (2017) asks whether advantages in cognitive processes (i.e. inhibition) related to language can be transferred to other areas of general cognition. That is, if bilinguals inhibit language related content over the course of their lifespan, does this mastered skill also transfer and serve as an advantage in other areas of life where inhibition is required?

Building on the above and given the crucial aspect of bilingualism involves language switching and its related cognitive processes, the CR advantage in this context may therefore be better represented by changes in metacognitive processes or in common cognitive processes acting upon semantic knowledge. In other words, bilingualism may offer a key advantage in this domain (Abu Rabia, 2019; Ordin et al., 2020). However, few studies have considered the consequences of bilingualism and cognitive control with respect to operations acting upon semantic memory.

In relation to this, the work of Hoffman (2018) (explained further below) may offer a useful framework for further investigation, as: (a) it provides guidance on some relevant experimental paradigms that can capture semantic control mechanisms, and (b) shows how patterns of cognitive performance linked to such mechanisms can be differentially impacted by ageing (Hoffman, 2018). Therefore, it offers a framework in which to explore whether age-related changes linked to the semantic control mechanism may in some way be moderated by bilingualism. Thus, the current work seeks to experimentally investigate the consequences of bilingualism on semantic memory itself and the key metacognitive mechanisms that operate on this system. For clarity, note that the metacognitive mechanisms that operate on this system have, and will continue to be referred to as semantic cognition (SC) in this thesis. However, before considering this broader picture, it is worth exploring the initial accounts of the organisation of semantic memory that helped determine the mechanisms that act on the information stored in this memory system. Please note that due to the lack of research on the effects of bilingualism on SC, this Chapter will first position the current study in the monolingual context before outlining the relevance of bilingualism.

3.1.1 Models of Semantic Organisation

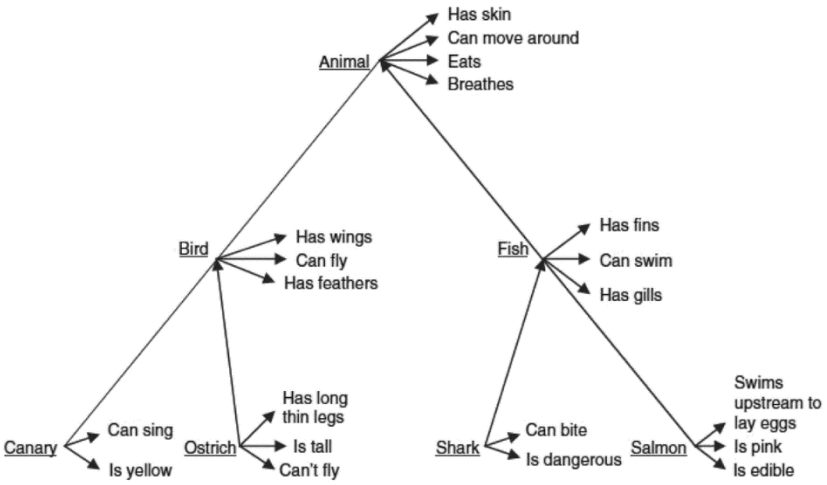
Meaning is crucial to all aspects of daily life and is the most apparent characteristic of memory (Jones et al., 2015; McRae, 2004). Despite this though, semantic memory is considered the most mysterious of the memory systems (Jones et al., 2015). Indeed, and compared to other aspects of cognition, relatively little is understood about the way in which humans store, access and retrieve the meaning of words, objects and experiences (Jones et al., 2015).

Various theories and models have attempted to explain how semantic memory is organised, with some proposing a hierarchical structure and others suggesting a more

distributed organisation (De Deyne et al., 2017; Yee et al., 2014). The three classic models of semantic memory that have received considerable attention in the literature are semantic networks, feature-list models, and spatial models (Jones et al., 2015). The two most influential models that fall under semantic networks and feature-list models, are the spreading activation theory and the feature-comparison theory, respectively (McNamara & Holbrook, 2003).

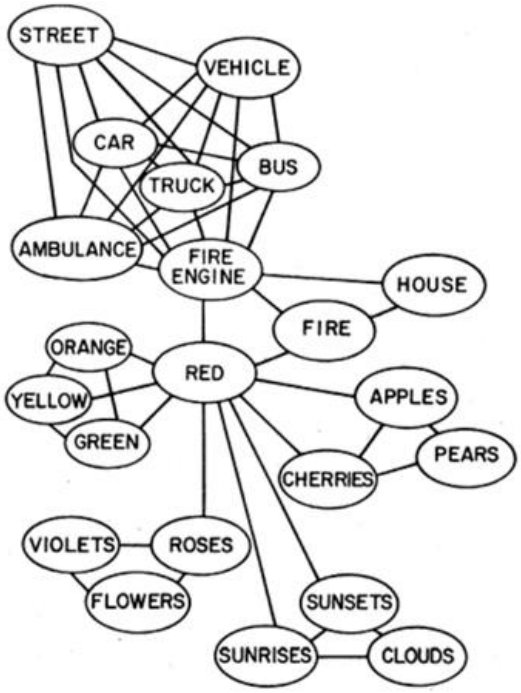
The semantic network model (Collins & Quillian, 1969) is a hierarchical framework in which concepts/mental representations (e.g., animals) are nodes and propositions (other things that are known about the concept) are labelled links (e.g., is a, has, can) (Jones et al., 2015). The lower and higher ranked structure of the links creates a hierarchical tree structure (animals are divided into fish, birds etc., and birds are further divided into robin, crow etc.), where knowledge is thought to be accessed via traversal of the tree to the required branch (Kenett et al., 2017). See Figure 6 for a diagram of the Semantic Network Model.

Figure 6. *Diagram of the Semantic Network Model Related to Animals originally proposed by Collins & Quillian, 1969. Version presented from Kumar (2021) – an open access article under the CC BY licence*



A second version of the semantic network model was later proposed by Collins and Loftus (1975). This later version (see Figure 7) considered the hierarchical nature of the network to be less important than previously thought, emphasising spreading activation throughout the network of links simultaneously (Jones et al., 2015). Put simply, the spreading activation suggested in this model proposes that the activation of a concept (e.g., bird) leads to the activation of both related concepts and related propositions (e.g., has wings). The strength of this activation depends on how strongly the concept and its propositions are related. In addition, the activation may be quicker for strongly related concepts and propositions, or they may be activated to a higher degree than weaker associated links (Lorch, 1982).

Figure 7. *Diagram of the Spreading Activation Model of Semantic Information originally proposed by Collins & Loftus, 1975. Version presented from McIntosh (2015) – an open access publication series*



Another influential model is the feature-comparison model (Rips et al., 1973), which suggests that the meaning of words is encoded as a list of descriptions (e.g., ‘has wings’,

'barks') and the appropriate descriptions are activated (e.g., 'has wings' activated for 'robin' but not for 'husky') (Jones et al., 2015). Two types of semantic features were suggested by Smith et al. (1974): (a) defining features (features that all concepts have), and (b) characteristic features (features that are typical of the concept but are not always present, for example, all birds have wings but not all birds can fly). The processing assumptions of this model are illustrated in Figure 8. Here, verification of a statement (e.g., a robin is a bird) is thought to involve a two-stage process (Smith & Estes, 1978). The first process is where all the features of the subject and predicate nouns are compared, and the similarity of the features of both terms assessed (Smith & Estes, 1978). True or false decisions can be made quickly if the featural similarity is very high or low, similar to the strength of association between the concept and proposition in the spreading activation model. However, if the featural similarity falls between the two criteria (e.g., chicken and bird or bat and bird), then a second stage is activated, where the defining features of both terms are isolated and those of the predicate and subject are compared. If all these features match, a 'true' response is activated (Smith & Estes, 1978). Additionally, the most dominant or most representative features of a concept are thought to be listed sooner than less dominant features (Montefinese et al., 2014), suggesting that they may be easier to access.

Thus, semantic memory models typically envisaged its organisation as being around structuring components or features with interconnections of some manner, where components/features may or may not be shared across concepts. Additionally, although these semantic memory models capture how 'meaning' may be *organised* in memory, they do not address the processes involved in *accessing* this knowledge from memory. Accessing knowledge from memory is important because memory contains a great deal of information that is not always required, meaning that individuals must select the relevant information for effective and efficient daily functioning, and for completion of specific cognitive tasks.

Figure 8. *Processing Assumptions of the Feature-Comparison Model taken from Smith & Estes (1978)*

3.1.2 Semantic Processing in Bilinguals

This section provides a brief overview of studies examining semantic processing in bilinguals. Understanding how meaning is stored and represented in each language of a bilingual is essential, as the interplay between languages may impact on memory, perception, and processes involved in increasing CR (Altarriba & Basnight-Brown, 2009). While some words may hold identical meaning across languages, others may have similar meaning but reference different things or concepts. Additionally, certain words may be exclusive to one language. For example, 'hiraeth' in Welsh has no equivalent meaning/translation in English. Described as a unique combination of homesickness, nostalgia, and longing, none of these English words fully capture its rich cultural nuances. Consequently, research has explored semantic representations of bilingualism, and how words are activated and processed in each language. Indeed, some research suggests that the meaning of a word can be activated in both languages of a bilingual, with greater activation observed for those who are highly proficient in their second language (Altarriba & Canary, 2004; Basnight-Brown & Altarriba, 2007). Consistent with this, findings from the Stroop Task, where participants are asked to complete

a task in one language whilst inhibiting the other, demonstrate that both languages are activated even when the task does not require information processing from both languages (Altarriba & Basnight-Brown, 2009). These findings hold particular relevance for the bilinguals studied in this thesis, who are highly proficient in both languages as balanced bilinguals. Notably, Zied et al. (2004) found that balanced bilinguals were faster to respond on a Stroop task compared to French dominant and Arabic dominant bilinguals. Zied et al. (2004) suggested that the superior performance stemmed from their ability to control both languages, where their equal proficiency in both had led to improved inhibitory abilities. Therefore, if both languages of a bilingual are active, and activation is stronger in those highly proficient in both languages, it seems sensible to suggest that the cognitive and semantic processes of balanced bilinguals differ from other bilinguals.

In keeping with the discussion on bilingual semantic processing, the nature of the task in question may also influence how bilinguals process meaning in their second language and activate words from both languages. For example, Grainger (1998) found that bilinguals might process words differently in their second language depending on whether they are completing a lexical decision task or a semantic categorisation task. Specifically, in a semantic categorisation task, they found that bilinguals performed better in their second language when primed with words from their first language. However, on a lexical decision task, participants did not exhibit enhanced performance when deciding word authenticity when primed with words from their first language (Grainger, 1998). These findings suggest that the nature and/or requirements of a linguistic task can impact the way in which words from the second language are processed in the bilingual brain. Further supporting this notion, Sheng et al. (2006) noted comparable performance between monolingual and bilingual children on a word association task, but bilingual children exhibited an advantage when required to connect words in a specific way, particularly with verbs. This suggests that

bilingualism may confer an advantage with some aspects of semantic organisation, such as paradigmatic organisation, wherein words are grouped based on their meanings, characteristics, or relationship with other words. In addition, Johns et al. (2016) investigated how bilinguals recognise words as they age by comparing younger and older bilinguals and monolinguals on a lexical decision task. They found that the variety of meanings connected to a word in different contexts had the greatest impact on word recognition speed, particularly among older adults and bilingual individuals. This may be attributed to the possibility that older adults and bilinguals may possess a larger lexicon due to their proficiency with two languages and accumulated life experience. However, it is essential to acknowledge that such findings may differ depending on individual differences such as type of bilingual, cultural context, and language use.

It is also important to note that while the above studies focus on bilingual individuals, they often lack detailed descriptions of the specific type of bilingualism. For instance, some groups are simply referred to as "highly skilled" in a particular language. As noted previously, the absence of such detail concerning participants' bilingual characteristics diminishes the meaningfulness of findings. Further, this is not only relevant to research investigating semantic processing, but also holds relevance for the broader literature on the potential bilingual advantage.

In summary, semantics in the bilingual brain seem to be interconnected and may undergo simultaneous processing across languages. However, this process may be influenced by variations in task demand and individual differences. Consequently, bilingual individuals may encounter additional challenges to monolinguals, who only control and manage information from a single linguistic system.

3.1.3 Semantic Cognition (SC)

As mentioned in the previous section, humans store a wealth of knowledge about everything in their environment, but usually only a small amount of this knowledge is required to complete a task or reach a goal. Thus, much of the stored information may be unnecessary to complete a specific task and can sometimes be unhelpful or unrelated (Noonan et al., 2010). For example, playing a piano requires an individual to access information relating to the gentle finger movements required, but that same individual would require knowledge about a different set of actions if they were required to move the piano to a different location (e.g., feet positioning, strength needed, hand placement) (Ralph et al., 2017; Saffran, 2000). Comparably, in word-picture matching tasks, semantic activation must prioritise the target while disregarding other semantically related options (Noonan et al., 2010). Moreover, during conversations or on picture naming tasks, individuals must retrieve and select the most appropriate label to ensure the accurate level of meaning for a given concept (e.g., animal, pet, cat, British Shorthair, or "Smokey" – Noonan et al., 2010). Consequently, the issues explored in this chapter also have consequences outside a laboratory or research setting, including the context of clinical assessment or day to day interactions with others.

In recent times, this cluster of metacognitive processes that allow the retrieval and selection of relevant information, has been given the label of SC – which is a metacognitive theory used to understand how knowledge is accessed and manipulated (Ralph et al., 2017). SC refers to the group of neurocognitive mechanisms that underpin our understanding of the meaning of objects, words, facts and people (Noonan et al., 2010). It is the ability to use meaning to make judgements about the environment and thus effectively direct behaviour in a contextually appropriate manner (Almaghyuli et al., 2012). It also refers to the ability to change, use and generalise knowledge to support both verbal and non-verbal behaviours. As a

result, SC provides the fundamentals for day-to-day life in a variety of ways (Ralph et al., 2017).

3.1.4 Semantic Cognition – The Mechanisms of Control that Operate on Semantic Memory

Effective SC (the ability to use semantic information to achieve goals/complete tasks) is thought to require at least two interacting elements: (a) semantic representations, and (b) cognitive control mechanisms (Hoffman, 2018; Wu & Hoffman, 2022) that are linked to processes *operating on* semantic representations (i.e., for appropriate meaning selection in ambiguity, inhibition of irrelevant information etc.). Semantic representations play a crucial role as they enable us to generalize knowledge acquired from previous experiences to novel situations. Thus, the way this information is processed and adapted to suit the specific context (context-appropriate) is influenced by cognitive control (Hoffman, 2018).

In relation to these two elements of effective SC, evidence suggests that different but interacting brain regions are used for semantic representation and control, because when their associated brain areas are damaged, different clinical presentations emerge (Jefferies & Lambon Ralph, 2006; Whitney et al., 2011). For example, damage to the anterior temporal cortices, a store of semantic representations (Pobric et al., 2007), leads to an extreme loss of semantic knowledge as seen in individuals with semantic dementia (Patterson et al., 2007). On the other hand, when damage involves regions typically activated during the controlled use of semantic knowledge (e.g., the inferior prefrontal cortex, posterior middle temporal gyrus, intraparietal sulcus), inhibition problems are commonly observed (Noonan et al., 2013). That is, difficulty ignoring irrelevant semantic information, difficulty in retrieving semantic associations that are less automatic, and difficulty in understanding highly ambiguous words (i.e., when a higher level of controlled problem solving is required –

Noonan et al., 2010; Hoffman et al., 2011). In support of these findings, Whitney et al. (2011) used transcranial magnetic stimulation in healthy control participants to disrupt processing in two brain areas thought to be involved in semantic cognition: the inferior frontal gyrus and posterior middle temporal cortex. Stimulation of both brain areas was found to particularly disrupt the processing of semantic judgements that were executively demanding, but semantic judgements that involved strong automatic associations were not affected by the stimulation. Further, Whitney et al. also found that performance in non-semantic tasks was not affected by the stimulation, regardless of how executively demanding the tasks were. Therefore, these findings are consistent with those mentioned above (e.g., Patterson et al., 2007; Pobric et al., 2007; Noonan et al., 2013) and suggest that prefrontal and posterior temporal regions support semantic control.

Further, evidence also suggests that the brain regions activated during tasks/activities that demand significant semantic control, exhibit some degree of overlap with those involved in the "multiple demand" network that responds to increased control demands across various cognitive domains (Duncan, 2010; Fedorenko et al., 2013). Tasks engaging the "multiple demand" network typically involve attentional control, working memory, inhibition, task switching, monitoring, and response selection. Such overlap suggests similar processes may be involved in the control of semantic knowledge and executive control in other domains (Hoffman, 2018). Indeed, engaging in tasks that require significant semantic control entails integrating language processing, conceptual knowledge, and semantic memory to solve problems or make decisions. Consequently, the demands related to the multiple demand network noted above (e.g., inhibition, monitoring, response selection) are therefore required during tasks that place high demands on semantic control. Supporting such a view, functional magnetic resonance imaging (fMRI) studies have shown that semantic control can be divided

into two components which activate neighbouring areas of the left inferior prefrontal cortex – the semantic-specific and the domain-general (Badre et al., 2005; Nagel et al., 2008).

The first (*semantic-specific*) is a controlled retrieval mechanism that is used when automatic activation of semantic knowledge is not enough to complete the task at hand. For example, when completing a semantic association task, participants may be asked whether 'bee' is associated with 'tree' or 'flower' (Hoffman, 2018). Automatic activation of the cue may lead to the activation of things that are most associated with bees (honey or hive) but not the response options available. In this situation, and when relevant past experiences or facts about the world do not come to mind automatically, participants may adopt a goal-directed controlled search to find the required information from the semantic store (Badre & Wagner, 2007). This controlled process is considered to be specific to semantic knowledge because functional neuroimaging studies demonstrate that when higher control is required, activation in the anterior, ventral portion of the left inferior prefrontal cortex is increased (Badre et al., 2005; Dobbins & Wagner, 2005; Krieger-Redwood et al., 2015) – an area connected to the regions where semantic representations are coded (Jackson et al., 2016; Von Der Heide et al., 2013).

The second (*domain-general*) is a semantic selection mechanism which is used when automatic activation of semantic knowledge leads to competition/conflict between numerous representations. That is, when more associated information comes to mind than what is required or relevant to complete the task (Badre & Wagner, 2007; Hoffman, 2018). The Feature association task is routinely used to study this mechanism. In brief, participants are asked to match concepts based on particular shared properties (e.g., colour or size). For example, Hoffman (2018) used the example – whether a wasp or hive is the same size as a bee. In this instance, the two response options are similar and are semantically related with the cue, meaning that they compete for selection. Thus, semantic decisions are made in the

context of having to contend with a complex semantic problem-solving situation against responses that must be ignored/inhibited to successfully select the target-correct response.

In general, when numerous representations come to mind, a process of semantic selection is required to eliminate/inhibit the competition between the representations and to allow the required representation to influence decision and action (Fletcher et al., 2000; Moss et al., 2005). This means that other irrelevant semantic links that are not required in that particular context need to be inhibited (e.g., bee-hive) (Hoffman, 2018). Of course, the need to inhibit certain responses to complete a goal/task is also needed in other areas of executive function, not only semantic processing (Hoffman, 2018). For instance, the area of the brain that is activated during increased semantic selection demands on semantic tasks, the posterior portion of the left inferior prefrontal cortex (Badre et al., 2005; Thompson-Schill et al., 1997), is also activated by high competition in working memory tasks (Badre & Wagner, 2005), the Stroop task (January et al., 2009), and various other inhibition tasks (Hazeltine et al., 2003). It is assumed, therefore, that instances of decisions made under *high semantic control*, may also be strongly connected with the 'domain-general' multiple demand network (Jung et al., 2017). In short, this suggests that certain, and largely difficult, semantic selection decisions may be linked to a domain-general executive selection system (Hoffman, 2018).

To recap, the literature suggests that semantic representations and control are governed by two different but interacting brain regions (Jackson, 2021). Semantic control is also suggested to be multifaceted, comprising semantic-specific processes used in controlled retrieval of information from memory, as well as domain-general selection mechanisms (Jackson, 2021; Hoffman, 2018). Importantly, it is possible that these mechanisms may be differentially impacted by the ageing process and therefore, contribute to age related changes in CR. This chapter will next discuss the crucial work that indicates evidence for the impact of ageing on

these two key SC processes – research that acted as the inspiration and underpinning framework for the current study.

3.1.5 Hoffman's (2018) Aims

The first study to investigate differences in these key semantic abilities in young and older adults, as well as relationships between semantic knowledge, semantic control and 'domain-general' executive function, was conducted by Hoffman (2018). First, Hoffman attempted to determine the extent to which the 'depth' and 'breadth' of semantic knowledge (i.e., the content and sophistication of semantic memory) has an impact on performance on semantic control tasks.

The aforementioned neuroimaging literature (e.g., Pobric et al., 2007; Noonan et al., 2013; Whitney et al., 2011) suggests that the amount of information one has in memory and the ability to use it effectively, in a goal-directed manner, may be independent of each other (i.e., they may be governed by different but interacting brain regions). On the other hand, findings from research exploring episodic memory suggests a possible negative relationship between the amount of semantic knowledge in memory and the capacity to retrieve information from memory (Hoffman, 2018). For example, Anderson (1974) argued that with the accumulation of more knowledge, individuals then also have more concepts and propositions to search to reach required information. That is, those who have a larger semantic store may experience increased interference between competing representations which may lead to more demand on the control mechanism. Consistent with this line of thinking, ageing could therefore be expected to have an impact on both the size of the semantic store, as well as the way in which such populations control this knowledge.

Second, Hoffman (2018) also investigated the relationship between semantic control and 'domain-general' executive control in other areas of cognition in healthy young and older adults. Some studies involving stroke survivors experiencing semantic impairments have

found a positive correlation between semantic control ability and performance in non-semantic tests of executive function (Allen et al., 2012; Jefferies & Lambon Ralph, 2006; Hoffman, 2018). However, for stroke survivors with large lesions spanning more than one site (i.e., parietal, frontal), findings may reflect greater impairment in numerous, independent systems (Hoffman, 2018).

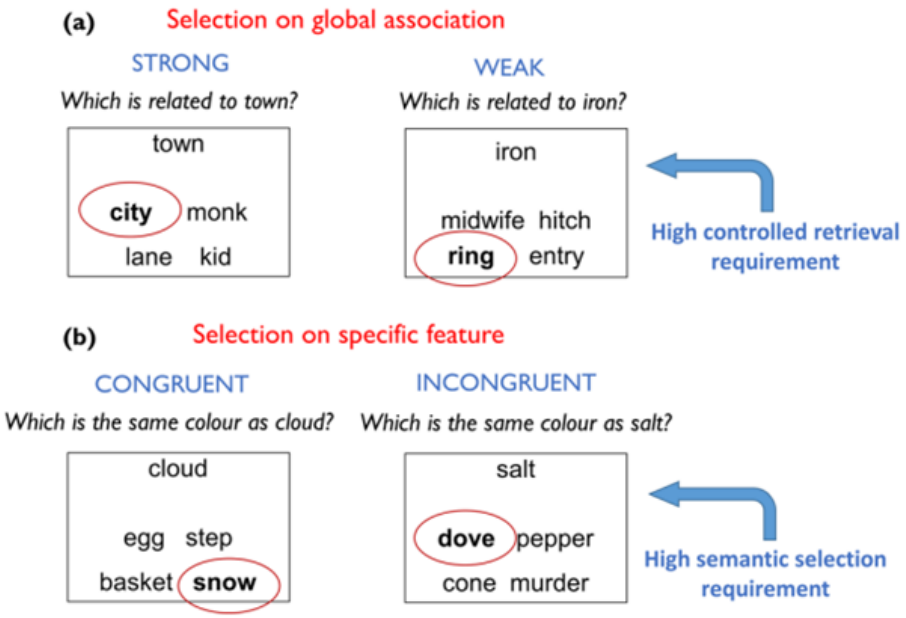
Finally, Hoffman also examined the impact of healthy ageing on semantic abilities. In contrast to other cognitive abilities, semantic tests are typically seen to be relatively stable across adulthood (Nyberg et al., 1996; Salthouse, 2004a). However, meta-analysis indicates how older adults (over 60 years) actually perform significantly better on tests of vocabulary, used as a measure of semantic ability, than young adults below 30 years of age (Verhaeghen et al., 2003). One explanation for this finding is the continuing growth of the store of semantic knowledge throughout life (Hoffman, 2018). That is, because older adults have lived longer and had more experiences, they may have a higher crystallised intelligence (Salthouse, 2004b). However, Hoffman (2018) argues that vocabulary tests do not measure the impact of ageing on semantic control abilities as they do not probe aspects of semantic processing that are executively demanding. Rather, they provide an indication of the *breadth and richness* of the store of semantic information. Therefore, it may be that the *controlled* use of semantic information declines with age.

3.1.6 The Semantic Control Paradigm – Probing Different Semantic Control Mechanisms

Having established the rationale underpinning the idea that mechanisms of semantic control may be differentially impacted by ageing, this section outlines the experimental task used in the current study to probe different semantic control mechanisms. First established by Badre et al. (2005), and then subsequently used by Hoffman (2018) and further validated

(feature task items) by Wu and Hoffman (2022), the key elements of this semantic control task are presented in Figure 9 below.

Figure 9. *Key Elements of the Semantic Control Task (Hoffman, 2018 – original article under the CC BY 4.0 International licence)*



In the Global association task, participants are asked to make semantic decisions based on *Global* semantic association (section 'a' in Figure 9) (Hoffman, 2018). A probe word is presented, and participants are asked to select its semantic associate from either two or four options (a task akin to synonym judgement). The associate is either strongly (e.g., town-city) or weakly associated (e.g., iron-ring) with the probe (Hoffman, 2018). In the other element of the task, participants are required to make decisions based on specific *Features* like colour and size (section 'b' in Figure 9). A key aspect of both elements of this task is that the "weak" condition in the *Global* task and the "incongruent" condition in the *Feature* task are both instances where *High* semantic control is required. That is, they are assumed to be more demanding with respect to semantic control because making a decision involves a higher

level of problem solving compared to the global-strong condition (which is more automatic and requires less semantic control), and often the inhibition of a stronger or more easily accessed response. In addition, it is assumed that the Global association element, being a more 'natural' semantic problem-solving task (akin to thinking about the synonyms of a particular word), would also be somewhat easier in general than the feature-based task.

From the four conditions (Global strong, Global weak, Feature congruent, Feature incongruent), it is also assumed that the *incongruent* trials on the specific *Feature* task would be the most challenging because one of the foils is highly associated (semantically) with the target item (e.g., salt-pepper) (Hoffman, 2018). This means that the participant must ignore/inhibit the response that may come to mind automatically (e.g., pepper). Therefore, the incongruent trials on the specific Feature task have two elements, with participants required to: (a) look for the option that is associated with the probe based on a specific feature (colour or size) and, (b) inhibit highly automatic semantic information that is not consistent with the feature. Again, it is assumed that this task requires a high level of semantic control.

Therefore, there are two kinds of intertwined elements to this semantic selection/decision process. One is moderated by how strong the semantic link between things might be, and the other is a process of inhibition in instances when the irrelevant items (i.e., the foils) may interfere with the required decision (because one of them might have a strong semantic link with the target). In the *incongruent* conditions on the specific *Feature* task, both of these things are occurring.

3.1.7 Findings from Hoffman (2018)

Hoffman (2018) was interested in determining whether the interaction between the kind of task (Global and Feature association) and the two key levels of semantic control (High vs Low – note, in one case [Global task] high/low = Strong/Weak associations and in the other

case [Feature task], high/low = Congruent/Incongruent) had implications across age groups (younger - 18-30 years and older adults - 61-91 years). In addition, a variety of other cognitive measures (explained further in the Methods section) were used to probe the degree to which semantic specific and 'domain-general' mechanisms are implicated in performance on elements of this paradigm. In terms of the Global association task (see 'a' examples in Figure 9 above, and Methods section below), Hoffman's (2018) findings demonstrated that the manipulation of control (High/Low – or Strong vs Weak trials, in Hoffman's terms) affected both age groups similarly for reaction times (i.e., the difference in the drop in performance from low to high control across both groups was similar), but older adults were more accurate (i.e., higher number of correct responses). This implied that for the Global association task, older adults were less affected by trials with High control demands (weak pairs). In short, the negative effect of High control was moderated by age to the benefit of older adults. One reason for this might be that on the Global association task, the likely higher crystalized knowledge (i.e., greater vocabulary knowledge) of older adults allows them to meet the task requirements more effectively than younger participants (Park et al., 2010; Spreng & Turner, 2019). In sum, when the semantic task involves making Global associations (akin to synonym judgement), older adults have an advantage based on their longer lifetime experience acquiring a breadth of semantic knowledge.

In the Feature association task (see 'b' examples in Figure 9 above, and Methods section below), the reverse is demonstrated. That is, older participants performed more poorly on the incongruent (High control) condition compared to the younger participants. Therefore, in one task (Global association) the interaction with age shows an *advantage* to older adults and in the other (Feature association), the interaction with age shows a *disadvantage* for older adults. Consequently, Hoffman's (2018) research confirmed that different forms of semantic control are differentially impacted by ageing. Specifically, older adults appear to have an

advantage or disadvantage depending on the nature of the task and its level of semantic control (i.e., a three-way interaction). A possible explanation for this older adult finding in the High control trials of the Feature task, is that inhibition seems to be adversely affected by ageing (Askey & Playfoot, 2018; Burke & MacKay, 1997; Kramer et al., 1994). For example, The Inhibition Deficit Hypothesis proposes that ageing impacts negatively on the inhibition of irrelevant information (Burke & MacKay, 1997). That is, older adults may be less able to overcome dominant responses or to ignore distracting information when performing tasks (Rey-Mermet & Gale, 2018). Using this hypothesis then, High control trials of the Feature task, where participants must ignore a foil that is highly associated with the probe and target and focus on the relevant information (e.g., Colour or Size), may therefore be more challenging for older adults.

Next, and having established the key pattern of an interaction between the consequences of high semantic control across tasks and age groups, Hoffman (2018) sought to understand whether performance on these two key high semantic control conditions was linked to performance on a number of key cognitive tests. These tests were selected to determine whether executive 'domain-general' tasks (i.e., planning/selection/cognitive processes assumed to involve the frontal lobes), or tasks tapping semantic breadth (i.e., how sophisticated one's semantic lexicon may be), might be related to performance on the elements of the semantic control paradigm described above – particularly performance on the two conditions that imply high semantic control (i.e., the incongruent condition in the Feature task and the weak association condition in the Global association task).

Hoffman (2018) found that semantic selection abilities were linked to performance on a non-semantic executive function task (the Wisconsin Card Sorting task), in both young and older adults. With respect to semantic breadth, older adults were found to have larger stores of semantic knowledge (i.e., greater crystalised knowledge) and performed better on tests of

knowledge for low frequency words (Mill Hill task) compared to the younger adult group. Combined, these findings suggest that the following contribute to SC processes: (a) a store of representations which increases throughout life and is thus positively impacted by ageing (linked to the concept of crystallised intelligence discussed earlier); (b) a controlled retrieval process unaffected by age which scans this 'store' of representations and is sensitive to the degree to which concepts being retrieved may require a high degree of semantic control (i.e., weakly associated concepts), and (c) a selection process that is more meta-cognitive, which seems to be where negative age-related change occurs.

3.1.8 Potential Effects of Bilingualism on Hoffman's Paradigm and Findings

Considering that the bilingual lexico-semantic system may be different from that of monolinguals, bilinguals' performance on the semantic control task may be impacted by the way in which they process and store the meaning of words. As mentioned earlier in this chapter, research suggests that both languages of a bilinguals are active to some extent, and the level of this activation may be affected by factors such as language proficiency (Altarriba & Canary, 2004; Basnight-Brown & Altarriba, 2007). This suggests that bilinguals may experience greater interference whilst completing the semantic control task due to competing lexical and semantic information from the other language.

As noted previously, findings from bilingual performance on the Stroop Task may offer insights into the bilingual effects in the semantic control task. Roselli et al. (2002) found that Spanish-English bilinguals performed similarly across both Spanish and English versions of the Stroop Task but were slower overall compared to monolinguals. For balanced bilinguals, the language in which the Stroop Task was administered did not significantly affect performance, whereas unbalanced bilinguals performed better when the task was conducted in their most proficient language. This may suggest that balanced bilinguals are better

rehearsed at controlling both languages and inhibiting competing information from their other language. That is, if both languages of a balanced bilingual are more active than those of an unbalanced bilingual (due to higher proficiency) (Altarriba & Canary, 2004; Basnight-Brown & Altarriba, 2007), balanced bilinguals may be better able to control their languages and more efficiently manage the interference from the other language.

Suarez et al. (2014) also found that higher proficiency in a second language leads to an advantage in suppressing competing information from the native language. Similarly, Zied et al. (2004), mentioned previously, found that balanced bilinguals were faster to respond across all Stroop conditions compared to unbalanced bilinguals, suggesting superior interference control. Given the interconnected lexico-semantic processes of the bilingual brain and the notion that the languages of balanced bilinguals are simultaneously active to a greater extent, balanced bilinguals may achieve correct answers more rapidly, benefiting from the dual confirmation provided by both languages. In other words, if the meaning of a word occurs to balanced bilinguals in both of their languages to a greater extent than unbalanced bilinguals, it is logical to infer that they can more quickly confirm the correct answer.

While there is some literature on the bilingual lexico-semantic system, research on SC in bilingual populations is lacking. Existing literature on SC and the organisation of semantic memory is primarily based on studies on the monolingual brain (Hoffman, 2018; Riehl, 2020; Wu & Hoffman, 2022). Therefore, it is unclear how the findings of Hoffman (2018) may generalise to bilingual populations. The only studies that offer some insight into the performance of bilinguals on the semantic control task are those investigating the effects of the bilingual lexico-semantic system on the Stroop Task. However, while both tasks require individuals to selectively attend to specific information while inhibiting interference from irrelevant stimuli, the Stroop Task focuses solely on colour and does not encompass the

extensive breadth and depth of semantic knowledge examined in the semantic control task. Therefore, it is unclear whether the mechanisms used in both tasks are the same.

Based on the findings from the literature on the Stroop Task, it could be hypothesised that bilinguals may be slower to respond on the semantic control task than monolinguals, due to increased interference from having knowledge of a second language. Conversely, it is also possible that balanced bilinguals, who are proficient in both languages, may be better able to inhibit interference and thus perform comparably on the semantic control task to monolinguals (Suarez et al., 2014). Additionally, in situations requiring increased control, such as incongruent trials in the Feature task of the semantic control task, balanced bilinguals or those highly proficient in both languages may have an advantage because of their superior inhibitory abilities (Suarez et al., 2014).

In relation to ageing, Hoffman's (2018) study confirmed that the ageing process appears to have varying effects on different forms of semantic control within monolingual populations. For example, older participants performed more poorly on the High control trials of the Feature task compared to the younger participants. As mentioned previously, one reason for this may be the impaired inhibition abilities that have been found to accompany the ageing process (Askey & Playfoot, 2018; Burke & MacKay, 1997; Kramer et al., 1994). With this in mind, the idea of inhibiting responses to reach a goal is highly relevant to bilinguals because they routinely manage language information *conflict,* inhibit responses from one language to successfully use another, and may also need to switch between languages which increases cognitive demand (Bialystok et al., 2004).

This is relevant here because if bilinguals are well rehearsed at inhibiting unwanted information in day-to-day life, then it may be assumed that this may impact the system operating on semantic information, which may increase one's level of CR. That is, bilinguals may differ with respect to the way in which they retrieve, manipulate and store semantic

information, meaning that their metacognitive system may be more efficient to some degree. Consequently, this might be a fruitful domain to explore as a proxy for a possible bilingual advantage.

In terms of the potential impact of bilingualism on the semantic control task, bilinguals have significantly greater reserves of semantic knowledge as they speak more than one language. Therefore, bilinguals might be expected to perform better than monolinguals on tasks that require this controlled retrieval process *regardless of age*. In addition, Hoffman's (2018) study indicated that older adults perform more poorly on the High control trials of the Feature task, and this is interpreted as reflecting greater difficulties with dealing with the irrelevant information signalled by the highly associated foil items (i.e., in Figure 9 – one must ignore the highly associated word 'pepper' in favour of the target item 'dove'). This has interesting implications with respect to how older bilinguals may perform, because if older bilinguals have more efficient inhibition abilities, then they might better cope with these High control trials than monolingual older adults. Importantly, semantic selection in this context is argued to be linked to domain-general executive tasks, and thus again, if an older adult bilingual advantage is observed in this context, it would suggest measures of semantic cognition may be a suitable 'proxy' for capturing the bilingual advantage and testing CR.

3.1.9 Study Aims

Considering the above, the aim of the current study was to replicate the work of Hoffman (2018) while also examining the impact of language status (i.e., bilingual/monolingual) on semantic control to provide insight into the role of Welsh-English bilingualism in cognitive ageing, specifically in relation to SC and CR. On the whole, this study sought clarity on: (a) whether those who know more (bilinguals) have better control over their knowledge; (b) whether aspects of cognitive control over semantic information are strongly associated with

executive (domain-general) control ability in other areas, and (c) whether the consequences of ageing on semantic representation and control are moderated by being bilingual (that is, the key three-way interaction described above may interact with bilingual status). To the author's knowledge, no such investigation has been conducted previously.

It was hypothesised that:

1. There would be a main effect of age, with older adults generally being slower but more accurate.
2. There would be a main effect of task. That is, the Global task would generally be easier because the Feature task contains a condition that is extremely challenging (i.e., incongruent trials) and the Global task is also more akin to familiar semantic decisions such as synonym judgement.
3. There would be a main effect of control. That is, judgements demanding high semantic control would be more challenging (i.e., strong > weak on the Global task and Congruent > Incongruent on the Feature task).
4. There would be a three-way interaction between task x control x age. This would reflect the following pattern: On the Global task, young participants would fare worse for High control trials (i.e., older adult advantage). On the Feature task, the High control pattern would be exaggerated for older adults compared to younger adults (i.e., the impact of age would 'flip').
5. Several of the above patterns would be influenced by language status due to the potential impact of bilingualism on the processes related to SC and the development of CR.

3.2 Methods

3.2.1 Ethical Approval Procedure

Ethical approval was obtained from the ethics committee of the Psychology Department at Swansea University. One ethics application was submitted for study two (Chapter 3) and study three (Chapter 4) as they were conducted concurrently and treated as a unified study (Appendix BB). Participants read a participant information sheet before agreeing to take part.

3.2.2 Participants

The study was advertised as 'Bilingualism & The Rumsfeld Effect' because two studies (the current study and the study presented in Chapter 4) were run together as one larger study. Participants were drawn from the general adult population and were recruited via social media adverts, by contacting various organisations/groups (e.g. University of the Third Age), via the researcher's own social and personal networks, by word of mouth (e.g. via participants), the Swansea University Dementia Research Group's registered participant database, via media interviews (e.g., radio) about the research, by the Swansea University Psychology Participant Pool, and from the initial group of participants recruited for the first study of this thesis. Participation was on an entirely voluntary basis, and initially there were no participant compensation arrangements in place. However, due to recruitment challenges, compensation arrangements were introduced part way through the study and 37 participants were reimbursed with a £15 Amazon voucher for their time.

Individuals were eligible to participate if they were bilingual (Welsh – English) speakers, with Welsh being their first language OR English monolingual and not fluent in any other language. Participants also had to be 18-35 (it was initially 18-30 but due to practicalities and recruitment challenges, this upper age limit was increased to 35) or 61-91 years of age. Participants also had to report no history of neurological or psychiatric illness, or substance abuse.

3.2.3 Data Exclusions

In total, 105 participants took part in the study. However, 13 participants were subsequently excluded because they withdrew part way through the study, leading to partial/incomplete data. After these exclusions, the final sample consisted of 92 participants (42 bilinguals, 50 monolinguals). See Table 19 for standard demographic characteristics.

Table 19

Summary of Demographic Characteristics

	n	***M* Age (*SD*)**	**Age Range**	**Gender**	
				Female n (%)	Male n (%)
Older Bilinguals	22	67.1 (4.75)	61-78	16 (73%)	6 (27%)
Young Bilinguals	20	25.55 (5.58)	18-35	16 (80%)	4 (20%)
Older Monolinguals	23	68.04 (4.84)	61-79	13 (57%)	10 (43%)
Young Monolinguals	27	23.37 (4.13)	19-32	17 (63%)	10 (37%)

Note: Age did not significantly differ across the young monolinguals and bilinguals, or across the older monolinguals and bilinguals (p >.05). The proportion of males and females was also similar across the four groups (p >.05)

In terms of participants' bilingual status (LexTALE; see Table 20), the older bilingual group scored comparably to the older normative sample reported in the first study of this thesis (Chapter 2), and higher compared to other normative data. The younger bilinguals scored lower than the older bilinguals in this study but performed comparably to the sample in Chapter 2 and other normative samples (normative data was obtained from an unpublished

PhD thesis – Kyle Jones, Swansea University). Thus, participants' self-reported bilingual status was confirmed.

Table 20

Performance Comparison on the Welsh LexTALE across Current and Normative Data

	Older Bilinguals	Young Bilinguals	Older Adults Study 1	Welsh Normative Data	
n	22	20	70	44	22
Age Range (years)	61-78	18-35	55-84	40-69	70-87
M Accuracy (%)	81	68	70.14	65.95	69.32
SD (%)	10.49	11.56	10.75	12.21	9.93

Note: normative data was obtained from an unpublished PhD thesis – Kyle Jones, Swansea University

3.2.4 Materials

In addition to the Lexical Test for Advanced Learners of English (LexTALE; Welsh and English versions) which has been described previously (see Chapter 2, Section 2.2.4.3), participants also completed the following tasks in English.

3.2.4.1 General Cognitive Tasks

In line with Hoffman (2018), the present study included several cognitive tasks to determine the degree to which performance on key semantic control conditions within the semantic control paradigm may be linked to other cognitive processes. This was in order to determine whether executive 'domain-general' tasks (i.e., planning/selection/processes that apply to both language and non-language tasks), or tasks that examine semantic breadth (i.e., how sophisticated one's semantic lexicon might be), might be related to performance on the elements of the semantic control paradigm described above. Particularly, performance on the

two conditions that require *high* semantic control (i.e., the weak condition in the global association task, and the incongruent condition in the feature association task). Tasks tapping into 'domain-general' executive control were:

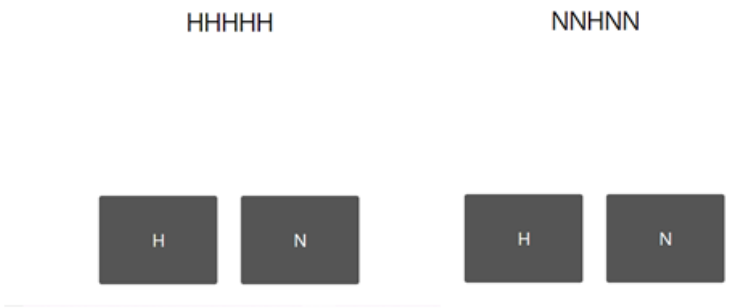
3.2.4.1.1 Verbal Fluency Tasks

Three categories of verbal fluency were administered via Zoom video call. Participants were given one minute to produce as many English words as possible that fit a specific criterion. The criteria included two semantic categories (animals and household objects) and one letter of the alphabet (words beginning with 'F'). Verbal fluency tasks require participants to access verbal-semantic knowledge, but they also place demands on executive functions (Hoffman, 2018; Henry et al., 2004). They require an organised retrieval process, as well as the self-monitoring elements of cognition because the participant must remember the responses already given as well as inhibiting these responses if they come to mind for a second time (Henry et al., 2004).

3.2.4.1.2 Eriksen Flanker Task – Executive Function (Eriksen, 1995)

Replicating the procedure employed by Hoffman (2018), participants were presented with a row of five letters and were required to click a button on screen to indicate whether the central letter was a H or an N. Participants were instructed to ignore the flanking letters at each side of the central target letter (H or N). The task consisted of 60 trials, 30 congruent and 30 incongruent. No practice trials were presented. In congruent trials, the flanking letters were identical to the target letter (e.g., HHHHH), whereas in the incongruent trials, the flanking letters were inconsistent with the target (e.g., NNHNN) (See Figure 10). Consequently, during incongruent trials, participants need to address the conflict arising from the two potential responses and make a decision (Hoffman, 2018).

Figure 10. *Congruent and incongruent trials of the Eriksen Flanker Task. Figure created by Hanna Thomas using the Gorilla Experiment Builder*

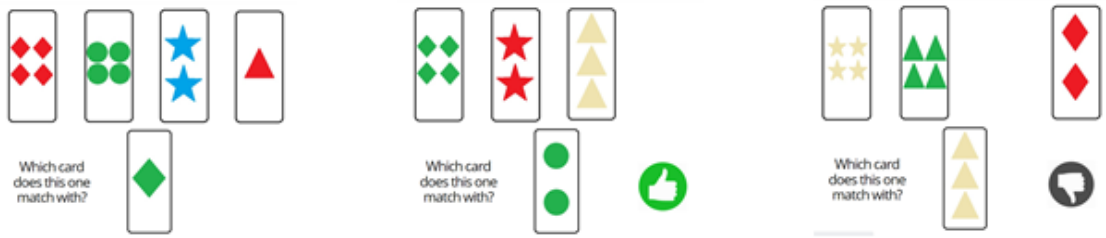


3.2.4.1.3 Adapted version of the Wisconsin Card-Sorting Task (WCST) – Complex Executive Function

The WCST is widely used in neuropsychological evaluations to assess abstract reasoning and cognitive flexibility in response to changing environmental contingencies (Grant & Berg, 1948). In the current study, an adapted and shorter version was used – a fully accessible and computerised version available via Gorilla Experiment Builder (www.gorilla.sc). This version retains all the features of the original WCST, is psychometrically comparable to the standard WCST (e.g., Smith-Seemiller et al., 1997), and contains the same number of response cards as used by Hoffman (2018) although there is a difference in the way that participants progress through the task as explained below. The task consists of four reference cards and a set of 64 response cards that show one to four identical patterns (stars, circles, diamonds and triangles) all in the same colour (red, blue, green and peach). Participants were required to match each of the 64 response cards to the four reference cards by clicking the card that matched the reference card, from a choice of four at the top of the screen, based on number, shape or colour. The rule changed after every 4th card (as opposed to changing after 10 consecutive

correct responses, in Hoffman 2018) and feedback was given in the form of a 'thumbs up' or 'thumbs down' sign on the screen to help participants to deduce the sorting principle (Eling et al., 2008) (See Figure 11 below for diagram and a note to explain each image). To be successful therefore, participants are required to learn and consistently use a correct sorting principle, but then flexibly shift to another sorting principle when the previous rule is no longer correct. No practice trials were included in this task.

Figure 11. *WCST Stimuli and Feedback. Fully accessible image of the WCST from Gorilla Experiment Builder (www.gorilla.sc)*



Note: The first image (left) shows the trial before the participant makes a selection, with four cards displayed on the screen. The second image (centre) shows the 'thumbs up' feedback after a correct response, where the chosen card disappears upon selection. The third image (right) displays the 'thumbs down' feedback after an incorrect response, with the chosen card also removed from the options upon selection.

Please note that the Trail-Making Task was also used by Hoffman (2018) as an additional measure of executive control, but due to the online administration format of the present study, it was not possible to include this task.

3.2.4.2 Breadth of Semantic Knowledge

All participants were asked to complete two tasks that aimed to examine the size of their store of semantic representations:

3.2.4.2.1 Spot-the-Word Task (Lexical Decision)

Taken from the Speed and Capacity of Language Processing battery (Baddeley et al., 1992), the Spot-the-Word task is a short test used to identify cognitive slowing due to brain injury/disease. It consists of 60 trials (plus 15 practice trials), with each presenting a pair of letter strings (Saxton et al., 2001). On each trial, participants were presented with a real word and a non-word and were asked to select the real word (Hoffman, 2018; Figure 12). Non-words were similar to real words in terms of orthography and phonology, meaning that participants had to rely on semantic knowledge to make a decision (Hoffman, 2018). Thus, as the task requires participant to use crystallized verbal knowledge (e.g., vocabulary or information about the world) to complete the task, it is less likely to be impacted by the effects of ageing (Saxton et al., 2001).

Please note that investigation into false friends between Welsh and English spellings was conducted after data collection and analysis. Only one false friend was identified: the English word ‘clod’, meaning a lump of earth, while in Welsh, ‘clod’ denotes laudation, commendation, praise or credit. Please refer to the limitations section in Chapter 5 for a discussion on how researchers could reduce the risk of cognates in future studies.

Figure 12. *Spot-the-Word Stimuli Example. Figure created by Hanna Thomas using the Gorilla Experiment Builder*

3.2.4.2.2 Mill Hill Vocabulary Test - modified version from Hoffman (2018). (Synonym Judgement)

This is a multiple-choice task in which participants must select the synonyms of certain words (Raven & Court, 1998). This study used the Senior Form 1, Set B synonyms task, where participants were presented with 33 reference words, one at a time, with a choice of 6 response options at the bottom of the screen. No practice items were included in this task. In the present study, one of the six options had the same meaning as the reference word and participants were required to click on that word (Figure 13). Unpublished data from a previous study by members of the supervisory team for this thesis indicated that ceiling effects were unlikely in this task.

Figure 13. *Mill Hill Stimuli Example. Figure created by Hanna Thomas using the Gorilla Experiment Builder*

It is important to note that Hoffman (2018) used a version of the task containing 44 target words, combining the hardest 22 trials from Set A and Set B in order to increase task difficulty and prevent ceiling effects. Hoffman (2018) also reduced the number of response options on each trial to four. Therefore, and in spite of the fact that the same cognitive processes are being assessed, the group mean performance in the current study is not directly comparable to Hoffman (2018).

After data collection and analysis, the stimuli were examined for false friends between Welsh and English spellings. No false friends were identified, which was also the case for the Test of Semantic Control presented next.

3.2.4.3 Test of Semantic Control

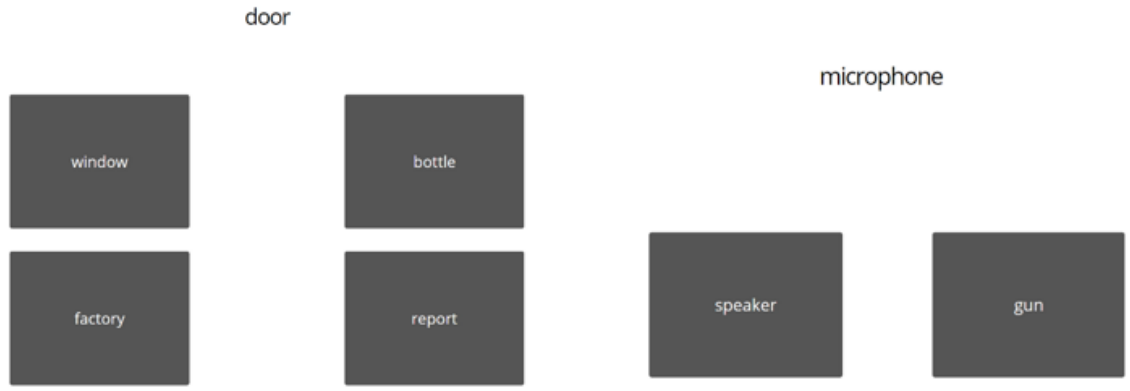
Using the same design employed by Badre et al. (2005) and Hoffman (2018), participants completed a task consisting of eight blocks of 12 trials examining two forms of semantic control. No practice trials were included in this task. In four blocks, participants were asked to make semantic decisions based on g*lobal* semantic associations (Hoffman, 2018). Participants were shown a probe word and tasked with choosing its corresponding semantic associate from a set of either two or four response options (Figure 14). The associate was either strongly associated with the probe (e.g., table-chair) or more weakly associated (e.g., apricot-stone), with the 24 weak (high control) tasks envisaged to create conditions that required participants to exert greater effort in retrieving semantic information (i.e., because automatic spreading of activation in the network would be insufficient to identify the correct response).

Figure 14. *Examples of Global Semantic Association Trials with Four and Two Response Options. Figure created by Hanna Thomas using the Gorilla Experiment Builder*

In the other four blocks, participants were asked to select items that matched on particular features: Colour or Size. Thus, these trials place high demands on semantic control as attention needs to be diverted away from global semantics associations in favour of focussing on specific item features. At the beginning of each block, participants were asked to focus on colour or size (i.e., a specific item feature). A probe was then presented, and participants had to select from a choice of either two or four options, which item was most like the specified feature (Figures 15 and 16). In total there were 24 congruent and 24 incongruent trials. In the congruent trials, the probe and target item were similar on other features, in addition to matching on the specified feature (e.g., tea-coffee). On these trials, the foils *did not* match with the specified feature, and they also did not share a global semantic association with the target. Consequently, these trials did not place high demands on the mechanisms of semantic selection (and involved less inhibition) as the correct response reflected the response option that had the strongest relationship with the target. In contrast, the incongruent trials consisted of a probe and target that shared no meaningful association other than matching on the specified feature (e.g., crow-blackboard). However, one of the foils *did* have a strong relationship with the target, even though it did not match on the currently relevant feature (e.g., salt-pepper). Therefore, such trials place significant demands on the selection mechanisms because there is no pre-existing semantic relationship between the probe and the target, and because participants need to ignore the strong but irrelevant relationship between the probe and foil (Hoffman, 2018).

Figure 15. *Examples of Colour Trials with Four and Two Response Options. Figure created by Hanna Thomas using the Gorilla Experiment Builder*

Figure 16. *Examples of Size Trials with Four and Two Response Options. Figure created by Hanna Thomas using the Gorilla Experiment Builder*



3.2.5 Design

This study used a 2x2x2x2 mixed design. There were four IVs in this study. Two IVs were within subjects: (a) kind of task (Global association or Feature task) and (b) semantic control required (High/Low), and two IVs were between subjects: (a) age (older/younger adults), and (b) language status (bilingual/monolingual. The DVs were the participants' accuracy and reaction times on the cognitive tasks. Participants were allocated to groups based on their age and language status.

3.2.6 Procedure

Both sessions were completed online via Gorilla. Participants were sent a participant information sheet via email and were asked to complete a consent form prior to completing any tasks (Appendix U and V). Bilingual participants were sent links via email to the Welsh and English LexTALE and were asked to complete a digital consent form which appeared on the screen before each LexTALE (*note:* the bilingual older adults (n = 10) who also completed the first study of this thesis were not asked to complete the LexTALE again). All groups were also sent a link via email to the semantic cognition tasks, which also included a digital consent form. Participants successively completed a concepts task related to the third study of the thesis first (see Chapter 4), followed by the general cognitive assessments, tests of breadth of semantic knowledge, and the semantic control task. Participants also completed the concepts task related to the third study of the thesis for a second time at the end of the session (see Chapter 4). Participants then completed the concepts task related to the third study of this thesis a week later (i.e., their second session on Gorilla). Zoom video calls were used to collect verbal fluency data, which took place between or any time after the two sessions on Gorilla. In terms of Gorilla itself, it was possible to see whether a participant was 'live' on the tasks and how long they spent completing each task.

3.3 Results

This study aimed to replicate the findings of Hoffman (2018) and to investigate the effect of bilingualism on semantic cognition.

3.3.1 General Cognitive Assessments (Effects of age/bilingualism)

Consistent with Hoffman (2018), data was collected for several key general cognitive measures falling across two broad sub-types:

1) ‘Domain-general’ executive control – processes of control that are expected to be relevant across a variety of decision-making contexts. In this study, these included: (a) the Eriksen Flanker task; (b) the WCST; and (c) verbal fluency tasks including semantic (animals and household objects) and letter fluency (letter F).
2) Tasks that tap ‘breadth of semantic knowledge,’ including the Spot-The-Word (lexical decision task) and adapted Mill Hill task (synonym judgement task). That is, processes that act as a proxy of individual semantic memory capacity/complexity.

3.3.2 Differences in Performance on General Cognitive Tasks and Breadth of Semantic Knowledge

This section will determine whether cognitive performance differed according to age (young versus old) or language status (bilingual versus monolingual). See Table 21 for descriptive statistics for both ‘domain-general’ and ‘breadth of knowledge tasks.

Factorial 2 (age) x 2 (language status) between-participants ANOVAs were carried out to determine whether there were any significant differences in the average performance on each task according to age and language status. The only task in which there was a significant main effect of language background was the Spot-the-Word task, where monolingual participants were significantly more accurate, $F(1, 88) = 4.89$, $p = .030$, $\eta p^2 = .036$. There was also a significant main effect of age in this task, $F(1, 88) = 39.48$, $p < .001$, $\eta p^2 = .294$, such that older adults were more accurate. There was no significant interaction between age and language status on this task, $F(1, 88) = 1.74$, $p = .190$, $\eta p^2 = .013$. The assumption of equal variance was violated in the Spot-the-Word task meaning that the risk of type 1 error is increased. Caution is needed when interpreting the results relating to this task. The assumption of normal distribution was also violated in the Spot-the-Word task (see Appendix W for descriptives table).

There were also significant main effects of age on the number of words beginning with F that participants generated in the verbal fluency task, $F(1, 88) = 13.84$, $p < .001$, $\eta p^2 = .134$, and on correct responses to the Mill Hill synonyms task, $F(1, 88) = 32.89$, $p < .001$, $\eta p^2 = .297$. In both instances, older adults outperformed younger adults. There were no significant main effects of language background in these tasks (both $F < 1$), nor were the interactions significant (verbal fluency - $F < 1$; Mill Hill - $F(1, 88) = 1.03$, $p = .314$, $\eta p^2 = .008$).

No other significant effects were observed on any of the tasks. For full details of these analyses, refer to Appendix W.

Table 21

Descriptive Statistics for the 'Domain-general' and Breadth of Knowledge Tasks according to Age (young versus old) and Language Status (bilingual versus monolingual)

	Age		**Language**	
	Young	**Old**	**Bilingual**	**Monolingual**
Task	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Domain-General				
WCST/64	29.85 (6.72)	28.56 (5.32)	28.74 (6.13)	29.62 (6.06)
Flanker Congruent Trials/30	29.26 (3.66)	29.31 (1.91)	29.26 (1.95)	29.30 (3.55)
Flanker Incongruent Trials/30	28.53 (5.31)	29.87 (1.55)	29.74 (1.62)	28.72 (5.18)
Animals	23.00 (6.15)	21.60 (5.11)	22.07 (5.13)	22.52 (5.19)
Household Objects	22.66 (5.40)	23.42 (4.68)	24.05 (5.19)	22.18 (4.81)
Letter F	11.87 (4.73)	16.09 (5.93)	13.64 (5.29)	14.18 (6.11)
Breadth of Knowledge				
STW Lexical Decision/60	47.97 (5.19)	53.82 (3.81)	49.93 (4.79)	51.60 (5.81)
Mill Synonyms/33	16.98 (4.35)	22.60 (4.27)	19.45 (5.04)	19.96 (5.25)

Note: maximum possible scores are indicated alongside test names where appropriate (e.g., WCST/64 indicates that the maximum possible score was 64) and performance on the Flanker appears close to ceiling. STW = Spot-the-Word task

3.3.3 Semantic Control Task: Effects of Age and Language Status on Semantic Control via Analysis of Accuracy Data

Consistent with Hoffman (2018), participants' scores on the semantic control tasks (i.e., Global and Feature association tasks) were analysed using mixed effects models to predict accuracy at the level of individual trials. A 2 x 2 x 2 x 2 factorial model was first specified, including age (older adults/younger adults) and language status (bilingual/monolingual group) as between-subjects factors, and task (Global vs. Feature judgements) and semantic control demands (High vs. Low) as within-subject factors. A linear mixed model approach was adopted with the four IVs described above entered as fixed effects and participants were entered as a random effect. To reiterate, the weak association condition of the Global judgements and the incongruent Feature condition were considered to have *High control* demands. A Q-Q plot of residuals indicated that there was a non-normal distribution of the errors for both accuracy and reaction time (see Appendix X1 and Y1 for plots). However, both Knief & Forstmeier (2021) and Schielzeth et al. (2020) have indicated that such models are robust against non-normality.

These analyses determined that there were significant main effects of: (a) Task Type $F(1, 8728) = 150.263$, $p < 0.001$, such that the accuracy was generally higher on the Global ($M = 0.94$) than Feature ($M = 0.86$) task (as predicted), and (b) Semantic Control $F(1, 8728) = 465.823$, $p < 0.001$, with lower accuracy on the High semantic control conditions ($M = 0.83$) than the Low semantic control conditions ($M = 0.96$), also as predicted. The main effects of age, $F(1, 88) = 3.379$, $p = 0.069$, and language status (bilingual/monolingual), $F(1, 88) = 0.919$, $p = 0.340$, were both not significant (see ANOVA summary table Appendix X). However, Task Type interacted with: (a) semantic control $F(1, 8728) = 116.436$, $p < 0.001$, such that the impact of semantic control (Low > High) was greater on the Feature task; (b) age, $F(1, 8728) = 91.267$, $p < 0.001$, such that the difference in accuracy across tasks (Global

> Feature) was greater for older participants. However, there was no significant interaction between Task Type and language status, $F(1, 8728) = 0.703$, $p = 0.402$.

There was also a significant interaction between semantic control and age group, $F(1, 8728) = 13.911$, $p < 0.001$, where the negative impact of a High semantic control trial was greater for older than younger participants. Additionally, semantic control also significantly interacted with language status, $F(1, 8728) = 6.208$, $p = 0.013$, in that the impact of semantic control (Low > High) was greater for monolinguals than bilinguals. Therefore, language status appears to moderate the main effects of semantic control. Finally, there was a three-way interaction of Task x Semantic Control x Age, $F(1, 8728) = 68.664$, $p < 0.001$, and a four-way interaction of Task x Semantic Control x Age x Language Status, $F(1, 8728) = 8.889$, $p = 0.003$.

The average accuracy scores relating to the three-way interaction are presented in Figure 17. On the Feature judgement task, the impact on accuracy in the High semantic control condition (incongruent trials) was much greater for the older ($M = 0.69$) than younger participants ($M = 0.84$; $z = 8.263$, $p < 0.001$). Meanwhile, the disadvantage for older participants on the high semantic control condition in the Feature judgement task is not present in the Global judgement task. In this case, older participants ($M = 0.93$) do better generally than younger participants ($M = 0.88$) on the high control (weak associates) condition, ($z = -2.845$, $p = 0.004$). That is, this key pattern of an age-related impact on high semantic control that appears to 'flip' across semantic tasks.

Figure 17. *Average Accuracy across the Four Experimental Conditions and Age*

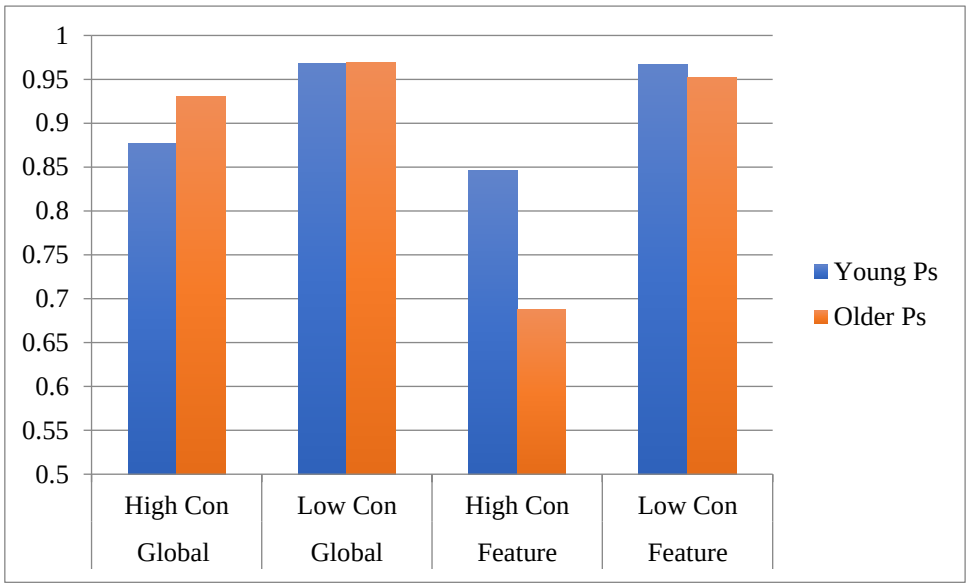
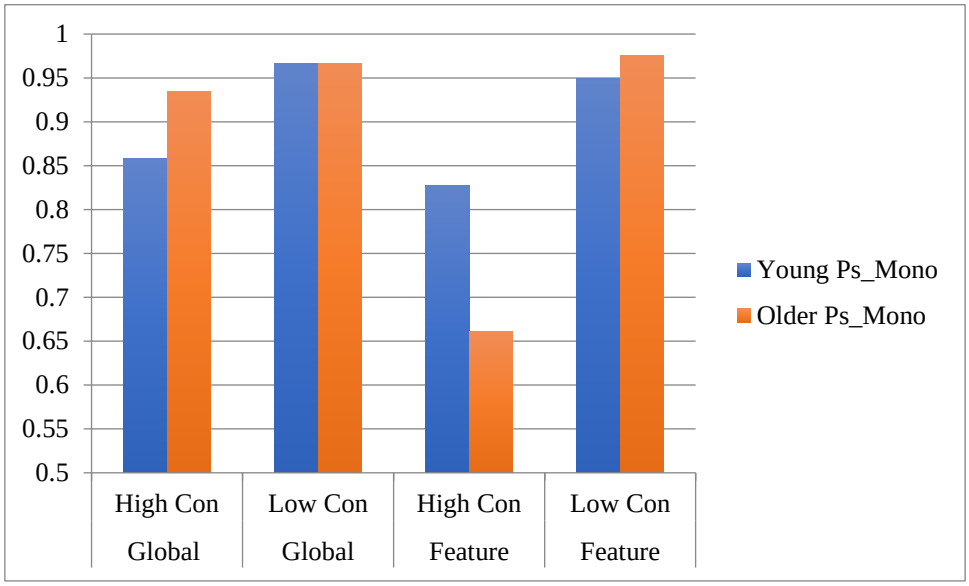


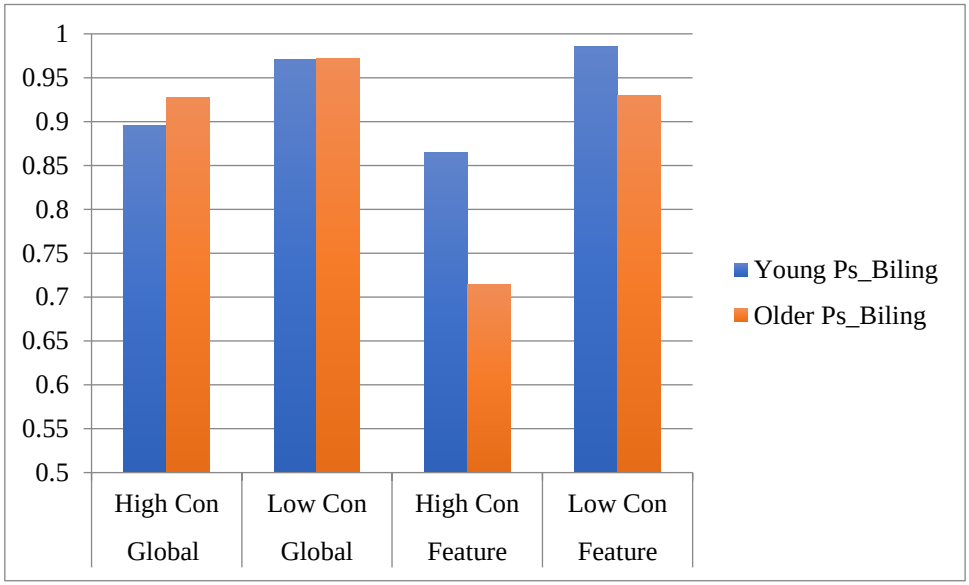
Figure 18 breaks down this pattern for just the monolingual groups, and it shows an age-related pattern consistent with the monolingual sample in Hoffman (2018). On the Feature association task, the impact of High semantic control (incongruent) on accuracy was much greater for the older participants ($M = 0.66$) than younger participants ($M = 0.83$; $z = 6.404$, $p < 0.001$. However, this critical age-related pattern changes on the Global task. In this case, older adults ($M = 0.935$) perform better than younger adults ($M = 0.86$) on the High control (weak associates) condition ($z = -2.963$, $p = 0.003$). To reiterate, for the monolingual sample only, the consequences of ageing on tasks that require High semantic control are not the same across tasks. That is, on the Global judgement task, older participants outperform younger participants on weakly associated items but this age effect 'flips' on the Feature judgement task. In this case, younger participants outperform older adults on incongruent trials.

Figure 18. *Average Accuracy for Younger and Older Monolinguals*



In Figure 19, the same key patterns are presented but for the bilingual groups only. As with the monolingual groups, and on the Feature association task, the impact of High semantic control (incongruent) on accuracy was much greater for the older bilinguals ($M = 0.71$) than younger bilinguals ($M = 0.86$; $z = 5.337$, $p < 0.001$). However, on the Global task, the older ($M = 0.93$) and younger ($M = 0.90$) bilinguals perform *similarly* in the High semantic control condition (weak associates) ($z = -1.141$, $p = 0.254$), meaning there was no statistically significant difference across the two groups. Moreover, younger bilingual participants ($M = 0.90$) outperformed their monolingual counterparts ($M = 0.86$) on the same condition (weak associates of the Global judgement task) ($z = -2.365$, $p = 0.018$), where the difference was significant (comparison of Figure 18 and 19). This demonstrates that the consequences of bilingualism appear to boost younger participants' performance on the high semantic control condition of the Global task, such that no older adult advantage remains.

Figure 19. *Average Accuracy for Younger and Older Bilinguals*



Additionally, comparing the effect of high control trials (incongruent) on the Feature association task across older monolinguals ($M = 0.66$) and bilinguals ($M = 0.71$) shows that there was evidence that they performed more poorly relative to their age matched bilingual counterparts ($z = -1.938$, $p = 0.053$), suggesting that the consequences of bilingualism appear to boost older participant performance on the high semantic control condition of the Feature task, such that the impact is somewhat moderated. In simple terms, the nature of the four-way interaction reported above, is that for bilinguals, the key patterns described by Hoffman (2018) regarding age effects and semantic control (i.e., the 'flip' pattern mentioned above) are moderated by being bilingual with respect to accuracy analyses.

3.3.4 Reaction Time Data Analysis

Before undertaking analyses, and in line with Hoffman (2018), reaction times (RTs) were screened for outliers by winsorizing any RTs more than two standard deviations from a participant's mean (a total of 3.44% of trials were thus removed). Analyses were then undertaken on correct trials only. The same 2 x 2 x 2 x 2 factorial model described earlier was then specified, and a linear mixed model analysis approach undertaken. The four IVs

described above were entered as fixed effects and participants were entered as a random effect. Plotting the residuals of these models against the fitted values indicated that the data was heteroscedastic (see Appendix Y2).

Significant main effects were found for:

(a) Task Type, $F(1, 7654.62) = 590.381$, $p < .001$, such that the reaction times were generally faster on the Global (M = 2951ms) than Feature (M =3780ms) task;

(b) Semantic Control, $F(1, 7653.05) = 627.637$, $p < .001$, with generally slower reaction times on the High semantic control conditions (M = 3791ms) than the Low semantic control conditions (M = 2940ms);

(c) Age, $F(1, 88.17) = 23.250$, $p < .001$, with generally slower reaction times for older (M = 3722ms) versus younger participants (M = 3008ms). See Figure 20.

There was no main effect of language status, with speed of reaction time similar across monolingual (M = 3452) and bilingual (M = 3279) participants.

In addition, Task Type significantly interacted with:

(a) Semantic control, $F(1, 7652.57) = 12.003$, $p < .001$, such that responses were generally slower on the High control (incongruent) condition of the Feature task;

(b) Age, $F(1, 7654.62) = 22.646$, $p < .001$, such that the difference in reaction times across tasks (Global > Feature) was greater for older participants.

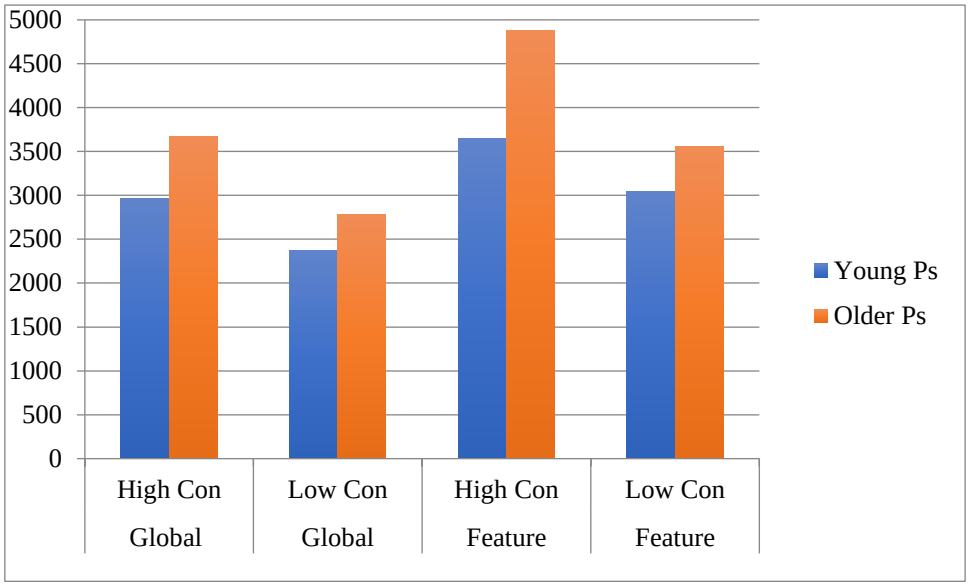
(c) Language status, $F(1, 7654.62) = 5.290$, $p = 0.021$, such that monolinguals were significantly slower to complete the Feature task compared to the Global task. There was no statistically significant difference in the bilinguals' reaction times across tasks.

A significant interaction was also found between semantic control and age group, $F(1, 7653.05) = 55.993$, $p < .001$, indicating that the negative impact of a High semantic control trial was greater for older than younger participants.

Finally, there was a three-way interaction between Task x Semantic Control x Age, $F(1, 7652.57) = 10.495$, $p = 0.001$, which will be described in detail below (Figure 20). No other interactions reached statistical significance (see Appendix Y for ANOVA summary).

The three-way interaction can be characterised as reflecting that, while older adults were slower to respond in all tasks and at both levels of semantic control, the magnitude of the difference in reaction time versus the younger adults varies from condition to condition. As can be seen in Figure 20, in the Feature task, the average reaction time in the High semantic control condition (incongruent trials) was significantly longer for the older participants (M = 4901ms) than the younger participants (M = 3650ms; $z = 7.644$, $p < .001$). The reaction time disadvantage for older participants in the High semantic control condition of the Global task was much smaller, although still significant (older adult M = 3673ms; younger adult M = 2972ms, $z = 4.371$, $p < .001$). In respect of the low semantic control conditions, the size of the reaction time difference between older and younger adults was, on average, 410ms in the Global task and 515ms in the Feature task. Although, both of these differences were also significant (Global task $z = 2.581$, $p = .02$; Feature task $z = 3.233$, $p = .002$) it is clear that the age effect is larger when the task is more difficult and the semantic control demands are higher.

Figure 20. *Average RTs across the Four Experimental Conditions and Age*



3.3.5 Relationships Between the Semantic Control Task and General Cognitive Tasks

This next section will explore whether performance on the key semantic High control conditions (i.e., incongruent and weak association trials) was related to performance on any of the general cognitive assessments, using a principal components approach. To reiterate, these general cognitive measures fall into two broad sub-types. On the one hand, there are tasks that tap 'domain-general' executive control (e.g., WCST and letter fluency) – processes of control that are expected to be relevant across a variety of decision-making contexts. These tasks were described earlier, and analyses across groups were undertaken (see Section 3.3.2 above). For these domain-general tasks, there was no significant main effect of language status (bilingual versus monolingual), but older participants outperformed younger participants on the letter fluency task. Therefore, there was no evidence that older participants demonstrated deficits across 'domain-general' executive control tasks. On the other hand, there were tasks that tap 'breadth of semantic knowledge', including Lexical Decision and Synonym judgement tasks – that is processes that act as a proxy of individual semantic

memory capacity/complexity. Again, these tasks were described earlier, and analyses across groups were undertaken which determined a general older adult participant advantage, but no effects of language status (see Section 3.3.2 above).

Hoffman (2018) conducted two principal component analyses (PCA), one with the younger participant sample and one with the older participant sample, in order to examine the commonalities between the general cognitive tasks and the high semantic control conditions of both the Global and the Features task. The aim of these analyses was to attempt to identify whether responses on the high control trials were more aligned with measures of the breadth of semantic knowledge or with domain general executive function. A similar approach has been taken in the current study. Bilingualism has been ignored for the purpose of the PCA analyses below for two main reasons. Firstly, the dataset included insufficient observations to allow for individual PCAs to be conducted in age x language status groupings, which would mean that any findings at this level of granularity would be unreliable. Therefore, it was necessary to group participants for the PCA according to a higher order factor. In the preceding section, it was demonstrated that bilingualism did not exert significant main effects on performance in the semantic control tasks, and in the majority of cases, interactions including bilingualism were not significant either. In contrast. significant age effects were more ubiquitous. Thus, the second reason for ignoring bilingualism rather than age was because it would be more likely for any differences between PCA findings in monolingual versus bilingual datasets to be attributable to age-related patterns than for differences between young and old datasets to be attributable to bilingualism.

Both PCAs used the accuracy measures for the two High control semantic conditions (Feature semantic task – incongruent trials and Global semantic task – weak associate trials), to determine which of these might be linked with the key cognitive measures. In addition, given the little variation on accuracy for the Flanker task (largely ceiling effects) the decision

was reached to not use this measure in this case (mirroring the approach taken by Hoffman, 2018).

3.3.6 Principal Component Analysis (PCA) – Older Participants

As described earlier, to explore the underlying commonalities among tests in more detail, test scores were entered into a PCA, in which components with eigenvalues greater than one were extracted and varimax rotated. The following accuracy scores were included: lexical decision (i.e., Spot-The-Word), synonyms (i.e., Mill Hill), WCST, semantic category fluency (Animals/Household objects) and letter fluency, to determine if any had a relation to performance on the two key High control semantic conditions (Incongruent Features and Weak Associations). There were three components with eigenvalues greater than one, which together accounted for 63.4% of the variance in scores (see Table 22).

Table 22

Principal Component Analysis - Older Participants

Component Loadings	PCI	PC2	PC3	Uniqueness
Synonyms	**0.895**			0.169
Lexical Decision	**0.861**			0.153
Animals	0.496		**0.587**	0.398
WCST		**0.850**		0.268
Weak Association Accuracy		**0.742**		0.363
Household Objects			**0.803**	0.355
Incongruent Accuracy			**0.568**	0.538
Letter F				0.685

Note: Applied rotation method is varimax. Component loadings >0.5 are shown in bold

The lexical decision and synonyms tasks loaded strongly on a first component - PC1 - which would appear to index the ‘breadth’ of participants’ semantic knowledge. It is

noteworthy that this had no relation to performance on the high control semantic conditions. The weak association task (Global) and the WCST loaded onto a second component, and a third component captured performance on the incongruent trials and loaded with the two verbal fluency tasks (Animals/Household objects).

On balance, the PCA with older participants suggests that different cognitive measures appear to be linked to variability of performance on our two key semantic control conditions. For incongruent trials on the feature task, performance appears associated with verbal fluency, which requires flexible, controlled retrieval of semantic knowledge. For weak associate trials on the Global task, performance appears linked to performance with the WCST, suggesting that it captures aspects of domain-general executive function that are not specific to semantics.

3.3.7 PCA – Younger Participants

As before, and to explore the underlying commonalities among tests in more detail, test scores with the younger participants were entered into a PCA, in which components with eigenvalues greater than one were extracted and varimax rotated. As with the earlier analysis, a three-component solution emerged, which together accounted for 66.3% of the variance in scores (see Table 23). In this case, a very different pattern emerged, such that the two critical High semantic control conditions (incongruent/weak associate) load together with two key measures of 'breadth' of participants' semantic knowledge (i.e., lexical decision/synonym judgment) - see PC1. Interestingly this contrasts with the pattern seen with older participants, for whom measures of semantic 'breadth' loaded together as a separate component. A second component saw all three verbal fluency tasks (Animals/Household objects/Letter F) loaded together. The third and final component saw WCST and letter fluency loaded together (and to a smaller degree performance on weak associate trials). In other words, the measures of

'domain-general' executive control, for the younger participants at least, appear to largely load onto components that are separate from the performance on the key high semantic control conditions.

Table 23

Principal Component Analysis - Younger Participants

Component Loadings	PCI	PC2	PC3	Uniqueness
Lexical Decisions	**0.899**			0.186
Synonyms	**0.879**			0.203
Weak Association Accuracy	**0.594**		0.409	0.479
Incongruent Accuracy	**0.571**			0.417
Animals		**0.879**		0.189
Household Objects		**0.771**		0.396
Letter F		**0.536**	**0.625**	0.321
WCST			**0.699**	0.502

Note: Applied rotation method is varimax. Factor loadings >0.5 are shown in bold

3.4 Discussion

The aim of the current study was to replicate the work of Hoffman (2018) while also investigating the impact of bilingualism on semantic control in an attempt to determine a more suitable proxy of CR. Previous literature focuses on the organisation of the mental lexicon (Dong et al., 2005) in bilinguals, and previous semantic memory models tend to focus on the organisation of semantic memory in monolingual populations (Collins and Loftus, 1975; Collins and Quillian, 1969; De Deyne et al., 2017; Rips et al., 1973; Yee et al., 2014). In addition, when considering a possible bilingual advantage in the context of CR, most studies investigate the impact of bilingualism on executive control measures, such as inhibition and attention switching (Bialystok et al., 2012; Miyake et al., 2000). However, metacognitive processes in a particular aspect of knowledge like semantic memory, may

constitute a more suitable proxy for testing the existence of a CR advantage in the bilingual context (Abu Rabia, 2019; Ordin et al., 2020). This notion has not been considered widely in the literature, where it could be that increased CR in bilinguals reflects a more efficient cognitive system that operates on semantic information.

For clarity and simplicity, the discussion firstly reiterates the key patterns of performance found in the context of previous literature and theory, before discussing the bilingual findings in the wider context of CR and their contribution to the literature.

3.4.1 General Cognitive Tasks

The results demonstrated that older adults performed better in the breadth of semantic knowledge tasks (i.e., Spot-The-Word and Mill Hill). This is largely consistent with the assumption that as individuals age, semantic knowledge increases. That is, there is an increase in crystallized knowledge, and this is often reported as an age-related advantage in the literature (Hoffman, 2019; Grady, 2012; Park et al., 2002). This age-related advantage of crystallized knowledge also seems to play a role in the older adult advantage found in the letter fluency task. Literature suggests that older adults tend to perform better in letter fluency tasks compared to semantic fluency. Gordon et al. (2018) suggest that semantic fluency relies on speed of lexical retrieval which may decline with age, whereas letter fluency relies on the amount of vocabulary knowledge available which increases with age as a result of life experience (crystallized knowledge). In keeping with the performance of older adults, but in contrast to Hoffman (2018), they were also found to perform similarly to the young adults on the WCST. One reason for this finding may be that the rule changed every four trials in the version of the WCST used in the present study, whereas the rule changed after ten consecutive correct trials in the version used by Hoffman. Given that the correct categorisation of a card can be based on shape, colour, or number, it could take a participant

three trials to deduce the correct rule. This means that the version used in the present study is more challenging because participants must learn the rule more quickly, and even after they have determined the correct sorting rule, there will be fewer consecutive scoring trials available. As a consequence, the performance of both age groups are quite low and the lack of a difference between their average scores reflects an artificial ceiling effect caused by the difficulty of the task.

In terms of the effects of language status on performance on the general cognitive tasks, the performance of the bilinguals and monolinguals was comparable. In terms of the verbal fluency task, and as mentioned in the first study of this thesis (Chapter 2), some semantic categories may generate more English words, compared to Welsh, because of the contextual language use differences across various categories (i.e., more English words may be used for certain categories compared to others). Consistent with this, it is likely that fewer cross-language intrusion errors occur when completing a task in English because the English words may be frequently accessed in day-to-day life. Thus, English words may have been coming to mind more automatically in the verbal fluency task, helping to explain why the performance of bilinguals and monolinguals was comparable. Alternatively, it may be more likely that bilingual participants experienced more Welsh words coming to mind that begun with the same letter, because both Welsh and English words beginning with the same letter would be more commonly practiced in daily life. This may lead to a greater need for inhibition under time pressure and in turn, no superior performance. Even so, the lack of a superior bilingual performance is in contrast with some literature. For example, Patra et al. (2020) found that bilinguals perform better than monolinguals in letter fluency tasks due to the increased daily demands on executive control (i.e. well-rehearsed executive functions due to managing two languages). However, because of the Welsh-English bilinguals' balanced dominance and frequent switching in day-to-day life, words from both languages may have been coming to

mind (as mentioned), meaning that despite potential increased control abilities, performance was in line with that of the monolinguals.

Finally, and as with all other measures, tasks that tapped into breadth of semantic knowledge (Spot-the-Word and Mill Hill) were conducted in English. Therefore, as the bilinguals in this study were highly proficient in English, the lack of performance differences between bilingual and English monolinguals may not be surprising and are also in line with the pattern of findings throughout this thesis. That is, Welsh-English bilingualism may not lead to an improvement in general cognitive processes.

3.4.2 Semantic Control Task Findings and Interpretations – Accuracy

As a reminder, this study aimed to examine the impact of age (a replication of Hoffman, 2018) and bilingualism on semantic control. Key findings were that accuracy was generally higher on the less challenging task (Global), with lower accuracy observed when High semantic control was required, regardless of task type (Global/Feature). However, the impact of High semantic control was more pronounced in the more challenging Feature task. Regarding age effects, the impact of High semantic control was greater for older adults, where there was larger variation in accuracy across tasks (Global/Feature). Moreover, the impact of age on semantic control also varied depending on the type of semantic task, as discussed further below. In terms of bilingualism, semantic control had a greater effect on monolinguals compared to bilinguals. Bilingualism appeared to enhance accuracy under conditions requiring High control. Additionally, comparison of performance indicated that older monolinguals were less accurate than older bilinguals when High control was needed for Feature-based associations, implying that bilingualism improves older adult performance in the most challenging task, moderating the impact of High control on accuracy. What follows, is a thorough examination of all findings relating to accuracy.

Overall, the findings revealed that accuracy was generally higher on the Global task compared to the Feature task, and lower on the High control conditions across the tasks (i.e., weak / incongruent). That is, decisions that required a High level of semantic control were less accurate in both semantic tasks. However, the impact of High control demands was greater in the Feature task. This was an expected finding as the Feature task is designed to be more challenging because the probe and target were not associated with each other except for on the specified feature (e.g., colour or size). In addition, the High control trials of the Feature task also included a foil that was strongly associated with the probe and target which requires a high level of inhibition to ignore the obvious association and focus on the required information.

In relation to task type, across both Global and Feature tasks, the difference in accuracy was greater for older participants. That is, the drop in performance from the Global task to the Feature task was greater for the older adults compared to the younger group, replicating a similar pattern to that of Hoffman (2018). Additionally, the negative impact of the High control trials on the Feature task was greater for the older participants and for monolingual participants compared to the younger and bilingual groups. This suggests that bilingualism moderates the effects of control demands, whereas age interacts with both task type and control demands.

It may be logical to suggest that the younger group and the bilingual group were able to better manage the demands of High control trials of the Feature task because of better inhibition abilities in younger adulthood and in bilingualism (Green, 1998; MacLeod, 2007; McDowd et al., 1995; Rubio-Fernández, 2017). However, there was no difference in performance on the Eriksen Flanker Task in this study which measures response inhibition. This may mean that different cognitive processes might be more central to managing the High control demands of the Feature task for younger participants and bilinguals, or that for

the bilinguals, the Flanker Task is not measuring the correct cognitive process. Indeed, as the Flanker Task is a non-linguistic/non-semantic inhibition task, it may be measuring a different type of inhibition that has less of a role in inhibiting semantic information. Another possible explanation is that bilinguals are also argued to be able to cope better with high cognitive load (Bialystok et al., 2004). To support this claim, Doroud et al. (2020) found that bilinguals' performance was maintained on the Stroop Task even when the cognitive load increased. In the present study, the cognitive load was increased by increasing the difficulty of the task. For example, the High control trials of the Feature task were more challenging because the probe and target word shared no meaningful association other than matching on the specified feature. However, one of the foils had a strong relationship with the target, even though it did not match on the currently relevant feature. This meant that participants were required to problem-solve as well as ignore the strong but redundant relationship between the probe and foil. Consequently, this task included the highest cognitive load compared to the other aspects of the semantic control task, and it is noteworthy that younger participants and bilingual participants were more accurate.

Moving to the three-way interaction found between task, control demands, and age, on the Feature task, the impact of High control was found to be greater for the older participants. However, this was not the case for the Global task. In this task, older adults outperformed younger adults on the High control trials. That is, the impact of age on performance 'flips' across both tasks which is consistent with Hoffman (2018). On the one hand, older adults appear to be much more detrimentally impacted by High semantic control when making Feature-based judgements (incongruent trials). On the other hand, older adults appear to be much less detrimentally impacted by High semantic control when making Global based judgments (weak association trials). The superior performance of older adults in this respect (Global) may reflect the fact that they have the experience of life to draw upon (and thus have

an advantage). This seems plausible as having more life experience is associated with having a richer vocabulary and more general knowledge, which in turn, may help to draw associations and find patterns (Buades-Sitjar et al., 2022). Further, a richer vocabulary may have been helpful in the Global task as having more information means that older participants may have had access to a larger array of connections/links between concepts, meaning that they were more likely to know something about the Global associations and in turn, perform more accurately.

A four-way interaction was also found between task, control demands, age, and language status. It was apparent that the impact of High control on the Feature task was greater for older bilinguals compared to young bilinguals. However, on the Global task, performance was comparable across older and younger bilinguals in the High control trials. Importantly, the young bilingual group also outperformed the young monolingual group on the High control trials of the Global task. This suggests that there may be a bilingual advantage that boosts the accuracy of younger participants on the High control trials of the Global task. In addition, older monolinguals were less accurate on the High control trials of the Feature task compared to the age matched bilingual group. Therefore, this implies that there may be a bilingual advantage that boosts the accuracy of older adults on the High control trials of the Feature task. This means that the negative effects of High control demands are reduced in the young bilingual group on the Global task and for the older bilingual group on the Feature task. Put simply, the four-way interaction indicates that the key three-way interaction pattern that replicated the work of Hoffman, (2018) is generally moderated by bilingualism, such that on the one hand the age advantage on the Global task is reduced for young bilinguals and the negative consequences of the high control condition on the Feature task is reduced for older bilinguals.

3.4.3 Semantic Control Task Findings and Interpretations – Reaction Time

As with the findings on accuracy above, a brief summary of the key findings relating to reaction time is provided before these are discussed in greater detail below. Participants exhibited faster reaction times when completing the simpler task (Global), while they were generally slower under conditions requiring High control in the most challenging task (Feature). With respect to age related effects, older adults were slower to respond compared to younger participants – a finding that is typical in the literature. The age effect was larger when the task was more challenging and when semantic control demands were higher. Bilinguals responded similarly across tasks, but monolinguals were slower to respond on the more challenging task (Feature). Further discussion and interpretation of the findings relating to reaction time is provided below.

The findings demonstrated that participants were generally faster to respond on the Global task compared to the Feature task. Again, and as mentioned previously, the Feature task is designed to be more challenging compared to the Global task which may lead to slower reaction times. Specifically, responses were slower for High control trials across tasks – that is, instances where the decision-making process is more challenging with inhibition of responses sometimes required leads to slower reaction times. Reaction times were particularly slower for the High control trials of the Feature task. Older adults were also slower to respond than younger adults which is a well-established finding in the literature on reaction times (Birren & Fisher, 1995; Der & Deary, 2006; Glass et al., 2000; McDowd et al., 2020). Indeed, when people age, reaction times decrease in most cognitive tasks (Ratcliff et al., 2001). As with accuracy, the difference in reaction times between the Global and Feature tasks was greater for the older participants and the impact of the High control trials was greater for older adults, which reflected the three-way interaction between task, control, and age.

Notably, bilingualism also impacted on reaction times. Monolinguals were slower to react on the Feature task compared to the Global task. This suggests that bilingualism helped to maintain reaction time even under increased difficulty. This is consistent with some previous literature exploring reaction times on various cognitive tasks in bilingual populations. For example, Martin-Rhee & Bialystok (2008) found faster reaction times for bilingual children on both congruent and incongruent trials of the Simon and Stroop tasks. In contrast, Bialystok (2007) found that adult bilinguals between the ages of 30 and 60 showed a reaction time advantage compared to age-matched monolinguals on the Simon task, but that reaction times slowed after the age of 60 for both groups, with a more rapid decline for the monolinguals. However, the findings of the present study suggest that Welsh-English bilingualism may positively affect reaction times across the lifespan, as opposed to only certain points across the lifespan, considering the participants' ages (i.e. the sample consisted of both younger and older participants). Interestingly, it may be that bilingualism impacts reaction time across the lifespan but for different reasons at each stage. For instance, reaction times may be faster in bilingual children (as reported by Martin-Rhee & Bialystok (2008) because bilingualism may be boosting their underdeveloped executive function (EF) abilities, compared to monolinguals. Then, between the ages of 30 and 60, a bilingual reaction time advantage may remain (as reported by Bialystok, 2007) because bilingualism may be boosting their established EF abilities. It is also important to note that in later life, although reaction times generally decrease across cognitive tasks, Bialystok (2007) found that the decline in reaction time is steeper for monolinguals. That is, there may be no protection against cognitive ageing for the monolingual group in the present study, which is in line with the findings on accuracy where the monolinguals were not at accurate as the bilinguals.

3.4.4 Relationship Between the Semantic Control Task and General Cognitive Tasks

For clarity, the relationship between the semantic control task and general cognitive tasks was examined for several reasons. Firstly, this analysis aimed to replicate Hoffman (2018). Secondly, it was important because different age groups may rely on different processes or knowledge to complete the semantic control task, especially when a High level of control is required. For example, older adults may have more stored knowledge, or may use their knowledge more efficiently due to life experience. In contrast, younger adults may be more efficient at using their knowledge due to more effective executive processes. Therefore, this section aids in understanding whether responses on the High control trials of the semantic control task are more aligned with breadth of semantic knowledge or with domain-general executive function in both older and younger adults. The key findings indicate that for older adults, performance on the more challenging task (Feature) when High control is required, may be associated with controlled retrieval of semantic knowledge. In contrast, performance on the Global task with the same level of control may be associated with domain-general executive function. However, this pattern was not observed for younger adults; when High semantic control is needed regardless of the task type, they may draw on semantic specific knowledge. These findings are discussed in more detail below.

For older adults, it seemed that two verbal fluency measures (household objects and animals) were related to the High control trials of the Feature task, and that the WCST was related to the High control trials of the Global task. This suggests that, for older adults, accuracy on the High control trials is associated with more domain-general executive function as opposed to breadth of semantic knowledge. In relation to these findings, verbal fluency may be related to the High control trials of the Feature task because of the similarity in the cognitive processes of both tasks. That is, during verbal fluency tasks, participants are

required to search through their semantic network (Henry et al., 2004; Shao et al., 2014) for all things related to the category (e.g., animals). Similarly, in the High control trials of the Feature task, participants are also required to search their semantic network for all things related to the target word in order to find its features before selecting the feature that meets the task requirements (e.g., Colour or Size). That is, the task requires the retrieval of information based on specific semantic criteria (Henry & Crawford, 2004).

In contrast, the WCST was related to the High control trials of the Global task. The decisions in this task involved Global semantic association and although the High control trials placed greater demands on the controlled retrieval of semantic information (because automatic spreading of activation in the network would not identify the correct response), the correct response is still more obvious than the High control trials of the Feature task. On that basis, it seems logical that the WCST is related to the High control trials of the Global task, as opposed to the Feature task because it taps more general attention type processes. That is, on both the WCST and the High control trials of the Global task, there is a limited number of options as to what the answer may be. For example, although the target of 'mountain' is not strongly associated with the two response options of 'mound' or 'director', the correct answer remains obvious. Moreover, the decisions are more surface level and do not require as much control or inhibition as the High control trials of the feature task or verbal fluency.

Interestingly, these findings show a reversed pattern to that reported by Hoffman (2018), where the WCST loaded with High control trials of the Feature task and verbal fluency with that on the Global task. In any case, it is noteworthy that both the present study and that of Hoffman (2018) suggest some connections between the key High semantic control conditions and other domain-general executive function measures for older participants.

In relation to the younger adults, Spot-the-Word (lexical decision) and Mill Hill (synonym judgement) performance was related to the High control trials on both the Global and Feature

Tasks. Meanwhile, the WCST and letter fluency (letter 'F') were weakly related to the High control trials on the Global task. The three verbal fluency measures loaded together as a separate component. Therefore, comparing the younger and older adults, it appears across the two analyses that: (a) for younger participants, variability in High semantic control trials, across both Global and Feature tasks, was linked to general measures of breadth of semantic knowledge, and (b) for older participants, the two key High semantic control measures were linked to performance on domain-general executive tasks.

It may be worth emphasising however that in the original work (Hoffman, 2018), not all participants were given the WCST for assessment (i.e., approximately only half of both young/older participant groups). It is not entirely clear why this occurred, and thus caution should be exercised when interpreting their PCA findings. Semantic control for younger participants is linked to semantic 'breadth' across all tasks, and in one case, this lower 'breadth' is detrimental relative to older adults – that is on weak associate trials for the Global task. In any case, the author would argue that the pattern of performance found in this study is quite straightforwardly interpreted. Namely, that for young participants, measures of semantic control are more generally associated with measures tapping semantic knowledge (broadly speaking). One reason for this may be because younger adults are likely to have less semantic knowledge than older adults (due to less life experience), meaning that it may be easier for them to retrieve the required information, as opposed to having to inhibit other responses that are found in a large semantic memory store (as with older adults).

On the other hand, if it is assumed that older adults have a much broader amount of this knowledge already (as indicated by their generally superior performance on measures of 'breadth' of semantic knowledge), for them, domain general 'executive function' is a key index of variability in semantic control (i.e., WCST and verbal fluency tasks). Moreover, this latter semantic 'breadth' advantage for older adults can also explain the pattern of their

superior performance on the high control trials for the Global semantic task. Again, such judgements may more obviously draw on the accumulation of semantic memory across the lifespan (an assumption also suggested by Hoffman, 2018).

3.4.5 The Effects of Bilingualism on SC in the Wider Context of CR

This study shows for the first time, that language status may have a key role in SC. It demonstrated that the two key patterns described above in each association task may both be moderated by bilingual ability. Arguably, this may well reflect the lifelong bilingual experience of managing the information processing demands of speaking and understanding two languages. Additionally, and as the PCA analyses suggested that semantic control is broadly linked to measures of semantic 'breadth' for younger adults and to measures of domain general 'executive' function for older adults, it is possible to tentatively suggest that the bilingual moderation effects reflect 'boosts' to each respectively. However, further work is required to determine and to fully elucidate the potential moderating effect of bilingualism in this context.

Further, the findings from the current study could be considered surprising given previous work. For instance, most studies comparing the performance of Welsh-English bilinguals to English monolinguals on cognitive tasks have found no clear evidence of an advantage in performance, including the first study of this thesis (Clare et al., 2016b; Gathercole et al; 2014; Hindle et al., 2015). However, it is important to point out that these previous studies have attempted to determine a bilingual advantage using tasks that tap similar areas of cognition (e.g., task switching, standard inhibition). Thus, the current study may be early evidence that systems *operating on* semantic memory are where the bilingual advantage may reside. Thus, such findings may offer a framework in which to further explore why lifelong

bilingualism appears to contribute to CR and offers some protection against the decline from neurodegenerative diseases like dementia (e.g., Guzmán-Vélez & Tranel, 2015).

In addition, this work contributes to understanding a key 'paradox' of Welsh-English balanced bilingual performance described earlier. Namely, that in the everyday context of communication, language switching occurs often and appears effortless. Clearly, if Welsh-English bilinguals do have superior levels of semantic control (as suggested by this study), then their capacity to deal with such regular switching may underpin this apparently effortless performance. Consequently, this work clearly raises an important question about how other bilingual populations may perform in a similar study, and whether the bilingual advantage identified in this study may be a unique property and consequence of being a 'balanced' bilingual. Therefore, further research is required to better understand the effects of different types of bilingualism on semantic control processes.

3.4.6 Strengths and Limitations

The present study has made significant contributions to the existing body of knowledge about Welsh-English bilingualism and the cognitive advantages that may be associated with it. It has identified a possible bilingual advantage in a largely unexplored type of bilingualism, group of bilinguals, and aspect of cognition in the context of CR. That is, to the author's knowledge, there are no other studies that have examined the effect of bilingualism on SC. The findings gained from this study provide a foundation for more theoretical developments within this specific population of bilinguals, as well as exploring SC mechanisms with other bilingual populations.

Even so, the study is not without limitations. In addition to the inherent complexities encountered when exploring phenomena when multiple variables are potentially involved and/or there are insufficient anchor points within existing evidence (discussed further in

Chapter 5), the current study was also not an exact replication of Hoffman (2018). Slightly different versions of the WCST and Mill-Hill task were used, and the age bracket differed for the younger adults. Even so, the current study replicated the findings of Hoffman (2018). Second, as the current study focused on balanced bilingual Welsh-English speakers, findings may not generalise to other bilingual populations (e.g., less balanced populations). For example, compared to those who rarely switch languages and for whom inhibition may be more effortful (e.g., Greek-English), it may be that the systems of those who switch and inhibit responses regularly (e.g., Welsh-English) are different and also lead to different consequences. Thus, further research is needed to explore such possibilities and to facilitate understanding of the particular aspects of bilingualism that may lead to a bilingual advantage.

3.4.7 Conclusion

To the author's knowledge, this is the first time that the role of Welsh-English bilingualism in the metacognitive process of semantic cognition (SC), specifically cognitive control over semantic information, has been investigated. The major contribution of this study is that it has provided the first piece of evidence that Welsh-English bilingualism may play a role in SC and provides direction for future work (e.g., investigation in bilingual populations that may differ with respect to balance/dominance). The findings suggest that the bilingual advantage in the context of CR may lie in these metacognitive processes, and thus it would be also interesting to consider future work that examines variability of performance on measures of semantic control with bilingual dementia populations.

Even so, much remains unknown about whether the impact of bilingualism on *accuracy* found in this study extends to the *consistency* with which people respond to different kinds of statements that have varying degrees of uncertainty involved. Therefore, the next study will explore the consistency with which Welsh-English bilinguals and English monolinguals

respond to different types of statements. Specifically, if Welsh bilinguals have an advantage with respect to measures that tap High semantic 'control' (i.e., more demands), as found in this study, how might they perform in contexts of decision-making under uncertainty, since such contexts clearly are ones of high difficulty and complexity? This experimental chapter may provide more information on the processes that might increase the CR of bilinguals.

Key Messages

- Most studies investigating the role of bilingualism in CR focus on executive control measures such as inhibition and attention switching (Miyake et al., 2000; Bialystok et al., 2012).
- This study argues that a more suitable bilingual proxy of CR may be the more sophisticated metacognitive processes of bilinguals that operate on semantic memory.
- This study utilised the experimental paradigm used by Hoffman (2018), extending to explore the impact of bilingualism on semantic cognition. Specifically, the study investigated whether accuracy and reaction times would vary on a semantic control task across a monolingual and bilingual sample of younger and older adults.
- The main findings were that the consequences of age and language status (monolingual or bilingual) are different across different tasks (when the semantic control demands are changed). Bilingualism also influenced reaction times.
- On the Global judgement task, older adults outperformed younger adults on weakly associated items but this age effect 'flips' on the Feature judgement task. In this case, younger participants outperformed older adults on incongruent trials. This pattern replicated the findings previously reported by Hoffman (2018).
- Most importantly, bilingualism appears to have a key role in semantic cognition. For the Global association semantic task, younger bilinguals performed better than young monolinguals on the high semantic control (weak association) trials, performing at an accuracy level equivalent to older participants. Similarly, it demonstrated that on the Feature semantic task, older bilinguals perform better on the high semantic control (incongruent) trials than their monolingual counterparts.
- Future research is required to study the performance of bilingual dementia populations on the semantic control task.

Chapter 4 – Study 3 – Bilingualism & The Rumsfeld Effect

4.0 Chapter Overview

Chapter 3 examined meaning in the brain and how the operations on this system may differ for bilinguals. Specifically, the study explored how people answer questions that have obvious answers – a process that must involve decision making based on learnt/stored cognitive information, characterised in this thesis as a metacognitive process. In this specific case, the questions were largely about the relationship between two things (i.e., synonyms) or particular features of things – in some cases the selection process was made more difficult, and it may be that the challenge to select the correct answer was more demanding on metacognitive processes (Hoffman, 2018). The study demonstrated that Welsh-English bilingualism can have an impact on reaction times and on the metacognitive processes linked to making more *accurate* judgements. To foreshadow the work to follow in this chapter, the study presented herein seeks to replicate the work of Hampton (2012) but extends it further by investigating whether bilingualism has an impact on the *consistency* of participants' evaluation or assessment of statements, particularly in the face of greater complexity and uncertainty about what evaluation to make. To clarify from the outset, 'consistency' does *not* refer to the consistency of actual responses, but rather to the consistency of appraisal regarding whether a statement is true or false. Therefore, from this point onward, 'consistency' or 'consistency of responses/responding' will be referring to the consistency of appraisal.

In order to examine this, and as explained further below, category membership decisions and general knowledge will be used as tasks. With this in mind, it is therefore important to emphasise that although the questions posed often have straightforward objectively correct answers (consistent with most laboratory-based research on semantic memory), in the 'real

world', decisions made in many contexts are not as obvious, and in some instances, individuals may be faced with decisions where the answer may be unclear. It seems sensible to suggest therefore, that this uncertainty may be further along on the dimension of difficulty in decision making and constitute a more difficult and taxing metacognitive processes.

If, as in the previous chapter, it is assumed that the efficiency of these metacognitive processes is linked to CR in some way, and that bilingualism plays a role in CR, it is possible that bilingual individuals may perform differently to monolinguals. Thus, the focus in this current study remains at the metacognitive level but moves from being merely about how accurate people answer questions, to the consistency with which individuals answer questions and whether the nature of responses vary across monolingual and bilingual populations.

4.1 Background

Before discussing the study, it is first necessary to provide some background information as a framework of context. This is a means of determining what has already been learnt about how people answer more complex questions where the answer is more uncertain, and whether those answers provided are always consistent.

Early research on semantic memory and concept formation focused on *typicality* as a key organisational dimension, and where natural categories are well defined (Bourne et al., 1971; McCloskey & Glucksberg, 1978). That is, it is clear what objects are category members and non-members. For example, which items presented may be classified as members of the key super-ordinate categories; 'vegetable' or a 'fruit.' However, it soon became clear that decisions by participants to such questions were often not so clear cut (e.g., is a tomato a fruit or a vegetable?). That is, some items have membership for which participants were 'fuzzy' about answers. Consequently, natural categories were considered to be too vague, and an alternative characterisation of natural categories was suggested (McCloskey & Glucksberg,

1978). Under this alternative theoretical framing, the assumption was that if natural categories are vague or fuzzy sets (Zadeh, 1996), then category membership is a matter of degree (i.e., more or less) as opposed to all or none. In this case, the suggestion is that objects that are highly *typical* of a category are considered to have a high degree of membership in the category (e.g., robin for the category of birds), whereas those that are less typical of the category have a lower degree of membership (McCloskey & Glucksberg, 1978). This notion of fuzzy sets is important because it introduced the idea that some objects in the real world have an ambiguous status, which means that clearly defined membership criteria do not exist for some categories in the minds of many participants (Zadeh, 1996). This implies that participants must resolve this uncertainty when deciding to answer key questions. For example, the category of animals obviously includes dogs, cats and birds and does not include plants, furniture or rocks. However, when considering objects like starfish or bacteria, decisions around category membership become more challenging (Zadeh, 1996). Thus, it is the idea of decision making under uncertainty or ambiguity that is the key theme of the experimental approach in this chapter, and moreover the degree to which individuals are *consistent* in the resolution of this uncertainty.

With this idea of vagueness in mind, individuals make, read, or hear numerous vague statements during day-to-day life. Key research by McCloskey & Glucksberg (1978) revealed that individuals' comprehension of prevalent semantic categories, such as fruit, sport, and vehicles, is vague in that there were definitely category members for which participants were unsure of the answer. In their study, participants were presented with lists containing potential members of various categories (e.g., orange, pineapple, pear, apple, pumpkin, and olive for the fruit category), and were required to indicate whether each item should be considered a member of the specified category by responding with a 'yes' or 'no', and this was then repeated a week later. McCloskey & Glucksberg found that there was poor consensus on

borderline items, and when participants did not agree on category membership, they were also likely to give inconsistent responses across the two sessions (McCloskey & Glucksberg, 1978). For example, when participants were asked whether an 'olive' is a 'fruit', it was clear that participants were uncertain. This 'borderline' item therefore provoked inconsistent responses from both across and within participants. Put simply, it appeared that participants did not have a strong and consistent opinion about the categorisation.

Following this paper on vagueness in categorisation, Hampton et al. (2012) investigated the extent of people's awareness around the uncertainty and possible inconsistency in their responses to statements. A key point made by Hampton et al. (2012), and the primary rationale for scoring the data in line with the thresholds described in the next section, is that traditional categorisation tasks often limit participants to 'true' and 'false' response options (i.e., an answer that is essentially binary and as observed in the study by McCloskey & Glucksberg, 1978 above). Consequently, when someone selects 'true' as their answer, it is unclear how *certain* they are in this selection, or whether they are 'guessing'. That is, the individual is unable to signal their level of certainty (the possibility that they are certain or less so). Therefore, the typical paradigm masks the finer grained decision being made. Here, the question becomes, if the response options used allow for this level of certainty to be signalled, would this also have a consequence on participants' *consistency* of responding? In essence, are individuals aware of their level of certainty, such that they can signal this in circumstances that allow them to select an answer that indicates such very high certainty or otherwise? Moreover, when they are initially very certain of their appraisal, do they remain as certain when asked again? At the same time, the cognitive process underpinning this entire chain of decision making, might also be linked to metacognitive processes akin to those discussed in the previous chapter (i.e., domain general processes that tap CR and are modulated by bilingualism).

With this in mind, Hampton et al. (2012) argued that people's consistency in their responses on a categorisation task should improve at retest if they are given response options that allow them to select an answer that spanned their individual subjective level of certainty: clearly (100%) true, unclear (generally true or generally false), clearly (100%) false. In other words, the use of response options that span a more continuous level of certainty should enable participants to select responses that reflect their personal level of certainty (e.g., "*I am very certain London is the capital of the UK"* to *"I am not really sure varnish is a tool"*). Thus, Hampton et al. (2012) sought to determine whether using a more 'fine-tuned' response method would result in participants providing more consistent responses across testing sessions. However, before discussing their findings, theories that attempt to explain how participants make decisions in the first place will be reviewed.

4.1.1 Thresholds – How are Decisions Made?

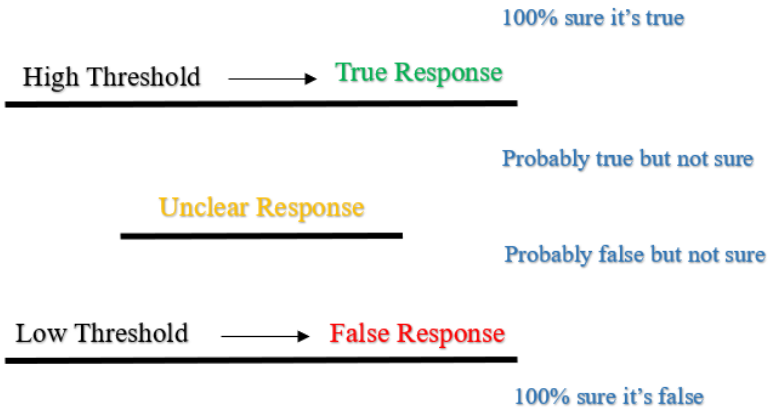
So far, this chapter has introduced the idea of vagueness and the uncertainty that people may face in day-to-day life when making decisions or responding to questions, and that testing of responding should enable individuals to be able to signal when they are very certain of a response (100% True or 100% False) versus less so. This section will focus on how various decisions (e.g., 100% True, 100% False, unclear) are made using a threshold metaphor. That is, how people reach a 'True' response.

With the suggestion (see above discussion) of a more continuous response option in mind (Hampton et al., 2012), as opposed to the traditional response options of 'True' and 'False', it is important to mention that the traditional threshold model of categorisation (Verheyen et al., 2010) largely does not make any clear distinction about the *conviction* individuals may place in their responses across these situations. This model suggests that an individual makes a category membership decision by simply comparing the similarity of the item's

representation to the category's representation (Hampton, 2007; Hampton et al., 2012; Verheyen et al., 2010). If the similarity surpasses a particular threshold, it is approved as a category member (e.g., 'true'); otherwise, a 'false' response is given (Verheyen et al., 2010). Thus, there is no consideration relating to the *conviction* someone may have when they have selected their response.

However, Hampton et al. (2012) argue that this model can be adapted to understand how individuals would make a decision when given the option to signal very clear certainty or otherwise. The suggestion is that individuals may adopt two thresholds, one higher threshold to distinguish a 'clearly or 100% true' response from an 'unclear' response, and a lower threshold to distinguish an 'unclear 'response from a 'clearly or 100% false' response (Hampton et al., 2012). See Figure 21 for a diagram of this kind of threshold model.

Figure 21. *Thresholds Proposed by Hampton et al. (2012). Figure created by Hanna Thomas*



Imagine an experiment where participants are given items and are asked to judge category membership via selecting answers across four different options: (1) 100% True, (2) 100% False, (3) probably True and (4) probably False (see Figure 21). In this paradigm, participants can clearly signal their level of certainty, such that they can either say they are *very* sure something is True (by selecting '1') or *very* certain something is False (by selecting '2').

They can also signal that they are less certain of it being true or false (by selecting '3' or '4'). To provide a concrete example, if a participant is being asked about whether particular items are members of the category 'Fruits', in one case an item presented might be an 'Orange'. In this case, a particular participant is very certain the item is a Fruit (based on stored information in semantic memory), and thus the accumulation of knowledge is sufficient for a response that exceeds the criteria threshold for 100% sure it is 'True' (it may be assumed that they will consistently do so, though to foreshadow what is to come, this remained to be explored). Similarly, an item presented might be a 'Chicken', and again the accumulation of knowledge for that participant is sufficient for a response that exceeds the criteria threshold for 100% 'False' (and again it might be assumed that the participant would consistently select this answer option). However, later - the participant is presented with the item 'Olive'. In this situation, they are uncertain of the answer and could signal this uncertainty by choosing either of the 'probably' answers (and again it might be assumed that they will consistently do so, but perhaps less consistently than in the other two instances since here they are already less certain).

As discussed previously, a key empirical question is raised by the utilisation of this alternative response paradigm. Namely, are people more *consistent* in giving answers when they are given the option to signal their subjective level of certainty of the answer? In particular, are people more consistent when they have made responses that meet the threshold of high conviction of certainty (Figure 21 – 100% sure) versus otherwise? Using the above threshold metaphor, a concrete example of how to score this consistency comparison would be the following: at initial testing, a participant may decide that 20 out of 100 questions/statements are 100% True and that the remaining 80 are not. That is, they choose another response option amongst those displayed in Figure 21. The second time the participant completes the task, the question is, how many of the 20 questions *remain* 100%

True for that particular participant, and how many of the answers to the other 80 questions have now moved to reach the threshold for a 100% True response (i.e., now deemed to be 100% True by the participant)? The same approach can be taken for answers that a participant gives that are deemed 100% False. In both cases therefore, it is possible to get an idea of the average level of consistency a participant has for responses where they have signalled a very high level of *certainty*. To determine if this proportion of responses is generally higher for less certain responses, and thus the *inclusion* of these options has some utility, it is possible to compare consistency scored in a third manner. In this case, it is possible to adopt the approach used in earlier experimental paradigms by essentially scoring all responses selected by an individual as either 'True' or 'False' (i.e., the typical binary response option). As a concrete example: at initial testing, a participant may select 60 as 'True' (collapsed across 100% True and generally True options – see Figure 21), and 40 as 'False' (collapsed across 100% False and generally False response options). On re-test, how many stay in one form of response or the other are then reviewed, and the average level of consistency is calculated (i.e., how many of the 60 answers given earlier as 'True' remain as 'True' etc.). In this case, the participant's average level of response consistency is obtained by using either a typical True/False (binary) response pattern or their level of consistency where they have signalled a very high level of certainty (i.e., 100% True or 100% False).

In sum, the question being explored here is the degree to which the inclusion of the very certain (100% True/False) and 'unclear' (probably True/False) responses has an impact on how consistent participants are in how they respond to different kinds of knowledge statements. Hampton et al. (2012) therefore argued that while the change in the number of responses may seem significant, it does not fundamentally alter the underlying factors that contribute to probabilistic responses and inconsistency in categorisation. In other words, the way similarities are calculated, and the criteria used to make categorization decisions, are not

changed by providing an additional response option. Consequently, the challenge of maintaining consistency will remain regardless of the number of response options given (Hampton et al., 2012).

4.1.2 Domains of Knowledge

Having established that traditional categorisation models largely ignore the impact of an individual's conviction or certainty in the responses they select, this section will now discuss the degree to which responses may vary across different domains of knowledge representation (i.e., judgements of semantic category versus general knowledge).

Hampton et al. (2012) investigated whether people are consistent in their responses to statements over time and whether this changes across different domains of knowledge, not only in semantic categorisation tasks. The statements used represented three *distinct* types of knowledge: (1) semantic knowledge (e.g., *"A rake is a tool"* [semantic categorisation]; (2) factual ('general') knowledge (e.g., *"London has more than 11 airports*"), and (3) autobiographical details/memories (e.g., *"I eat more sweets during the weekend"*). Across three experiments, the consistency of responses in a group of participants presented with a two-response option (true or false) versus a group of participants presented with a three-response option (definitely true, uncertain, definitely false) was examined. Then, a fourth experiment used a within-subject version of the above procedure with four-response options (100% sure it is true, Probably true but not sure, Probably false but not sure, 100% sure it is false). Note, it is the latter response paradigm that was subsequently used in the current study.

General Knowledge (GK) is typically considered a more 'crystallised' intelligence as it is based upon facts. Indeed, crystallized intelligence is generally thought of as language, experience-based knowledge or general knowledge about the world (Baghaei & Tabatabaee, 2015; Cattell, 1987). It is accumulated through life experiences and the storage of information throughout one's lifetime (Salthouse, 2004b). As has been discussed earlier (see

Chapter 3), it seems reasonable to anticipate that older adults will have a higher crystallized intelligence because they have acquired more knowledge over time (Chamorro-Premuzic & Furnham, 2004). However, whilst ageing seems to be related to crystallized intelligence (Horn & Cattell, 1967), little is known about the consequences of bilingual ageing on this construct. However, it may be possible to make predictions based on neuroanatomical correlates of intelligence. Research suggests that the total cortical surface area and average gyrification is positively correlated with crystallized intelligence (Tadayon et al., 2020). Early bilinguals have an extended cortical surface area over certain brain regions like the left pars opercularis and the right superior temporal gyrus (Hämäläinen et al., 2018), and preserved gyrification exclusive to bilinguals has been detected in brain areas associated with normal and pathological ageing (Del Maschio et al., 2019). As a result, it may be that bilingual ageing may lead to higher crystallized intelligence for a specific kind of bilingual compared to that of monolinguals, but there is a lack of literature studying this directly. However, it may be hypothesised that if crystallized intelligence is higher in bilinguals as well as older adults, it may be the case that these groups of people would perform differently to monolingual and younger participants on such GK questions.

In terms of one of the other types of knowledge statements used in Hampton's study, semantic knowledge is one's understanding of facts, objects and the meaning of words (Ralph et al., 1997). Therefore, it is much more closely aligned to the kind of knowledge discussed in the previous experimental chapter on semantic memory and semantic cognition. However, in Chapter 3, the focus was on semantic knowledge and around response accuracy, and so was naturally less well positioned to pose questions for which participants' knowledge might be 'vague'. In the current study, the focus is exactly on this uncertainty and the degree to which participants are *consistent* in responses (e.g., the current study includes statements for which an answer cannot be determined merely by scanning an existing external knowledge base).

The degree to which consistency for semantic questions in this situation may be moderated by bilingual status is an entirely open question. Given that the study in the previous chapter found some evidence for a bilingual 'boost' for semantic knowledge *accuracy* and reaction time, it may well be that a similar benefit is also found in the context of *consistency* of certainty. Indeed, there is some suggestion in the literature that it may be possible to recruit specific aspects of category knowledge which are most relevant to the semantic categorisation task (Yeh & Barsalou, 2006). For example, Ellis and Oakes (2006) found that young children can flexibly change the criteria used to evaluate and categorise objects when the task requirements change (e.g., from shape to colour). Research also suggests that the more rehearsed individuals may be at being able to switch to make an evaluation based on other criteria, the more likely they are to have numerous representations to use for future judgements (Schyns & Rodet, 1997; Verheyen et al., 2019). Therefore, it may be argued that this might be relevant to bilingual populations. As found in Chapter 3, balanced Welsh-English bilinguals were more accurate in controlling semantic information. The frequent language switching and enhanced semantic control processes of bilinguals may play a role in the consistency of semantic membership decisions. That is, they may be more rehearsed at adjusting their response to meet the requirements/demands of a task or conversation, meaning they may have more routes available to access useful information.

4.1.3 Findings from Hampton et al (2012) – The 'Rumsfeld Effect'

As was mentioned earlier, Hampton et al. (2012) explored the impact of using a variable response method across questions tapping either semantic knowledge, GK or autobiographical knowledge. In particular, they contrasted instances where participants could signal a high degree of certainty (by selecting a 100% True response or 100% False response) versus a paradigm in which participants could only select 'True' or 'False' (and thus

participants might select ‘True’ or ‘False’ with varying degrees of certainty that are essentially obscured). As covered previously, the latter binary approach of seeking responses from participants was generally the ‘standard’ approach in previous research (Hampton et al., 2012; Rips et al., 1973; Smith et al., 1974). However, this scoring approach does not allow an individual to signal their level of certainty (the possibility that they are certain or less so).

Their key finding was that the utility of these different response methods differed across the knowledge domains. In relation to semantic and autobiographical knowledge, they reported that participants were generally no more likely to change their responses in a paradigm that enabled people to choose high certainty options (see Figure 21) than when responses were limited to only two options (i.e., True or False). In other words, even when participants were given the opportunity to signal their absolute certainty with response options like ‘100% True’ or ‘100% False’, they were no more consistent in their appraisal (Hampton et al., 2012). This is a finding that Hampton dubbed the ‘Rumsfeld Effect’ - referencing the famous quote by Rumsfeld that *“there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know.”* That is, it appears in this case, that individuals were no more certain for either of these two thresholds of knowledge that the late Donald Rumsfeld described, even when they had signalled 100% certainty in a ‘known known’. This may imply that participants may either be unaware of their inconsistency in their evaluations or may fail to discern when they are merely guessing.

However, although this pattern of inconsistency was clear for the semantic and autobiographical statements, this was not the case for GK type statements. In this case, participants benefited from being provided with a paradigm that enabled them to signal high certainty, such that they gave much more consistent responses across the different thresholds outlined in Figure 21. That is, they were more certain of ‘known knowns’, such that being

able to select ‘100% True’ or ‘100% False’ was more consistent and therefore a beneficial option to provide.

Overall, the findings reported by Hampton et al. suggest that people are unaware of the degree of their uncertainty around some domains of knowledge, including semantic and autobiographical statements. This lack of awareness is indicated by the pattern of consistency observed for GK statements, suggesting that if individuals were aware of their uncertainty, consistency would have improved across all statement types (Hampton et al., 2012). Consequently, some explanation needs to be provided for this variable performance. Hampton et al. (2012) suggested that for most domains of knowledge (e.g., semantic and autobiographical), participants will always face some degree of uncertainty because they will almost always have some knowledge, albeit ‘vague’ to consider. As a result, there will always be some degree of inconsistency in responding (e.g., “I know something about ‘olives’ but not enough to be certain one way or another, so I will change my mind from time to time”) and moreover, people *learn* more over time, meaning that their response will adjust to this new information and in turn impact responding. Based on these findings, it was hypothesised that the Rumsfeld Effect (that is, the lack of certainty of known knowns or known unknowns), occurs mainly in domains of knowledge that involve a degree of personal information variability.

However, in the case of GK statements, it was argued that there are likely to be more instances when a participant is asked something for which their search through memory leads to the retrieval of no information about the given statement (e.g., Chromophobia is the fear of colour), particularly if the topic is unfamiliar (e.g., “I really have no idea what the answer is”). As a result, the distinction between very certain and less certain answers is likely to be much more consistent (Perfect & Hollins, 1996). Put simply, participants may be better able to judge ‘known unknowns’ from ‘known knowns’ for GK statements because in many cases

they have no knowledge whatsoever to draw upon (i.e., GK is related to facts and some individuals may not have any knowledge of certain facts – for example, not all individuals will know what colours make up the Ghanaian flag if they have never seen it before). Therefore, there is no 'vagueness' present to provoke the kind of uncertainty seen in other domains of knowledge (e.g., semantic and autobiographical), and this then boosts the consistency of responses to such GK statements (Perfect & Hollins, 1996).

A key implication of the work by Hampton et al. (2012) is that participants do not always respond consistently to questions about the world around them. In fact, depending on the kinds of questions/statements, inconsistent responding may be the most typical pattern for the general population even when someone has signalled that they are *very certain* of the answer. This finding clearly goes against 'perceived wisdom' in general, that for the most part, people tend to take high certainty responses at face value (e.g., if someone says that the answer to a statement is 100% true, it would be assumed that they would respond in the same way if asked again a short period later). However, the work discussed above clearly pushes against such assumptions, and moreover, the way changes in responses over time are *interpreted*. This has implications for testing in the 'real world', particularly in a clinical setting because it is often assumed that answers provided by patients follow the assumption that people 'know what they know', and so either respond with the answer they think is correct or do not answer at all (like a computer-type response). Clearly, this is not the case, and participants are often 'vague' about what they do or do not know, such that they may respond inconsistently. As a key clinical example, although patients may not be asked a precise yes/no or true/false question (e.g., is a tomato a fruit?), patients may be asked a categorisation type question as part of an assessment. For example, the Wechsler Adult Intelligence Scale (fourth edition) contains a 'similarities' subtest in which patients are asked to specify the relationship between two words (e.g., how is a table and chair similar?). In addition, the Cambridge

Semantic Memory test battery contains a picture sorting task in which patients are required to sort pictures at the superordinate level (e.g., living/non-living), and into basic categories (e.g., household objects versus tools, fruit versus vegetables etc.) (Adlam et al., 2010). Consequently, if patients were set a task like the above example and responded differently on each testing occasion, this could be interpreted as being clinically abnormal, when the evidence described above suggests that this may not be the case. Additionally, the inconsistency could be perceived as variable *effort* when in fact it may merely reflect an individual's honest uncertainty and inconsistency around a body of information, which is rather common.

In sum, more research around inconsistent patterns of responding is required to be able to transfer these findings to the interpretations of performance in other 'real world settings.' Additionally, it is unclear whether inconsistent performance may change with older participants (since previous research is largely focused on younger participant samples). If older adults possess greater crystallized knowledge (Zaval et al., 2015), it may mean that they are also more consistent in responding to various statements because they are sure of what they know. In line with this, research has found that older adults are better at making decisions, with more accurate evaluations of their own knowledge and greater recognition of the limitations of their decision-relevant knowledge (Bruine de Bruin et al., 2012; Kovalchik et al., 2005). Higher accuracy may lead to greater consistency in this context. In addition to older adults, it is also unclear whether other populations such as bilinguals, who are potentially more adept linguistically, may be more or less consistent. With the previous study (Chapter 3) in mind, Welsh-English bilinguals performed differently on particular aspects of the semantic control task, such that there was a 'bilingual boost'. Therefore, there may also be differences with respect to the consistency with which this population answer more

realistic questions. In sum, perhaps bilinguals are more consistent in knowing and responding to ‘known knowns’ relative to ‘known unknowns’.

4.1.4 Present Study

While the study by Hampton et al. (2012) demonstrates that consistency differs across domains of knowledge, such that only GK type statements indicate an advantage of being able to signal high certainty, it is important to note that their work exclusively focuses on young, monolingual participants. Moreover, the consistency patterns identified in their work have yet to be replicated in other work. Overall, therefore, the present study sought to understand whether age and bilingualism would impact the consistency with which people respond to statements.

Thus, the aim of the current study was to:

1. Determine whether the key pattern of consistency of responding across domains of knowledge could be *replicated* by asking participants either GK or semantic category questions. Based on the findings of Hampton et al. (2012), it was expected that only GK statements would lead to more consistent responses across instances of signalling clear certainty (see discussion above).
2. For the first time, investigate the key consistency pattern reported by Hampton et al. (2012), determining whether older adults show a similar pattern to younger participants or whether ageing leads to a difference.
3. For the first time, investigate whether bilingualism would change this general pattern of responding. Therefore, this work will provide entirely novel findings in this context.
4. Investigate for the first time whether the consistency patterns reported by Hampton et al. (2012) change *over time* by testing participants’ second responses at a very short

re-test interval (45 minutes - 1 hour) versus a longer re-test interval (1 week). This manipulation will be discussed in further detail below.

The study by Hampton et al. (2012) only used one re-test interval which was one week after the first session. However, the potential for consistency of responses to change over time is an interesting variable, where the inclusion of another time point may lead to a better understanding of consistency across time. That is, are people more or less consistent with a small or larger interval between exposures, and might re-test interval interact with the key patterns of interest – namely the interaction between response threshold and statement type? This will enable examination of the key interaction of interest at the time period used in Hampton's original study, but also help determine whether the findings of Hampton et al. (2012) were simply an artefact of the specific time period (re-test interval) originally used. In the present study, the pattern of consistency across a short (1 hour) and a longer time frame (1 week – as used in the original work) will be investigated, and the reason for doing so was two-fold.

First, examining a different and shorter time frame may be important in terms of change of knowledge over time, and might therefore be a purer measure of the metacognitive processes used during decision making under uncertainty. That is, it is impossible for a participant to learn new information between the first exposure and an hour later (when they have remained in an experiment setting for the duration). Therefore, if participants are asked to respond to a statement and then again after a shorter re-test interval without the opportunity to learn new information, they are likely searching the *same* existing information in memory to respond to the statements, leading to more consistent responses. However, with that said, using an hour re-test interval also means that consistency might increase merely due to recent exposure leading to ease of retrieval. On the other hand, a one-week re-test interval would allow the possibility that a participant learnt new information, and lead to less consistency because new

information that was not available at initial responding would now be retrieved from memory. Although, if the general pattern of consistency is not impacted by time/re-test interval, this might then rule out the suggestion that changes in level of consistency are possibly impacted by things such as incidental learning over time.

In addition, and with bilingualism, semantic cognition and CR in mind, the metacognitive processes involved in the ability to use information effectively and in a goal-directed manner, may be better examined using an hour interval because no *new* information would be found in semantic memory. That is, the consistency pattern at the one-week interval may be influenced by factors such as: (a) the participant having learned new information or (b) the strategy used to search through knowledge is not the same as the previous instance. Put simply, the one-week interval does not help determine whether participants are more consistent because of differences in the contents of the knowledge store, or because of differences in the way that the knowledge store was searched. However, the one-hour interval does help to understand this because no new information should have been gained, meaning that the metacognitive processes would be searching the same store of information. If the metacognitive processes of bilinguals are more sophisticated, studying their consistency an hour later may reveal the bilingual advantage - where they may exhibit greater consistency in their responses as they are more certain in the first place.

The second reason for selecting an additional re-test interval of an hour as opposed to another additional (longer) interval (e.g., a day), was because it was deemed practically achievable as it allowed the collection of re-test responses to occur across the single first session. Thus, the decision to include a short re-test interval (45 minutes to one hour) and a longer re-test interval (one week) was motivated by both practical and theoretical reasons.

Overall, this study will contribute to our understanding of why different types of knowledge statements lead to less or more consistent responses, and whether the

metacognitive processes used to answer such questions are in some way different in the bilingual and ageing brain. Importantly, it also enables further exploration of the findings from Chapter 2 which will help to determine the generalisability of the bilingual advantage seen in one 'meta-cognitive' context to another. The study will also help determine the degree to which decision making under uncertainty and its relation to bilingualism status is a fruitful 'proxy' for measuring CR.

4.2 Methods

4.2.1 Participants

This study utilised the same recruitment process and inclusion/exclusion criteria as the previous study (Chapter 3), and ethical approval was granted concurrently with the previous study and is detailed in Appendix BB. The study was advertised as 'Bilingualism and the Rumsfeld Effect'. Participants were drawn from the general adult population and were recruited via social media adverts, by contacting various organisations/groups (e.g. University of the Third Age), via the researcher's own social and personal networks, by word of mouth (e.g. via participants), the Swansea University Dementia Research Group's registered participant database, via media interviews (e.g., radio) about the research, by the Swansea University Psychology Participant Pool, and from the initial group of participants recruited for the first study of this thesis. Participation was on an entirely voluntary basis, and initially there were no participant compensation arrangements in place. However, due to recruitment challenges, compensation arrangements were introduced part way through the study. As a result, 37 participants were reimbursed with a £15 Amazon voucher for their time. Individuals were eligible to participate if they were bilingual (Welsh – English) speakers, with Welsh being their first language OR English monolingual and not fluent in any other language. Participants also had to be 18-35 (it was 18-30 initially but due to practicalities and

recruitment challenges, the upper age limit was increased to 35) or 61-91 years of age. Participants also had to report no history of neurological or psychiatric illness, or substance abuse.

4.2.2 Data Exclusions

The following exclusion criteria were applied prior to data analysis. Participants were excluded if they:

1. Chose to withdraw part way through the study, leading to partial/incomplete data.
2. Completed the third time point less or more than a week (7 calendar days) after the second time point.
3. Completed an earlier version of the experiment on Gorilla. Due to an error identified in the original version, the experiment was subsequently changed to resolve the issue.

In total, 105 participants took part in the study. Using the above data exclusions, 18 participants were subsequently excluded for the following reasons: completed the initial version of the experiment containing the error (n = 1); withdrew from the study (n = 13), and completed the final time point more or less than 7 calendar days after the second time point (n = 4).

After these exclusions, the final sample consisted of 87 participants. See Table 24 for summary of demographic characteristics of each group. The final sample of bilinguals were the same bilinguals who participated in the previous study (Chapter 3). The bilinguals' scores on the Welsh version of the LexTALE supported their bilingual status. Please refer to Table 19 in the previous chapter for a comparison of scores across current and normative data.

Table 24

Summary of Demographic Characteristics of each Group.

	n	*M* Age (*SD*)	Age Range	Gender	
				Female n (%)	Male n (%)
Old Bilinguals	22	67 (4.85)	61-78	16 (73)	6 (27)
Young Bilinguals	20	26 (5.68)	18-35	16 (80)	4 (20)
Old Monolinguals	22	68 (4.91)	61-79	13 (59)	9 (41)
Young Monolinguals	23	23 (4.35)	19-32	16 (70)	7 (30)

4.2.3 Materials

Since data for Chapter 2 and the current study were collected at the same time, it is important to note that participants completed a range of tasks. The materials specific to the current study are described below.

4.2.3.1 LexTALE

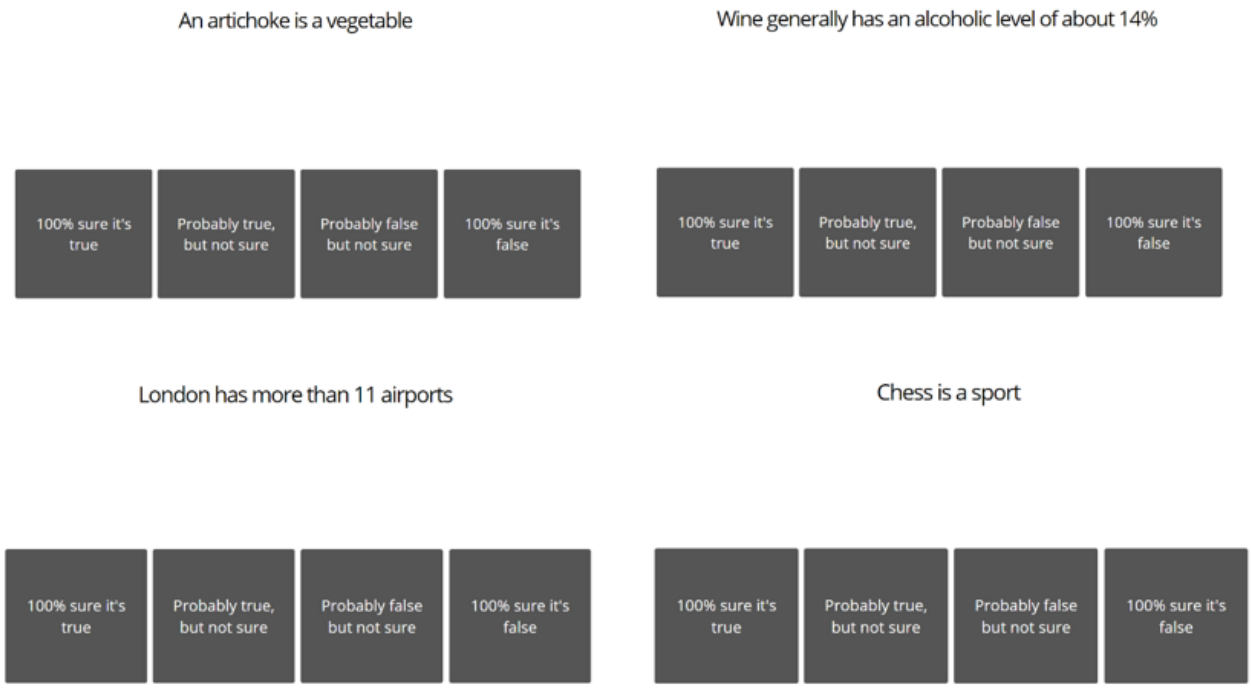
Please refer to the Materials section in Chapter 2 (Section 2.4.4.3) for a description of the Welsh and English versions of the LexTALE.

4.2.3.2 Knowledge Judgement Task

This task examines the consistency of decisions made by participants about GK and category membership questions/statements when they are permitted to place items or concepts into a highly certain response category (Hampton et al., 2012). Participants were presented with 100 statements (one at a time), constructed by Hampton et al. 2012, in random order, and were asked to make category membership decisions or answer questions based on GK.

Of the 100 questions/statements, 50 were GK statements and 50 were category membership statements. For each set of 50 statements, approximately 15 were clearly true (so should provoke a 100% True response), 15 were clearly false (so should provoke a 100% False response), and 20 uncertain (so should provoke an 'uncertain' response signalled by selecting 'Probably True but not sure' or 'Probably False but not sure'. Thus, there were 4 response options for each statement: "100% sure it's true", "Probably true, but not sure", "Probably false, but not sure" and "100% it's false." Participants were required to click/tap one of the four response options depending on the device they were using. No practice items were included in this task. See Figure 22 for examples of both GK and category membership statements and response options.

Figure 22. *Examples of Questions/Statements and Response Options from the Knowledge Judgement Task. Figure created by Hanna Thomas using the Gorilla Experiment Builder*



4.2.4 Design

This study used a mixed 2x3x2x2x2 design. There were 5 IVs in this study:

IV 1 (within subjects): statement type – GK (i.e., "*London has more than 11 airports"*) or Category (i.e., *"Chess is a sport"*).

IV 2 (within subjects): threshold of response type - as participants could choose from the following four options: 100% sure it's true, Probably true but not sure, Probably false but not sure, 100% sure it's false; this enabled analysis of consistency across instances of responses in which participants signalled high certainty to be differentiated. Following the work of Hampton et al. (2012), this study utilised three different measures of consistency which reflected: (a) how consistent participants were in selecting a'100% sure it's True' response or otherwise (Figure 23); (b) how consistent participants were in selecting a '100% sure it is False' response or otherwise (Figure 24), and (c) how consistent people were in choosing 'True' or 'False' responses in general (i.e., responses that fall into 'all true' responses collapsed versus responses that fall into 'all false' responses collapsed; Figure 25). This enabled the degree to which consistency improved when a high certainty response was selected relative to the standard approach of simply asking participants to select 'True' versus 'False' to be examined (see Section 4.1.1 - 'Thresholds – How are Decisions Made?' for a detailed discussion of these three means of measuring average consistency of threshold of response type).

Figure 23. *Analysis of Consistency across Instances when Initial 100% True Responses Remain 100% True and when Initial Responses Outside of the 100% True Box Remain Outside of this Box*[4]

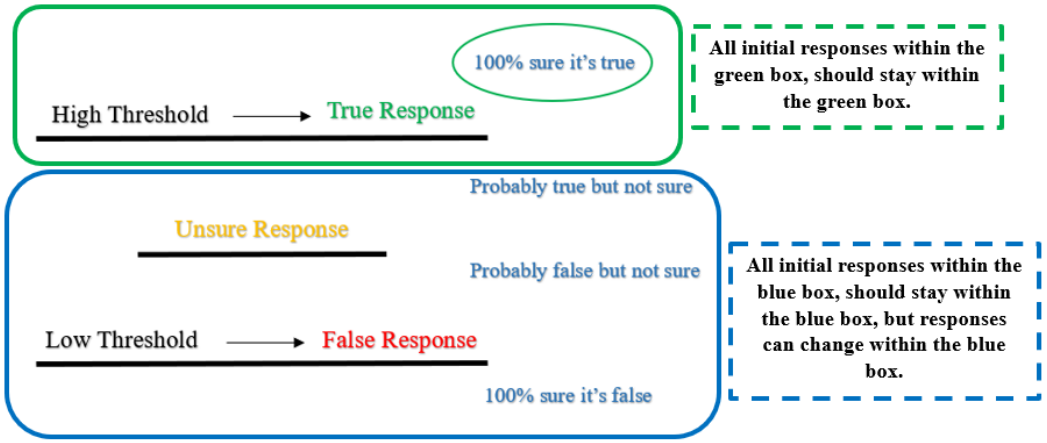


Figure 24. *Analysis of Consistency when Initial 100% False Responses Remain 100% False and when Initial Responses Outside of the 100% False Box Remain Outside of this Box*[4]

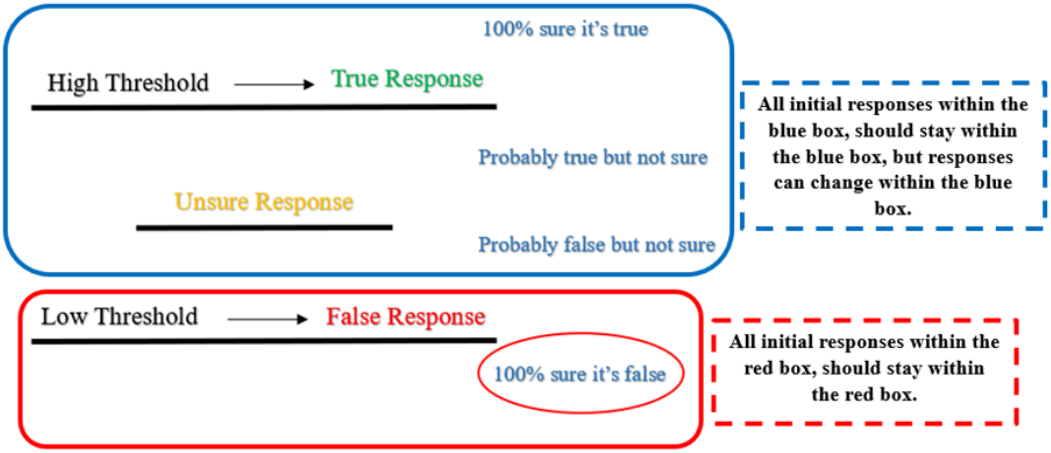
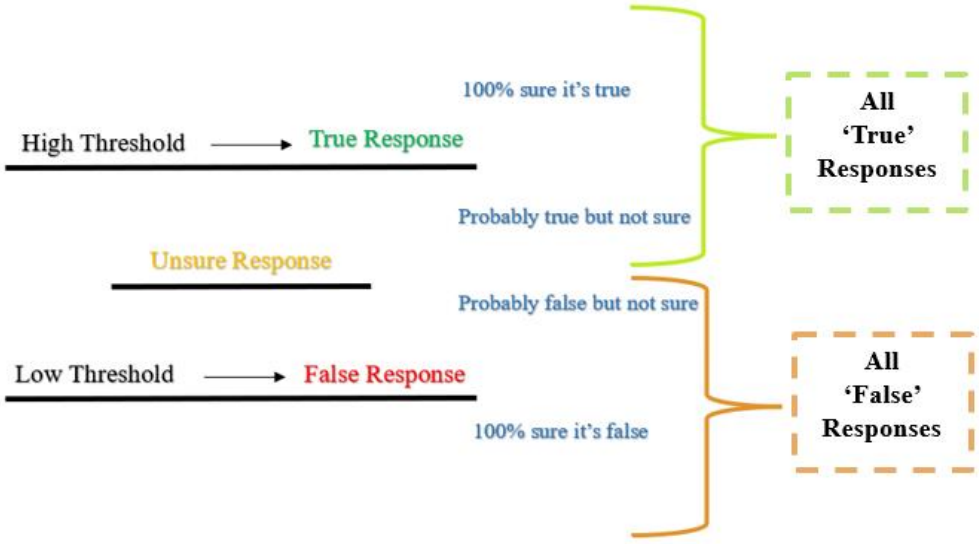


Figure 25. *Analysis of Consistency when 'All True' and 'All False' Responses are Collapsed*[4]



[4] Figures 23, 24, and 25 were created by Hanna Thomas.

IV 3 (within subjects): re-test interval - measuring the degree to which responses selected in the first attempt of the Knowledge Judgement Task were consistent with those given approximately 45-60 minutes later or one week later (the interval used in the study by Hampton et al. (2012).

IV 4 (between subjects): age of participant - young or old.

IV 5 (between subjects): participant language status - bilingual or monolingual.

It is important to emphasise again that unlike the previous study (Chapter 3), this study is not focused on accuracy, but rather the key DV is the consistency with which people respond to different types of statements. Participants were allocated to groups based on their age (young or old) and language status (bilingual or monolingual).

4.2.5 Procedure

Both sessions were completed online via Gorilla. Prior to session one, bilingual participants completed the Welsh and English LexTALE task online. In session one, participants read a participant information page and completed a consent form (Appendix U and V) before commencing the Knowledge Judgement Task. Participants then completed a series of tasks for Study 2 (Chapter 3) before the same Knowledge Judgement Task was completed approximately 45 minutes to one hour later. Participants were informed - *"we are not as interested in whether you are correct in your decision about the statement as we are in whether you know what you know."* One week later, participants were asked to complete the same Knowledge Judgement Task for a third time.

4.3 Results

This section will explore consistency of responding across domains of knowledge (e.g., GK versus Category knowledge), and whether consistency is impacted by re-test interval, age of participant, or language status.

4.3.1 Descriptive Statistics

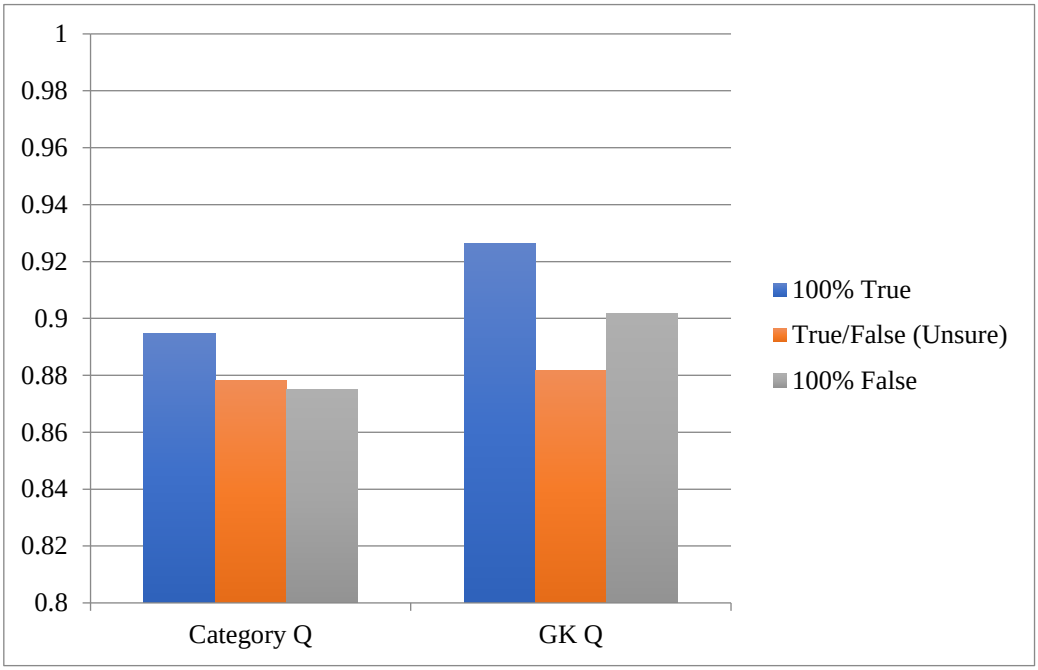
Average consistency levels were calculated across: (a) re-test interval – one hour versus one week, (b) statement type – GK or Category, (c) threshold of response type – 100% True (High), True/False (Unsure), 100% False (Low), (d) age of participant - young or old, and (e) participant language status – bilingual or monolingual (see Appendix Z for full descriptive information).

Responses were scored according to each of the three thresholds of response types (i.e., Figures 23-25) separately and in turn. All three of these scores are presented across the figures that follow. Therefore, all pairs of responses contribute to each of the three scoring thresholds but may not always be considered consistent when scored by each of the three threshold of response types. This is consistent with the approach employed by Hampton et al. (2012).

For GK questions/statements and as shown in Figure 26, it appears that participants were more consistent when scored according to both the 100% sure it is True and 100% sure it is False thresholds relative to the True/False (Unsure) threshold. In contrast, there appears to be less of a difference in consistency scores across the three threshold of response types for Category statements/questions, albeit consistency is somewhat higher when scored according to the 100% sure it is True threshold. Finally, participants also appear to be generally more consistent on GK and Category questions/statements when scored according to the 100% it is

True threshold compared to both the 100% sure it is False and True/False (Unsure) thresholds.

Figure 26. *Mean Consistency of Response Selection across Question/Statement Type and Response Thresholds*



The above general pattern of consistency across GK versus Category questions/statements, also seems to be evident across both monolinguals versus bilinguals (see Figures 27 and 28, respectively), and across young versus old participants (see Figures 29 and 30, respectively). Although, there appears to be less of a difference between the 100% sure it is False threshold and the True/False (Unsure) thresholds in both bilingual (Figure 28), and older (Figure 30) participants for GK questions/statements.

Figure 27. *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Monolinguals only*

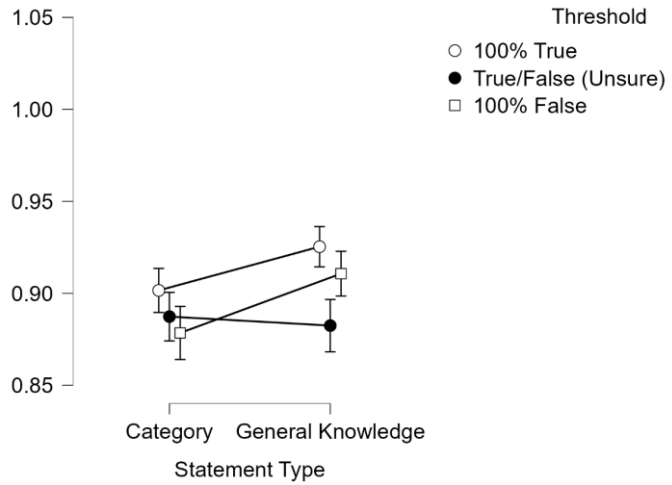


Figure 28. *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Bilinguals only*

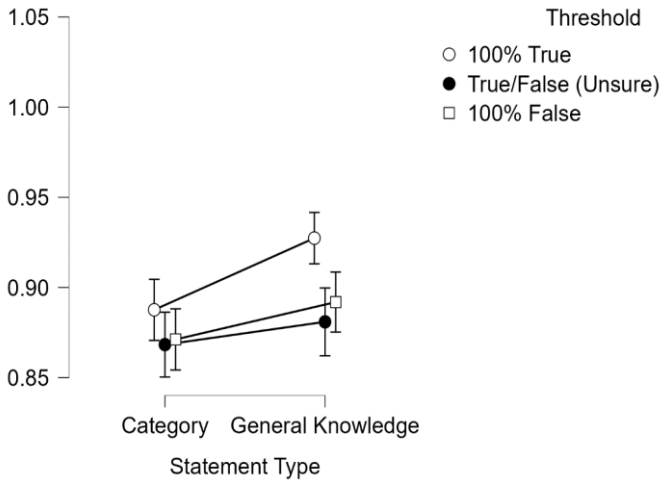


Figure 29. *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Young Participants only*

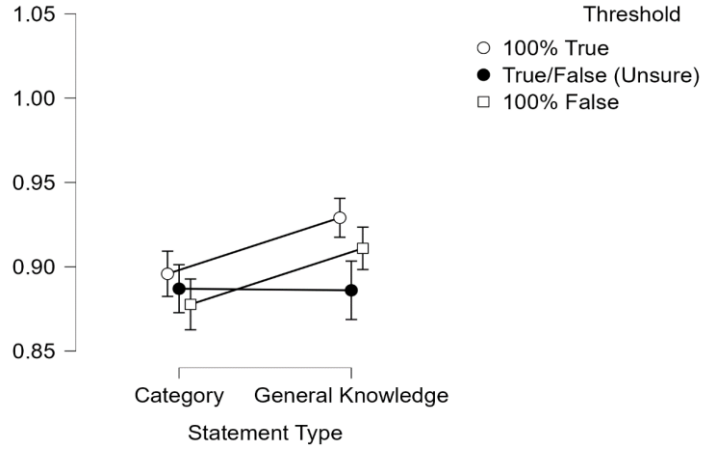
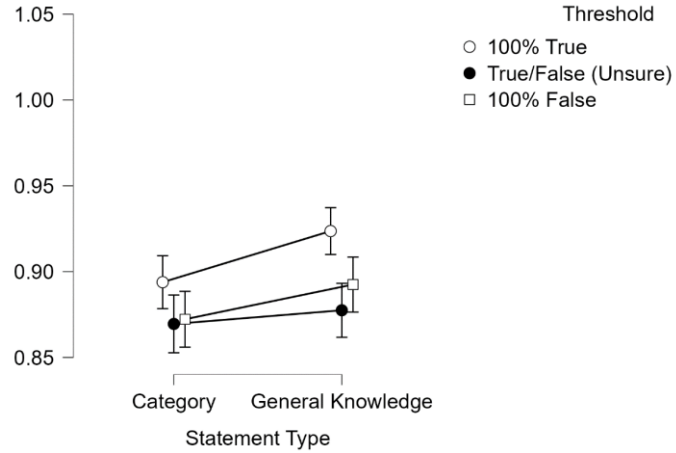


Figure 30. *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – Old Participants only*



Finally, Figures 31 and 32 also suggest that consistency was generally higher at the one hour than one week re-test interval, and that participants were again generally more consistent for GK questions/statements when scored according to both the 100% sure it is True and 100% sure it is False thresholds relative to the True/False (Unsure) threshold. Although, this difference appears to be more somewhat more pronounced at the one-week re-test interval.

Figure 31. *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – One Hour Re-test Interval*

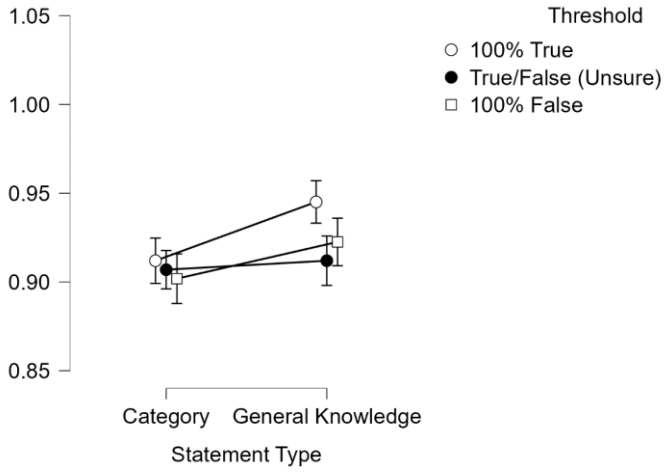
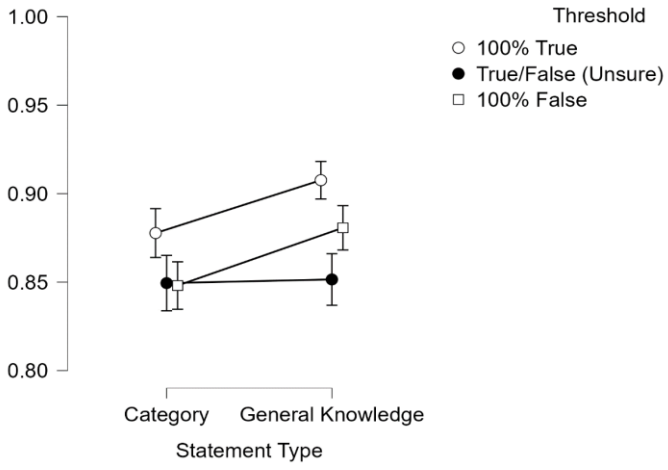


Figure 32. *Mean Consistency of Response Selection across Question/Statement Type and Response Threshold – One Week Re-test Interval*



4.3.2 Inferential Statistical Analyses

As a reminder, the design of the study involved five IVs. Three of these were within subjects: (a) statement type: GK or Category; (b) threshold of response type: 100% True, 100% False and True/False (unsure), and (c) re-test interval: one hour versus one week, and two were between subjects: (a) participant language status: bilingual or monolingual, and (b) age of participant: young or old. To reflect this, a factorial 2x3x2x2x2 mixed ANOVA was undertaken. Please note that given the complexity of this design and analyses, and for the purposes of clarity, the findings are outlined in parts. Firstly, the impact of the three within subject variables (as main effects and their interactions) are considered, followed by the two between subject variables (as main effects and interactions). The sphericity assumption was violated for the threshold of response variable, so *p* values for all effects including this IV are reported following the application of the Greenhouse-Geisser correction. There was equality of variance in all but one condition (general knowledge, at one hour, using the collapsed threshold) where $p = 0.049$ (see Appendix Z). Violations of the equality of variance assumption are not a concern for the findings in this case because there are a similar number of participants in the between-subjects groups. There were no significant two- or three-way interactions between IVs that were varied within and between participants (see Appendix Z for full details). In all post-hoc analyses reported below, the Bonferroni correction was also applied.

4.3.2.1 Within Participants IV Analyses – Statement Type, Threshold of Response Type, and Re-Test Interval

A statistically significant main effect of statement type was found, $F(1, 83) = 26.05$, $p < 0.001$, $\eta p^2 = 0.239$, where average response consistency was generally higher for GK ($M = 0.903$) than Category ($M = 0.883$) question/statements. There was also a significant main

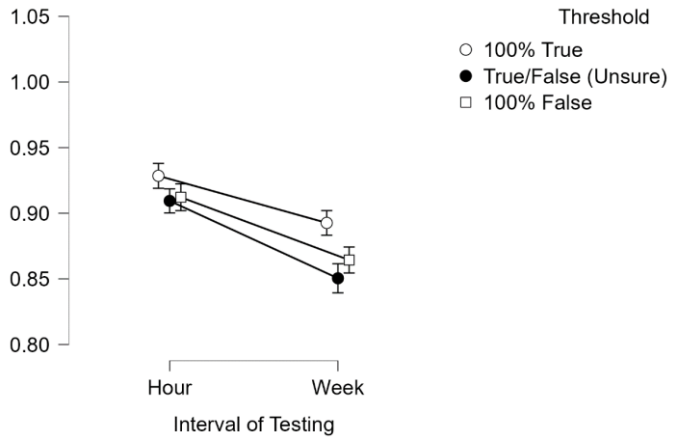
effect of threshold of response type, $F(2, 166) = 19.374$, $p < 0.001$, $\eta p^2 = 0.189$, reflecting that participants were generally more consistent when scored according to the 100% sure it is True threshold ($M = 0.911$), compared to both the 100% sure it is False ($M = 0.888$) and True/False (Unsure) ($M = 0.88$) thresholds. Post-hoc statistical analyses confirmed this pattern to be the case (see Appendix Z for all outputs).

Importantly, there was a significant interaction between statement type and threshold of response type, $F(2, 166) = 9.925$, $p < 0.001$, $\eta p^2 = 0.107$. For GK questions/statements, post-hoc analysis indicated that participants were less consistent when scored on the GK True/False (Unsure) threshold ($M = 0.882$ - see Figure 26) relative to both the 100% sure it is True ($M = 0.927$; $t = 7.33$, $p < 0.001$) and 100% sure it is False ($M = 0.902$) thresholds, $t = 3.206$, $p = 0.01$. Further, responses yielded a higher level of overall consistency on GK statements/questions when scored according to the 100% sure it is True than the 100% sure it is False threshold, $t = 4.214$, $p < 0.001$. In effect, and in relation to GK statements/questions, the general pattern of consistency was 100% sure it is True > 100% sure it is False > True/False (Unsure). Additionally, a higher degree of consistency for the 100% sure it is True threshold was also found on the category questions/statements. Specifically, post-hoc analyses determined that the pattern of responses was more consistent when scored according to the 100% sure it is True ($M = 0.895$) than both the 100% sure it is False ($M = 0.875$; $t = 3.251$, $p < 0.01$) and True/False (Unsure) thresholds ($M = 0.878$; $t = 2.759$, $p = 0.037$). Finally, no other key interactions were present across statement type (GK or Category) or threshold of response type (see Appendix Z for all outputs).

In relation to re-test interval (one hour versus one week), a significant main effect was found, $F(1, 83) = 29.631$, $p < 0.001$, $\eta p^2 = 0.263$, such that participants were more consistent at the one-hour ($M = 0.917$) than one week ($M = 0.869$) re-test interval. Further, a significant interaction was also found between re-test interval and threshold of response type, $F(2, 166)$

= 10.273 $p < 0.001$, $\eta p^2 = 0.11$ (see Figure 33). Whilst participants were less consistent across all response thresholds at the one-week re-test interval, post hoc analysis showed that the average rate of consistency declined at a steeper rate for both the 100% sure it is False and True/False (Unsure) thresholds. Put simply, the change between the one hour and one week interval was smaller for the 100% sure it is True threshold (See Figure 33). Finally, no other key interactions were found, including the three-way interaction between re-test interval, statement type and threshold of response type (see Appendix Z for all outputs).

Figure 33. *Mean Consistency of Response Selection across Re-test Intervals*



4.3.2.2 Between Participants IV Analyses – Age of Participant and Language Status

No significant main effect of age (young or old) was found, $F(1, 83) = 0.793$, $p = 0.376$, $\eta p^2 = 0.009$, indicating that average consistency scores were not affected by age of participant. Similarly, no significant interactions were found between age of participant and the other IVs (e.g., statement type, threshold of response type). Therefore, findings suggest that age is not a key moderating factor in the average consistency of response patterns discussed throughout.

Similarly, no significant main effect of participant language status (bilingual or monolingual) was found, $F(1, 83) = 0.705$, $p = 0.404$, $\eta p^2 = 0.008$, and no significant

interactions were observed between participants' language status and the other IVs (see Appendix Z for all outputs). Thus, participants language status does not appear to be a key moderating factor in the average consistency of response patterns discussed throughout.

4.3.3 Overall Summary

In summary, the analyses above have indicated that consistency of repeated responses to the same question is higher when participants are more certain of their initial response, when the re-test interval is shorter, and when the question/statement is related to GK rather than Category membership. Responses to GK questions/statements are more consistent when scored using either of the "100% sure" thresholds than the True/False (Unsure) threshold of response type; the pattern is not the same for Category membership questions/statements. Importantly, this interaction does not appear to be moderated by re-test interval (one hour versus one week), age of participant (young or old), or participants language status (bilingual or monolingual).

4.4 Discussion

This study sought to replicate the findings of Hampton et al. (2012) and determine whether age and bilingualism influence people's consistency in making decisions under uncertainty. As a reminder, 'consistency' in this study referred to the consistency of appraisal (or conviction) as opposed to the consistency of actual responses from participants. Previous research has focused on the organisation of semantic memory, peoples' understanding of certain categories of things (e.g., natural categories), and how decisions are made based on models of semantic memory (Glucksberg & McCloskey, 1981; McCloskey & Glucksberg, 1978 and 1979). Importantly, the *consistency* of decision making under ambiguity and uncertainty has rarely been a topic of investigation, so the evidence base is extremely limited.

In addition, of the little research that has been conducted (Hampton et al., 2012; Verheyen et al., 2019), only young and/or monolingual participants have been considered meaning that little is also known about how consistent bilinguals and older adults respond to various statements.

With the central theme of bilingualism and CR in mind, this study was also interested in understanding whether a bilingual advantage may be found in tasks that explore the consistency with which people respond to statements. That is, increased CR may be a consequence of a more efficient cognitive system that operates on semantic information. More specifically, the metacognitive processes linked to making more *consistent* judgements might be more sophisticated in the bilingual brain which may increase CR. This is unexplored in the literature, and this was the motivation for this experimental chapter. Accordingly, this study investigated whether Welsh-English bilinguals are more consistent in their beliefs about various statements (e.g., GK and category membership) compared to English monolinguals.

Replicating the findings of Hampton et al. (2012), consistency for GK statements was found to be higher for responses scored according to the 100% True than 100% False threshold. That is, participants were more consistent in their responses to GK statements and when scored according to the 100% sure it is True threshold. Hampton et al. (2012) suggested that people might be more consistent with GK statements because it is likely that people are presented with statements for which they have no information in memory to retrieve, and so are more likely to answer in the same way each time (Perfect & Hollins, 1996).

However, this study also found something that was not reported by Hampton et al. (2012). That is, consistency was generally higher for responses scored according to the 100% sure it is True threshold across the board, compared to the other scoring options. Thus, when given the option to signal their high degree of certainty about the truth of a statement, participants

were more consistent. This finding may have materialised in the current study but not in Hampton et al. (2012) because the present study used a larger sample size and was more diverse in terms of age and language status. In addition, the idea that people tend to overestimate the number of answers they get correct on a quiz might also explain this finding (Fischhoff et al., 1977). This reference to the overconfidence bias offers a plausible explanation for why more participants opted to signal high certainty.

Indeed, research on calibration of judgements has shown that people tend to overestimate what they know and be overconfident in what they know is correct (Lichtenstein et al., 1977). Calibration is a common proxy of metacognition which has been defined as the extent to which an individual's assessment of their performance aligns with their actual performance (Bol & Hacker, 2012; Keren, 1991). Consequently, if people are overconfident in their awareness about what they know, it may make sense that they will also be more consistent for responses scored according to the high certainty 100% True threshold. However, more research is needed to confirm whether people are indeed more consistent in their appraisals scored according to this threshold (100% True) or whether it was a characteristic of the sample here. It is also important to restate that this study was not concerned with the accuracy with which participants responded to questions, but rather the consistency with which they responded. That is, participants' confidence in their answer to a question/statement. Studies investigating calibration in educational settings have found that greater precision in calibration leads to improved academic performance, including higher grades on exams (Callender et al., 2016; Zhao & Ye, 2020). In the current study, this may mean that participants who were better at calibration and were also unsure about the response to a statement, were very aware that they were unsure and therefore remained outside of the 100% certainty threshold on each occasion (i.e., remained in the blue box in Figure 23). This would also contribute to the consistency of responding in this threshold.

Further, the current study also found that participants were more consistent for responses scored according to the 100% True threshold for semantic category statements – another finding not reported by Hampton et al. (2012). The lack of higher consistency for semantic categories in the study by Hampton et al. (2012), was thought to be because people will always have *some* amount of information about semantic categories to retrieve from memory which may always lead to some degree of inconsistency. However, the present study provides early evidence that this may not be the case and that it may be less to do with the amount of evidence required by the individual before a decision is made. Indeed, an explanation for the higher consistency for appraisals scored according to the 100% True threshold for semantic categories may lie in literature exploring criterial vagueness (Verheyen et al., 2019). It may be that people are clearer about what parts of concepts they focus on or evaluate to make some category membership decisions than others. This clarity would mean that they are likely to consistently use the same evaluation process each time they are required to make a category membership decision. For example, if an individual is of the opinion that the method of categorising activities into 'sports' *must* involve the extent of physical activity involved as opposed to team playing, they are more likely to always use this same method of evaluation each time and therefore respond consistently (Verheyen et al., 2019). It is also important to note that individuals may have several criteria which they use to make evaluations (e.g., in the context of fruit, some may categorise based on colour, the presence of seeds, culinary usage, taste and flavour), with the importance of each not equivalently weighted in making the decision. For example, some of these criteria may be more meaningful or influential in the decision-making process, meaning that they may be more consistently adopted, leading to more consistent responses. Unlike the previous study (Chapter 3), no impact of bilingualism was found in the present study suggesting that bilingualism does not provide the same 'boost' for consistency of responding. Consequently, language switching and enhanced semantic

control processes in bilinguals are unlikely to influence the consistency of responding and therefore may not be related to a bilingual advantage in CR.

Another explanation for the difference in findings for the consistency of semantic categorisation may be linked to the samples used across studies. That is, and as mentioned previously, the sample size used in the original study by Hampton et al. (2012) was smaller and only included young, monolingual speakers.

It is also important to note that the data were grouped according to the thresholds approach described earlier in this chapter, primarily because this study sought to replicate the methodology of Hampton et al. (2012). While the data could have been categorised according to whether each participant chose the exact same response at different time points (e.g., 100% sure it is true at both time 1 and 2), this approach was not necessary to investigate participants' appraisals or evaluations of statements. Thus, in this study, 'consistency of responding' refers to participants' appraisals regarding whether a statement is true or false, rather than whether they chose the same response option each time.

In terms of the consistency across re-test intervals and as expected, participants were more consistent at the one-hour compared to one-week interval. A possible explanation for this finding is that participants may have remembered their chosen response at the one-hour re-test interval and were therefore more likely to choose the same response again. In contrast, participants may be less likely to remember their chosen responses a week later, and therefore respond inconsistently at that re-test interval. In addition, it is possible for participants to learn new information during the one-week interval which means they may respond differently, leading to inconsistent responding. Interestingly, the results demonstrated that participants were more consistent for responses scored according to the 100% True threshold across both testing intervals. Therefore, people's certainty and conviction in what they know (mentioned above) (Lichtenstein et al., 1977) seems to be stable across time.

As previously stated, one of the key aims of this study and one of its novel contributions to the literature, was to explore the impact of Welsh-English bilingualism and age on the consistency of responding under uncertainty. That is, it sought to understand whether bilingualism and age would modulate the interaction between statement type and threshold of responding. The findings suggest that age had no effect on participants' consistency of responding, which implies that the cognitive processes involved in responding to statements consistently, are unlikely to be affected by age. This was also the case for language status, suggesting that bilingualism is unlikely to play a role in the *consistency* with which people respond to statements. The findings of this study suggested that the more effective bilingual processes of SC found in the previous chapter, are unlikely to be involved in this new context of consistency of responding. That is, the factors that influence the consistency of responding may not be metacognitive and may go beyond cognition. However, more research is required to confirm this.

In terms of factors that go beyond cognition, it may be that personality influences decision making processes in this context. For example, by quantifying responses to a novel food and mild threat, Carter et al. (2013) explored how the personality traits boldness and anxiety affected decision making under conditions of uncertainty in 55 chacma baboons. They assessed foraging decisions in two scenarios: natural foraging with reliable information about their environment, and during in-situ foraging experiment where they had access to experimental food patches with unreliable information about their contents (e.g., when the contents of the food patch was unknown). They found that personality affected foraging decisions only when information was unreliable. Drawing parallels between decision-making in animals and humans, such findings therefore suggest that individual differences may also influence decision making in humans when faced with uncertain information. That is, when

individuals lack reliable information, whether originating from their external environment or from within memory, their responses may become more inconsistent.

Another factor that may play a role in the consistency of decision making is people's thoughts about the opinion of the general population (Acevedo & Krueger, 2004). That is, when asked to respond to Category statements for example, participants may think about how other people would respond to such statements. This process in itself can be considered a metacognitive process though – thinking about the opinions of the general population to guide one's own responses about a statement for which the individual may not have any information in memory to guide a response. In sum, there may be a host of factors that influence the consistency of responding that may be unrelated to language and processes of SC.

4.4.1 Strengths and Limitations

Given the one-hour re-test interval, participants may have been more consistent at this re-test interval phase merely because they were able to remember their previous response to the same statements. Future research is required to examine other re-test intervals to determine whether length of the interval (time) is an important factor in the consistency of responding. However, determining an appropriate re-test interval is challenging because it requires a compromise between avoiding a lengthy interval that would provide participants with an opportunity to learn new information (otherwise, artificially increasing inconsistency), and an insufficient interval where participants may remember their previous responses (artificially reducing inconsistency). One day or three to four *supervised* hours may be a sensible compromise, but this was not possible in the present study because of the nature of remote data collection.

In addition, and as mentioned previously, research investigating the consistency with which people respond to statements is extremely limited, with the current study introducing

new variables into this context (i.e., bilingualism). Equally however, due to the limited research available on this subject and the inherent complexity of investigating multiple variables without established reference points in the literature (discussed further in Chapter 5), it is challenging to find valid and well-researched explanations for the findings revealed in this study. It is not possible to draw any conclusions as to what factors play a more central role in this context. More research is required that examines domains of knowledge, individuals' awareness of knowledge, the effect of ageing and different types of bilingualism on the consistency of responding, and other various factors that may influence consistency.

4.4.2 Conclusion

To the author's knowledge, this is the first study to investigate the impact of bilingualism, age, and re-test interval on the consistency with which people respond to various statements. The main contribution of this study is that it has provided initial evidence that age and bilingualism, at least in individuals without dementia, is unlikely to play a role in cognitive processes involved in the consistency of responding to specific kinds of knowledge statements. However, no firm conclusions can be made until this is replicated by further research. In addition, it remains unclear whether the processes involved in making decisions under uncertainty is associated with cognition or factors that are more associated with the individual (e.g., personality).

Key Messages

- This study aimed to replicate the findings of Hampton et al. (2012) and explore whether age, bilingualism, and re-test interval would influence participants' consistency in responding to various statements.
- This study also sought to understand whether consistency of responding under uncertainty, is where a bilingual advantage that potentially contributes to CR may be found.
- Previous research has proposed models of semantic memory (Collins & Quillian, 1969; Collins & Loftus, 1975; Rips et al., 1973) to understand how meaning is organised in the brain, how concepts are connected, and to investigate people's understanding of semantic categories (e.g., animals). Research demonstrated that participants were unsure about the category membership of certain items (McCloskey & Glucksberg, 1978).
- However, most of the studies in this area focus on monolinguals, and little research has also explored how consistent people are in responding to different kinds of statements (e.g., GK or semantic categories).
- Replicating the findings of Hampton et al. (2012), participants were more consistent in their responses to GK than Category statements. Neither age nor language status impacted consistency of responding. The study also found some unexpected findings in relation to Category type statements – consistency was higher for responses scored according to the 100% sure it is True responses for this statement type.
- More research is required to replicate these findings before conclusions can be made about the factors that might influence consistency of responding in this context. This will allow a better understanding of whether decision making under uncertainty is indeed a metacognitive process.

Chapter 5 – General Discussion

5.0 Chapter Overview

This final chapter will provide closure to the thesis by re-iterating the main objectives, summarising the key findings of each study, and highlighting their contributions, limitations and implications. This chapter will also outline potential directions for future research before providing a thesis conclusion.

5.1 Summary of Thesis Aims

This thesis aimed to study the impact of a unique and largely ignored type of bilingualism on cognition across the lifespan, particularly cognitive reserve (CR). This was attempted in three different ways by: 1) investigating within population performance differences across Welsh and English versions of general cognitive assessment batteries, as well as comparing this performance with that of normative data, 2) examining key metacognitive mechanisms that operate on semantic memory systems, and 3) exploring the consistency of decision-making under uncertainty – another process considered to be 'metacognitive'.

5.2 Summary of Each Study: Their Contribution to Knowledge and Practical and Theoretical Implications

The first study (Chapter 2) investigated whether native Welsh speakers (Welsh-English bilinguals) without dementia would perform differently on general cognitive assessment batteries – the RBANS and ACE-III. That is, this study focused on gross cognitive performance differences across general and relatively well researched cognitive domains (i.e., attention, language, memory, visuospatial abilities). Despite finding no difference in performance across the Welsh and English versions of the assessments, the practical implications are significant. The study was the first to collect data from healthy Welsh-

English bilinguals on the RBANS and ACE-III. This data led to the early evidence that native Welsh speakers are unlikely to be placed under a disadvantage if assessed using the English versions of the assessments. However, it must be noted that these findings may have been different if a clinical sample of individuals living with dementia were assessed, but this was beyond the scope of the thesis due to the COVID-19 pandemic. This will be discussed in more detail below.

Despite no performance differences being found across the Welsh and English versions of the assessments, and if this remains true for dementia populations, it is important to remember that the consequences of undergoing an assessment in the individual's non-preferred language go beyond cognition. As discussed in Chapter 2, assessing an individual in their preferred language may be beneficial for their experience, is aligned with providing person-centred care (Marulappa et al., 2022; Older People's Commissioner for Wales, 2016; Welsh Language Commissioner, 2012), could mitigate against the risk of test related stress and anxiety (Dorenkamp & Vik, 2018), and is also a legal requirement in Wales. (Welsh Government, 2016). Therefore, even though further research is required to collect data with dementia populations to validate the Welsh versions of the RBANS and ACE-III, this study has recognised the importance of improved Welsh language provision in older adult services and has played a vital role in responding to the needs of Welsh speaking populations.

One of the central themes of this thesis is CR in the context of bilingualism. The lack of strong evidence of a bilingual advantage when comparing the bilinguals' performance on the English assessments with that of normative data suggests that a bilingual advantage may not reflect a general improvement across generic cognitive processes (at least in the Welsh context). However, and as mentioned previously, drawing comparisons with previously published normative data is difficult owing to differences in sampling and/or design (e.g., inclusion/exclusion criteria, language of the assessments, language status of the sample).

Consequently, conducting the same study with appropriate age-matched English monolinguals would be a helpful next step to provide more concrete evidence of whether Welsh-English bilingualism leads to a general uplift of cognitive processes.

Adoption of this approach would also mean that test-retest correlations could be compared within and between languages in well-matched groups, which would inform clinical evidence-based practice. For instance, comparing test-retest correlations within language is important to evaluate the reliability and consistency of language-specific assessment tools. In turn, this would help to determine the validity of such assessments in monitoring cognitive decline over time and for guiding treatment and management plans. Also, comparing test-retest correlations between languages would determine the reliability and consistency of the assessments across linguistic contexts. Importantly, between-language comparison could also help determine whether a difference in performance across languages is due to language-specific issues or cognitive impairment. In addition, if CR helps bilinguals to complete tasks in the face of injury or disease and/or is employed to complete more challenging tasks, then it may be important to be mindful of bilinguals' scores on cognitive assessments because CR may be 'masking' the extent of their cognitive impairment. That is, CR may act as a 'buffer' against the negative impacts of ageing and therefore cloud assessment scores.

For bilingual individuals, the above may mean drawing on other means of measuring the early signs of cognitive decline, as total/global scores may not reach the cut-off for an 'abnormal' classification. Indeed, if CR is more frequently employed for challenging tasks, it seems reasonable to suggest that there may be a stronger correlation between subtests in healthy bilinguals. The score on a particularly challenging subtest could reflect a genuine skill in the specific cognitive process being measured, or the use of CR to bolster performance on a difficult task. Consequently, when a bilingual individual is affected by cognitive decline, the correlation between challenging subtests may become weaker than it is

in an unimpaired bilingual (in that the CR cannot fully compensate for the decline in the cognitive process being tested) while the overall scores remain above the cut-off score that indicates abnormal performance. Therefore, consideration of consistency over time in this way may transpire to be a more effective and reliable measure of cognitive decline for bilinguals, although this is still an open question. To further investigate a possible bilingual advantage and in terms of testing an appropriate age-matched English monolingual sample, it would also be possible to analyse whether a correlation between the two scores of a bilingual population is of a significantly different strength or direction of that observed in another population (e.g., monolinguals). However, moving away from general cognitive domains, a key argument that is clear throughout this thesis is that a better proxy of CR might be found if studies move away from the traditional domains of cognition that are widely studied in the literature, and focus on new and unexplored cognitive processes in this context.

Consequently, the second study presented in this thesis (Chapter 3) explored the impact of Welsh-English bilingualism and ageing on processes of semantic cognition (SC) to identify a more effective proxy of CR. Specifically, whether bilinguals and younger and older adults may differ in the way in which they retrieve and manipulate semantic information. Few studies have researched the impact of cognitive control processes in semantic memory, and those that have tend to focus on monolingual individuals (Hoffman, 2018; Wu & Hoffman, 2022). This study was the first to the author's knowledge to investigate the impact of bilingualism on these processes and the first to find a possible role of bilingualism in SC.

Age was found to impact on processes of SC, such that the findings replicated the original work of Hoffman (2018). That is, the age effect was different on the Global and Feature judgement tasks. In the first instance, older adults were more accurate than younger adults on items that demanded High control, but in the latter instance, accuracy was higher for younger adults on trials with significant semantic control demands.

In relation to bilingualism, the key takeaway of this study is that a bilingual advantage in the Welsh context may be related to controlling the information stored in memory and inhibiting redundant information. The study also found that bilinguals maintained a more rapid reaction time even when cognitive demands were increased (Feature task), compared to monolinguals. Therefore, the ability to accurately and rapidly select the task-appropriate information whilst inhibiting task-irrelevant information seems to be more efficient in Welsh-English bilinguals. This addresses an important gap in the literature by increasing the small number of studies exploring cognition in Welsh-English bilinguals, and by challenging existing reports of no performance differences between this population and English monolinguals on various cognitive tasks (e.g., Clare et al. 2016b; Gathercole et al., 2014). In addition, and in order to investigate the differences in cognitive performance between bilinguals and monolinguals, this thesis has highlighted the importance of moving away from established and well researched cognitive processes to more specific and unexplored areas of cognition that tap the main difference between bilinguals and monolinguals: language processing.

The findings from this study have practical implications for the way in which semantic memory is assessed in Welsh-English bilingual populations. That is, this early evidence suggests that clinicians may need to ensure that semantic memory assessments are sufficiently challenging to detect cognitive decline in this population of bilinguals because they are likely to outperform monolinguals. The consequences of not using a sensitive measure may lead to challenges in understanding the extent of the individual's impairment as well as a reduced ability to track disease progression over time. For example, in the Pyramids & Palm Trees Test, individuals are asked to generate a word that may be related to the stimulus presented, which assesses their ability to make semantic associations and identify the overall theme whilst ignoring other semantic connections (Klein & Buchanan, 2009).

Therefore, in this instance, it may be possible to increase task difficulty by ensuring that the items used require the individual to make feature-based associations as opposed to general or global semantic associations. As mentioned in Chapter 3, feature-based associations are more challenging because they are less automatic. That is, the individual must attend to particular aspects of an item rather than make semantic links based on an overall theme, which may be more automatic. Utilising such tasks would then increase the likelihood of detecting cognitive impairment in bilingual populations when it occurs, as opposed to it going undetected because of the lack of task difficulty. In addition, developing language specific versions of various semantic assessments so that their semantic abilities are assessed in each language separately may also help obtain a more valid understanding of the individual's level of cognitive decline. Such assessments could also include words that only exist in one language but not in the other so that bilinguals are required to make semantic associations that are specific to the linguistic context. Time constraints could also be an additional method of increasing task difficulty by increasing the cognitive demands of the individual. In sum, the key point is that it cannot be assumed that the assessments routinely used for semantic memory are appropriate for some bilingual populations. In addition, considering the language needs of this bilingual population is also crucial given the importance of person-centred care (Marulappa et al., 2022) and ensuring Welsh-English bilinguals receive the same standard of care as English monolinguals in Wales. Therefore, whilst Welsh language versions of semantic dementia tests are required, it is equally important that English versions are fit for purpose for bilingual populations and can account for their superior performance.

Having outlined the practical implications, it is also important to note the theoretical implications of Study 2 (Chapter 3). This thesis examined a distinctive group of bilingual individuals proficient in both Welsh and English, who possess the ability to effortlessly switch between the two languages without experiencing any cognitive disadvantages. It

seems logical to suggest that one of the main differences between these types of bilinguals and less balanced bilinguals is language switching. On that basis, it is plausible that this seamless language switching may contribute to their superiority in controlling semantic knowledge, compared to monolinguals. Moreover, the nature of their language switching also suggests a high degree of cognitive flexibility which may be an essential component of semantic control. Consequently, this study provides theoretical implications for the field of bilingual language processing by highlighting the potential advantages of seamless language switching and its potential relationship to semantic control. More specifically, it highlights the potential impact of language-switching on the ability to select the appropriate information from memory to complete a task. To support the idea that frequent language switching may play a central role in the bilingual advantage, Verreyt et al. (2016) found that Dutch-French balanced bilinguals who frequently switched languages displayed smaller congruency effects on the Simon arrow task and Flanker task compared to Dutch-French balanced bilinguals who did not frequently switch languages. It is also important to note that even after controlling for second language proficiency, the frequency of switching was correlated with performance on the tasks (Verreyt et al., 2016). These findings, along with those of this thesis, suggest that a control advantage may be observed when the lexical representations of both languages are frequently and simultaneously active (i.e., balanced bilinguals in dual-language environments who switch languages within conversations). However, an important consideration is that there may be two parts to the act of language switching, one is being able to select the correct language, and the other is the ability to ignore/inhibit the other language (see Chapter 1 for a detailed discussion). Importantly, because of the difference in language use between bilinguals, different types of bilinguals may be stronger at *one or both* of these processes.

Consequently, if the bilingual advantage is more related to the frequency of switching, it is unclear whether the findings around semantic control would persist for bilinguals who do not operate in a dual-language context. In other words, it is not yet clear whether language switching (or which aspect of this process) is responsible for increased semantic control. As mentioned previously, not all bilinguals are the same in terms of their level of proficiency, situational language use, language switching, and language identity (e.g., de Bruin, 2019). Therefore, different classifications of bilingualism may lead to different consequences on processes operating on semantic memory. Indeed, it may be that inhibition plays more of a central role in the bilingual advantage, meaning that less balanced types of bilingualism (e.g., Greek-English) leads to even more of an advantage in semantic control than found in the Welsh context. Less balanced bilinguals may not necessarily operate in a dual-language environment, meaning that they are required to more frequently inhibit responses from one language or the other. For example, during a conversation with an English monolingual, a Greek-English bilingual is more likely to have to inhibit responses from their first language compared to Welsh-English bilinguals living in a dual-language context. However, it may also depend on factors such as the amount of experience the individual has in managing the demands of both languages (Macnamara & Conway, 2014). That is, even if less balanced populations would need to inhibit responses due to the lack of a dual-language context, it may depend on how often they are in situations where they have the opportunity to exercise inhibition. Consequently, conducting the same study with a group of less balanced bilinguals would help elucidate whether the semantic control advantage is unique to Welsh-English (balanced bilinguals), or whether it is a consequence of bilingualism in general. Put simply, it remains unknown whether a bilingual advantage in the context of semantic control is due to frequent language switching, well-exercised inhibitory processes, or whether both factors contribute to the bilingual advantage and in turn CR. Therefore, this study opens up some

potentially fruitful lines of enquiry with Welsh-English bilinguals as well as other populations of bilinguals in this context, to understand whether the impact of bilingualism on semantic control is only relevant to a certain type of bilingual. Exploring such avenues is important because the more that is understood about bilingualism and CR, the more that will be understood about the interindividual variability and acceleration of cognitive decline (Stern, 2002). This is particularly important considering the ageing population and rapidly increasing rates of dementia (Ahmadi-Abhari et al., 2017; Musselwhite & Scott, 2019).

Having found a possible role of bilingualism in SC, it was then considered whether this superior bilingual performance in terms of *accuracy* would extend to contexts in which increased decision-making difficulty may be even more taxing on metacognitive processes. To explore this, the final experimental chapter investigated the *consistency* with which bilinguals/monolinguals and older and younger adults respond to different kinds of statements. The findings suggested neither age nor bilingualism impact on the consistency of responding. Here, the main message is that consistency of responding may not involve metacognitive processes, but rather, may be due to individual differences such as personality (Carter et al., 2013). That is, factors that go beyond cognitive processes may be more central to this decision-making process. This may be a fruitful direction for further research to better understand whether factors such as personality or other individual differences (e.g., social desirability) might be responsible for the consistency with which people respond to statements generally, or whether such factors only play a role when someone is unsure of their own opinion. Nevertheless, this study was the first to the author's knowledge to study the consequences of bilingualism in this context and one of a few studies investigating the consistency with which people respond to statements. The replication of Hampton et al. (2012) in itself is a significant contribution to the literature because it confirms that generally, people are inconsistent in how they answer questions. This may have implications in the real

world, such as cognitive testing in the context of semantic dementia where people progressively lose the meaning of words, objects and concepts (Davies et al., 2005). For example, during repeat cognitive testing with semantic dementia patients, assessments may contain subtests that ask categorisation type questions (e.g., WAIS-IV or the Cambridge Semantic Memory Test Battery) (Adlam et al., 2010) and the individual may respond differently each time. This may be perceived as a cognitive abnormality when in reality, the evidence suggests that the general population are also inconsistent in their responding. In sum, this study demonstrated that the processes involved in making decisions under uncertainty are unlikely to be an effective measure of CR because neither age nor language status impacted consistency of responding.

In addition to the limitations and practical and theoretical implications already discussed, one of the broader questions arising from the studies presented in this thesis pertains to the generalisability of the findings to other populations of balanced bilinguals. Whilst the findings may provide valuable insights into certain aspects of cognition within a group of balanced bilinguals, the findings may not be directly transferable to *all* other balanced bilinguals. Indeed, there may be several factors that make various populations of balanced bilinguals different from each other. For example, the specific languages involved may entail different linguistic and cognitive processes to the Welsh-English population studied in this thesis. Some groups of balanced bilinguals may speak languages from different language families, resulting in different structural features and characteristics between the two languages. For example, bilinguals such as French-Arabic speakers may read from left-to-right in one language and right-to-left in the other, potentially influencing their cognitive processes in different ways compared to the Welsh-English bilinguals studied in this thesis.

Moreover, the socio-cultural context also varies among balanced bilinguals, further impacting the transferability of the findings to other cultural contexts. For example, cultural

attitudes can influence individuals' language use and maintenance of balanced bilingualism. Additionally, differences in educational systems across countries can shape the bilingual experience of balanced bilinguals. Consequently, even though balanced bilingualism exists outside of the Welsh-English context, specific characteristics and broader contextual factors need to be considered to determine whether the findings of this thesis are applicable to other balanced bilinguals.

Furthermore, and as mentioned previously, research seeking to understand the bilingual advantage and its potential role in CR often investigates the same cognitive processes to locate this advantage, regardless of the population of bilinguals in question. However, it is possible that varying types of bilingualism, along with individual differences in language use, may yield different types of advantages. For example, an advantage in inhibition observed in one group of bilinguals may not necessarily extend to another group with distinct language usage patterns or unique social cultural contexts. Additionally, bilingualism from birth and bilingualism from an older age may also potentially lead to different advantages (de Bot, 2017), yet research studies frequently categorise participants into a generic 'bilingual group', overlooking the nuances of their bilingual experience. Recognising that bilingualism is not a one-size-fits-all phenomenon, paying attention to the different characteristics of bilingual groups may provide insights into the conditions under which a bilingual advantage in CR is more likely to manifest.

Similarly, and in line with the discussion in Chapter 1 regarding the definition of bilingualism, the level of proficiency required in each language to influence CR is also largely unexplored (de Bot, 2017). It remains unclear whether the most significant 'boost' stems from being fluent in both languages or from fluency in one language coupled with intermediate skills/proficiency in another (de Bot, 2017). Additionally, the impact of bilingualism on CR may be dependent, at least in part, on factors such as the frequency and

nature of bilingual language usage. Thus, it is vital that the characteristics (e.g., language use, proficiency, type of bilingual) of bilingual samples in research studies are fully considered and documented clearly.

With the bilingual 'boost' in mind, it is also important to acknowledge that CR is thought to be influenced by other factors beyond bilingualism. These factors include occupation, level of education, SES, and social interaction. However, and as raised initially in Chapter 1, it remains unclear whether there is an upper limit to the number of factors that can contribute to CR before the potential for further enhancement becomes unlikely. For example, if an individual's CR is already bolstered by factors such as levels of education and SES, there may not be additional capacity for further enhancement from other factors such as bilingualism. Put simply, there may be a threshold beyond which the contribution of bilingualism to CR decreases (Gollan et al., 2011).

Building on de Bot's (2017) argument regarding the level of language proficiency required to affect CR, it remains unclear whether similar principles apply to other potential contributing factors. Specifically, it is unclear whether specific levels of lifestyle factors contribute incrementally to CR, or if multiple factors enhance it but to varying degrees. For example, in terms of education level, it is unclear whether CR undergoes gradual incremental enhancement with each educational stage (i.e., each stage provides some 'boost'), or if significant boosts occur primarily at higher levels, such as obtaining a university level qualification. This may explain the lack of bilingual effects in some studies in the literature.

It is also important to note, as mentioned in the first chapter of this thesis, that while evidence suggests that bilingualism leads to cognitive benefits, it may also be associated with disadvantages depending on the cognitive skill in question (e.g., memory, attention, switching; de Bot, 2017). However, it is crucial to recognise the broader perspective that despite potentially offering both advantages and disadvantages, bilingualism can also

contribute to fostering positive neuroplasticity. Future research should focus on both the advantages *and* disadvantages of different types of bilingualism across various cognitive tasks (de Bot, 2017). This would provide a more nuanced and comprehensive understanding of the relationship between bilingualism and cognition.

Furthermore, another point of consideration is the definition of the 'bilingual advantage' itself. The age of participants included in research studies is important in terms of what constitutes the 'bilingual advantage' and its impact on CR. For example, the advantage may be the maintenance of certain cognitive abilities in healthy older adults, whereas the advantage may be more closely related with a delay in disease progression in those with dementia (de Bot, 2017). Another consideration is that the impact of CR (i.e., the bilingual advantage) may not become apparent until the onset of cognitive decline. That is, the reserve may build up over time but may not activate or become observable until it is required by the individual. Consequently, the interpretation of the bilingual advantage, or the specific 'advantage' being investigated in different studies, is unclear. Such ambiguity could lead to researchers not referring to the same 'advantage' in their studies, potentially impacting on the quality and progress of research in the field, as a shared understanding of a concept is fundamental. It could also pose challenges for replication efforts, as researchers may be examining different concepts, resulting in the failure of the literature to establish a robust body of knowledge. This could lead researchers to question the validity of the concept, potentially resulting in a reduced volume of literature on the subject, despite its potential significance.

Another issue to consider is that most neuropsychological assessments' normative standards are based on monolingual populations, despite bilingualism being the norm in most countries (Grosjean, 2008). Given the extensive body of research indicating potential differences in the bilingual brain and its cognitive processes compared to those of

monolinguals, comparing the performance of bilingual individuals with that of monolinguals may not be accurate. In other words, relying on the performance of monolinguals as a benchmark for cognitive function may not be appropriate, especially considering that a significant proportion of the population is bilingual. To support this argument, Suarez et al. (2014) found that assessing a decline in executive function in individuals who speak a second language using standards derived from monolinguals may lead to inaccurate assessment outcomes, potentially underestimating the decline in their abilities. With this in mind, the risk of misdiagnosis in bilingual individuals may also increase if monolingual norms are used.

However, this argument leads to questions regarding the identification of a bilingual advantage, as establishing such an advantage typically involves comparing the performance of bilingual individuals with those without proficiency in a second language. Nevertheless, it may be possible to use an alternative approach that does not rely on direct comparisons with monolinguals. For example, researchers could compare cognitive performance between different bilingual populations, considering variations in language backgrounds and levels of exposure. In addition, neuroimaging tools offer a valuable approach for exploring neural differences associated with bilingualism.

5.3 Strengths, Limitations, and Future Directions

One of the key strengths of this thesis is its originality. Each study contributes novel information and perspectives to the evidence base concerning Welsh-English bilingualism and cognition. Another important strength is the clarity around the type of bilingualism investigated. That is, considerable consideration has been given to a relatively uncommon type of bilingualism (e.g., balanced dominance, language switching, contextual language use) as well as the understanding that this type of bilingualism may differentially impact on cognitive processes. Although there are other populations of balanced bilinguals, few bilingual situations exhibit similar characteristics to those described above. Therefore,

investigating the Welsh-English context can serve as a model of this type of bilingualism, especially given its relative lack of attention in the literature. This thesis also makes it clear that the findings of each study are relevant to this particular population of bilinguals, unlike some studies that do not acknowledge the type of bilingualism investigated in their studies. This will be discussed further below.

One of the main limitations attached to the thesis, is that due to the COVID-19 pandemic, and as briefly introduced in Chapter 1, it was not possible to recruit a clinical sample of participants to understand whether cognitive performance across the cognitive assessment batteries would be different in some way. In terms of future research, more studies are needed that investigate the impact of Welsh-English bilingualism in the contexts described in this thesis in clinical populations as well as healthy individuals. This would determine whether different patterns of performance emerge depending on the cognitive status of the bilingual. Much is unknown about when and under what conditions the possible advantageous effects of bilingualism may emerge across the lifespan. For example, testing a clinical population in the context of the first study of this thesis would enable the understanding of: (a) whether performance across Welsh and English versions of the ACE-III and RBANS would remain comparable or whether scores would be higher in one language, and (b) a more general uplift in cognitive performance may be found when comparing the performance of bilingual and monolingual individuals living with a cognitive impairment. It may be that a bilingual advantage in the 'general' cognitive domain context does not emerge until the onset of pathological ageing. Therefore, comparing the performance of a bilingual clinical sample with that of a monolingual sample on general cognitive assessment batteries would also improve understanding around the impact of Welsh-English bilingualism in CR. That is, whether Welsh-English bilingualism protects against or slows the cognitive decline when affected by a neurodegenerative disease. In addition, and in the context of semantic control,

future research examining a clinical population would determine whether the bilingual advantage in this area persists into cognitive decline. That is, whether the bilingual superiority in semantic control protects somewhat from the losses that accompany pathological ageing by increasing CR.

Furthermore, many studies in the literature, which serve as the background for this thesis suffer from weaknesses in their design. As covered briefly previously, many studies referenced in this thesis fail to define the type of bilingualism under study or use an objective measure of bilingualism, potentially leading to misleading findings. The absence of an objective measure to confirm participants' bilingual status, overlooks key differences between different types of bilingualism, such as language switching and AoA. These variations among bilinguals cannot be underestimated, as cognitive system differences may emerge as a result. Indeed, complex skills like bilingualism require thorough exploration and a clear definition to facilitate further study and understanding. However, the broader limitations in the design of many studies in the literature also extend to the current work and other research in the broader field, where synthesising findings coherently becomes challenging. That is, the ability to draw clear conclusions and identify consistent patterns is limited. For example, comparing the performance of the samples in this thesis with those in the literature is challenging given uncertainties regarding the similarity of the samples used. In the first study of this thesis, the comparison with normative samples was flawed due to a lack of information about the normative samples themselves. The language in which assessments were completed was often not explicitly stated, and the language status of the samples was unclear. Consequently, the conclusions drawn from this thesis would be more robust if the evidence base were of higher quality, with studies being transparent about the characteristics of their samples.

Another important point to highlight, which is closely related to the aforementioned issues is the exploration of multiple variables, which can present several challenges. This challenge occurs throughout this thesis, emphasising the inherent complexity of investigating phenomena when there are few established anchor points in the literature. Some of the challenges stemming from this issue include ambiguity in establishing causality, difficulty in generalising findings, limited availability of data, risk of biased interpretation, and difficulty in establishing theoretical frameworks. Whilst statistical analyses can provide insights into the associations between variables, which offers valuable indications of potential causal links, it is essential to recognise that statistical methods alone may only partly address the underlying complexity of exploring phenomena in the absence of sufficient anchor points. Indeed, relying on statistical analyses alone without contextual or theoretical understanding may lead to oversimplification of the intricacies of real-world experiences such as bilingualism. Consequently, it is vital that researchers adopt a nuanced approach wherever possible, where statistical approaches are used wisely, and quantitative findings are complemented with qualitative insights or theoretical frameworks.

It is also important to recognise a broader point of consideration that has not been raised thus far. That is, the challenge of finding appropriate or valid objective measures of language proficiency. For example, whilst the LexTALE may be the most suitable measure of language proficiency available, it is not without limitations. For example, the Welsh version of the LexTALE used in this thesis has limitations, where task difficulty is not necessarily well matched across the Welsh and English versions. However, the findings remain useful when considered in conjunction with subjective proficiency measures and when compared with the performance of other Welsh-English bilingual samples. However, studies should endeavour to include both objective and subjective measures of proficiency to obtain a comprehensive understanding of language proficiency.

Another point to recognise, which has not been discussed thus far, is the importance of considering possible contamination from cognates or false friends in linguistic research. Failing to identify cognates could lead to confounding effects in cognitive tasks. For instance, performance differences between certain conditions on cognitive tasks may be attributable to cognates as opposed to the cognitive process in question. Indeed, one of the limitations of this thesis is that consideration was not given to the possibility of false friends in the materials used in study 2. However, having since checked for cognates, as mentioned in chapter 3, no false friends were identified in any of the measures except for one in the Spot-the-Word lexical decision task. While cognates are important to consider in lexical decision tasks, the impact of the one cognate found in the Spot-the-Word task is unlikely to have affected the overall findings and the research question in study 2. Nevertheless, future researchers should carefully consider the role of cognates in their studies and select stimuli based on their research goals. This could be achieved by conducting a thorough literature review to find studies or resources that have identified false friends, or by consulting with native speakers who may be aware of words that appear similar but have different meanings.

As mentioned previously, research on cognitive processes and bilingualism across the board fails to consider important factors about the sample of participants used in the studies, including the language status of the individuals (e.g., monolinguals/bilinguals), as well as the unique and finer details such as language switching and contextual language use. This is also the case for studies that are not focused on bilinguals but fail to explicitly state the language status of their sample. The above may partly explain the mixed evidence of a bilingual advantage in the literature. That is, the studies reporting evidence for and against a bilingual advantage are problematic in terms of the smaller but crucial characteristics that differentiate various types of bilinguals. Further, when such characteristics are considered in relation to a possible bilingual advantage, it is imperative that all outcomes are shared irrespective of the

direction of findings. Indeed, recent research has highlighted how the literature base can also be positioned within a broader context of publication bias and related issues in psychological research. De Bruin et al. (2015) found that research supporting the bilingual advantage is most likely to be published, followed by research with mixed evidence. Studies challenging the bilingual advantage were the least commonly published. With this in mind, and to facilitate understanding and knowledge around the possible bilingual advantage, it is vital that all study findings (evidence for, against and mixed findings) are reported in the literature. This will aid understanding about the specific circumstances in which the bilingual advantage is more or less likely to emerge (De Bruin et al., 2015).

Finally, and as covered in greater detail previously, another important avenue for future research is to conduct the same studies as presented in this thesis but with different *types* of bilinguals. As mentioned previously, this thesis focused on an uncommon population of bilinguals and therefore it remains unclear whether there are significant cognitive system differences between balanced bilinguals and more typical bilingual populations (e.g., Greek-English), or those who do not operate in a dual-language environment.

5.4 Key Standards for Researching the Bilingual Advantage

Considering the critical review presented above, it is evident that research investigating the complexities of the bilingual advantage is plagued by several methodological shortcomings and inconsistencies. To address such issues and advance understanding of this complex phenomenon, establishing key standards that serve as guiding principles for conducting future research would be beneficial. Such key standards represent the essential aspects necessary for conducting more rigorous and reliable research into the bilingual advantage. By synthesising the material and key points discussed above, it becomes clear that careful consideration of these standards is imperative for shaping the trajectory of future

research endeavours and informing both theory and practice in the field. The recommended key standards for researching the bilingual advantage, include:

1) Definition of 'Advantage' in the Bilingual Context: Research studies should explicitly define what is meant by the term 'advantage' when referring to bilingualism. This definition should be clearly stated to ensure consistency and understanding across studies.
2) Characteristics of the Bilingual Group: Studies should provide detailed information about the characteristics of the bilingual sample used, including AoA of each language, language proficiency levels, language dominance, and patterns of language use. Such information is crucial for understanding the diversity within the bilingual population and for interpreting key findings accurately.
3) Measures of Language Proficiency: Researchers should ideally employ a variety of measures to assess language proficiency in both languages of a bilingual. These measures should include both objective proficiency assessments, such as standardised language tests, and subjective evaluations, such as self-reported proficiency ratings. Using multiple measures would help capture the complexity of language proficiency and ensure a more comprehensive understanding.
4) Assessing the Bilingual Experience: To provide valuable insights, gathering more information about the bilingual experience of participants would be valuable. Requisite examples include participants language exposure, contexts of language use (e.g., home, school, community), frequency of language switching, and proficiency in different language skills (e.g., speaking, listening, reading, writing).
5) Translation of Measures and Tasks: When using translated measures or tasks in bilingual research, researchers must ensure that these translations are appropriate and accurate to avoid confounding effects. This includes considering factors such as

cognates (words with similar meanings in both languages) or false friends, and also ensuring equivalence of task difficulty across languages (e.g., cognitive assessments, language proficiency measures). Additionally, in instances where languages may have unique linguistic and grammatical characteristics (e.g., mutations in the Welsh context), special attention should be paid to ensure that translations adequately capture these differences and do not bias results.

5.5 Thesis Conclusion

In conclusion, this thesis has contributed significantly to the literature investigating the impact of Welsh-English bilingualism on various aspects of cognition across the lifespan. It has presented three novel studies which led to significant theoretical and practical implications. To answer the main research question, it seems that Welsh-English bilingualism may impact on, and lead to an advantage in particular aspects of cognition in young and older adults, like semantic control processes, as opposed to an uplift of general cognitive domains. Consequently, a more effective proxy of cognitive reserve might be processes of semantic cognition (SC). However, more research is required to (a) replicate these findings, and (b) study the impact of Welsh-English bilingualism in the cognitive contexts presented in this thesis but with clinical populations (e.g., dementia populations) and with other types of bilinguals. This will help elucidate whether (a) a general improvement across cognitive processes only manifests after the onset of cognitive decline, and (b) whether the bilingual advantage found in the control processes of SC remains evident after the onset of cognitive impairment, and in less balanced bilinguals.

References

Abu Rabia, S. (2019). The effect of degrees of bilingualism on metacognitive linguistic skills. *International Journal of Bilingualism, 23*(5), 1064-1086. https://doi.org/10.1177/1367006918781060

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition, 19*(4), 689-698. https://doi.org/10.1017/S1366728916000225

Abutalebi, J., Canini, M., Della Rosa, P. A., Green, D. W., & Weekes, B. S. (2015). The neuroprotective effects of bilingualism upon the inferior parietal lobule: a structural neuroimaging study in aging Chinese bilinguals. *Journal of Neurolinguistics, 33*, 3-13. https://doi.org/10.1016/j.jneuroling.2014.09.008

Abutalebi, J., Canini, M., Della Rosa, P. A., Sheung, L. P., Green, D. W., & Weekes, B. S. (2014). Bilingualism protects anterior temporal lobe integrity in aging. *Neurobiology of Aging, 35*(9), 2126-2133. https://doi.org/10.1016/j.neurobiolaging.2014.03.010

Acevedo, A., Loewenstein, D. A., Barker, W. W., Harwood, D. G., Luis, C., Bravo, M., Jurwitz, D. A., Aguero, L., Greenfield, L., & Duara, R. (2000). Category fluency test: normative data for English-and Spanish-speaking elderly. *Journal of the International Neuropsychological Society, 6*(7), 760-769. https://doi.org/10.1017/s1355617700677032

Acevedo, M., & Krueger, J. I. (2004). Two egocentric sources of the decision to vote: The voter's illusion and the belief in personal relevance. *Political Psychology, 25*(1), 115-134. https://doi.org/10.1111/j.1467-9221.2004.00359.x

Adlam, A. L. R., Patterson, K., Bozeat, S., & Hodges, J. R. (2010). The Cambridge Semantic Memory Test Battery: Detection of semantic deficits in semantic dementia and

Alzheimer's disease. *Neurocase, 16*(3), 193-207. https://doi.org/10.1080/13554790903405693

Ahmadi-Abhari, S., Guzman-Castillo, M., Bandosz, P., Shipley, M. J., Muniz-Terrera, G., Singh-Manoux, A., Kivimaki, M., Steptoe, A., Capewell, S., O'Flaherty, M., & Brunner, E. J. (2017). Temporal trend in dementia incidence since 2002 and projections for prevalence in England and Wales to 2040: modelling study. BMJ, 358. https://doi.org/10.1136/bmj.j2856

Ahmed, R. M., Paterson, R. W., Warren, J. D., Zetterberg, H., O'Brien, J. T., Fox, N. C., Halliday, G. M., & Schott, J. M. (2014). Biomarkers in dementia: clinical utility and new directions. J*ournal of Neurology, Neurosurgery & Psychiatry, 85*(12), 1426-1434. https://doi.org/10.1136/jnnp-2014-307662

Aitchison, J., & Carter, H. (Eds.) (2004). *Spreading the word: The Welsh language 2001* (pp. 6-149). Ceredigion: Y Lolfa.

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., Chaudhuri, J. R., & Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology, 81*(22), 1938-1944. https://doi.org/10.1212/01.wnl.0000436620.33155.a4

Allen, C. M., Martin, R. C., & Martin, N. (2012). Relations between short-term memory deficits, semantic processing, and executive function. *Aphasiology, 26*(3-4), 428-461. https://doi.org/10.1080/02687038.2011.617436

Almaghyuli, A., Thompson, H., Ralph, M. A. L., & Jefferies, E. (2012). Deficits of semantic control produce absent or reverse frequency effects in comprehension: evidence from neuropsychology and dual task methodology. *Neuropsychologia, 50*(8), 1968-1979. https://doi.org/10.1016/j.neuropsychologia.2012.04.022

Altarriba, J. & Basnight-Brown, D. (2009). An Overview of Semantic Processing in Bilinguals: Methods and Findings. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches* (pp. 79-98). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781847691262-006

Altarriba, J., & Canary, T. M. (2004). The influence of emotional arousal on affective priming in monolingual and bilingual speakers. *Journal of Multilingual and Multicultural Development, 25*(2-3), 248-265. https://doi.org/10.1080/01434630408666531

Anderson, J. A. E., Grundy, J. G., & Bialystok, E. (2019). *Lifelong bilingualism leads to cognitive reserve in older adults.* Annual Meeting of the Psychonomic Society. Montreal, Canada.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6*(4), 451-474. https://doi.org/10.1016/0010-0285(74)90021-8

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition & Emotion, 8*(6), 503-514. https://doi.org/10.1080/02699939408408956

Anstey, K. J., Cherbuin, N., & Herath, P. M. (2013). Development of a new method for assessing global risk of Alzheimer's disease for use in population health approaches to prevention. P*revention Science, 14*(4), 411-421. https://doi.org/10.1007%2Fs11121-012-0313-2

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

Ardila, A., Rosselli, M., & Puente, A. E. (1994). *Neuropsychological evaluation of the Spanish speaker.* Plenum Press, New York.

Askey, C., & Playfoot, D. (2018). Examining theories of cognitive ageing using the false memory paradigm. *Quarterly Journal of Experimental Psychology, 71*(4), 931-939. https://doi.org/10.1080/17470218.2017.1307433

Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (1992). *The speed and capacity of language-processing test.* Thames Valley Test Company.

Badre, D., & Wagner, A. D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex, 15*(12), 2003-2012. https://doi.org/10.1093/cercor/bhi075

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907-918. https://doi.org/10.1016/j.neuron.2005.07.023

Baghaei, P., & Tabatabaee, M. (2015). The C-Test: An integrative measure of crystallized intelligence. *Journal of Intelligence, 3*(2), 46-58. https://doi.org/10.3390/jintelligence3020046

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

Barros, C. P., Caporale, G. M., & Gil-Alana, L. A. (2009). Basque terrorism: Police action, political measures and the influence of violence on the stock market in the Basque Country. *Defence and Peace Economics, 20*(4), 287-301. https://doi.org/10.1080/10242690701750676

Bartolotti, J., & Marian, V. (2012). Bilingual memory: structure, access, and processing: Theoretical and applied approaches to bilingualism. In J. Altarriba, & L. Isurin (Eds.), *Memory, Language, and Bilingualism: Theoretical and Applied Approaches* (pp. 7-47). Cambridge University Press. doi:10.1017/CBO9781139035279.002

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition, 35*(5), 953-965. https://doi.org/10.3758/BF03193468

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics, 35*(5), 857. doi:10.1017/S0142716414000174

Beardsmore, H. B. (1986). *Bilingualism: basic principles* (2nd ed.). Multilingual Matters.

Beatty, W. W., Ryder, K. A., Gontkovsky, S. T., Scott, J. G., McSwan, K. L., & Bharucha, K. J. (2003). Analyzing the subcortical dementia syndrome of Parkinson's disease using the RBANS. *Archives of Clinical Neuropsychology, 18*(5), 509-520. https://doi.org/10.1016/S0887-6177(02)00148-8

Beishon, L. C., Batterham, A. P., Quinn, T. J., Nelson, C. P., Panerai, R. B., Robinson, T., & Haunton, V. J. (2019). Addenbrooke's Cognitive Examination III (ACE-III) and mini-ACE for the detection of dementia and mild cognitive impairment. *Cochrane Database of Systematic Reviews, 12*(12), CD013282. https://doi.org/10.1002/14651858.cd013282.pub2

Berkes, M., & Bialystok, E. (2022). Bilingualism as a contributor to cognitive reserve: What it can do and what it cannot do. *American Journal of Alzheimer's Disease & Other Dementias*, *37*. https://doi.org/10.1177/15333175221091417

Bhatia, T. K. (2017). Bilingualism and multilingualism from a socio-psychological perspective. *Oxford Research Encyclopedia of Linguistics.* Retrieved 23 Jun. 2023,

from https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-82.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*(3), 636-644. https://doi.org/10.1111/1467-8624.00046

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism, 10*(3), 210-223. https://doi.org/10.2167/beb441.0

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11. https://doi.org/10.1017/S1366728908003477

Bialystok, E. (2010). Global–local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology, 46*(1), 93. https://psycnet.apa.org/doi/10.1037/a0015466

Bialystok, E. (2011). Reshaping the mind: the benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 65*(4), 229. https://doi.org/10.1037/a0025406

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives, 9*(2), 117-121. https://doi.org/10.1111/cdep.12116

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin, 143*(3), 233. https://doi.org/10.1037/bul0000099

Bialystok, E. (2021). Bilingualism: Pathway to cognitive reserve. *Trends in Cognitive Sciences, 25*(5), 355-364. https://doi.org/10.1016/j.tics.2021.02.003

Bialystok, E. (2021). Bilingualism: Pathway to cognitive reserve. *Trends in Cognitive Sciences, 25*(5), 355-364. https://doi.org/10.1016/j.tics.2021.02.003

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., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition, 15*(2), 397-401. https://doi.org/10.1017/S136672891100040X

Bialystok, E., Abutalebi, J., Bak, T. H., Burke, D. M., & Kroll, J. F. (2016). Aging in two languages: Implications for public health. *Ageing Research Reviews, 27*, 56-60. https://doi.org/10.1016/j.arr.2016.03.003

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240-250. https://doi.org/10.1016/j.tics.2012.03.001

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459-464. https://doi.org/10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological science in the public interest: a journal of the American Psychological Society, 10*(3), 89-129. https://doi.org/10.1177/1529100610387084

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging, 19*(2), 290. https://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging, 29*(3), 696. https://psycnet.apa.org/doi/10.1037/a0037254

Birren, J. E., & Fisher, L. M. (1995). Aging and speed of behavior: Possible consequences for psychological functioning. *Annual Review of Psychology, 46*(1), 329-353. 10.1146/annurev.ps.46.020195.001553

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2018). Effect of variance ratio on ANOVA robustness: Might 1.5 be the limit?. *Behavior Research Methods, 50*, 937-962. https://doi.org/10.3758/s13428-017-0918-2

Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition, 213,* 104625. https://doi.org/10.1016/j.cognition.2021.104625

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

Bloomfield, L. (1933) *Language.* Holt, Rinehart & Winston: New York.

Bluvshtein, M. (2004). Repeatable battery for the assessment of neuropsychological status (RBANS): Russian language adaptation. USA: Capella University, Minneapolis, MN, Unpublished Dissertation.

Bol, L., & Hacker, D. J. (2012). Calibration research: Where do we go from here? *Frontiers in Psychology, 3*, 229. https://doi.org/10.3389/fpsyg.2012.00229

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation

variables in a large patient population. *Archives of Clinical Neuropsychology, 22*(3), 355-365. https://doi.org/10.1016/j.acn.2007.01.010

Bourne, L. E., Ekstrand, B. R., & Dominowski, R. L. (1971). The psychology of thinking. Englewood Cliffs, N.J: Prentice-Hall.

Braun, K. L., Takamura, J. C., & Mougeot, T. (1996). Perceptions of dementia, caregiving, and help-seeking among recent Vietnamese immigrants. *Journal of Cross-Cultural Gerontology, 11*(3), 213-228. https://doi.org/10.1007/BF00122702

Braun, K. L., Takamura, J. C., Forman, S. M., Sasaki, P. A., & Meininger, L. (1995). Developing and testing outreach materials on Alzheimer's disease for Asian and Pacific Islander Americans. *The Gerontologist, 35*(1), 122-126. https://doi.org/10.1093/geront/35.1.122

Brown, P., Heirene, R. M., John, B., & Evans, J. J. (2019). Applicability of the ACE-III and RBANS cognitive tests for the detection of alcohol-related brain damage. *Frontiers in Psychology, 10*, 2636. https://doi.org/10.3389%2Ffpsyg.2019.02636

Bruine de Bruin W., Parker A. M., Fischhoff B. (2012). Explaining adult age differences in decision-making competence. *Journal of Behavioral Decision Making, 25*(4), 352-360. https://doi.org/10.1002/bdm.712

Bruno, D., & Vignaga, S. (2019). Addenbrooke's cognitive examination III in the diagnosis of dementia: a critical review. *Neuropsychiatric Disease and Treatment, 15,* 441-447. https://doi.org/10.2147/NDT.S151253

Brysbaert, M. (2013). Lextale_FR a fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*(1), 23-37. https://doi.org/10.5334/pb-53-1-23

Buades-Sitjar, F., Boada, R., Guasch, M., Ferré, P., Hinojosa, J. A., & Duñabeitia, J. A. (2022). The predictors of general knowledge: Data from a Spanish megastudy.

Behavior Research Methods, 54(2), 898-909. https://doi.org/10.3758/s13428-021-01669-4

Burke, D. M., & MacKay, D. G. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1363), 1845-1856. https://doi.org/10.1098/rstb.1997.0170

Butler, Y. G. (2012). Bilingualism/multilingualism and second-language acquisition. In T. K. Bhatia, & W. C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingua*lism (2nd ed., pp. 109-136). Blackwell Publishing, Ltd. https://doi.org/10.1002/9781118332382.ch5

Butler, Y. G., & Hakuta, K. (2006). Bilingualism and second language acquisition. In T. K. Bhatia, & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 114-144). Blackwell Publishing, Ltd.

Callender, A. A., Franco-Watkins, A. M., & Roberts, A. S. (2016). Improving metacognition in the classroom through instruction, training, and feedback. *Metacognition and Learning, 11*, 215-235. DOI 10.1007/s11409-015-9142-6

Calvo, N., García, A. M., Manoiloff, L., & Ibáñez, A. (2016). Bilingualism and cognitive reserve: a critical overview and a plea for methodological innovations. *Frontiers in Aging Neuroscience, 7*, 249. https://doi.org/10.3389/fnagi.2015.00249

Carlozzi, N. E., Horner, M. D., Yang, C., & Tilley, B. C. (2008). Factor analysis of the Repeatable Battery for the Assessment of Neuropsychological Status. *Applied Neuropsychology, 15*(4), 274-279. https://doi.org/10.1080/09084280802325124

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

Carter, A. J., Marshall, H. H., Heinsohn, R., & Cowlishaw, G. (2013). Personality predicts decision making only when information is unreliable. *Animal Behaviour, 86*(3), 633-639. https://doi.org/10.1016/j.anbehav.2013.07.009

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action.* North-Holland.

Cenoz, J. (2008). Achievements and challenges in bilingual and multilingual education in the Basque Country. *AILA Review, 21*(1), 13-30. https://doi.org/10.1075/aila.21.03cen

Cenoz, J., & Gorter, D. (2017). Minority languages and sustainable translanguaging: Threat or opportunity? J*ournal of Multilingual and Multicultural Development, 38*(10), 901-912. https://doi.org/10.1080/01434632.2017.1284855

Chamorro-Premuzic, T., & Furnham, A. (2004). A possible model for understanding the personality-intelligence interface. *British Journal of Psychology, 95*(2), 249-264. https://doi.org/10.1348/000712604773952458

Chen, P., Bobb, S. C., Hoshino, N., & Marian, V. (2017). Neural signatures of language co-activation and control in bilingual spoken word comprehension. *Brain Research, 1665*, 50-64. https://doi.org/10.1016/j.brainres.2017.03.023

Chen, Q., Kong, Y., Gao, W., & Mo, L. (2018). Effects of socioeconomic status, parent–child relationship, and learning motivation on reading ability. *Frontiers in Psychology, 9,* 1297. https://doi.org/10.3389/fpsyg.2018.01297

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: evidence from a bilingual community. *Alzheimer Disease & Associated Disorders, 24*(2), 118-125. https://doi.org/10.1097/wad.0b013e3181ca1221

Chou, K. R., Jiann-Chyun, L., & Chu, H. (2002). The reliability and validity of the Chinese version of the caregiver burden inventory. *Nursing Research, 51*(5), 324-331. https://doi.org/10.1097/00006199-200209000-00009

Christoffels, I. K., de Haan, A. M., Steenbergen, L., van den Wildenberg, W. P., & Colzato, L. S. (2015). Two is better than one: bilingual education promotes the flexible mind. *Psychological Research, 79*(3), 371-379. https://doi.org/10.1007/s00426-014-0575-3

Clare, L., Whitaker, C. J., Craik, F. I., Bialystok, E., Martyr, A., Martin-Forbes, P. A., Bastable, A. J. M., Pye, K. L., Quinn, C., Thomas, E. M., Gathercole, V. C. M., & Hindle, J. V. (2016a). Bilingualism, executive control, and age at diagnosis among people with early-stage Alzheimer's disease in Wales. *Journal of Neuropsychology, 10*(2), 163-185. https://doi.org/10.1111/jnp.12061

Clare, L., Whitaker, C. J., Martyr, A., Martin-Forbes, P. A., Bastable, A. J., Pye, K. L., Quinn, C., Thomas, E. M., Gathercole, V. C. M., & Hindle, J. V. (2016b). Executive control in older Welsh monolinguals and bilinguals. *Journal of Cognitive Psychology, 28*(4), 412-426. https://doi.org/10.1111/jnp.12061

Clare, L., Wu, Y. T., Teale, J. C., MacLeod, C., Matthews, F., Brayne, C., Wood, R., & CFAS-Wales Study Team. (2017). Potentially modifiable lifestyle factors, cognitive reserve, and cognitive function in later life: A cross-sectional study. *PLoS Medicine, 14*(3), e1002259. https://doi.org/10.1371/journal.pmed.1002259

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407. https://doi.org/10.1037/0033-295X.82.6.407

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour, 8*(2), 240-247. https://doi.org/10.1016/S0022-5371(69)80069-1

Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLoS Biology, 10*(3), e1001293. https://doi.org/10.1371/journal.pbio.1001293

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., 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., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135-149. https://doi.org/10.1016/j.cognition.2009.08.001

Craik, F. I., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: Bilingualism as a form of cognitive reserve. *Neurology, 75*(19), 1726-1729. https://doi.org/10.1212%2FWNL.0b013e3181fc2a1c

Crane, P. K., Gruhl, J. C., Erosheva, E. A., Gibbons, L. E., McCurry, S. M., Rhoads, K., Nguyen, V., Arani, K., Masaki, K., & White, L. (2010). Use of spoken and written Japanese did not protect Japanese-American men from cognitive decline in late life. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 65*(6), 654-666. https://doi.org/10.1093/geronb/gbq046

Cummins, J. (1984). Bilingualism and cognitive functioning. In S. Shapson, & V. D'Oyley (Eds.), *Bilingual and Multicultural Education: Canadian Perspectives* (pp. 55-70). Multilingual Matters.

Cuyvers, E., van der Zee, J., Bettens, K., Engelborghs, S., Vandenbulcke, M., Robberecht, C., Dillen, L., Merlin, C., Geerts, N., Graff, C., Thonberg, H., Chiang, H. H., Pastor, P., Ortega-Cubero, S., Pastor, M. A., Diehl-Schmid, J., Alexopoulos, P., Benussi, L., Ghidoni, R., Binetti, G., … BELNEU Consortium and of the EU EOD Consortium (2015). Genetic variability in SQSTM1 and risk of early-onset Alzheimer dementia: a

European early-onset dementia consortium study. *Neurobiology of Aging, 36*(5), 2005.e15–2005.e2.005E22. https://doi.org/10.1016/j.neurobiolaging.2015.02.014

Daker-White, G., Beattie, A. M., Gilliard, J., & Means, R. (2002). Minority ethnic groups in dementia care: a review of service needs, service provision and models of good practice. *Aging & Mental Health, 6*(2), 101-108. https://doi.org/10.1080/13607860220126835

Darcy, N. T. (1953). A review of the literature on the effects of bilingualism upon the measurement of intelligence. *The Pedagogical Seminary and Journal of Genetic Psychology, 82*(1), 21-57. https://doi.org/10.1080/08856559.1953.10533654

Dassa, A., & Amir, D. (2014). The role of singing familiar songs in encouraging conversation among people with middle to late stage Alzheimer's disease. *Journal of Music Therapy, 51*(2), 131-153. https://doi.org/10.1093/jmt/thu007

Davies, J. (2014). T*he Welsh language: a history* (pp. 1-193). Cardiff: University of Wales Press.

Davies, R. R., Hodges, J. R., Kril, J. J., Patterson, K., Halliday, G. M., & Xuereb, J. H. (2005). The pathological basis of semantic dementia. *Brain, 128*(9), 1984-1995. https://doi.org/10.1093/brain/awh582

de Bot, K. (2017). The future of the bilingual advantage. In S. Pfenniger & J. Navracsics (Eds.), *Future Research Directions for Applied Linguistics* (pp. 15-32). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781783097135-003

de Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences, 9*(3), 33. https://doi.org/10.3390/bs9030033

De Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26*(1), 99-107. 10.1177/0956797614557866

De Deyne, S., Kenett, Y. N., Anaki, D., Faust, M., & Navarro, D. (2017). Large-scale network representations of semantics in the mental lexicon. In M. N. Jones (Ed.), *Big Data in Cognitive Science* (pp. 174–202). Routledge/Taylor & Francis Group.

Del Maschio, N., Fedeli, D., Sulpizio, S., & Abutalebi, J. (2019). The relationship between bilingual experience and gyrification in adulthood: A cross-sectional surface-based morphometry study. *Brain and Language, 198*, 104680. https://doi.org/10.1016/j.bandl.2019.104680

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging, 21*(1), 62. https://doi.org/10.1037/0882-7974.21.1.62

Deuchar, M., Donnelly, K., & Piercy, C. (2016). 'Mae pobl monolingual yn minority': Factors Favouring the Production of Code Switching by Welsh–English Bilingual Speakers. In M. Durham, & J. Morris, (Eds.), *Sociolinguistics in Wales* (pp. 209-239). Palgrave Macmillan, London. https://doi.org/10.1057/978-1-137-52897-1_8

Dewaele, J. M. (2015). Bilingualism and multilingualism. *The international Encyclopedia of Language and Social Interaction, 1-11*. DOI:10.1002/9781118611463/wbielsi108

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

Digard, B. G., Davis, R., Stanfield, A., Sorace, A., & Fletcher-Watson, S. (2022). "The Languages That You Know Draw the Boundary of Your World": A Thematic Analysis of the Experiences of Autistic Bilingual Adults Living in the United Kingdom. *Autism in Adulthood, 4*(4), 328-339. https://doi.org/10.1089/aut.2021.0077

Dijkstra, T. O. N., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition, 22*(4), 657-679. https://doi.org/10.1017/S1366728918000287

Dijkstra, T., & Van Heuven, W. J. (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

Dobbins, I. G., & Wagner, A. D. (2005). Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cerebral Cortex, 15*(11), 1768-1778. https://doi.org/10.1093/cercor/bhi054

Dong, Y., Gui, S., & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8*(3), 221-238. https://doi.org/10.1017/S1366728905002270

Dong, Y., Thompson, C. L., Tan, S. H. J., Lim, L. B. S., Pang, W., & Chen, C. L. H. (2013). Test-retest reliability, convergent validity and practice effects of the RBANS in a memory clinic setting: a pilot study. *Open Journal of Medical Psychology, 2*(04), 11. http://dx.doi.org/10.4236/ojmp.2013.24B003

Dorenkamp, M. A., & Vik, P. (2018). Neuropsychological assessment anxiety: A systematic review. *Practice Innovations, 3*(3), 192. https://doi.org/10.1037/pri0000073

Doroud, S., Saeedi, Z., & Radman, N. (2020). Cognitive resilience in bilinguals and its potentiality in dealing with cognitive load. *bioRxiv*, 2020-09. https://doi.org/10.1101/2020.09.17.301606

Duff, K., Beglinger, L. J., Schoenberg, M. R., Patton, D. E., Mold, J., Scott, J. G., & Adams, R. L. (2005). Test-retest stability and practice effects of the RBANS in a community

dwelling elderly sample. *Journal of Clinical and Experimental Neuropsychology, 27(5)*, 565-575. https://doi.org/10.1080/13803390490918363

Duff, K., Hobson, V. L., Beglinger, L. J., & O'Bryant, S. E. (2010). Diagnostic accuracy of the RBANS in mild cognitive impairment: limitations on assessing milder impairments. *Archives of Clinical Neuropsychology, 25*(5), 429-441. https://doi.org/10.1093/arclin/acq045

Duff, K., Patton, D., Schoenberg, M. R., Mold, J., Scott, J. G., & Adams, R. L. (2003). Age- and education-corrected independent normative data for the RBANS in a community dwelling elderly sample. *The Clinical Neuropsychologist, 17*(3), 351-366. https://doi.org/10.1076/clin.17.3.351.18082

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited. *Experimental Psychology, 61*(3), 234-251. https://psycnet.apa.org/doi/10.1027/1618-3169/a000243

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences, 14*(4), 172-179. https://doi.org/10.1016/j.tics.2010.01.004

Eling, P., Derckx, K., & Maes, R. (2008). On the historical and conceptual background of the Wisconsin Card Sorting Test. *Brain and Cognition, 67*(3), 247-253. https://doi.org/10.1016/j.bandc.2008.01.006

Ellis, A. E., & Oakes, L. M. (2006). Infants flexibly use different dimensions to categorize objects. *Developmental Psychology, 42*(6), 1000. https://doi.org/10.1037/0012-1649.42.6.1000

Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals: Evidence from bimodal bilinguals. *Psychological Science, 19*(12), 1201-1206. https://doi.org/10.1111%2Fj.1467-9280.2008.02224.x

Encinas, S. M., Gramunt-Fombuena, N., Guia, D. B., Hernanz, L. C., & Barbera, M. A. (2012). Spanish translation and adaptation of the repeatable battery for the assessment of neuropsychological status (RBANS) Form A in a pilot sample. *Neurología (English Edition), 27*(9), 531-546. https://doi.org/10.1016/j.nrl.2011.07.006

Engel de Abreu, P. M. (2011). Working memory in multilingual children: Is there a bilingual effect?. *Memory, 19*(5), 529-537. https://doi.org/10.1080/09658211.2011.590504

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition, 2*(2-3), 101-118. https://doi.org/10.1080/13506289508401726

Ervin, S. M., & Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal and Social Psychology, 49*(supp), 139-145.

Escobar, J. I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J., & Golding, J. M. (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity: Cultural and linguistic artifacts. *Journal of Nervous and Mental Disease, 174*(10), 607–614. https://doi.org/10.1097/00005053-198610000-00005

Ethnologue. (2020). *Explore the World's Languages.* Retrieved June 23, 2023, from https://www.ethnologue.com/

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

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

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, 110*(41), 16616-16621. https://doi.org/10.1073/pnas.1315235110

Ferré, P., Anglada-Tort, M., & Guasch, M. (2018). Processing of emotional words in bilinguals: Testing the effects of word concreteness, task type and language status. *Second Language Research, 34*(3), 371-394. https://doi.org/10.1177/0267658317744008

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 552-564. https://doi.org/10.1037/0096-1523.3.4.552

Fishman, J. (1977). The social science perspective. In Fishman, J. (Ed.), *Bilingual Education: Current Perspectives, Vol 1: Social Science* (pp.14-62). Arlington: Center for Applied Linguistics.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the response space"—an account of left prefrontal activation at encoding. *Neuroimage, 12*(4), 404-417. https://doi.org/10.1006/nimg.2000.0633

Fratiglioni, L., & Wang, H. X. (2007). Brain reserve hypothesis in dementia. *Journal of Alzheimer's Disease, 12*(1), 11-22. https://doi.org/10.3233/jad-2007-12103

Gallo, F., & Abutalebi, J. (2023). The unique role of bilingualism among cognitive reserve-enhancing factors. *Bilingualism: Language and Cognition,* 1-8. https://doi.org/10.1017/S1366728923000317

Galusha-Glasscock, J. M., Horton, D. K., Weiner, M. F., & Cullum, C. M. (2016). Video teleconference administration of the repeatable battery for the assessment of neuropsychological status. *Archives of Clinical Neuropsychology, 31*(1), 8-11. https://doi.org/10.1093/arclin/acv058

Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007). Language of administration and neuropsychological test performance in neurologically intact Hispanic American bilingual adults. *Archives of Clinical Neuropsychology, 22*(8), 991-1001. https://doi.org/10.1016/j.acn.2007.08.003

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas-Guasch, N., Roberts, E. J., Highes, 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*, Article 11. https://doi.org/10.3389/fpsyg.2014.00011

Giebel, C. M., & Challis, D. (2017). Sensitivity of the Mini-Mental State Examination, Montreal Cognitive Assessment and the Addenbrooke's Cognitive Examination III to everyday activity impairments in dementia: an exploratory study. *International Journal of Geriatric Psychiatry, 32*(10), 1085-1093. https://doi.org/10.1002/gps.4570

Glass, J. M., Schumacher, E. H., Lauber, E. J., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., & Meyer, D. E. (2000). Aging and the psychological refractory period: task-coordination strategies in young and old adults. *Psychology and Aging, 15*(4), 571-595. https://doi.org/10.1037/0882-7974.15.4.571

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. Journal of Experimental Psychology: *Human Learning and Memory, 7*(5), 311-325. https://doi.org/10.1037/0278-7393.7.5.311

Gold, B. T., Johnson, N. F., & Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia, 51*(13), 2841-2846. https://doi.org/10.1016/j.neuropsychologia.2013.09.037

Gold, B. T., Johnson, N. F., & Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia, 51*(13), 2841-2846. https://doi.org/10.1016/j.neuropsychologia.2013.09.037

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%2Fa0014981

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society, 13*(2), 197-208. https://doi.org/10.1017/s1355617707070038

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220-1234. https://doi.org/10.3758/bf03193224

Gollan, T. H., Salmon, D. P., Montoya, R. I., & Galasko, D. R. (2011). Degree of bilingualism predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated Hispanics. *Neuropsychologia, 49*(14), 3826-3830. https://doi.org/10.1016/j.neuropsychologia.2011.09.041

Gordon, J. K., Young, M., & Garcia, C. (2018). Why do older adults have difficulty with semantic fluency? *Aging, Neuropsychology, and Cognition, 25*(6), 803-828. https://doi.org/10.1080/13825585.2017.1374328

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience, 13*(7), 491-505. https://doi.org/10.1038/nrn3256

Grainger, J. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes*, *13*(6), 601-623. https://doi.org/10.1080/016909698386393

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. *Advances in Psychology, 83*, 207-220. https://doi.org/10.1016/S0166-4115(08)61496-X.

Grant, A., Legault, J., & Li, P. (2019). What do bilingual models tell us about the neurocognition of multiple languages? In J. W. Schwieter, & M. Paradis (Eds.), The Handbook of the Neuroscience of Multilingualism (pp. 48-74). John Wiley & Sons LTd. https://doi.org/10.1002/9781119387725.ch3

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*(4), 404. https://doi.org/10.1037/h0059831

Green, A., Garrick, T., Sheedy, D., Blake, H., Shores, A., & Harper, C. (2008). Repeatable battery for the assessment of neuropsychological status (RBANS): preliminary Australian normative data. *Australian Journal of Psychology, 60*(2), 72-79. https://doi.org/10.1080/00049530701656257

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. (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

Griffiths, A. (1994). National and cultural identity in a Welsh-language soap opera. In R. Allen (Ed.), *To Be Continued... Soap Operas around the World* (1st ed., pp. 91-97). Routledge. https://doi.org/10.4324/9780203131855

Grosjean, F. (1992). The bilingual & the bicultural person in the hearing & in the deaf world. *Sign Language Studies, 77*(1), 307-320. http://www.jstor.org/stable/26204700

Grosjean, F. (1994). Individual bilingualism. In R. E. Asher, & J. M. Y. Simpson (Eds,), *The Encyclopaedia of Language and Linguistics* (pp. 1656-1660). Oxford: Pergamon Press.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One Mind, Two Languages: Bilingual Language Processing* (pp.1-22). Oxford: Blackwell.

Grosjean, F. (2008). *Studying Bilinguals.* Oxford, UK: Oxford University Press.

Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean, & L. Ping (Eds,). *The Psycholinguistics of Bilingualism* (pp. 5-25). John Wiley & Sons.

Gureje, O., Ogunniyi, A., Kola, L., & Abiona, T. (2011). Incidence of and risk factors for dementia in the Ibadan study of aging. *Journal of the American Geriatrics Society, 59*(5), 869-874. https://doi.org/10.1111/j.1532-5415.2011.03374.x

Guzmán-Vélez, E., & Tranel, D. (2015). Does bilingualism contribute to cognitive reserve? Cognitive and neural perspectives. *Neuropsychology, 29*(1), 139-150. https://doi.org/10.1037/neu0000105

Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: A critical discussion and some new longitudinal data. In K. E. Nelson (Ed.), *Children's Language* (pp. 319-344). Lawrence Erlbaum Associates, Inc.

Hämäläinen, S., Joutsa, J., Sihvonen, A. J., Leminen, A., & Lehtonen, M. (2018). Beyond volume: A surface-based approach to bilingualism-induced grey matter changes. *Neuropsychologia, 117*, 1-7. https://doi.org/10.1016/j.neuropsychologia.2018.04.038

Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science, 31*(3), 355-384. https://doi.org/10.1080/15326900701326402

Hampton, J. A., Aina, B., Andersson, J. M., Mirza, H. Z., & Parmar, S. (2012). The Rumsfeld effect: The unknown unknown. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(2), 340-355. https://doi.org/10.1037/a0025376

Harris, R. J., & Nelson, E. M. M. (1992). Bilingualism: Not the exception any more. *Advances in Psychology, 83*, 3-14. https://doi.org/10.1016/S0166-4115(08)61485-5

Hassett, A., George, K., & Harrigan, S. (1999). Admissions of elderly patients from English-speaking and non-English-speaking backgrounds to an inpatient psychogeriatric unit. *Australian & New Zealand Journal of Psychiatry, 33*(4), 576-582. https://doi.org/10.1080/j.1440-1614.1999.00537.x

Haugen, E. (1953). *The Norwegian language in America: A study of bilingual behavior, Volume 2: The American Dialects of Norwegian.* Philadelphia: University of Pennsylvania Press. https://doi.org/10.9783/9781512820522.

Hazeltine, E., Bunge, S. A., Scanlon, M. D., & Gabrieli, J. D. (2003). Material-dependent and material-independent selection processes in the frontal and parietal lobes: an event-related fMRI investigation of response competition. *Neuropsychologia, 41*(9), 1208-1217. https://doi.org/10.1016/s0028-3932(03)00040-x

Hendrie, H. C., Albert, M. S., Butters, M. A., Gao, S., Knopman, D. S., Launer, L. J., Yaffe, K., Cuthbert, B. N., Edwards, E., Wagster, M. V. (2006). The NIH cognitive and emotional health project: report of the critical evaluation study committee. *Alzheimer's & Dementia, 2*(1), 12-32. https://doi.org/10.1016/j.jalz.2005.11.004

Henley, A., & Schott, J. (1999). *Culture, religion and patient care in a multi-ethnic society: A handbook for professionals.* London: Age Concern.

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia, 42*(9), 1212-1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*, 625-658. https://doi.org/10.3758/s13423-011-0116-7

Hindle, J. V., Martin-Forbes, P. A., Bastable, A. J., Pye, K. L., Martyr, A., Whitaker, C. J., Craik, F. I. M., Bialystok, E., Thomas, E. M., Mueller, V. C., & Clare, L. (2015). Cognitive reserve in Parkinson's disease: The effects of Welsh-English bilingualism on executive function. *Parkinson's Disease, 2015*, 943572. https://doi.org/10.1155/2015/943572

Hoffman, P. (2018). An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *Scientific Reports, 8*(1), 8145. https://doi.org/10.1038/s41598-018-26569-0

Hoffman, P. (2019). Divergent effects of healthy ageing on semantic knowledge and control: Evidence from novel comparisons with semantically impaired patients. *Journal of Neuropsychology, 13*(3), 462-484. https://doi.org/10.1111/jnp.12159

Hoffman, P., Rogers, T. T., & Ralph, M. A. L. (2011). Semantic diversity accounts for the "missing" word frequency effect in stroke aphasia: Insights using a novel method to quantify contextual variability in meaning. *Journal of Cognitive Neuroscience, 23*(9), 2432-2446. https://doi.org/10.1162/jocn.2011.21614

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107-129. https://doi.org/10.1016/0001-6918(67)90011-X

Houston, J. R., Allen, P. A., Rogers, J. M., Lien, M. C., Allen, N. J., Hughes, M. L., Bapuraj, J. R., Eppelheimer, M. S., Loth. F., Stoodley, M. A., Vorster, S. J., & Luciano, M. G. (2019). Type I Chiari malformation, RBANS performance, and brain morphology: Connecting the dots on cognition and macrolevel brain structure. *Neuropsychology, 33*(5), 725-738. https://doi.org/10.1037/neu0000547

Hsieh, S., Schubert, S., Hoon, C., Mioshi, E., & Hodges, J. R. (2013). Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 36*(3-4), 242-250. https://doi.org/10.1159/000351671

Hualde, J. I., Lakarra, J. A., & Trask, R. L. (1996). T*owards a history of the Basque language.* The Netherlands: John Benjamins Publishing.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica, 35*(1), 49-66.

Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., Shaw, L. M., Vemuri, P., Wiste, H. J., Weigand, S. D., Lesnick, T. G., Pankratz, V. S., Donohue, M. C., & Trojanowski, J. Q. (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet. Neurology, 12*(2), 207-216. https://doi.org/10.1016/S1474-4422(12)70291-0

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet. Neurology*, 9(1), 119-128. https://doi.org/10.1016/S1474-4422(09)70299-6

Jackson, R. L. (2021). The neural correlates of semantic control revisited. *NeuroImage, 224*, 117444. https://doi.org/10.1016/j.neuroimage.2020.117444

Jackson, R. L., Hoffman, P., Pobric, G., & Ralph, M. A. L. (2016). The semantic network at work and rest: differential connectivity of anterior temporal lobe subregions. *Journal of Neuroscience, 36*(5), 1490-1501. https://doi.org/10.1523/jneurosci.2999-15.2016

Jacobson, P. F., & Schwartz, R. G. (2002). Morphology in incipient bilingual Spanish-speaking preschool children with specific language impairment. *Applied Psycholinguistics, 23*(1), 23-41. doi:10.1017/S0142716402000024

Jafari, Z., Esmaili, M., Toufan, R., & Aghamollaei, M. (2015). Bilingual proficiency and cognitive reserve in Persian–English bilingual older adults. *Aging Clinical and Experimental Research, 27*, 351-357. https://doi.org/10.1007/s40520-014-0288-x

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience, 21*(12), 2434-2444. https://doi.org/10.1162/jocn.2008.21179

Jedrziewski, M. K., Lee, V. M. Y., & Trojanowski, J. Q. (2007). Physical activity and cognitive health. *Alzheimer's & Dementia, 3*(2), 98-108. https://doi.org/10.1016/j.jalz.2007.01.009

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain, 129*(8), 2132-2147. https://doi.org/10.1093/brain/awl153

Johns, B. T., Sheppard, C. L., Jones, M. N., & Taler, V. (2016). The role of semantic diversity in word recognition across aging and bilingualism. *Frontiers in Psychology, 7*, 195083. https://doi.org/10.3389/fpsyg.2016.00703

Jones, H. M. (2012). *A statistical overview of the Welsh language* (pp. 1-137). Cardiff: Welsh Language Board. https://orca.cardiff.ac.uk/id/eprint/67149

Jones, M. N., Willits, J., Dennis, S., & Jones, M. (2015). Models of semantic memory. In J. R. Busemeyer, Z, Wang, J. T. Townsend, & A. Eidels (Eds,), *The Oxford Handbook of Mathematical and Computational Psychology* (pp. 232-254). Oxford University Press.

Jones, R. N., Manly, J., Glymour, M. M., Rentz, D. M., Jefferson, A. L., & Stern, Y. (2011). Conceptual and measurement challenges in research on cognitive reserve. *Journal of the International Neuropsychological Society, 17*(4), 593-601. https://doi.org/10.1017/s1355617710001748

Jones, R., & Merriman, P. (2009). Hot, banal and everyday nationalism: Bilingual road signs in Wales. *Political Geography, 28*(3), 164-173. https://doi.org/10.1016/j.polgeo.2009.03.002

Jung, J., Cloutman, L. L., Binney, R. J., & Ralph, M. A. L. (2017). The structural connectivity of higher order association cortices reflects human functional brain networks. *Cortex, 97*, 221-239. https://doi.org/10.1016%2Fj.cortex.2016.08.011

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging, 23*(1), 70-83. https://doi.org/10.1037/0882-7974.23.1.70

Kempler, D. (1995). Language changes in dementia of the Alzheimer type. In R. Lubinski (Ed.). *Dementia and Communication* (pp. 98-114). Philadelphia: B.C. Decker.

Kenett, Y. N., Levi, E., Anaki, D., & Faust, M. (2017). The semantic distance task: Quantifying semantic distance with semantic network path length. J*ournal of Experimental Psychology: Learning, Memory, and Cognition, 43*(9), 1470-1489. https://doi.org/10.1037/xlm0000391

Keren, G. (1991). Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica, 77*(3), 217-273. https://doi.org/10.1016/0001-6918(91)90036-Y

Khleif, B. B. (1979). Language as an ethnic boundary in Welsh-English relations. *International Journal of the Sociology of Language, Vol.1979*(20), 59-74. https://doi.org/10.1515/ijsl.1979.20.59

Kisser, J. E., Wendell, C. R., Spencer, R. J., & Waldstein, S. R. (2012). Neuropsychological performance of native versus non-native English speakers. *Archives of Clinical Neuropsychology, 27*(7), 749-755. https://doi.org/10.1093/arclin/acs082

Klein, L. A., & Buchanan, J. A. (2009). Psychometric properties of the pyramids and palm trees test. *Journal of Clinical and Experimental Neuropsychology, 31*(7), 803-808. doi.org/10.1080/13803390802508926

Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods, 53*(6), 2576-2590. https://doi.org/10.3758/s13428-021-01587-5

Kong, E. H., Kim, H., & Kim, H. (2022). Nursing home staff's perceptions of barriers and needs in implementing person-centred care for people living with dementia: A qualitative study. *Journal of Clinical Nursing, 31*(13-14), 1896-1906. https://doi.org/10.1111/jocn.15729

Kovalchik, S., Camerer, C. F., Grether, D. M., Plott, C. R., & Allman, J. M. (2005). Aging and decision making: A comparison between neurologically healthy elderly and young individuals. *Journal of Economic Behavior & Organization, 58*(1), 79-94. https://doi.org/10.1016/j.jebo.2003.12.001

Kowoll, M. E., Degen, C., Gladis, S., & Schröder, J. (2015). Neuropsychological profiles and verbal abilities in lifelong bilinguals with mild cognitive impairment and Alzheimer's

disease. *Journal of Alzheimer's Disease, 45*(4), 1257-1268. https://doi.org/10.3233/jad-142880

Kramer, A. F., Humphrey, D. G., Larish, J. F., & Logan, G. D. (1994). Aging and inhibition: beyond a unitary view of inhibitory processing in attention. *Psychology and Aging, 9*(4), 491-512. https://doi.org/10.1037/0882-7974.9.4.491

Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia, 76*, 92-107. https://doi.org/10.1016%2Fj.neuropsychologia.2015.02.030

Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletin & Review, 28*(1), 40-80. https://doi.org/10.3758/s13423-020-01792-x

Laditka, S. B., Corwin, S. J., Laditka, J. N., Liu, R., Tseng, W., Wu, B., Beard, R. L., Sharkey, J. R., & Ivey, S. L. (2009). Attitudes about aging well among a diverse group of older Americans: Implications for promoting cognitive health. *The Gerontologist, 49*(S1), S30-S39. https://doi.org/10.1093/geront/gnp084

Lasagabaster, D. (2007). Language use and language attitudes in the Basque Country. In D. Lasagabaster, & A. Huguet (Eds,), *Multilingualism in European Bilingual Contexts: Language Use and Attitudes* (pp. 65-89). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781853599316

Lawton, D. M., Gasquoine, P. G., & Weimer, A. A. (2015). Age of dementia diagnosis in community dwelling bilingual and monolingual Hispanic Americans. *Cortex, 66,* 141-145. https://doi.org/10.1016/j.cortex.2014.11.017

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behaviour Research Methods, 44*(2), 325-343. https://doi.org/10.3758%2Fs13428-011-0146-0

Leonard, M. K., Torres, C., Travis, K. E., Brown, T. T., Hagler Jr, D. J., Dale, A. M., Elam, J. L., & Halgren, E. (2011). Language proficiency modulates the recruitment of non-classical language areas in bilinguals. *PLoS One, 6*(3), e18240. https://doi.org/10.1371/journal.pone.0018240

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1977). Calibration of probabilities: The state of the art. In H. Jungermann, G. & De Zeeuw (Eds,), *Decision Making and Change in Human Affairs: Theory and Decision Library, vol 16.* Springer, Dordrecht. https://doi.org/10.1007/978-94-010-1276-8_19

Liedtke, W. W., & Nelson, L. D. (1968). Concept formation and bilingualism. *Alberta Journal of Educational Research, 14*(4), 225-232.

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon, 3*(3), 349-374. https://doi.org/10.1075/ml.3.3.06lin

Ljungberg, J. K., Hansson, P., Adolfsson, R., & Nilsson, L. G. (2016). The effect of language skills on dementia in a Swedish longitudinal cohort. *Linguistic Approaches to Bilingualism, 6*(1-2), 190-204. https://doi.org/10.1075/lab.14031.lju

López, E., Steiner, A. J., Hardy, D. J., IsHak, W. W., & Anderson, W. B. (2016). Discrepancies between bilinguals' performance on the Spanish and English versions of the WAIS Digit Span task: Cross-cultural implications. *Applied Neuropsychology: Adult, 23*(5), 343-352. https://doi.org/10.1080/23279095.2015.1074577

Lorch Jr, R. F. (1982). Priming and search processes in semantic memory: A test of three models of spreading activation. J*ournal of Verbal Learning and Verbal Behavior, 21*(4), 468-492. https://doi.org/10.1016/S0022-5371(82)90736-8

Loughan, A. R., Braun, S. E., & Lanoye, A. (2019). Repeatable battery for the assessment of neuropsychological status (RBANS): preliminary utility in adult neuro-oncology. *Neuro-Oncology Practice, 6*(4), 289-296. https://doi.org/10.1093/nop/npy050

Luk, G. (2015). Who are the bilinguals (and monolinguals)? *Bilingualism: Language and Cognition, 18*(1), 35-36. https://doi.org/10.1017/S1366728914000625

Lyons, J. (1981). *Language and Linguistics. An Introduction.* Cambridge: Cambridge University Press.

Macizo, P., Bajo, T., & Paolieri, D. (2012). Language switching and language competition. *Second Language Research, 28*(2), 131-149. https://doi.org/10.1177/026765831143489

MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. S. Gorfein, & C. M. MacLeod (Eds.), *Inhibition in Cognition* (pp. 3-23). American Psychological Association. https://doi.org/10.1037/11587-004

Macnamara, B. N., & Conway, A. R. (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

Marian, V., & Hayakawa, S. (2021). Measuring bilingualism: The quest for a “bilingualism quotient”. *Applied Psycholinguistics, 42*(2), 527-548. https://doi.org/10.1017/S0142716420000533

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition, 6*(2), 97-115. https://doi.org/10.1017/S1366728903001068

Marra, D. E., Hamlet, K. M., Bauer, R. M., & Bowers, D. (2020). Validity of teleneuropsychology for older adults in response to COVID-19: A systematic and critical review. *The Clinical Neuropsychologist, 34*(7-8), 1411-1452. https://doi.org/10.1080%2F13854046.2020.1769192

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

Marulappa, N., Anderson, N. N., Bethell, J., Bourbonnais, A., Kelly, F., McMurray, J., Rogers, H. L., Vedel, I., & Gagliardi, A. R. (2022). How to implement person-centred care and support for dementia in outpatient and home/community settings: Scoping review. *BMC Health Services Research, 22*(1), 1-15. https://doi.org/10.1186/s12913-022-07875-w

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition, 6*(4), 462-472. https://doi.org/10.3758/BF03197480

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11*(1), 1-37. https://doi.org/10.1016/0010-0285(79)90002-1

McDowd, J. M., Oseas-Kreger, D. M., & Filion, D. L. (1995). Inhibitory processes in cognition and aging. In. F. M. Dempster, & C. J. Brainerd (Eds,), Interference and inhibition in cognition (pp. 363-400). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-012208930-5/50012-X

McDowd, J., Vercruyssen, M., & Birren, J. E. (2020). Aging, divided attention, and dual-task performance. In D, Damos (Ed.), *Multiple-task performance* (pp. 387-414). CRC Press. https://doi.org/10.1201/9781003069447

McGuire, C., Crawford, S., & Evans, J. J. (2019). Effort testing in dementia assessment: A systematic review. *Archives of Clinical Neuropsychology, 34*(1), 114-131. https://doi.org/10.1093/arclin/acy012

McIntosh, V. (2015). Treatment of Phonological Anomia in Aphasia: Some Suggestions for Users of Signed Language. *Newcastle and Northumbria Working Papers in Linguistic 21*(2). https://www.ncl.ac.uk/media/wwwnclacuk/crills/files/importedresearch/volume-21.2.pdf#page=44

McKay, C., Casey, J. E., Wertheimer, J., & Fichtenberg, N. L. (2007). Reliability and validity of the RBANS in a traumatic brain injured sample. *Archives of Clinical Neuropsychology, 22*(1), 91-98. https://doi.org/10.1016/j.acn.2006.11.003

McMurtray, A., Saito, E., & Nakamoto, B. (2009). Language preference and development of dementia among bilingual individuals. *Hawaii Medical Journal, 68*(9), 223-226. https://pubmed.ncbi.nlm.nih.gov/19842364

McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. In A. F. Healy, R. W. Proctor, & I. B. Weiner (Eds.), *Handbook of Psychology: Experimental Psychology* (Vol. 4, pp. 447-474). John Wiley & Sons, Inc.

McRae, K. (2004). Semantic memory: Some insights from feature-based connectionist attractor networks. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol 4, pp. 41-86). Elsevier Academic Press.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature, 431*(7010), 757-757. https://doi.org/10.1038/431757a

Mendez, M. F., Perryman, K. M., Pontón, M. O., & Cummings, J. L. (1999). Bilingualism and dementia. *The Journal of Neuropsychiatry and Clinical Neurosciences, 11*(3), 411-412. https://doi.org/10.1176/jnp.11.3.411

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*(1), 25-40. https://doi.org/10.1006/jmla.1998.2602

Mindt, M. R., Arentoft, A., Germano, K. K., D'Aquila, E., Scheiner, D., Pizzirusso, M., Sandoval, T. C., & Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review, 18*(3), 255-268. https://doi.org/10.1007/s11065-008-9069-7

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

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). Semantic significance: a new measure of feature salience. *Memory & Cognition, 42*, 355-369. https://doi.org/10.3758/s13421-013-0365-y

Moore, R. C., Davine, T., Harmell, A. L., Cardenas, V., Palmer, B. W., & Mausbach, B. T. (2013). Using the repeatable battery for the assessment of neuropsychological status (RBANS) effort index to predict treatment group attendance in patients with schizophrenia. *Journal of the International Neuropsychological Society, 19*(2), 198-205. https://doi.org/10.1017/s1355617712001221

Moradi, H. (2014). An investigation through different types of bilinguals and bilingualism. *International Journal of Humanities & Social Science Studies, 1*(2), 107-112. https://www.ijhssnet.com/

Morgan, T., & Crowder, R. (2003). Mini Mental State Examinations in English: are they suitable for people with dementia who are Welsh speaking? *Dementia, 2*(2), 267-272. https://doi.org/10.1177/1471301203002002009

Morris, J. (2014). The influence of social factors on minority language engagement amongst young people: an investigation of Welsh-English bilinguals in North Wales. *International Journal of the Sociology of Language, 2014*(230), 65-89. https://doi.org/10.1515/ijsl-2014-002

Mortimer, J. A., Snowdon, D. A., & Markesbery, W. R. (2003). Head circumference, education and risk of dementia: findings from the Nun Study. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 671-679. https://doi.org/10.1076/jcen.25.5.671.14584

Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., & Tyler, L. K. (2005). Selecting among competing alternatives: selection and retrieval in the left inferior frontal gyrus. *Cerebral Cortex, 15*(11), 1723-1735. https://psycnet.apa.org/doi/10.1093/cercor/bhi049

Mueller Gathercole, V. C., Mon Thomas, E., & Hughes, E. (2008). Designing a normed receptive vocabulary test for bilingual populations: A model from Welsh. I*nternational Journal of Bilingual Education and Bilingualism, 11*(6), 678-720. https://doi.org/10.1080/13670050802149283

Mueller, K. D., Koscik, R. L., Turkstra, L. S., Riedeman, S. K., LaRue, A., Clark, L. R., Hermann, B., Sagar, M. A., Johnson, S. C., & Johnson, S. C. (2016). Connected

language in late middle-aged adults at risk for Alzheimer's disease. *Journal of Alzheimer's Disease, 54*(4), 1539-1550. https://doi.org/10.3233/jad-160252

Müllers, P., Taubert, M., & Müller, N. G. (2019). Physical exercise as personalized medicine for dementia prevention? *Frontiers in Physiology, 10*, 672. https://doi.org/10.3389/fphys.2019.00672

Murphy, A. (2017). Children's perceptions of their own national identity and what it means to be Welsh. UK: Cardiff Metropolitan University. Unpublished Doctoral Dissertation. https://doi.org/10.25401/cardiffmet.20259093.v1

Musselwhite, C., & Scott, T. (2019). Developing A Model of Mobility Capital for An Ageing Population. *International Journal of Environmental Research and Public Health, 16*(18), 3327. https://doi.org/10.3390/ijerph16183327

Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage, 43*(4), 801-807. https://doi.org/10.1016%2Fj.neuroimage.2008.07.017

Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech, Language, and Hearing Research, 28*(3), 405-410. https://doi.org/10.1044/jshr.2803.405

Nielsen, T. R., Vogel, A., Riepe, M. W., de Mendonça, A., Rodriguez, G., Nobili, F., Gade, A., & Waldemar, G. (2011). Assessment of dementia in ethnic minority patients in Europe: A European Alzheimer's Disease Consortium survey. I*nternational Psychogeriatrics, 23*(1), 86-95. https://doi.org/10.1017/s1041610210000955

Noonan, K. A., Jefferies, E., Corbett, F., & Lambon Ralph, M. A. (2010). Elucidating the nature of deregulated semantic cognition in semantic aphasia: evidence for the roles of prefrontal and temporo-parietal cortices. *Journal of Cognitive Neuroscience, 22*(7), 1597-1613. https://doi.org/10.1162/jocn.2009.21289

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience, 25*(11), 1824-1850. https://doi.org/10.1162/jocn_a_00442

Noone, P. (2015). Addenbrooke's cognitive examination-III. *Occupational Medicine, 65*(5), 418-420. https://doi.org/10.1093/occmed/kqv041

Nour, S., Struys, E., & Stengers, H. (2019). Attention network in interpreters: The role of training and experience. *Behavioral Sciences, 9*(4), 43. https://doi.org/10.3390/bs9040043

Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive Reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging Clinical and Experimental Research, 24*(3), 218-226. https://doi.org/10.3275/7800

Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U., & Nilsson, L. G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 51*(4), P234-P240. https://doi.org/10.1093/geronb/51b.4.p234

Obler, L.K. and Albert M. L. (1984). Language in aging. In M. L. Albert (Ed.), *Clinical Neuropsychology of Aging* (pp.245-253). New York: Oxford University Press.

Office for National Statistics. (2020). *Population estimates for the UK, England and Wales, Scotland and Northern Ireland, provisional: mid-2019*. Retrieved June 23, 2023 from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2021

Okanda, M., Moriguchi, Y., & Itakura, S. (2010). Language and cognitive shifting: Evidence from young monolingual and bilingual children. *Psychological Reports, 107*(1), 68-78E. https://doi.org/10.2466/03.10.28.PR0.107.4.68-78

Olaithe, M., Weinborn, M., Lowndes, T., Ng, A., Hodgson, E., Fine, L., Denise, P., Pushpanathan, M., Bayliss, D., Anderson, M., & Bucks, R. S. (2019). Repeatable battery for the assessment of neuropsychological status (RBANS): normative data for older adults. *Archives of Clinical Neuropsychology, 34*(8), 1356-1366. https://psycnet.apa.org/doi/10.1093/arclin/acy102

Older People's Commissioner for Wales. (2016). *Dementia: More than just memory loss.* Retrieved June 23, 2023, from https://socialcare.wales/cms-assets/documents/Dementia-more-than-just-memory-loss.pdf

Olsen, R. K., Pangelinan, M. M., Bogulski, C., Chakravarty, M. M., Luk, G., Grady, C. L., & Bialystok, E. (2015). The effect of lifelong bilingualism on regional grey and white matter volume. *Brain Research, 1612*, 128-139. https://doi.org/10.1016/j.brainres.2015.02.034

Ordin, M., Polyanskaya, L., & Soto, D. (2020). Metacognitive processing in language learning tasks is affected by bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(3), 529. https://doi.org/10.1037/xlm0000739

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232-258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Sawi, O. M., Dalibar, C., Darrow, J., & Johnson, H. A. (2014). The brain mechanisms underlying the cognitive benefits of bilingualism may be extraordinarily difficult to discover. *AIMS Neuroscience, 1*(3), 245-56. https://doi.org/10.3934/Neuroscience.2014.3.245

Paradis, M. (2008). Bilingualism and neuropsychiatric disorders. *Journal of Neurolinguistics, 21*(3), 199-230. https://doi.org/10.1016/j.jneuroling.2007.09.002

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17*(2), 299-230. https://doi.org/10.1037/0882-7974.17.2.299

Park, J., Carp, J., Hebrank, A., Park, D. C., & Polk, T. A. (2010). Neural specificity predicts fluid processing ability in older adults. J*ournal of Neuroscience, 30*(27), 9253-9259. https://doi.org/10.1523%2FJNEUROSCI.0853-10.2010

Parks, A. C., Davis, J., Spresser, C. D., Stroescu, I., & Ecklund-Johnson, E. (2021). Validity of in-home teleneuropsychological testing in the wake of COVID-19. *Archives of Clinical Neuropsychology, 36*(6), 887-896. https://doi.org/10.1093/arclin/acab002

Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition, 23*(1), 204-218. https://doi.org/10.1017/S1366728918001098

Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition, 23*(1), 204-218. doi:10.1017/S1366728918001098

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976-987. https://doi.org/10.1038/nrn2277

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied, 76*(27), 1-23. https://doi.org/10.1037/h0093840

Peal, E., & Lambert, W. E. (1967). The relation of bilingualism to intelligence. In J. Michael (ed.), *Foreign Language Teaching: An anthology* (pp. 143-191). New York: Macmillan. https://doi.org/10.1037/h0093840

Penhallurick, R. (1993). Welsh English: a national language? *Dialectologia et Geolinguistica, 1993*(1), 28-46.

Penninkilampi, R., Casey, A. N., Singh, M. F., & Brodaty, H. (2018). The association between social engagement, loneliness, and risk of dementia: a systematic review and meta-analysis. *Journal of Alzheimer's Disease, 66*(4), 1619-1633. https://doi.org/10.3233/jad-180439

Perani, D., Farsad, M., Ballarini, T., Lubian, F., Malpetti, M., Fracchetti, A., Magnani, G., March, A., & Abutalebi, J. (2017). The impact of bilingualism on brain reserve and metabolic connectivity in Alzheimer's dementia. *Proceedings of the National Academy of Sciences, 114*(7), 1690-1695. https://doi.org/10.1073/pnas.1610909114

Perfect, T. J., & Hollins, T. S. (1996). Predictive feeling of knowing judgements and postdictive confidence judgements in eyewitness memory and general knowledge. *Applied Cognitive Psychology, 10*(5), 371-382. https://doi.org/10.1002/(SICI)1099-0720(199610)10:5%3C371::AID-ACP389%3E3.0.CO;2-O

Perquin, M., Vaillant, M., Schuller, A. M., Pastore, J., Dartigues, J. F., Lair, M. L., Diedrich, N. & on behalf of the MemoVie Group. (2013). Lifelong exposure to multilingualism: new evidence to support cognitive reserve hypothesis. *PLoS One, 8*(4), e62030. https://doi.org/10.1371/journal.pone.0062030

Pigliautile, M., Ricci, M., Mioshi, E., Ercolani, S., Mangialasche, F., Monastero, R., Croce, M. F., Federici, S., & Mecocci, P. (2011). Validation study of the Italian Addenbrooke's Cognitive Examination Revised in a young-old and old-old population. *Dementia and Geriatric Cognitive Disorders, 32*(5), 301-307. https://doi.org/10.1159/000334657

Pitta, P., Marcos, L. R., & Alpert, M. (1978). Language switching as a treatment strategy with bilingual patients. *American Journal of Psychoanalysis, 38*(3), 255-258. https://doi.org/10.1007/BF01252975

Pobric, G., Jefferies, E., & Ralph, M. A. L. (2007). Anterior temporal lobes mediate semantic representation: mimicking semantic dementia by using rTMS in normal participants. *Proceedings of the National Academy of Sciences, 104*(50), 20137-20141. https://doi.org/10.1073/pnas.0707383104

Polidori, M. C., Nelles, G., & Pientka, L. (2010). Prevention of dementia: focus on lifestyle. *International Journal of Alzheimer's Disease, vol. 2010,* 393579. https://doi.org/10.4061/2010/393579

Ponjoan, A., Garre-Olmo, J., Blanch, J., Fages, E., Alves-Cabratosa, L., Martí-Lluch, R., Comas-Cufi, M., Parramon, D., Garcia-Gill, M., & Ramos, R. (2019). Epidemiology of dementia: prevalence and incidence estimates using validated electronic health records from primary care. *Clinical Epidemiology, 4*(11), 217-228. https://doi.org/10.2147/clep.s186590

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*(3), 415-422. https://doi.org/10.1016/j.acn.2007.01.015

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's & Dementia, 9*(1), 63-75. https://doi.org/10.1016/j.jalz.2012.11.007

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

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. Nature Reviews Neuroscience, 18(1), 42. https://doi.org/10.1038/nrn.2016.150

Ralph, M. A. L., Patterson, K., & Hodges, J. R. (1997). The relationship between naming and semantic knowledge for different categories in dementia of Alzheimer's type. *Neuropsychologia, 35*(9), 1251-1260. https://doi.org/10.1037/a0029106

Ramírez-Esparza, N., & García-Sierra, A. (2014). The bilingual brain: Language, culture, and identity. In V. Benet-Martinez, & Y. Y. Hong (Eds.), *The Oxford Handbook of Multicultural Identity* (pp. 35-56). Oxford University Press.

Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *Journal of Clinical and Experimental Neuropsychology, 20*(3), 310-319. https://doi.org/10.1076/jcen.20.3.310.823

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging, 16*(2), 323. https://doi.org/10.1037/0882-7974.16.2.323

Raven, J. C., & Court, J. H. (1998). *Raven's progressive matrices and vocabulary scales* (Vol. 759). Oxford: Oxford Psychologists Press Ltd.

Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review, 25*(5), 1695-1716. https://doi.org/10.3758/s13423-017-1384-7

Riehl, C. M. (2010). The mental representation of bilingualism. *Wiley Interdisciplinary Reviews: Cognitive Science, 1(5)*, 750-758. https://doi.org/10.1002/wcs.74

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*(1), 1-20. https://doi.org/10.1016/S0022-5371(73)80056-8

Rivera Mindt, M., Arentoft, A., Kubo Germano, K., D'Aquila, E., Scheiner, D., Pizzirusso, M., ... & Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review, 18,* 255-268. https://doi.org/10.1007/s11065-008-9069-7

Rosli, R., Tan, M. P., Gray, W. K., Subramanian, P., & Chin, A. V. (2016). Cognitive assessment tools in Asia: a systematic review. *International Psychogeriatrics, 28*(2), 189-210. https://doi.org/10.1017/s1041610215001635

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology, 7*(1), 17-24. https://doi.org/10.1017/s1041610215001635

Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., LUIS, M., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. International Journal of Neuroscience, 112(6), 759-776.

Rosselli, M., Ardila, A., Santisi, M. N., Arecco, M. D. R., Salvatierra, J., Conde, A., & Lenis, B. O. N. I. E. (2002). Stroop effect in Spanish–English bilinguals. *Journal of the International Neuropsychological Society, 8*(6), 819-827. https://doi.org/10.1017/S1355617702860106

Rubio-Fernández, P. (2017). Why are bilinguals better than monolinguals at false-belief tasks? *Psychonomic Bulletin & Review, 24*(3), 987-998. https://doi.org/10.1080/00207450290025752

Saffran, E. M. (2000). The organization of semantic memory: In support of a distributed model. *Brain and Language, 71*(1), 204-212. https://doi.org/10.1006/brln.1999.2251

Salthouse, T. A. (2004a). What and when of cognitive aging. *Current Directions in Psychological Science, 13*(4), 140-144. https://doi.org/10.1111/j.0963-7214.2004.00293.x

Salthouse, T. A. (2004b). Localizing age-related individual differences in a hierarchical structure. *Intelligence, 32*(6), 541-561. https://doi.org/10.1016/j.intell.2004.07.003

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, 13*(2), 231-252. doi:10.1017/S1366728909990514

Sasaki, S., Sato, A., Tanabe, Y., Matsuoka, S., Adachi, A., Kayano, T., Yamazaki, H., Matsuno, Y., Miyake, A., & Watanabe, T. (2021). Associations between socioeconomic status, social participation, and physical activity in older people during the COVID-19 pandemic: a cross-sectional study in a northern Japanese city. *International Journal of Environmental Research and Public Health, 18*(4), 1477. https://doi.org/10.3390/ijerph18041477

Sattler, C., Toro, P., Schönknecht, P., & Schröder, J. (2012). Cognitive activity, education and socioeconomic status as preventive factors for mild cognitive impairment and Alzheimer's disease. *Psychiatry Research, 196*(1), 90-95. https://doi.org/10.1016/j.psychres.2011.11.012

Saxton, J. A., Ratcliff, G., Dodge, H., Pandav, R., Baddeley, A., & Ganguli, M. (2001). Speed and capacity of language processing test: normative data from an older American community-dwelling sample. *Applied Neuropsychology, 8*(4), 193-203. https://doi.org/10.1207/s15324826an0804_1

Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 625-633. https://doi.org/10.1076/jcen.25.5.625.14576

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., ... & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141-1152. https://doi.org/10.1111/2041-210X.13434

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex, 48*(8), 991-996. https://doi.org/10.1016/j.cortex.2011.04.009

Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(3), 681. https://doi.org/10.1037/0278-7393.23.3.681

Shannon, K., Bail, K., & Neville, S. (2019). Dementia-friendly community initiatives: An integrative review. *Journal of Clinical Nursing, 28*(11-12), 2035-2045. https://doi.org/10.1111/jocn.14746

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers In Psychology, 5*, 772. https://doi.org/10.3389%2Ffpsyg.2014.00772

Sheng, L., McGregor, K. K., & Marian, V. (2006). Lexical–semantic organization in bilingual children: Evidence from a repeated word association task. J*ournal of Speech, Language, and Hearing Research, 49*(3), 572-587. https://doi.org/10.1044/1092-4388(2006/041)

Silverberg, N. D., Wertheimer, J. C., & Fichtenberg, N. L. (2007). An effort index for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). *The*

Clinical Neuropsychologist, 21(5), 841-854. https://doi.org/10.1080/13854040600850958

Smith, E. E., & Estes, W. K. (1978). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol 6. Linguistic Functions in Cognitive Theory* (pp. 1-56). Hillsdale, NJ: Erlbaum.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*(3), 214-241. https://doi.org/10.1037/h0036351

Smith-Seemiller L, Franzen MD, & Bowers D (1997). Use of Wisconsin card sorting test short forms in clinical samples. *Clinical Neuropsychologist, 11*(4), 421–427. https://doi.org/10.1080/13854049708400472

Snowdon, D. A. (2003). Healthy aging and dementia: findings from the Nun Study. Annals of Internal Medicine, 139(5 Part 2), 450-454. https://doi.org/10.7326/0003-4819-139-5_part_2-200309021-00014

So, M., Foxe, D., Kumfor, F., Murray, C., Hsieh, S., Savage, G., Ahmed, R. M., Burrell, J. R., Hodges, J. R., Irisj, M., Piguet, O. (2018). Addenbrooke's cognitive examination III: psychometric characteristics and relations to functional ability in dementia. *Journal of the International Neuropsychological Society, 24*(8), 854-863. https://doi.org/10.1017/s1355617718000541

Soltani, M., Moradi, N., Rezaei, H., Hosseini, M., & Jasemi, E. (2019). Comparison of verbal fluency in monolingual and bilingual elderly in Iran. *Applied Neuropsychology: Adult, 28*(1), 80-87. https://doi.org/10.1080/23279095.2019.1594234

Spreng, R. N., & Turner, G. R. (2019). The shifting architecture of cognition and brain function in older adulthood. *Perspectives on Psychological Science, 14*(4), 523-542. https://doi.org/10.1177/1745691619827511

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society, 8*(3), 448-460. https://doi.org/10.1017/S1355617702813248

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology, 11*(11), 1006-1012. https://doi.org/10.1016/s1474-4422(12)70191-6

Stern, Y., Barnes, C. A., Grady, C., Jones, R. N., & Raz, N. (2019). Brain reserve, cognitive reserve, compensation, and maintenance: operationalization, validity, and mechanisms of cognitive resilience. *Neurobiology of Aging, 83*, 124-129. https://doi.org/10.1016/j.neurobiolaging.2019.03.022

Stilwell, B. L., Dow, R. M., Lamers, C., & Woods, R. T. (2016). Language changes in bilingual individuals with Alzheimer's disease. *International Journal of Language & Communication Disorders, 51*(2), 113-127. https://doi.org/10.1111/1460-6984.12190

Suarez, P. A., Gollan, T. H., Heaton, R., Grant, I., Cherner, M., & HNRC Group. (2014). Second-language fluency predicts native language Stroop effects: Evidence from Spanish–English bilinguals. *Journal of the International Neuropsychological Society, 20*(3), 342-348. https://doi.org/10.1017/S1355617714000058

Suraweera, C. U., Anandakumar, D., Dahanayake, D., Subendran, M., Perera, U. T., Hanwella, R., & de Silva, V. A. (2016). Validation of the Sinhala version of the Repeatable Battery for Assessment of Neuropsychological Status (RBANS). Ceylon Medical Journal, 61(4), 167-170. https://doi.org/10.4038/cmj.v61i4.8383

Tadayon, E., Pascual-Leone, A., & Santarnecchi, E. (2020). Differential contribution of cortical thickness, surface area, and gyrification to fluid and crystallized intelligence. *Cerebral Cortex, 30*(1), 215-225. https://doi.org/10.1093/cercor/bhz082

Tare, M., & Linck, J. (2011, November). *Exploring bilingual cognitive advantages when controlling for background variables* (Poster Presentation). 52nd Annual Meeting of the Psychonomic Society, Seattle, WA.

The British Psychological Society. (2020). *Psychological assessment undertaken remotely.* Retrieved June 23, 2023, from https://cms.bps.org.uk/sites/default/files/2022-06/Psychological%20assessment%20undertaken%20remotely.pdf

The Welsh Language Standards (No.7) Regulations 2018. (2018). Retrieved June 23, 2023, from https://www.legislation.gov.uk/wsi/2018/441/made

Thiery, C. (1978). True Bilingualism and Second-language Learning. In D, Gerver, & H. W. Sinaiko (Eds.), *Language Interpretation and Communication* (pp. 145-153). Boston: Springer

Thomas, A. (1982). Change and decay in language. In C. David (Ed.), *Linguistic controversies: Essays in linguistic theory and practice in honour of F. R. Palmer* (pp. 209-219). London, UK: Edward Arnold.

Thomas, A. R. (1992). The Welsh language. In M. MacAulay, *The Celtic Languages* (pp. 251-345). Cambridge: Cambridge University Press.

Thomas, M. S., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language, 43*(1), 44-66. https://doi.org/10.1006/jmla.1999.2700

Thomas, M. S., & Van Heuven, W. J. (2005). Computational models of bilingual comprehension. In J. Kroll, & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 202-225). Oxford University Press.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation.

Proceedings of the National Academy of Sciences, 94(26), 14792-14797. https://doi.org/10.1073/pnas.94.26.14792

Tsatali, M., Fotiadou, F., Giaglis, G., & Tsolaki, M. (2019). The repeatable battery for the assessment of the neuropsychological status (RBANS): a diagnostic validity study in Greek elderly. *Aging Clinical and Experimental Research, 31*(9), 1305-1312. https://doi.org/10.1007/s40520-018-1076-9

Tucker, A. M., & Stern, Y. (2011). Cognitive reserve in aging. *Current Alzheimer Research, 8*(4), 354-360. https://doi.org/10.2174/156720511795745320

United Nations. (2015). World Population Ageing. Retrieved June 23, 2023, from https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf

Urla, J. (2012). *Reclaiming Basque: Language, nation, and cultural activism.* University of Nevada Press. https://doi.org/10.1111/amet.12070_25

Van Den Noort, M., Struys, E., & Bosch, P. (2019). Individual variation and the bilingual advantage—Factors that modulate the effect of bilingualism on cognitive control and cognitive reserve. *Behavioral Sciences, 9*(12), 120. https://doi.org/10.3390/bs9120120

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458-483. https://doi.org/10.1006/jmla.1998.2584

Van Heuven, W. J., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral cortex, 18*(11), 2706-2716. https://doi.org/10.1093/cercor/bhn030

Vance, D. E., & Crowe, M. (2006). A proposed model of neuroplasticity and cognitive reserve in older adults. *Activities, Adaptation & Aging, 30*(3), 61-79. https://doi.org/10.1300/J016v30n03_04

Vance, D. E., Roberson, A. J., McGuinness, T. M., & Fazeli, P. L. (2010). How neuroplasticity and cognitive reserve protect cognitive functioning. *Journal of Psychosocial Nursing & Mental Health Services, 48*(4), 23-30. https://doi.org/10.3928/02793695-20100302-01

Velayudhan, L., Ryu, S. H., Raczek, M., Philpot, M., Lindesay, J., Critchfield, M., & Livingston, G. (2014). Review of brief cognitive tests for patients with suspected dementia. *International Psychogeriatrics, 26*(8), 1247-1262. https://doi.org/10.1017/s1041610214000416

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: a meta-analysis. *Psychology and Aging, 18*(3), 443-460. https://doi.org/10.1037/0882-7974.18.3.443

Verheyen, S., Hampton, J. A., & Storms, G. (2010). A probabilistic threshold model: Analyzing semantic categorization data with the Rasch model. *Acta Psychologica, 135*(2), 216-225. https://doi.org/10.1016/j.actpsy.2010.07.002

Verheyen, S., White, A., & Égré, P. (2019). Revealing criterial vagueness in inconsistencies. *Open Mind: Discoveries in Cognitive Science, 3*, 41-51. https://doi.org/10.1162/opmi_a_00025

Verreyt, N., Woumans, E. V. Y., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition, 19*(1), 181-190. DOI: https://doi.org/10.1017/S1366728914000352

Von Der Heide, R. J., Skipper, L. M., Klobusicky, E., & Olson, I. R. (2013). Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis. *Brain, 136*(6), 1692-1707. https://doi.org/10.1093/brain/awt094

Wadsworth, H. E., Dhima, K., Womack, K. B., Hart Jr, J., Weiner, M. F., Hynan, L. S., & Cullum, C. M. (2018). Validity of teleneuropsychological assessment in older patients with cognitive disorders. *Archives of Clinical Neuropsychology, 33*(8), 1040-1045. https://doi.org/10.1093/arclin/acx140

Wald, B. (1974). Bilingualism. *Annual Review of Anthropology, 3*(1), 301-321. https://doi.org/10.1146/annurev.an.03.100174.001505

Walker, A. J., Batchelor, J., Shores, E. A., & Jones, M. (2010). Effects of cultural background on WAIS-III and WMS-III performances after moderate-severe traumatic brain injury. *Australian Psychologist, 4*5(2), 112-122. https://doi.org/10.1080/00050060903428210

Wang, B. R., Zheng, H. F., Xu, C., Sun, Y., Zhang, Y. D., & Shi, J. Q. (2019). Comparative diagnostic accuracy of ACE-III and MoCA for detecting mild cognitive impairment. *Neuropsychiatric Disease and Treatment, 15*, 2647-2653. https://doi.org/10.2147/NDT.S212328

Weinreich, U. (1953). *Languages in contact: Findings and problems.* New York: Linguistic Circle.

Welsh Government. (2015). Health and Care Standards. Retrieved June 23, 2023, from http://www.wales.nhs.uk/sitesplus/documents/1064/24729_Health%20Standards%20Framework_2015_E1.pdf

Welsh Government. (2016). *More than just words.* Retrieved 2022, from https://nworth-ctu.bangor.ac.uk/documents/morethanjustwordsen.pdf

Welsh Government. (2022). *Welsh Language in Wales (Census 2021)*. Retrieved 2022, from https://www.gov.wales/sites/default/files/pdf-versions/2022/12/3/1671609478/welsh-language-wales-census-2021.pdf

Welsh Language Commissioner. (2012). *My Language, My Health: The Welsh Language Commissioner's Inquiry into the Welsh Language in Primary Care.* Retrieved 2022,

from http://www.comisiynyddygymraeg.cymru/English/Publications%20List/Health%20inquiry%20full%20report.pdf

Welsh Language Health Measures. (n.d.). LLAIS: Language Awareness Infrastructure Support. Retrieved 2022 from https://nworth-ctu.bangor.ac.uk/llais/mesurau-iechyd.php.en

Whalley, L. J., Deary, I. J., Appleton, C. L., & Starr, J. M. (2004). Cognitive reserve and the neurobiology of cognitive aging. *Ageing Research Reviews, 3*(4), 369-382. https://doi.org/10.1016/j.arr.2004.05.001

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex, 21*(5), 1066-1075. https://doi.org/10.1093/cercor/bhq180

Williams, C. H. (2017). Policy review: Wake me up in 2050! Formulating language policy in Wales. Retrieved 2022 from http://www.meits.org/policy-papers/paper/wake-me-up-in-2050-formulating-language-policy-in-wales

Wilson, R. S., Yu, L., Lamar, M., Schneider, J. A., Boyle, P. A., & Bennett, D. A. (2019). Education and cognitive reserve in old age. *Neurology, 92(*10), e1041-e1050. https://doi.org/10.1212%2FWNL.0000000000007036

Wittenberg, R., Hu, B., Jagger, C., Kingston, A., Knapp, M., Comas-Herrera, A., King, D., Rehill, A., & Banerjee, S. (2020). Projections of care for older people with dementia in England: 2015 to 2040. *Age and Ageing, 49*(2), 264-269. https://doi.org/10.1093/ageing/afz154

Woumans, E. V. Y., Santens, P., Sieben, A., Versijpt, J. A. N., Stevens, M., & Duyck, W. (2015). Bilingualism delays clinical manifestation of Alzheimer's disease.

*Bilingualism: Language and Cognition, 1*8(3), 568-574.0. doi:10.1017/S136672891400087X

Wu, W., & Hoffman, P. (2022). Validated measures of semantic knowledge and semantic control: Normative data from young and older adults for more than 300 semantic judgements. *Royal Society Open Science, 9*(2), 211056. https://doi.org/10.1098/rsos.211056

Yee, E., Chrysikou, E. G., & Thompson-Schill, S. L. (2014). Semantic memory. In K. Ochsner, & S. Kosslyn (Eds.), T*he Oxford Handbook of Cognitive Neuroscience, Volume 1: Core Topics* (pp. 353-374). Oxford University Press.

Yegla, B., & Foster, T. C. (2022). Operationally defining cognitive reserve genes. *Neurobiology of Aging, 110*, 96-105. https://doi.org/10.1016/j.neurobiolaging.2021.08.015

Yeh, W., & Barsalou, L. W. (2006). The situated nature of concepts. *The American Journal of Psychology, 119*(3), 349-384. https://doi.org/10.2307/20445349

Yeung, C. M., John, P. D. S., Menec, V., & Tyas, S. L. (2014). Is bilingualism associated with a lower risk of dementia in community-living older adults? Cross-sectional and prospective analyses. *Alzheimer Disease & Associated Disorders, 28*(4), 326-332. https://doi.org/10.1097/wad.0000000000000019

Yoshida, H., Terada, S., Honda, H., Ata, T., Takeda, N., Kishimoto, Y., Etsuko, O., Takeshi, I., & Kuroda, S. (2011). Validation of Addenbrooke's cognitive examination for detecting early dementia in a Japanese population. *Psychiatry Research, 185*(1-2), 211-214. https://doi.org/10.1016/j.psychres.2009.06.012

Zadeh, L. A. (1996). Fuzzy sets. In G. J. Zadeh, & B. Yuan (Eds.), *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh. Advances in Fuzzy Systems –*

Applications and Theory: Volume 6 (pp. 394-432). https://doi.org/10.1142/9789814261302_0021

Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., & Manly, J. J. (2014). Bilingualism does not alter cognitive decline or dementia risk among Spanish-speaking immigrants. *Neuropsychology, 28*(2), 238-246. https://doi.org/10.1037/neu0000014

Zaval, L., Li, Y., Johnson, E. J., & Weber, E. U. (2015). Complementary contributions of fluid and crystallized intelligence to decision making across the life span. In T. M. Hess, J. Strough, & C. E. Löckenhoff (Eds.), *Aging and decision making: Empirical and applied perspectives* (pp. 149–168). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-417148-0.00008-X

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*(1), 37-63. https://doi.org/10.1016/S0885-2014(96)90027-1

Zhang, B. H., Tan, Y. L., Zhang, W. F., Wang, Z. R., Yang, G. G., Shi, C., Zhang, X. Y., & Zhou, D. F. (2008). Repeatable battery for the assessment of neuropsychological status as a screening test in Chinese: reliability and validity. *Chinese Mental Health Journal, 22*(12), 865-869.

Zhao, L., & Ye, C. (2020). Time and performance in online learning: Applying the theoretical perspective of metacognition. *Decision Sciences Journal of Innovative Education, 18*(3), 435-455. https://doi.org/10.1111/dsji.12216

Zhou, Z., Wang, P., & Fang, Y. (2018). Social engagement and its change are associated with dementia risk among Chinese older adults: a longitudinal study. *Scientific Reports, 8*(1), 1-7. https://doi.org/10.1038/s41598-017-17879-w

Zied, K. M., Phillipe, A., Karine, P., Valerie, H. T., Ghislaine, A., & Arnaud, R. (2004). Bilingualism and adult differences in inhibitory mechanisms: Evidence from a

bilingual Stroop task. *Brain and Cognition, 54*(3), 254-256. https://doi.org/10.1016/j.bandc.2004.02.036

Zuazo, K. (1995). The Basque Country and the Basque Language: An overview of the external history of the Basque. In J. I. Hualde, J. A. Lakarra, & R. L. Trask (Eds.), *Current Issues in Linguistic Theory [131]: Towards a history of the Basque language* (pp. 5-30). https://doi.org/10.1075/cilt.131.02zua

Appendices

Appendix A

Demographic Questionnaire

PNum:

Demographic Questionnaire

1) Please state your age:

---

2) Gender:

- □ Male
- □ Female
- □ Other
- □ Prefer not to specify

3) Nationality:

---

4) Handedness:

- □ Right
- □ Left

5) Employment:

- Employed
- Unemployed
- Self-Employed
- Retired
- Student
- Volunteering
- Other

6) What is your highest completed level of education?

- Primary school
- GCSE / O Level
- A Level
- Diploma
- Bachelor’s Degree
- Post-graduate Degree (Master’s, MPhil, Phd)
- Not applicable
- Other

7) What does/did your mother do for a living?

8) What does/did your father do for a living?

---

9) What does/did your caregiver do for a living? (OPTIONAL, if you had an additional caregiver other than your mother and/or father):

---

10) What is/was your mother's highest completed level of education?

- □ Primary School
- □ GCSE / O Level
- □ A Level
- □ Bachelor's Degree
- □ Post-graduate Degree (Master's, MPhil, PhD)
- □ Not applicable
- □ Other

11) What is/was your father's highest completed level of education?

- □ Primary School

- GCSE / O Level
- A Level
- Bachelor’s Degree
- Post-graduate Degree (Master’s, MPhil, PhD)
- Not applicable
- Other

12) What is/was your caregiver’s highest completed level of education?

(OPTIONAL)

- Primary School
- GCSE / O Level
- A Level
- Bachelor’s Degree
- Post-graduate Degree (Master’s, MPhil, PhD)
- Not applicable
- Other

13) Do you consider yourself a social person?

- Yes

- No

14) How often do you go out for social activities?

- On special occasions
- Once per month
- Two or three times a month
- Two or three times a week
- Most days

15) Do you take part in any sort of leisure activity?

- Hardly ever
- On holidays
- Few times per month
- Once a week
- Few times a week
- At least once every day
- Every day, several times a day

Appendix B

Language Proficiency Questionnaire

PNum:

Language Proficiency Questionnaire

Thank you for volunteering to participate in this research.

The questions below seek to determine how proficient you are in the languages you speak.

1) What is/are your native language/s?

   - English
   - Welsh
   - Both
   - Other

2) What language(s) do/did your parents/guardians speak with each other?

---

3) What language(s) do/did you speak with your parents/guardians?

---

4) What language do you use mostly when speaking to your siblings?

---

5) In what language(s) were your classes taught in primary school?

---

6) In what language(s) were your classes taught in secondary school?

---

7) Please state any foreign languages you were taught in school

---

8) What language(s) do/did you use with friends?

---

9) In what language are the majority of channels that you watch on TV?

---

10) What language(s) do you use at home?

---

11) Please rate your ability in English by ticking the appropriate boxes:

	Not at all	**Poor**	**Average**	**Good**	**Very good**	**Fluent**
Speaking						
Listening						
Reading						
Writing						

12) At what age did you learn English?

---

13) Please rate your ability in Welsh by ticking the appropriate boxes:

	Not at all	**Poor**	**Average**	**Good**	**Very good**	**Fluent**
Speaking						
Listening						
Reading						
Writing						

14) At what age did you learn Welsh?

---

15) How frequently do you use the Welsh language?

- □ Hardly ever
- □ On holidays
- □ Few times per month
- □ Once a week
- □ Few times a week
- □ At least once every day
- □ Every day, several times a day

16) Please select your level of education in Welsh (please tick all that are appropriate):

- □ Did not attend formal education in this language
- □ Primary school or equivalent
- □ GSCE or equivalent
- □ A-level or equivalent
- □ College or equivalent

- University degree
- Master's degree
- PhD

17) In which language do you feel more confident?

---

18) Other than English and Welsh, how many additional languages do you speak? (If none, please select 0):

- 0
- 1
- 2
- 3+

If you selected "0" above, do not continue.

Other than English and Welsh, please give details for your ***third most proficient*** language:

19) What is this language?

---

20) At what age did you learn this language?

---

21) Please rate your ability in this language by ticking the appropriate boxes:

	Not at all	**Poor**	**Average**	**Good**	**Very good**	**Fluent**
Speaking						
Listening						
Reading						
Writing						

22) How frequently do you use this language?

- □ Hardly ever
- □ On holidays
- □ Few times per month
- □ Once a week
- □ Few times a week
- □ At least once every day

□ Every day, several times a day

23) Please select your level of education in this language (please tick all that are appropriate):

□ Did not attend formal education in this language

□ Primary school or equivalent

□ Secondary school or equivalent

□ College or equivalent

□ University degree

□ Master's degree

□ PhD

Other than English and Welsh, please give details for your ***second most proficient*** language:

24)What is this language?

_____

25) At what age did you learn this language?

---

26) Please rate your ability in this language by ticking the appropriate boxes:

	Not at all	**Poor**	**Average**	**Good**	**Very good**	**Fluent**
Speaking						
Listening						
Reading						
Writing						

27) How frequently do you use this language?

- □ Hardly ever
- □ On holidays
- □ Few times per month
- □ Once a week
- □ Few times a week
- □ At least once every day

□ Every day, several times a day

28) Please select your level of education in this language (please tick all that are appropriate):

□ Did not attend formal education in this language

□ Primary school or equivalent

□ Secondary school or equivalent

□ College or equivalent

□ University degree

□ Master's degree

□ PhD

Other than English and Welsh, please give details for your ***most proficient language***:

29)What is this language?

_____

30) At what age did you learn this language?

_____

31) Please rate your ability in this language by ticking the appropriate boxes:

	Not at all	Poor	Average	Good	Very good	Fluent
Speaking						
Listening						
Reading						
Writing						

32) How frequently do you use this language?

- □ Hardly ever
- □ On holidays
- □ Few times per month
- □ Once a week
- □ Few times a week
- □ At least once every day
- □ Every day, several times a day

33) Please select your level of education in this language (please tick all that are appropriate):

- □ Did not attend formal education in this language
- □ Primary school or equivalent
- □ Secondary school or equivalent
- □ College or equivalent
- □ University degree
- □ Master’s degree
- □ PhD

Appendix C

Participant Information Sheet

PARTICIPANT INFORMATION SHEET (Control Sample)

Study Title – Validating clinical neuropsychological assessments for Welsh speaking populations.

You are being invited to take part in some research. Before you decide whether or not to participate, it is important for you to understand why the research is being conducted and what it will involve.

The first part of the Participant Information Sheet outlines the purpose of the study and what will happen if you take part. The second part presents more detailed information about the conduct of the study. Do ask if anything is unclear.

One of our team will also go through the contents of this information sheet with you to help you decide whether or not you would like to take part and to answer any questions you may have. Please feel free to talk to others about the study if you wish.

What is the purpose of this research?

The main objective of this research is to validate two major cognitive assessments for the purposes of testing Welsh speaking (bilingual) participants living with dementia.

Research conducted in Wales by organisations such as the Welsh Government and the Alzheimer's Disease Society has determined that there is growing desire amongst persons living with dementia (and their carers), to have the option for Welsh speaking people to be tested through the medium of Welsh. Therefore, there is an emerging demand from Welsh speakers to have dementia assessment materials validated in Welsh.
Additionally, research suggests that testing through the medium of English may hamper performance for bilingual versus monolingual English speakers. The Welsh language versions of these assessments have not yet been validated as a clinical tool (this means it is unknown whether the assessments measure what they aim to measure). By testing a large sample of persons living with dementia and healthy older adult control participants on two widely used cognitive clinical assessments at two different time points, we can undertake cross-validation of test materials, enabling clinicians to have flexibility in choice of cognitive assessments for use across different neuropsychological contexts.

Consequently, this research seeks to contribute to the scientific literature by evaluating performance on measures of memory, language and attention via extensive testing across Welsh and English in a population of dementia and healthy older controls. The significance of this work is thus two-fold – on the one hand, upon completion of the study we will have validated two major cognitive assessment batteries widely used in the clinical sphere so that they are available through the medium of Welsh (in line with the wishes of many persons living with dementia and their carers) and on the other, we will have a new experimental data-set that will demonstrate the consequences of bilingual proficiency across cognitive domains in healthy old adults and dementia populations.

Who is carrying out the research?

The data are being collected by Hanna Thomas, a PhD candidate from the Department of Psychology, College of Human and Health Sciences, Swansea University, supervised by Professor Jeremy Tree, Professor Claire Williams and Dr David Playfoot, Department of Psychology, College of Human and Health Sciences, Swansea University.

What happens if I agree to take part?

You will be asked to provide your written consent prior to taking part. If you consent to participate, you will be asked to undertake a series of cognitive tests (total testing time is likely to span across 2 sessions, 2 to 3 months apart).

In session 1 you will be asked to complete the Welsh or English version of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) and Welsh or English version of the Addenbrookes Cognitive Examination-III (ACE-III), plus the following:

- A demographic questionnaire to collect information such as age and gender.
- Language proficiency test (1): a subjective measure of language skill which involves filling in a questionnaire.
- Language proficiency test (2) – in English or in Welsh: an objective measure of language skill that involves undertaking a vocabulary and proficiency assessment.
- Welsh or English verbal fluency tasks.

In session 2, you will be asked to complete the Welsh or English version of the RBANS (whichever was not completed in the previous session), and the Welsh or English version of the ACE-III (whichever was not completed in the previous session). You will also be asked to complete the Welsh or English version of the language proficiency test (2) and some Welsh and English word verbal fluency tasks.

Please note that if you consent to participate, the testing sessions will either take place at your home, at the university, or via a video-call (e.g. Zoom). In addition, if you would prefer to have a family member (or similar) present during the testing sessions, this is at your discretion. Both testing sessions will be administered by PhD student Hanna Thomas.

Are there any risks and benefits associated with taking part?

The research has been approved by the College of Human and Health Sciences Research Ethics Committee and is funded by the Coleg Cymraeg Cenedlaethol. There are no significant risks associated with participation. There are no definite benefits to taking part although the findings could lead to Welsh speaking (bilingual) individuals having the option of being assessed using validated Welsh cognitive assessments in the future. The findings will also help clinicians to have flexibility in choice of cognitive assessments for use across different neuropsychological contexts. In addition your contribution to such work helps clinicians and researchers to better understand how language can impact performance on cognitive assessments.

Incidental Findings Care Pathway (Health Related findings)

There may be a small likelihood of uncovering health related findings during the course of the cognitive assessment. Therefore, the research study may reveal health related or incidental findings of which you were unaware of (clinically significant or otherwise).

For this reason we will ask you for consent to contact your GP so the research team can notify and make them aware.

Your GP will be able to advise if any further input is needed.

Data Protection and Confidentiality

Your data will be processed in accordance with the Data Protection Act 2018 and the General Data Protection Regulation 2016 (GDPR). All information about you will be kept strictly confidential. Your data will only be viewed by the research team.

We will create a two-letter acronym (initials) so that we can match up your answers from each of the different time points. This information will be digitally stored so it cannot be connected to you – this ensures that no one external to the study can identify your answers. All information presented in any reports or publications will also be anonymous and unidentifiable.

All electronic data will be stored on a password-protected computer file at the Department of Psychology, Swansea University. All paper records will be stored in a locked filing cabinet at the Department of Psychology, Swansea University. Your consent information will also be stored separately from your responses to minimise risk in the event of a data breach.

All the data we collect will relate to the testing scores that you obtain on the cognitive tests that we are intending to do. The data controller will be the lead researcher, Prof Jeremy Tree. As explained above, all your data will be kept strictly confidential. It may be helpful for you to provide information about your medical history if you wish, but this will be entirely up to you. If you wish, we can also share the information obtained in this research with your GP so they are fully informed.

Swansea University is the sponsor for this study based in the United Kingdom. This means that we are responsible for looking after your information and using it properly. Your rights to access, change or

move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate.

What will happen to the information I provide?

The results of this research will be used as the basis of a future research publication as a case study report – rest assured any data collected regarding you will be kept in strictest confidence. Results may also be presented to interested parties and published in scientific journals and related media - all information presented in any reports or publications will be anonymous and unidentifiable. If you would like to receive information about this possible outcome please feel free to ask.

Is participation voluntary and what if I wish to later withdraw?

Your participation is entirely voluntary – you do not have to participate if you do not want to. If you decide to participate, but subsequently wish to withdraw during the research study, then you are free to withdraw at any time by informing the researcher. Your data will then be removed from the research. However, if you wish to withdraw your data from the research after your participation is complete, you will need to inform us within two-weeks of your participation ending so we can identify your data for removal. It will not be possible to identify and remove your data beyond this point (the two week window) as then it would have been collated, anonymised, and analysed.

Data Protection Privacy Notice

The data controller for this project will be Swansea University. The University Data Protection Officer provides oversight of university activities involving the processing of personal data, and can be contacted at the Vice Chancellors Office.

Your personal data will be processed for the purposes outlined in this information sheet.

Standard ethical procedures will involve you providing your consent to participate in this study by completing the consent form that has been provided to you.

The legal basis that we will rely on to process your personal data will be processing is necessary for the performance of a task carried out in the public interest. This public interest justification is approved by the College of Human and Health Sciences Research Ethics Committee, Swansea University.

The legal basis that we will rely on to process special categories of data will be processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

How long will your information be held?

Data will be preserved and accessible **for a minimum of 10 years after completion of the research.** Records from studies with major health, clinical, social, environmental or heritage importance, novel intervention, or studies which are on-going or controversial should be retained for at least 20 years after

completion of the study. It may be appropriate to keep such study data permanently within the university, a national collection, or as required by the funder's data policy.

What are your rights?

You have a right to access your personal information, to object to the processing of your personal information, to rectify, to erase, to restrict and to port your personal information. Please visit the University Data Protection webpages for further information in relation to your rights. Any requests or objections should be made in writing to the University Data Protection Officer:

University Compliance Officer (FOI/DP)

Vice-Chancellor's Office

Swansea University

Singleton Park, Swansea

SA2 8PP

Email: dataprotection@swansea.ac.uk

How to make a complaint?

If you are unhappy with the way in which your personal data has been processed you may in the first instance contact the University Data Protection Officer using the contact details above. If you remain dissatisfied then you have the right to apply directly to the Information Commissioner for a decision. The Information Commissioner can be contacted at:

Information Commissioner's Office

Wycliffe House

Water Lane

Wilmslow

Cheshire

SK9 5AF

www.ico.org.uk

What if I have other questions?

If you have further questions about this study, please do not hesitate to contact us:

Researchers	**Supervisor**
Hanna Thomas	Professor Jeremy Tree

Department of Psychology
Swansea University
[redacted]

Department of Psychology
Swansea University
[redacted]

Professor Claire Williams
Department of Psychology
Swansea University
[redacted]

Dr David Playfoot
Department of Psychology
Swansea University
[redacted]

Appendix D

Consent Form

Participant Identification Number for this trial:

CONSENT FORM

Title of Project: Validating clinical neuropsychological assessments for Welsh speaking populations.

Name of Researcher: Hanna Thomas

Name of Lead Researcher: Professor Jeremy Tree

Please initial box

1. I confirm that I have read the information sheet dated 8th December 2020 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw during and up to two weeks after participating, without giving any reasons and without my medical care or legal rights being affected. ☐
3. I understand that relevant sections of my data collected during the study may be looked at by the research team from Swansea University. I give permission for the research team to have access to my records. ☐
4. I understand that the information collected about me may be used to support other research in the future, and may be shared anonymously with other researchers in academic papers and other formal research outputs. ☐
5. I agree to my General Practitioner being informed of my participation in the study if necessary. ☐
6. I agree to take part in the above study. ☐

Name of Participant	Date	Signature
Name of person taking consent (researcher)	Date	Signature

I WOULD like to participate in future research ☐

I WOULD NOT like to participate in future research ☐

Appendix E

RBANS ANOVA Table

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η²
Language	202.766	1	202.766	4.831	0.031	0.008
Language ✻ S1_order	1826.701	3	608.900	14.508	< .001	0.069
Residuals	3735.246	89	41.969			

Note. Type III Sum of Squares

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η²
S1_order	990.911	3	330.304	1.496	0.221	0.038
Residuals	19656.540	89	220.860			

Note. Type III Sum of Squares

Test for Equality of Variances (Levene's)

	F	df1	df2	p
RBANS_Total_Scale_W	1.411	3	89	0.245
RBANS_Total_Scale_E	1.533	3	89	0.211

Appendix F

RBANS Post-hoc Pairwise Comparisons

Post Hoc Comparisons - S1_order ✻ Language ▼

		Mean Difference	SE	t	p_{bonf}
1, Welsh	2, Welsh	−4.894	3.384	−1.446	1.000
	3, Welsh	4.648	3.384	1.374	1.000
	4, Welsh	7.794	3.419	2.280	0.682
	1, English	4.136	1.953	2.118	1.000
	2, English	−0.686	3.384	−0.203	1.000
	3, English	−3.186	3.384	−0.941	1.000
	4, English	−1.075	3.419	−0.314	1.000
2, Welsh	3, Welsh	9.542	3.309	2.883	0.130
	4, Welsh	12.688	3.345	3.793	0.007**
	1, English	9.030	3.384	2.669	0.242
	2, English	4.208	1.870	2.250	0.753
	3, English	1.708	3.309	0.516	1.000
	4, English	3.819	3.345	1.142	1.000
3, Welsh	4, Welsh	3.147	3.345	0.941	1.000
	1, English	−0.511	3.384	−0.151	1.000
	2, English	−5.333	3.309	−1.612	1.000
	3, English	−7.833	1.870	−4.189	0.002**
	4, English	−5.723	3.345	−1.711	1.000
4, Welsh	1, English	−3.658	3.419	−1.070	1.000
	2, English	−8.480	3.345	−2.535	0.350
	3, English	−10.980	3.345	−3.282	0.038*
	4, English	−8.870	1.910	−4.643	< .001***
1, English	2, English	−4.822	3.384	−1.425	1.000
	3, English	−7.322	3.384	−2.164	0.908
	4, English	−5.211	3.419	−1.524	1.000
2, English	3, English	−2.500	3.309	−0.755	1.000
	4, English	−0.389	3.345	−0.116	1.000
3, English	4, English	2.111	3.345	0.631	1.000

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

Note. P-value adjusted for comparing a family of 28

Appendix G

ACE-III ANOVA Table

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
Language	43.147	1	43.147	7.541	0.007	0.012
Language ✻ S1_order	145.253	3	48.418	8.462	< .001	0.040
Residuals	509.242	89	5.722			

Note. Type III Sum of Squares

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
S1_order	316.960	3	105.653	3.585	0.017	0.087
Residuals	2622.631	89	29.468			

Note. Type III Sum of Squares

Test for Equality of Variances (Levene's)

	F	df1	df2	p
ACE_Total_W	3.460	3	89	0.020
ACE_Total_E	1.628	3	89	0.189

Appendix H

ACE-III Post-hoc Pairwise Comparisons

Post Hoc Comparisons - S1_order ✻ Language ▼

		Mean Difference	SE	t	p_{bonf}
1, Welsh	2, Welsh	−2.114	1.238	−1.707	1.000
	3, Welsh	0.678	1.238	0.548	1.000
	4, Welsh	2.984	1.251	2.386	0.520
	1, English	1.364	0.721	1.891	1.000
	2, English	−2.072	1.238	−1.674	1.000
	3, English	−1.322	1.238	−1.068	1.000
	4, English	−0.277	1.251	−0.221	1.000
2, Welsh	3, Welsh	2.792	1.211	2.305	0.639
	4, Welsh	5.098	1.224	4.165	0.002**
	1, English	3.477	1.238	2.809	0.162
	2, English	0.042	0.691	0.060	1.000
	3, English	0.792	1.211	0.654	1.000
	4, English	1.837	1.224	1.501	1.000
3, Welsh	4, Welsh	2.306	1.224	1.884	1.000
	1, English	0.686	1.238	0.554	1.000
	2, English	−2.750	1.211	−2.271	0.697
	3, English	−2.000	0.691	−2.896	0.133
	4, English	−0.955	1.224	−0.780	1.000
4, Welsh	1, English	−1.621	1.251	−1.296	1.000
	2, English	−5.056	1.224	−4.131	0.002**
	3, English	−4.306	1.224	−3.518	0.017*
	4, English	−3.261	0.705	−4.623	< .001***
1, English	2, English	−3.436	1.238	−2.775	0.179
	3, English	−2.686	1.238	−2.169	0.896
	4, English	−1.640	1.251	−1.311	1.000
2, English	3, English	0.750	1.211	0.619	1.000
	4, English	1.795	1.224	1.467	1.000
3, English	4, English	1.045	1.224	0.854	1.000

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

Note. P-value adjusted for comparing a family of 28

Appendix I

ACE-III Logistic Regression Model with Order

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	87.194	89.194	91.337	62						
H_1	73.474	85.474	98.333	57	13.719	0.017	0.157	0.231	0.197	0.196

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−13.483	8.601	−1.568	2.457	1	0.117
LexTALE_W	0.042	0.055	0.757	0.572	1	0.449
LexTALE_E	0.146	0.125	1.174	1.377	1	0.241
Age_S1	0.064	0.048	1.331	1.771	1	0.183
S1_order (2)	−1.850	0.787	−2.351	5.525	1	0.019
S1_order (3)	−2.486	0.809	−3.073	9.443	1	0.002

Note. ACEPref level '3' coded as class 1.

Appendix J

ACE-III Logistic Regression Model with Education Level

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	87.194	89.194	91.337	62						
H_1	75.613	89.613	104.615	56	11.581	0.072	0.133	0.224	0.161	0.168

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−4.839	7.415	−0.653	0.426	1	0.514
LexTALE_W	0.067	0.059	1.133	1.284	1	0.257
LexTALE_E	0.014	0.108	0.129	0.017	1	0.898
Age_S1	0.044	0.044	1.001	1.002	1	0.317
education_level (3)	−2.256	1.233	−1.830	3.350	1	0.067
education_level (4)	−2.973	1.249	−2.381	5.671	1	0.017
education_level (Not applicable)	14.085	1455.398	0.010	9.366e−5	1	0.992

Note. ACEPref level '3' coded as class 1.

Appendix K

ACE-III Logistic Regression Model with Level of Education in Welsh

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	87.194	89.194	91.337	62						
H_1	81.352	99.352	118.640	54	5.842	0.665	0.067	0.118	0.083	0.089

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	12.043	1455.415	0.008	6.847e−5	1	0.993
LexTALE_W	−0.013	0.058	−0.216	0.047	1	0.829
LexTALE_E	0.008	0.108	0.074	0.006	1	0.941
Age_S1	0.050	0.045	1.119	1.251	1	0.263
Level of education in Welsh (2)	−15.012	1455.398	−0.010	1.064e−4	1	0.992
Level of education in Welsh (3)	−15.801	1455.398	−0.011	1.179e−4	1	0.991
Level of education in Welsh (4)	−15.836	1455.398	−0.011	1.184e−4	1	0.991
Level of education in Welsh (5)	−14.729	1455.398	−0.010	1.024e−4	1	0.992
Level of education in Welsh (6)	−15.546	1455.398	−0.011	1.141e−4	1	0.991

Note. ACEPref level '3' coded as class 1.

Appendix L

ACE-III Logistic Regression Model with Language Currently Used with Family

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	87.194	89.194	91.337	62						
H_1	84.004	96.004	108.863	57	3.190	0.671	0.037	0.056	0.050	0.049

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−1.122	7.074	−0.159	0.025	1	0.874
LexTALE_W	0.010	0.051	0.192	0.037	1	0.848
LexTALE_E	−0.064	0.110	−0.587	0.344	1	0.558
Age_S1	0.051	0.044	1.153	1.330	1	0.249
languagefamilytotal (2)	0.995	0.846	1.176	1.383	1	0.240
languagefamilytotal (3)	1.259	1.245	1.011	1.021	1	0.312

Note. ACEPref level '3' coded as class 1.

Appendix M

ACE-III Logistic Regression Model with Language of the Majority of TV Programmes Watched

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	87.194	89.194	91.337	62						
H_1	85.504	97.504	110.363	57	1.689	0.890	0.019	0.035	0.026	0.026

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−2.490	6.817	−0.365	0.133	1	0.715
LexTALE_W	0.011	0.050	0.211	0.045	1	0.833
LexTALE_E	−0.010	0.103	−0.095	0.009	1	0.924
Age_S1	0.039	0.045	0.882	0.778	1	0.378
Lang majority TV prog (2)	−0.250	0.665	−0.376	0.141	1	0.707
Lang majority TV prog (3)	−0.221	0.877	−0.252	0.064	1	0.801

Note. ACEPref level '3' coded as class 1.

Appendix N

ACE-III Logistic Regression Model with Language most Confident in Using

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	87.194	89.194	91.337	62						
H_1	78.793	92.793	107.795	56	8.400	0.210	0.096	0.137	0.118	0.125

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−4.109	7.288	−0.564	0.318	1	0.573
LexTALE_W	−0.029	0.056	−0.520	0.270	1	0.603
LexTALE_E	0.026	0.109	0.234	0.055	1	0.815
Age_S1	0.065	0.046	1.431	2.047	1	0.153
Lang most confident (2)	−2.038	1.243	−1.640	2.689	1	0.101
Lang most confident (3)	−1.004	0.619	−1.622	2.631	1	0.105
Lang most confident (NO RESPONSE)	15.579	1455.398	0.011	1.146e−4	1	0.991

Note. ACEPref level '3' coded as class 1.

Appendix N1

ACE-III Logistic Regression Models – Including Counterbalancing Order 4

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	111.641	113.641	116.072	83						
H_1	81.701	95.701	112.717	77	29.940	< .001	0.268	0.408	0.307	0.300

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−17.770	8.727	−2.036	4.147	1	0.042
S1_order (2)	−2.045	0.806	−2.538	6.440	1	0.011
S1_order (3)	−2.719	0.845	−3.218	10.352	1	0.001
S1_order (4)	−3.983	1.031	−3.863	14.926	1	< .001
LexTALE_W	0.055	0.052	1.062	1.129	1	0.288
LexTALE_E	0.163	0.126	1.300	1.690	1	0.194
Age_S1	0.106	0.046	2.276	5.182	1	0.023

Note. ACEPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
S1_order	0.643	1.554
LexTALE_W	0.923	1.084
LexTALE_E	0.762	1.312
Age_S1	0.797	1.255

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	111.641	113.641	116.072	83						
H_1	101.965	119.965	141.842	75	9.676	0.288	0.087	0.148	0.107	0.109

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	8.491	1455.412	0.006	3.404e−5	1	0.995
Level of education in Welsh (2)	−16.379	1455.398	−0.011	1.266e−4	1	0.991
Level of education in Welsh (3)	−16.183	1455.398	−0.011	1.236e−4	1	0.991
Level of education in Welsh (4)	−16.338	1455.398	−0.011	1.260e−4	1	0.991
Level of education in Welsh (5)	−15.232	1455.398	−0.010	1.095e−4	1	0.992
Level of education in Welsh (6)	−16.224	1455.398	−0.011	1.243e−4	1	0.991
LexTALE_W	0.027	0.050	0.541	0.293	1	0.588
LexTALE_E	0.024	0.099	0.243	0.059	1	0.808
Age_S1	0.063	0.038	1.658	2.748	1	0.097

Note. ACEPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
Level of education in Welsh	0.601	1.665
LexTALE_W	0.698	1.433
LexTALE_E	0.842	1.187
Age_S1	0.813	1.229

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	111.641	113.641	116.072	83						
H_1	99.900	115.900	135.346	76	11.741	0.109	0.105	0.177	0.124	0.130

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−24.717	1584.804	−0.016	2.433e−4	1	0.988
education_level (2)	15.904	1584.792	0.010	1.007e−4	1	0.992
education_level (3)	15.278	1584.792	0.010	9.263e−5	1	0.992
education_level (4)	14.610	1584.792	0.009	8.498e−5	1	0.993
education_level (Not applicable)	33.037	2875.653	0.011	1.320e−4	1	0.991
LexTALE_W	0.067	0.049	1.365	1.862	1	0.172
LexTALE_E	0.039	0.102	0.389	0.151	1	0.697
Age_S1	0.052	0.037	1.425	2.029	1	0.154

Note. ACEPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
education_level	0.817	1.224
LexTALE_W	0.811	1.232
LexTALE_E	0.970	1.031
Age_S1	0.951	1.051

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	111.641	113.641	116.072	83						
H_1	104.501	116.501	131.086	78	7.140	0.210	0.064	0.111	0.085	0.081

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−5.501	6.212	−0.886	0.784	1	0.376
languagefamilytotal (2)	1.004	0.748	1.343	1.803	1	0.179
languagefamilytotal (3)	1.762	1.220	1.444	2.085	1	0.149
LexTALE_W	0.056	0.041	1.360	1.850	1	0.174
LexTALE_E	−0.033	0.097	−0.341	0.116	1	0.733
Age_S1	0.049	0.035	1.386	1.921	1	0.166

Note. ACEPref level '2' coded as class 1.

Multicollinearity Diagnostics ▼

	Tolerance	VIF
languagefamilytotal	0.856	1.168
LexTALE_W	0.928	1.077
LexTALE_E	0.854	1.170
Age_S1	0.923	1.083

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	111.641	113.641	116.072	83						
H_1	107.079	119.079	133.664	78	4.562	0.472	0.041	0.072	0.053	0.053

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−7.047	6.121	−1.151	1.326	1	0.250
Lang majority TV prog (2)	−0.200	0.610	−0.327	0.107	1	0.743
Lang majority TV prog (3)	−0.290	0.780	−0.372	0.139	1	0.710
LexTALE_W	0.045	0.042	1.084	1.175	1	0.278
LexTALE_E	0.025	0.097	0.258	0.066	1	0.797
Age_S1	0.046	0.037	1.223	1.497	1	0.221

Note. ACEPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
Lang majority TV prog	0.737	1.358
LexTALE_W	0.919	1.088
LexTALE_E	0.877	1.141
Age_S1	0.799	1.251

Model Summary - ACEPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	111.641	113.641	116.072	83						
H_1	101.999	115.999	133.015	77	9.641	0.141	0.086	0.147	0.103	0.108

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−9.122	6.479	−1.408	1.982	1	0.159
Lang most confident (2)	−1.521	1.217	−1.249	1.561	1	0.212
Lang most confident (3)	−0.658	0.558	−1.180	1.391	1	0.238
Lang most confident (NO RESPONSE)	16.194	1455.398	0.011	1.238e−4	1	0.991
LexTALE_W	0.020	0.046	0.443	0.196	1	0.658
LexTALE_E	0.059	0.102	0.581	0.337	1	0.561
Age_S1	0.066	0.037	1.796	3.225	1	0.073

Note. ACEPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
Lang most confident	0.777	1.287
LexTALE_W	0.797	1.255
LexTALE_E	0.826	1.211
Age_S1	0.888	1.126

Appendix O

RBANS Logistic Regression Model with Order

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	92.747	94.747	96.952	66						
H_1	75.396	87.396	100.625	61	17.351	0.004	0.187	0.304	0.240	0.228

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	2.248	8.177	0.275	0.076	1	0.783
S1_order (2)	0.129	0.697	0.185	0.034	1	0.853
S1_order (3)	−2.190	0.754	−2.903	8.428	1	0.004
LexTALE_W	0.008	0.054	0.144	0.021	1	0.885
LexTALE_E	−0.076	0.126	−0.601	0.361	1	0.548
Age_S1	0.037	0.046	0.816	0.666	1	0.414

Note. RBANSPref level '3' coded as class 1.

Appendix P

RBANS Logistic Regression Model with Education Level

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	92.747	94.747	96.952	66						
H_1	88.255	102.255	117.688	60	4.492	0.610	0.048	0.087	0.061	0.065

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	2.659	7.056	0.377	0.142	1	0.706
LexTALE_W	0.021	0.053	0.390	0.152	1	0.696
LexTALE_E	−0.094	0.109	−0.861	0.742	1	0.389
Age_S1	0.037	0.040	0.914	0.836	1	0.361
education_level (3)	−0.543	0.867	−0.627	0.393	1	0.531
education_level (4)	−0.865	0.869	−0.996	0.992	1	0.319
education_level (Not applicable)	14.743	1455.398	0.010	1.027e−4	1	0.992

Note. RBANSPref level '3' coded as class 1.

Appendix Q

RBANS Logistic Regression Model with Level of Education in Welsh

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	92.747	94.747	96.952	66						
H_1	87.470	105.470	125.312	58	5.277	0.728	0.057	0.101	0.069	0.076

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	21.517	1455.417	0.015	2.186e−4	1	0.988
LexTALE_W	0.034	0.055	0.608	0.367	1	0.545
LexTALE_E	−0.165	0.121	−1.366	1.865	1	0.172
Age_S1	0.035	0.042	0.835	0.698	1	0.404
Level of education in Welsh (2)	−16.622	1455.398	−0.011	1.304e−4	1	0.991
Level of education in Welsh (3)	−15.590	1455.398	−0.011	1.147e−4	1	0.991
Level of education in Welsh (4)	−15.893	1455.398	−0.011	1.192e−4	1	0.991
Level of education in Welsh (5)	−16.180	1455.398	−0.011	1.236e−4	1	0.991
Level of education in Welsh (6)	−15.767	1455.398	−0.011	1.174e−4	1	0.991

Note. RBANSPref level '3' coded as class 1.

Appendix R

RBANS Logistic Regression Model with Language Currently Used with Family

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	92.747	94.747	96.952	66						
H_1	88.952	102.952	118.384	60	3.796	0.704	0.041	0.073	0.049	0.055

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	3.506	7.102	0.494	0.244	1	0.622
LexTALE_W	0.005	0.050	0.104	0.011	1	0.917
LexTALE_E	−0.105	0.115	−0.913	0.834	1	0.361
Age_S1	0.037	0.044	0.844	0.712	1	0.399
languagefamilytotal (2)	0.042	0.758	0.055	0.003	1	0.956
languagefamilytotal (3)	−0.551	1.190	−0.463	0.215	1	0.643
languagefamilytotal (4)	−15.212	1455.398	−0.010	1.092e−4	1	0.992

Note. RBANSPref level '3' coded as class 1.

Appendix S

RBANS Logistic Regression Model with Language of the Majority of TV Programmes Watched

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	92.747	94.747	96.952	66						
H_1	87.835	101.835	117.268	60	4.913	0.555	0.053	0.094	0.066	0.071

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	5.053	6.890	0.733	0.538	1	0.463
LexTALE_W	−0.001	0.051	−0.029	8.168e−4	1	0.977
LexTALE_E	−0.090	0.110	−0.818	0.670	1	0.413
Age_S1	0.011	0.044	0.249	0.062	1	0.803
Lang majority TV prog (2)	−0.661	0.664	−0.995	0.991	1	0.320
Lang majority TV prog (3)	0.132	0.886	0.149	0.022	1	0.882
Lang majority TV prog (4)	−15.856	1455.398	−0.011	1.187e−4	1	0.991

Note. RBANSPref level '3' coded as class 1.

Appendix T

RBANS Logistic Regression Model with Language most Confident in Using

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	92.747	94.747	96.952	66						
H_1	84.940	98.940	114.372	60	7.808	0.253	0.084	0.147	0.104	0.110

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	4.212	7.266	0.580	0.336	1	0.562
LexTALE_W	−0.036	0.055	−0.658	0.433	1	0.510
LexTALE_E	−0.090	0.115	−0.784	0.615	1	0.433
Age_S1	0.047	0.042	1.124	1.264	1	0.261
Lang most confident (2)	−2.035	1.238	−1.644	2.703	1	0.100
Lang most confident (3)	−0.452	0.571	−0.792	0.627	1	0.429
Lang most confident (NO RESPONSE)	15.633	1455.398	0.011	1.154e−4	1	0.991

Note. RBANSPref level '3' coded as class 1.

Appendix T1

All RBANS Logistic Regression Models – Including Counterbalancing Order 4

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	122.469	124.469	126.957	88						
H_1	97.436	111.436	128.857	82	25.032	< .001	0.204	0.328	0.261	0.245

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−4.773	6.721	−0.710	0.504	1	0.478
LexTALE_W	0.068	0.045	1.503	2.259	1	0.133
LexTALE_E	0.032	0.104	0.305	0.093	1	0.760
Age_S1	0.008	0.037	0.204	0.042	1	0.838
S1_order (2)	−0.172	0.693	−0.248	0.062	1	0.804
S1_order (3)	−2.492	0.769	−3.241	10.507	1	0.001
S1_order (4)	−2.015	0.713	−2.824	7.978	1	0.005

Note. RBANSPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
LexTALE_W	0.918	1.089
LexTALE_E	0.869	1.150
Age_S1	0.946	1.057
S1_order	0.850	1.176

Model Summary - RBANSPref ▼

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	122.469	124.469	126.957	88						
H_1	112.399	130.399	152.797	80	10.069	0.260	0.082	0.143	0.108	0.107

Coefficients

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−1.394	6.210	−0.224	0.050	1	0.822
LexTALE_W	0.085	0.050	1.698	2.884	1	0.089
LexTALE_E	−0.058	0.096	−0.598	0.358	1	0.550
Age_S1	0.018	0.035	0.516	0.266	1	0.606
Level of education in Welsh (2)	−2.267	1.680	−1.350	1.821	1	0.177
Level of education in Welsh (3)	−0.406	1.479	−0.275	0.075	1	0.784
Level of education in Welsh (4)	−0.190	1.742	−0.109	0.012	1	0.913
Level of education in Welsh (5)	−1.144	1.579	−0.725	0.525	1	0.469
Level of education in Welsh (6)	−0.643	1.655	−0.388	0.151	1	0.698

Note. RBANSPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
LexTALE_W	0.658	1.519
LexTALE_E	0.887	1.127
Age_S1	0.870	1.150
Level of education in Welsh	0.613	1.630

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	122.469	124.469	126.957	88						
H_1	115.478	131.478	151.387	81	6.990	0.430	0.057	0.101	0.065	0.076

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−18.479	1668.733	−0.011	1.226e−4	1	0.991
LexTALE_W	0.065	0.045	1.454	2.113	1	0.146
LexTALE_E	−0.016	0.093	−0.172	0.029	1	0.864
Age_S1	0.008	0.033	0.231	0.053	1	0.817
education_level (2)	15.542	1668.724	0.009	8.674e−5	1	0.993
education_level (3)	15.567	1668.724	0.009	8.703e−5	1	0.993
education_level (4)	15.429	1668.724	0.009	8.548e−5	1	0.993
education_level (Not applicable)	32.798	2922.748	0.011	1.259e−4	1	0.991

Note. RBANSPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
LexTALE_W	0.834	1.199
LexTALE_E	0.968	1.033
Age_S1	0.937	1.067
education_level	0.848	1.180

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	122.469	124.469	126.957	88						
H_1	117.233	131.233	148.653	82	5.236	0.514	0.043	0.076	0.054	0.057

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−2.418	5.734	−0.422	0.178	1	0.673
LexTALE_W	0.075	0.040	1.874	3.513	1	0.061
LexTALE_E	−0.026	0.093	−0.276	0.076	1	0.783
Age_S1	−0.002	0.033	−0.059	0.003	1	0.953
languagefamilytotal (2)	0.255	0.625	0.408	0.167	1	0.683
languagefamilytotal (3)	0.120	1.136	0.106	0.011	1	0.916
languagefamilytotal (4)	−15.581	1455.398	−0.011	1.146e−4	1	0.991

Note. RBANSPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
LexTALE_W	0.904	1.106
LexTALE_E	0.852	1.174
Age_S1	0.915	1.093
languagefamilytotal	0.859	1.164

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	122.469	124.469	126.957	88						
H_1	115.405	129.405	146.826	82	7.063	0.315	0.058	0.102	0.074	0.076

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−2.200	5.604	−0.393	0.154	1	0.695
LexTALE_W	0.065	0.040	1.606	2.579	1	0.108
LexTALE_E	0.014	0.092	0.152	0.023	1	0.879
Age_S1	−0.023	0.036	−0.643	0.413	1	0.521
Lang majority TV prog (2)	−0.681	0.585	−1.163	1.353	1	0.245
Lang majority TV prog (3)	0.119	0.740	0.161	0.026	1	0.872
Lang majority TV prog (4)	−16.398	1455.398	−0.011	1.269e−4	1	0.991

Note. RBANSPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
LexTALE_W	0.897	1.115
LexTALE_E	0.873	1.145
Age_S1	0.787	1.270
Lang majority TV prog	0.738	1.355

Model Summary - RBANSPref

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	122.469	124.469	126.957	88						
H_1	115.155	129.155	146.575	82	7.314	0.293	0.060	0.106	0.070	0.079

Coefficients ▼

				Wald Test		
	Estimate	Standard Error	z	Wald Statistic	df	p
(Intercept)	−3.261	5.741	−0.568	0.323	1	0.570
LexTALE_W	0.044	0.044	1.017	1.035	1	0.309
LexTALE_E	0.008	0.094	0.090	0.008	1	0.928
Age_S1	0.008	0.032	0.251	0.063	1	0.802
Lang most confident (2)	−1.517	1.192	−1.273	1.620	1	0.203
Lang most confident (3)	−0.187	0.497	−0.376	0.141	1	0.707
Lang most confident (NO RESPONSE)	15.746	1455.398	0.011	1.171e−4	1	0.991

Note. RBANSPref level '2' coded as class 1.

Multicollinearity Diagnostics

	Tolerance	VIF
LexTALE_W	0.799	1.251
LexTALE_E	0.838	1.194
Age_S1	0.941	1.062
Lang most confident	0.815	1.228

Appendix U

Participant Information Sheet

PARTICIPANT INFORMATION SHEET

Version number 1.1, dated 02/03/2021

Bilingualism and the Rumsfeld Effect

You are being invited to take part in some research. Before you decide whether or not to participate, it is important for you to understand why the research is being conducted and what it will involve. Please read the following information carefully.

What is the purpose of the research?

We are conducting research on the nature of knowledge, and how we access that knowledge from memory. Research has shown that, for each individual, there are some things which we definitely know to be true and some things which we definitely know to be false. There are also a large number of things which we *sort of* know but are not 100% certain about. This research is trying to examine knowledge that falls into the grey area of uncertainty, and how this uncertainty might change across the lifespan. We are also interested in how speaking more than one language (in this case Welsh and English) might affect the way in which participants are able to organise and access their memory. Your participation in this study will take around 45-55 minutes in the first session, and you will be asked to complete another 5-10 minute online task a week later.

Who is carrying out the research?

The data are being collected by Hanna Thomas (PhD candidate) under the supervision of Professor Jeremy Tree, Dr David Playfoot and Professor Claire Williams in the Department of Psychology at Swansea University. The research has been approved by the departmental Research Ethics Committee.

What happens if I agree to take part?

We will ask you to do a variety of tasks in this study, all of which are related to memory and the way you access the information stored there. One task is designed to assess your knowledge about the truth of a series of statements. You will be asked to indicate whether a statement is true or false and your level of certainty about your answer. *Note: we are not as interested in whether you are correct in your decision about the statement as we are in whether you know what you know.*

Other tasks consider what you know about the meanings of various words, the relationships between these meanings and the way you are able to choose between several options as a response to a question.

We will also ask you to take part in a short Zoom session where you will be asked to complete a series of verbal fluency tasks (you will be asked to name as many words as possible from a particular category in a given time). This session may take place before or after the online experiment, whichever is convenient for you.

Additionally we will ask for some background information including your age and sex. If you are a Welsh-English bilingual, you will be asked to complete an additional measure of your language proficiency.

Are there any risks associated with taking part?
The research has been approved by the Department of Psychology's Research Ethics Committee. There are no significant risks associated with participation.

Data Protection and Confidentiality

Your data will be processed in accordance with the General Data Protection Regulation 2016 (GDPR). All information collected about you will be kept strictly confidential. Your data will only be viewed by the researcher/research team.

All electronic data will be stored on a password-protected computer file in the Department of Psychology, Swansea University. All paper records will be stored in a locked filing cabinet in Dr David Playfoot's office and Hanna Thomas' office. Your consent information will be kept separately from your responses to minimise risk in the event of a data breach.

Please note that the data we will collect for our study will be made anonymous at the end of the experiment. Once the data is anonymised, it will not be possible to identify and remove your data at a later date. Consequently, if you decide to have your data withdrawn, please let us know within two weeks after the follow-up session. If you decide to withdraw after session one but before the follow-up session, you can simply choose not to complete the follow-up session and your data will then be automatically withdrawn from the research. All the data from today will be saved under a unique participant identifier. This will allow us to make the association between your answers today and your answers next week, but it will not be possible to link either to your identity.

What will happen to the information I provide?

An analysis of the information will form part of a PhD thesis (authored by Miss Thomas). In addition, we may present our findings to interested parties and/or publish these in scientific journals and related media. *Note that all information presented in any reports or publications will be anonymous and unidentifiable*.

Is participation voluntary and what if I wish to later withdraw?

Your participation is entirely voluntary – you do not have to participate if you do not want to. If you decide to participate, but later wish to withdraw from the study, you can do so without giving a reason and without penalty but you must notify Hanna Thomas within two weeks after the follow-up session. If you decide to withdraw after session one but before the follow-up session, you can simply choose not to complete the follow-up session and your data will then be automatically withdrawn from the research.

Data Protection Privacy Notice

The data controller for this project will be Swansea University. The University Data Protection Officer provides oversight of university activities involving the processing of personal data, and can be contacted at the Vice Chancellors Office: VCO@swanseauniversity.com. Swansea University's Data Protection Officer is Bev Buckley and she may be contacted at b.y.buckley@swansea.ac.uk**.**

Your personal data will be processed for the purposes outlined in this information sheet.

Standard ethical procedures will involve you providing your consent to participate in this study by completing the consent form that has been provided to you. However, the legal basis on which this task is being performed is public interest, approved by the departmental Research Ethics Committee.

If you are concerned about how your personal data is being processed, please contact Swansea University's Data Protection Officer at b.y.buckley@swansea.ac.uk. Details of your individual rights are available on the ICO website at: https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/individuals-rights/

How long will your information be held?
Data will be preserved and accessible for a minimum of 10 years after completion of the research. Records from studies with major health, clinical, social, environmental or heritage importance, novel intervention, or studies which are on-going or controversial should be retained for at least 20 years after completion of the study. It may be appropriate to keep such study data permanently within the university, a national collection, or as required by the funder's data policy.

What are your rights?
You have a right to access your personal information, to object to the processing of your personal information, to rectify, to erase, to restrict and to port your personal information. Please visit the University Data Protection webpages for further information in relation to your rights. Any requests or objections should be made in writing to the University Data Protection Officer:
University Compliance Officer (FOI/DP)
Vice-Chancellor's Office
Swansea University
Singleton Park, Swansea
SA2 8PP

Email: dataprotection@swansea.ac.uk

How to make a complaint?
If you are unhappy with the way in which your personal data has been processed you may in the first instance contact the University Data Protection Officer using the contact details above. If you remain dissatisfied then you have the right to apply directly to the Information Commissioner for a decision. The Information Commissioner can be contacted at:

Information Commissioner's Office
Wycliffe House
Water Lane
Wilmslow
Cheshire
SK9 5AF
www.ico.org.uk.

What if I have other questions?

If you have further questions about this study, please do not hesitate to contact us:

Hanna Thomas
Department of Psychology
Swansea University
[redacted]

Prof Jeremy Tree
Department of Psychology
Swansea University
[redacted]

Dr David Playfoot
Department of Psychology
Swansea University
[redacted]

Prof Claire Williams
Department of Psychology
Swansea University
c[redacted]

Appendix V

Consent Form

CONSENT FORM

Project Title: Bilingualism and the Rumsfeld Effect
Name and Contact Details of the Researcher: Hanna Thomas [redacted]
Supervisors: Prof Jeremy Tree, Department of Psychology, Swansea University, [redacted] Dr David Playfoot, Department of Psychology, Swansea University, [redacted] Prof Claire Williams, Department of Psychology, Swansea University, [redacted]

1	Rwyf wedi cael copi o'r Daflen Wybodaeth i Gyfranogwyr. I have been provided with a copy of the Participant Information Sheet.	Ydw Yes	Nac Ydw No
2	Rwyf wedi darllen a deall y Daflen Wybodaeth i Gyfranogwyr (fersiwn 1.1, dyddiedig 02/03/2021) I have read and understood the Participant Information Sheet (version number: 1.1, dated: 02/03/2021).	Ydw Yes	Nac Ydw No
3	Rwy'n deall bod fy nghyfranogiad yn wirfoddol a fy mod i'n rhydd i dynnu'n ôl o'r ymchwil am unrhyw reswm a heb ragfarn ond mae'n rhaid i mi hysbysu'r ymchwilydd cyn pen pythefnos ar ôl y sesiwn olaf. Os byddaf yn penderfynu tynnu'n ôl ar ôl sesiwn un ond cyn y sesiwn ddilynol, deallaf y gallaf ddewis peidio â chwblhau'r sesiwn ddilynol ac yna bydd fy nata yn cael ei dynnu'n ôl o'r ymchwil yn awtomatig. I understand that participation is voluntary and also that I am free to withdraw from the research for any reason and without prejudice but I must notify the researcher within two weeks after the final session. If I decide to withdraw after session one but before the follow-up session, I understand that I can simply choose not to complete the follow-up session and my data will then be automatically withdrawn from the research.	Ydw Yes	Nac Ydw No
4	Rwy'n deall beth fydd fy rôl yn yr ymchwil hon, ac mae fy holl gwestiynau wedi'u hateb i'm boddhad. I understand what my role will be in this research, and all my questions have been answered to my satisfaction.	Ydw Yes	Nac Ydw No

5	Rwy'n deall fy mod yn rhydd i ofyn unrhyw gwestiynau ar unrhyw adeg cyn ac yn ystod yr astudiaeth. I understand that I am free to ask any questions at any time before and during the study.	Ydw Yes	Nac Ydw No
6	Fe'm hysbyswyd y bydd y wybodaeth a roddaf yn cael ei diogelu. I have been informed that the information I provide will be safeguarded.	Ydw Yes	Nac Ydw No
7	Rwy'n hapus i'r wybodaeth a roddaf gael ei defnyddio (yn ddienw) mewn papurau academaidd ac allbynnau ymchwil ffurfiol eraill. I am happy for the information I provide to be used (anonymously) in academic papers and other formal research outputs.	Ydw Yes	Nac Ydw No
8	Rwy'n cytuno i'r ymchwilwyr brosesu fy nata personol yn unol â nodau'r astudiaeth a ddisgrifir yn y Daflen Wybodaeth i Gyfranogwyr. I agree to the researchers processing my personal data in accordance with the aims of the study described in the participant information sheet.	Ydw Yes	Nac Ydw No
9	Rwy'n cadarnhau nad wyf yn rhugl mewn unrhyw iaith heblaw Cymraeg (a Saesneg). I confirm that I am not fluent in any language other than English (and Welsh).	Ydw Yes	Nac Ydw No
10	Rwy'n cadarnhau nad oes gen i hanes o / wedi cael diagnosis o salwch niwrolegol neu seiciatrig. I confirm that I do not have a history of/have been diagnosed with a neurological or psychiatric illness.	Ydw Yes	Nac Ydw No
11	Rwy'n cadarnhau nad oes gen i hanes o gam-drin sylweddau. I confirm that I do not have a history of substance abuse.	Ydw Yes	Nac Ydw No

12	Rwy'n deall bod tasg fer arall ar-lein i'w chwblhau wythnos o nawr, ac efallai y bydd yr ymchwilydd yn anfon e-bost i'm hatgoffa'r diwrnod cynt. I understand that there is another short online task to complete a week from now, and that the researcher may email a reminder the day before.	Ydw Yes	Nac Ydw No
13	Rwy'n cytuno i gymryd rhan yn yr astudiaeth. I agree to take part in the study.	Ydw Yes	Nac Ydw No

Diolch i chi am gymryd rhan yn yr astudiaeth. Rydym yn gwerthfawrogi eich help yn fawr.

Thank you for your participation in this study. Your help is very much appreciated.

Name of participant

(Print) …….. Signed ……..…….. Date…….

Appendix W

ANOVA General Cognitive Tasks

Verbal Fluency – Animals

ANOVA - Animals

Cases	Sum of Squares	df	Mean Square	F	p
AgeGroup	44.593	1	44.593	1.360	0.247
Languages	3.006	1	3.006	0.092	0.763
AgeGroup ✻ Languages	1.740	1	1.740	0.053	0.818
Residuals	2886.103	88	32.797		

Note. Type III Sum of Squares

Verbal Fluency – Household Objects

ANOVA - Household_Objects

Cases	Sum of Squares	df	Mean Square	F	p
AgeGroup	5.657	1	5.657	0.230	0.632
Languages	74.268	1	74.268	3.025	0.085
AgeGroup ✻ Languages	69.162	1	69.162	2.817	0.097
Residuals	2160.540	88	24.552		

Note. Type III Sum of Squares

Verbal Fluency – Letter 'F'

ANOVA - Letter_F

Cases	Sum of Squares	df	Mean Square	F	p	η^2
AgeGroup	399.877	1	399.877	13.838	< .001	0.134
Languages	15.240	1	15.240	0.527	0.470	0.005
AgeGroup ✻ Languages	17.085	1	17.085	0.591	0.444	0.006
Residuals	2542.897	88	28.897			

Note. Type III Sum of Squares

Post Hoc Tests

Standard

Post Hoc Comparisons - AgeGroup

		Mean Difference	SE	t	p_{tukey}
Older	Young	4.194	1.127	3.720	< .001

Note. Results are averaged over the levels of: Languages

Eriksen Flanker Task – Incongruent Trials

ANOVA - Flanker_incon

Cases	Sum of Squares	df	Mean Square	F	p
AgeGroup	31.320	1	31.320	2.041	0.157
Languages	19.400	1	19.400	1.264	0.264
AgeGroup ✻ Languages	32.545	1	32.545	2.121	0.149
Residuals	1350.410	88	15.346		

Note. Type III Sum of Squares

Eriksen Flanker Task – Congruent Trials

ANOVA - Flanker_con ▼

Cases	Sum of Squares	df	Mean Square	F	p
AgeGroup	0.003	1	0.003	3.533e−4	0.985
Languages	0.058	1	0.058	0.007	0.935
AgeGroup ✻ Languages	15.463	1	15.463	1.793	0.184
Residuals	759.078	88	8.626		

Note. Type III Sum of Squares

WCST

ANOVA - WCST

Cases	Sum of Squares	df	Mean Square	F	p
AgeGroup	39.210	1	39.210	1.048	0.309
Languages	14.958	1	14.958	0.400	0.529
AgeGroup ✻ Languages	14.890	1	14.890	0.398	0.530
Residuals	3291.538	88	37.404		

Note. Type III Sum of Squares

Spot The Word

ANOVA - Spot_the_Word

Cases	Sum of Squares	df	Mean Square	F	p	η^2
AgeGroup	782.954	1	782.954	39.479	< .001	0.294
Languages	97.013	1	97.013	4.892	0.030	0.036
AgeGroup ✻ Languages	34.532	1	34.532	1.741	0.190	0.013
Residuals	1745.243	88	19.832			

Note. Type III Sum of Squares

Mill Hill

ANOVA - Mill_Hill

Cases	Sum of Squares	df	Mean Square	F	p	η^2
AgeGroup	712.482	1	712.482	38.290	< .001	0.297
Languages	17.591	1	17.591	0.945	0.334	0.007
AgeGroup ✻ Languages	19.104	1	19.104	1.027	0.314	0.008
Residuals	1637.475	88	18.608			

Note. Type III Sum of Squares

Descriptive Statistics

Descriptive Statistics

	Animals	Household_Objects	Letter_F	Flanker_incon	Flanker_con	Flanker_incon_meanRT	Flanker_con_meanRT	WCST	Spot_the_Word	Mill_Hill
Valid	92	92	92	92	92	92	92	92	92	92
Missing	0	0	0	0	0	0	0	0	0	0
Mean	22.315	23.033	13.935	29.185	29.283	1259.371	1046.776	29.217	50.837	19.728
Std. Deviation	5.680	5.048	5.726	3.983	2.918	4723.293	391.997	6.076	5.407	5.136
Skewness	0.037	0.007	0.258	−5.574	−7.125	9.538	2.639	0.162	−0.857	0.103
Std. Error of Skewness	0.251	0.251	0.251	0.251	0.251	0.251	0.251	0.251	0.251	0.251
Kurtosis	−0.090	−0.649	−0.221	32.592	55.850	91.696	13.874	−0.817	0.735	−0.781
Std. Error of Kurtosis	0.498	0.498	0.498	0.498	0.498	0.498	0.498	0.498	0.498	0.498
Shapiro-Wilk	0.993	0.988	0.986	0.255	0.242	0.099	0.814	0.975	0.946	0.978
P-value of Shapiro-Wilk	0.923	0.455	0.444	< .001	< .001	< .001	< .001	0.073	< .001	0.114
Minimum	9.000	12.000	2.000	2.000	5.000	455.516	494.167	17.000	34.000	10.000
Maximum	38.000	35.000	28.000	31.000	30.000	48035.300	3432.647	42.000	60.000	30.000

Appendix X

ANOVA Semantic Control Task – Accuracy

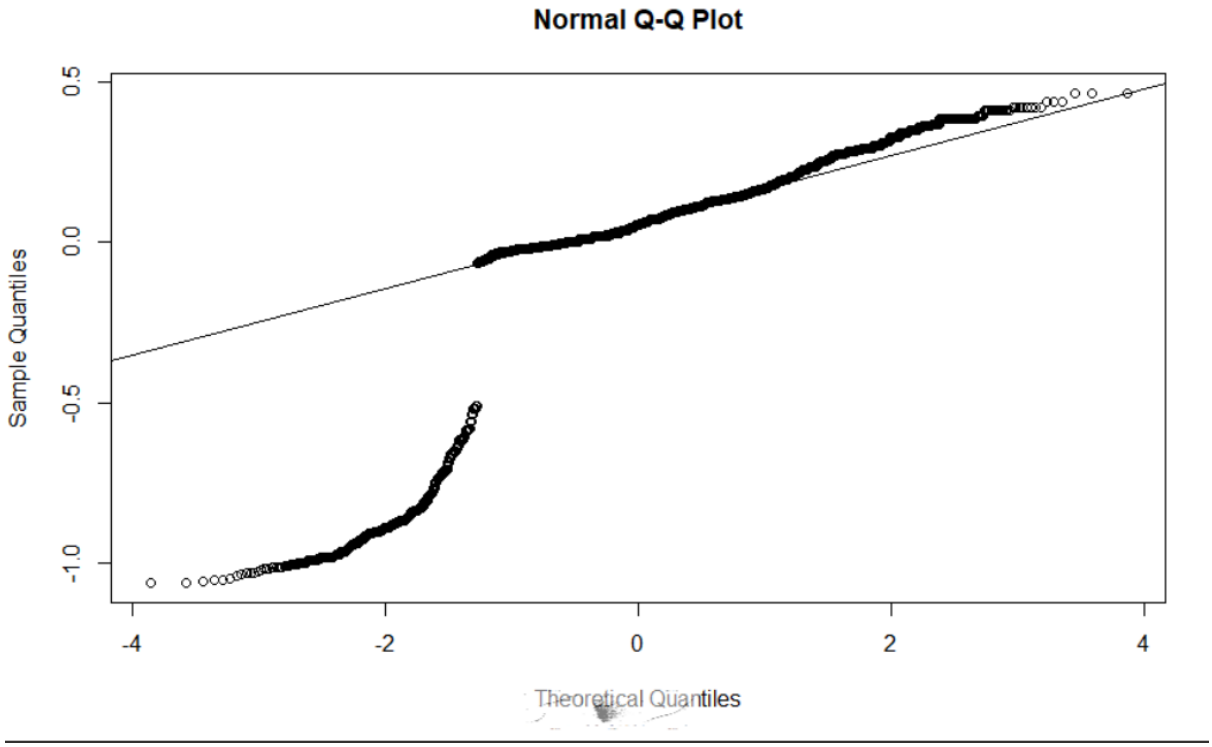
ANOVA Summary

Effect	df	F	p
Task_type	1, 8728.00	150.263	< .001
Semantic_control	1, 8728.00	465.823	< .001
Mono_biling	1, 88.00	0.919	0.340
Young_old	1, 88.00	3.379	0.069
Task_type ✻ Semantic_control	1, 8728.00	116.436	< .001
Task_type ✻ Mono_biling	1, 8728.00	0.703	0.402
Semantic_control ✻ Mono_biling	1, 8728.00	6.208	0.013
Task_type ✻ Young_old	1, 8728.00	91.267	< .001
Semantic_control ✻ Young_old	1, 8728.00	13.911	< .001
Mono_biling ✻ Young_old	1, 88.00	0.714	0.400
Task_type ✻ Semantic_control ✻ Mono_biling	1, 8728.00	2.696	0.101
Task_type ✻ Semantic_control ✻ Young_old	1, 8728.00	68.664	< .001
Task_type ✻ Mono_biling ✻ Young_old	1, 8728.00	0.220	0.639
Semantic_control ✻ Mono_biling ✻ Young_old	1, 8728.00	1.127	0.289
Task_type ✻ Semantic_control ✻ Mono_biling ✻ Young_old	1, 8728.00	8.889	0.003

Note. Model terms tested with Satterthwaite method.
Note. The following variable is used as a random effects grouping factor: 'Participant Public ID'.
Note. Type III Sum of Squares

Appendix X1

Q-Q plot for normality - Accuracy



Appendix Y

ANOVA Semantic Control Task – Reaction Time

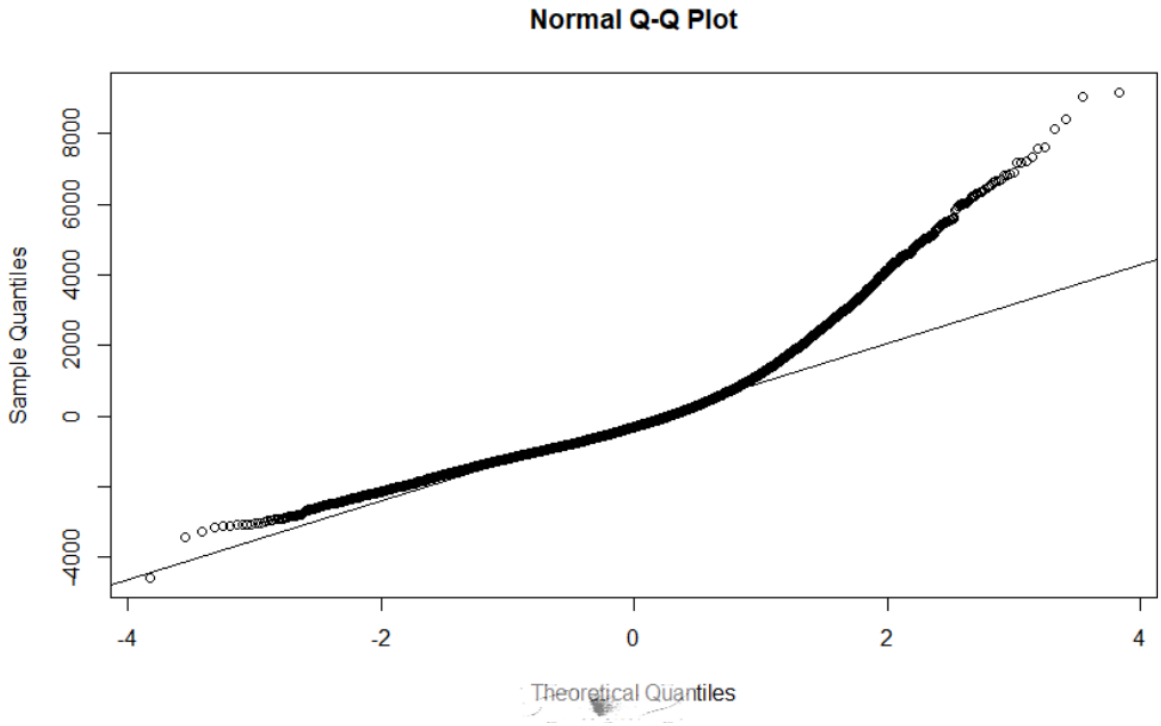
ANOVA Summary

Effect	df	F	p
Task_type	1, 7654.62	590.381	< .001
Semantic_control	1, 7653.05	627.637	< .001
Young_old	1, 88.17	23.250	< .001
Mono_biling	1, 88.17	1.356	0.247
Task_type ✻ Semantic_control	1, 7652.57	12.003	< .001
Task_type ✻ Young_old	1, 7654.62	22.646	< .001
Semantic_control ✻ Young_old	1, 7653.05	55.993	< .001
Task_type ✻ Mono_biling	1, 7654.62	5.290	0.021
Semantic_control ✻ Mono_biling	1, 7653.05	2.176	0.140
Young_old ✻ Mono_biling	1, 88.17	0.076	0.784
Task_type ✻ Semantic_control ✻ Young_old	1, 7652.57	10.495	0.001
Task_type ✻ Semantic_control ✻ Mono_biling	1, 7652.57	0.094	0.760
Task_type ✻ Young_old ✻ Mono_biling	1, 7654.62	4.006e−5	0.995
Semantic_control ✻ Young_old ✻ Mono_biling	1, 7653.05	1.050	0.305
Task_type ✻ Semantic_control ✻ Young_old ✻ Mono_biling	1, 7652.57	3.675e−5	0.995

Note. 191 observations were removed due to missing values.
Note. Model terms tested with Satterthwaite method.
Note. The following variable is used as a random effects grouping factor: 'Participant Public ID'.
Note. Type III Sum of Squares

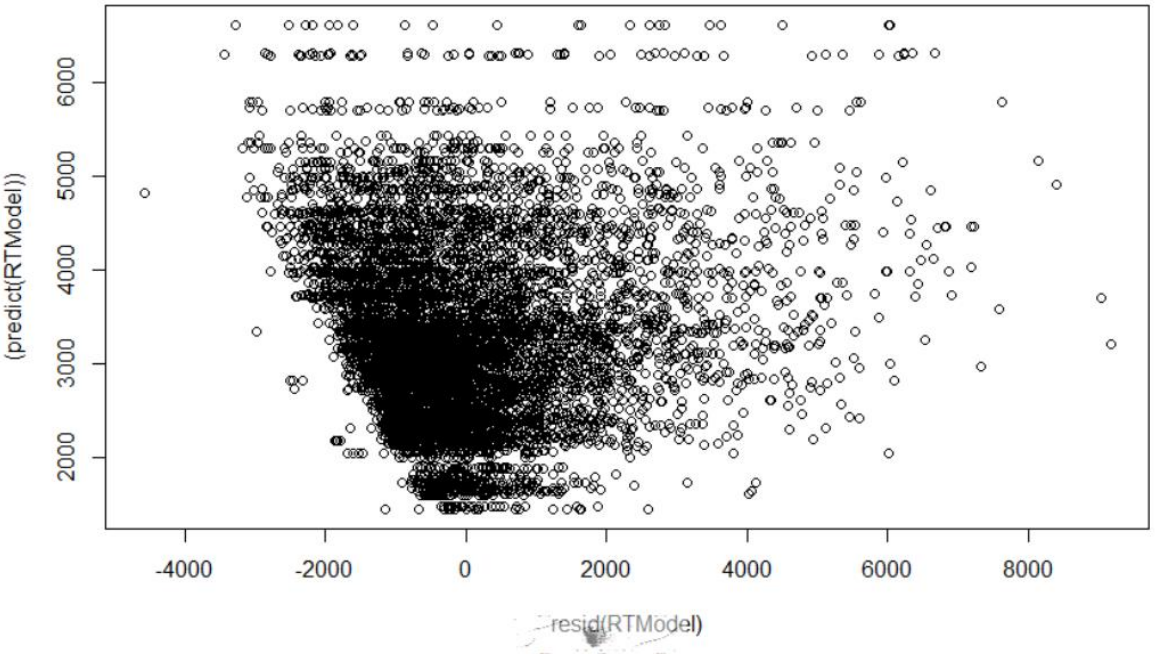
Appendix Y1

Q-Q plot for normality – Reaction Time



Appendix Y2

Homoscedasticity/Heteroscedasticity Plot – Reaction Time



Appendix Z

All Chapter 4, Study 3 Outputs

Descriptives ▼

Interval of Testing	Statement Type	Threshold	Age	Languages	Mean	SD	N
Hour	Category	100% True	Old	Bi	0.906	0.056	22
				Mono	0.920	0.062	22
			Young	Bi	0.916	0.041	20
				Mono	0.906	0.070	23
		True/False (Unsure)	Old	Bi	0.893	0.047	22
				Mono	0.910	0.049	22
			Young	Bi	0.911	0.047	20
				Mono	0.914	0.058	23
		100% False	Old	Bi	0.895	0.058	22
				Mono	0.904	0.053	22
			Young	Bi	0.910	0.055	20
				Mono	0.900	0.067	23
	General Knowledge	100% True	Old	Bi	0.946	0.053	22
				Mono	0.943	0.049	22
			Young	Bi	0.957	0.027	20
				Mono	0.936	0.058	23
		True/False (Unsure)	Old	Bi	0.905	0.057	22
				Mono	0.914	0.039	22
			Young	Bi	0.926	0.052	20
				Mono	0.905	0.072	23
		100% False	Old	Bi	0.896	0.074	22
				Mono	0.927	0.045	22
			Young	Bi	0.933	0.033	20
				Mono	0.934	0.041	23
Week	Category	100% True	Old	Bi	0.850	0.133	22
				Mono	0.899	0.099	22
			Young	Bi	0.880	0.088	20
				Mono	0.882	0.080	23
		True/False (Unsure)	Old	Bi	0.815	0.144	22
				Mono	0.860	0.098	22
			Young	Bi	0.857	0.092	20
				Mono	0.865	0.089	23
		100% False	Old	Bi	0.825	0.130	22
				Mono	0.866	0.085	22
			Young	Bi	0.858	0.065	20
				Mono	0.844	0.083	23
	General Knowledge	100% True	Old	Bi	0.891	0.106	22
				Mono	0.915	0.087	22
			Young	Bi	0.917	0.077	20
				Mono	0.909	0.073	23
		True/False (Unsure)	Old	Bi	0.828	0.126	22
				Mono	0.864	0.074	22
			Young	Bi	0.868	0.109	20
				Mono	0.848	0.093	23
		100% False	Old	Bi	0.860	0.125	22
				Mono	0.886	0.078	22
			Young	Bi	0.881	0.060	20
				Mono	0.895	0.067	23

Within Subjects Effects

Within Subjects Effects

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η^2_p
Interval of Testing	None	0.591	1.000	0.591	29.631	< .001	0.263
Interval of Testing ✻ Age	None	0.001	1.000	0.001	0.068	0.794	8.231e−4
Interval of Testing ✻ Languages	None	0.015	1.000	0.015	0.771	0.383	0.009
Interval of Testing ✻ Age ✻ Languages	None	0.005	1.000	0.005	0.249	0.619	0.003
Residuals	None	1.655	83.000	0.020			
Statement Type	None	0.111	1.000	0.111	26.050	< .001	0.239
Statement Type ✻ Age	None	4.911e−4	1.000	4.911e−4	0.115	0.736	0.001
Statement Type ✻ Languages	None	0.004	1.000	0.004	0.834	0.364	0.010
Statement Type ✻ Age ✻ Languages	None	1.533e−4	1.000	1.533e−4	0.036	0.850	4.316e−4
Residuals	None	0.355	83.000	0.004			
Threshold	None	0.174	2.000	0.087	19.374	< .001	0.189
	Greenhouse-Geisser	0.174	1.604	0.109	19.374	< .001	0.189
Threshold ✻ Age	None	0.004	2.000	0.002	0.483	0.618	0.006
	Greenhouse-Geisser	0.004	1.604	0.003	0.483	0.576	0.006
Threshold ✻ Languages	None	0.002	2.000	0.001	0.224	0.799	0.003
	Greenhouse-Geisser	0.002	1.604	0.001	0.224	0.750	0.003
Threshold ✻ Age ✻ Languages	None	2.798e−4	2.000	1.399e−4	0.031	0.969	3.748e−4
	Greenhouse-Geisser	2.798e−4	1.604	1.745e−4	0.031	0.945	3.748e−4
Residuals	None	0.746	166.000	0.004			
	Greenhouse-Geisser	0.746	133.104	0.006			
Interval of Testing ✻ Statement Type	None	2.141e−4	1.000	2.141e−4	0.118	0.732	0.001
Interval of Testing ✻ Statement Type ✻ Age	None	3.375e−4	1.000	3.375e−4	0.187	0.667	0.002
Interval of Testing ✻ Statement Type ✻ Languages	None	4.965e−4	1.000	4.965e−4	0.274	0.602	0.003
Interval of Testing ✻ Statement Type ✻ Age ✻ Languages	None	0.002	1.000	0.002	0.887	0.349	0.011
Residuals	None	0.150	83.000	0.002			
Interval of Testing ✻ Threshold	None	0.023	2.000	0.011	10.273	< .001	0.110
	Greenhouse-Geisser	0.023	1.800	0.013	10.273	< .001	0.110
Interval of Testing ✻ Threshold ✻ Age	None	0.002	2.000	0.001	0.970	0.381	0.012
	Greenhouse-Geisser	0.002	1.800	0.001	0.970	0.374	0.012
Interval of Testing ✻ Threshold ✻ Languages	None	0.002	2.000	8.519e−4	0.766	0.466	0.009
	Greenhouse-Geisser	0.002	1.800	9.463e−4	0.766	0.454	0.009
Interval of Testing ✻ Threshold ✻ Age ✻ Languages	None	5.806e−4	2.000	2.903e−4	0.261	0.771	0.003
	Greenhouse-Geisser	5.806e−4	1.800	3.224e−4	0.261	0.747	0.003
Residuals	None	0.185	166.000	0.001			
	Greenhouse-Geisser	0.185	149.438	0.001			
Statement Type ✻ Threshold	None	0.038	2.000	0.019	9.525	< .001	0.107
	Greenhouse-Geisser	0.038	1.888	0.020	9.525	< .001	0.107
Statement Type ✻ Threshold ✻ Age	None	0.005	2.000	0.002	1.194	0.305	0.014
	Greenhouse-Geisser	0.005	1.888	0.002	1.194	0.304	0.014
Statement Type ✻ Threshold ✻ Languages	None	0.011	2.000	0.006	2.968	0.054	0.035
	Greenhouse-Geisser	0.011	1.888	0.006	2.968	0.057	0.035
Statement Type ✻ Threshold ✻ Age ✻ Languages	None	0.003	2.000	0.002	0.930	0.400	0.011
	Greenhouse-Geisser	0.003	1.888	0.002	0.930	0.395	0.011
Residuals	None	0.315	166.000	0.002			
	Greenhouse-Geisser	0.315	156.664	0.002			
Interval of Testing ✻ Statement Type ✻ Threshold	None	0.003	2.000	0.002	1.572	0.211	0.020
	Greenhouse-Geisser	0.003	1.997	0.002	1.572	0.211	0.020
Interval of Testing ✻ Statement Type ✻ Threshold ✻ Age	None	1.598e−4	2.000	7.991e−5	0.085	0.918	0.001
	Greenhouse-Geisser	1.598e−4	1.997	8.004e−5	0.085	0.918	0.001
Interval of Testing ✻ Statement Type ✻ Threshold ✻ Languages	None	1.924e−4	2.000	9.620e−5	0.103	0.902	0.001
	Greenhouse-Geisser	1.924e−4	1.997	9.638e−5	0.103	0.902	0.001
Interval of Testing ✻ Statement Type ✻ Threshold ✻ Age ✻ Languages	None	0.002	2.000	0.001	1.313	0.272	0.016
	Greenhouse-Geisser	0.002	1.997	0.001	1.313	0.272	0.016
Residuals	None	0.156	166.000	9.382e−4			
	Greenhouse-Geisser	0.156	165.722	9.398e−4			

Note. Sphericity corrections not available for factors with 2 levels.
Note. Type II Sum of Squares.
▪ Mauchly's test of sphericity indicates that the assumption of sphericity is violated ($p < .05$).

Between Subjects Effects

Between Subjects Effects ▼

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Age	0.025	1	0.025	0.793	0.376	0.009
Languages	0.022	1	0.022	0.705	0.404	0.008
Age ✳ Languages	0.063	1	0.063	2.024	0.159	0.024
Residuals	2.591	83	0.031			

Note. Type III Sum of Squares

Test for Equality of Variances (Levene's)

	F	df1	df2	p
HOUR Cat HIGH BOUNDARY	1.219	3	83	0.308
HOUR Cat YES/NO	0.158	3	83	0.924
HOUR Cat LOW BOUNDARY	0.361	3	83	0.781
HOUR Know HIGH BOUNDARY	1.630	3	83	0.189
HOUR Know YES/NO	2.734	3	83	0.049
HOUR Know LOW BOUNDARY	1.925	3	83	0.132
WEEK Cat HIGH BOUNDARY	1.134	3	83	0.340
WEEK Cat YES/NO	1.466	3	83	0.230
WEEK Cat LOW BOUNDARY	1.718	3	83	0.170
WEEK Know HIGH BOUNDARY	0.928	3	83	0.431
WEEK Know YES/NO	0.855	3	83	0.468
WEEK Know LOW BOUNDARY	2.554	3	83	0.061

Post Hoc Tests ▼

Post Hoc Comparisons - Threshold

		Mean Difference	SE	t	p_{bonf}
100% True	Unsure	0.031	0.005	6.015	< .001
	100% False	0.022	0.005	4.396	< .001
Unsure	100% False	−0.008	0.005	−1.618	0.323

Note. P-value adjusted for comparing a family of 3
Note. Results are averaged over the levels of: Age, Languages, Interval, Type

Post Hoc Comparisons - Interval ✻ Threshold ▼

		Mean Difference	SE	t	p_{bonf}
Hour, 100% True	Week, 100% True	0.036	0.009	3.905	0.003
	Hour, Unsure	0.019	0.006	3.368	0.013
	Week, Unsure	0.078	0.010	7.636	< .001
	Hour, 100% False	0.016	0.006	2.891	0.053
	Week, 100% False	0.064	0.010	6.291	< .001
Week, 100% True	Hour, Unsure	−0.017	0.010	−1.651	1.000
	Week, Unsure	0.042	0.006	7.403	< .001
	Hour, 100% False	−0.020	0.010	−1.916	0.880
	Week, 100% False	0.028	0.006	4.982	< .001
Hour, Unsure	Week, Unsure	0.059	0.009	6.391	< .001
	Hour, 100% False	−0.003	0.006	−0.477	1.000
	Week, 100% False	0.045	0.010	4.419	< .001
Week, Unsure	Hour, 100% False	−0.062	0.010	−6.030	< .001
	Week, 100% False	−0.014	0.006	−2.421	0.243
Hour, 100% False	Week, 100% False	0.048	0.009	5.193	< .001

Note. P-value adjusted for comparing a family of 15
Note. Results are averaged over the levels of: Age, Languages, Type

Post Hoc Comparisons - Type ✻ Threshold ▼

		Mean Difference	SE	t	p_{bonf}
Category, 100% True	Knowledge, 100% True	−0.032	0.006	−5.694	< .001
	Category, Unsure	0.017	0.006	2.759	0.092
	Knowledge, Unsure	0.013	0.007	1.885	0.907
	Category, 100% False	0.020	0.006	3.251	0.019
	Knowledge, 100% False	−0.007	0.007	−0.985	1.000
Knowledge, 100% True	Category, Unsure	0.048	0.007	7.145	< .001
	Knowledge, Unsure	0.044	0.006	7.330	< .001
	Category, 100% False	0.051	0.007	7.585	< .001
	Knowledge, 100% False	0.025	0.006	4.124	< .001
Category, Unsure	Knowledge, Unsure	−0.004	0.006	−0.712	1.000
	Category, 100% False	0.003	0.006	0.492	1.000
	Knowledge, 100% False	−0.023	0.007	−3.454	0.010
Knowledge, Unsure	Category, 100% False	0.007	0.007	1.025	1.000
	Knowledge, 100% False	−0.019	0.006	−3.206	0.022
Category, 100% False	Knowledge, 100% False	−0.026	0.006	−4.743	< .001

Note. P-value adjusted for comparing a family of 15
Note. Results are averaged over the levels of: Age, Languages, Interval

Appendix AA

Ethical Approval Letters 0297

19 November 2018

Dear HANNA ELAN THOMAS, , , Associate Professor Jeremy Tree, David Playfoot,

Re: 0297 , Validating clinical neuropsychological assessments for Welsh speakers

Your application - https://swansea.review.ethicalreviewmanager.com/projectTimeline/Index/265 - has been reviewed and approved by the Department of Psychology Ethics Committee.

The conditions of this approval are as follows:

1. To conduct your study strictly in accordance with the proposal that has been approved by the committee, including any approved amendments
2. To advise the ethics committee chair of any complaints or other issues that may warrant ethical review of the project
3. To submit for approval any changes to the approved protocol before implementing any such changes
4. To keep any information obtained from your participants absolutely confidential

Please note that failure to comply with these conditions of approval may result in the withdrawal of approval for the project.

To advertise your study on the departmental Participant Pool: You will need to send send a request for your study to be made visible, via the link on the Experiment Management System website (see Researcher Documentation for details). Please ensure that you attach this letter to your request. (If you are unable to attach the Ethics approval, send it in a separate email to Dr. Phil Tucker p.t.tucker@swan.ac.uk).

For students: Please ensure that the signed copy of this Ethical Approval, together with any other paperwork associated with your research, is included in your final write up.

Yours Sincerely,

Dr IRENE Reppa (Reviewer of Application)

Associate Professor Andrew Kemp (Committee Chair)

10 December 2020

Dear HANNA ELAN THOMAS, , , Associate Professor Jeremy Tree, David Playfoot,

Re: 0297 , Validating clinical neuropsychological assessments for Welsh speakers

Your application - https://swansea.forms.ethicalreviewmanager.com/ProjectView/Index/297 - has been reviewed and approved by the Department of Psychology Ethics Committee.

The list of additional students (if any) are included in the table below:

Other student applicant - first name	Other student applicant - Surname	Other student applicant - email

additional researcher or student - first name	additional researcher or student - surname	additional researcher or student - email
Claire	Williams	[redacted]

The conditions of this approval are as follows:

1. To conduct your study strictly in accordance with the proposal that has been approved by the committee, including any approved amendments
2. To advise the ethics committee chair of any complaints or other issues that may warrant ethical review of the project
3. To submit for approval any changes to the approved protocol before implementing any such changes
4. To keep any information obtained from your participants absolutely confidential

Please note that failure to comply with these conditions of approval may result in the withdrawal of approval for the project.

To advertise your study on the departmental Participant Pool: You will need to send send a request for your study to be made visible, via the link on the Experiment Management System website (see Researcher Documentation for details). Please ensure that you attach this letter to your request. (If you are unable to attach the Ethics approval, send it in a separate email to Dr. Phil Tucker [redacted]).

For students: Please ensure that the signed copy of this Ethical Approval, together with any other paperwork associated with your research, is included in your final write up.

Yours Sincerely,

Dr IRENE Reppa (Reviewer of Application)

Dr Gabriela Jiga-Boy (Committee Chair)

Appendix BB

Ethical Approval Letters 2603

8 January 2020

Dear Dr David Playfoot, , Dr Jeremy Tree, Miss Hanna Thomas, ,

Re: 2603 , Bilingualism and the Rumsfeld effect

Your application - https://swansea.forms.ethicalreviewmanager.com/ProjectView/Index/2603 - has been reviewed and approved by the Department of Psychology Ethics Committee.

The list of additional students (if any) are included in the table below:

Other student applicant - first name	Other student applicant - Surname	Other student applicant - email

additional researcher or student - first name	additional researcher or student - surname	additional researcher or student - email

The conditions of this approval are as follows:

1. To conduct your study strictly in accordance with the proposal that has been approved by the committee, including any approved amendments
2. To advise the ethics committee chair of any complaints or other issues that may warrant ethical review of the project
3. To submit for approval any changes to the approved protocol before implementing any such changes
4. To keep any information obtained from your participants absolutely confidential

Please note that failure to comply with these conditions of approval may result in the withdrawal of approval for the project.

To advertise your study on the departmental Participant Pool: You will need to send send a request for your study to be made visible, via the link on the Experiment Management System website (see Researcher Documentation for details). Please ensure that you attach this letter to your request. (If you are unable to attach the Ethics approval, send it in a separate email to Dr. Phil Tucker [redacted].

For students: Please ensure that the signed copy of this Ethical Approval, together with any other paperwork associated with your research, is included in your final write up.

Yours Sincerely,

Dr Menna Price (Reviewer of Application)

Dr Gabriela Jiga-Boy (Committee Chair)

12 March 2021

Dear Dr David Playfoot, , , ,

Re: 2603 , Bilingualism and the Rumsfeld effect

Your application - https://swansea.forms.ethicalreviewmanager.com/ProjectView/Index/2603 - has been reviewed and approved by the Department of Psychology Ethics Committee.

The list of additional students (if any) are included in the table below:

Other student applicant - first name	Other student applicant - Surname	Other student applicant - email

additional researcher or student - first name	additional researcher or student - surname	additional researcher or student - email
Claire	Williams	[redacted]

The conditions of this approval are as follows:

1. To conduct your study strictly in accordance with the proposal that has been approved by the committee, including any approved amendments
2. To advise the ethics committee chair of any complaints or other issues that may warrant ethical review of the project
3. To submit for approval any changes to the approved protocol before implementing any such changes
4. To keep any information obtained from your participants absolutely confidential

Please note that failure to comply with these conditions of approval may result in the withdrawal of approval for the project.

To advertise your study on the departmental Participant Pool: You will need to send send a request for your study to be made visible, via the link on the Experiment Management System website (see Researcher Documentation for details). Please ensure that you attach this letter to your request. (If you are unable to attach the Ethics approval, send it in a separate email to Dr. Phil Tucker[redacted]).

For students: Please ensure that the signed copy of this Ethical Approval, together with any other paperwork associated with your research, is included in your final write up.

Yours Sincerely,

Dr Menna Price (Reviewer of Application)

Dr Gabriela Jiga-Boy (Committee Chair)

6 September 2022

Dear Dr David Playfoot, , , ,

Re: 2603 , Bilingualism and the Rumsfeld effect

Your application - https://swansea.forms.ethicalreviewmanager.com/ProjectView/Index/2603 - has been reviewed and approved by the Department of Psychology Ethics Committee.

The list of additional students (if any) are included in the table below:

Other student applicant - first name	Other student applicant - Surname	Other student applicant - email

additional researcher or student - first name	additional researcher or student - surname	additional researcher or student - email

The conditions of this approval are as follows:

1. To conduct your study strictly in accordance with the proposal that has been approved by the committee, including any approved amendments
2. To advise the ethics committee chair of any complaints or other issues that may warrant ethical review of the project
3. To submit for approval any changes to the approved protocol before implementing any such changes
4. To keep any information obtained from your participants absolutely confidential

Please note that failure to comply with these conditions of approval may result in the withdrawal of approval for the project.

To advertise your study on the departmental Participant Pool: You will need to send send a request for your study to be made visible, via the link on the Experiment Management System website (see Researcher Documentation for details). Please ensure that you attach this letter to your request. (If you are unable to attach the Ethics approval, send it in a separate email to Dr. Phil Tucker p.t.tucker@swan.ac.uk).

For students: Please ensure that the signed copy of this Ethical Approval, together with any other paperwork associated with your research, is included in your final write up.

Yours Sincerely,

Dr Menna Price (Reviewer of Application)

Dr Gabriela Jiga-Boy (Committee Chair)