

**The relationship between proficiency in multiple
languages and working memory:
a study of multilingual advantages in South Africa.**

**Gabriel Espi-Sanchis
(1778063)**

**A research project submitted in partial fulfilment of the requirements for the
degree of MA in Psychology in the Faculty of Humanities, University of the
Witwatersrand, Johannesburg, 20 June 2018**

**I declare that this research project is my own, unaided work. It has not been
submitted before for any other degree or examination at this or any other university**

Signed: _____

Date: _____

Abstract

This study explores the relationship between multilingualism and working memory. Multilingual advantages in various executive functions have been established, but little is known about whether multilingual advantages extend to working memory capacity and functioning, or about the effect of speaking more than two languages. In a sample of 189 multilingual young adults in South Africa, this study used a multiple regression design in which numerous aspects of multilingualism - balance in proficiency across and within languages, the age of acquisition of additional languages, and speaking a third language - could be compared with one another while controlling for socio-economic status. Four aspects of working memory (verbal storage, verbal processing, visuospatial storage and visuospatial processing), measured using the Automated Working Memory Assessment (Alloway, 2007), acted as the dependent variables in respective regressions while independent variables measuring multilingualism, including the continuous measures of balance in reading, speaking and understanding proficiency across languages, were based on self-report information from the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007). Balance in proficiency emerged as a strong predictor of the verbal processing component of working memory, while no aspect of multilingualism significantly predicted visuospatial working memory. Combined with other results, this finding suggested that the effect of multilingualism on working memory may not follow the pattern observed in other tasks where multilinguals are advantaged in domain-general executive functions (like inhibitory control) but disadvantaged in linguistic tasks. Multilinguals' experience in storing and processing linguistic information may lead to advantages (possibly through managing attention) that are specific to this kind of information.

Keywords: bilingual advantage, executive function, multilingual advantage, trilingualism, working memory

Acknowledgements

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Dedication

My sincere gratitude goes to my supervisor Professor Kate Cockcroft who has been so generous with her time and so interested and willing to share her wisdom and have extended conversations about the topics and themes addressed in this report.

Additional thanks go to Lisa Espi-Sanchis for her invaluable editing and proof-reading assistance, to Joshua M. Martin for sharing his passion for the mind, and to Abigail Dawson who has been so important a presence in my life for the duration of this work.

Table of Contents

Chapter One: Literature Review	6
Theoretical Review	6
Defining multilingualism	9
Working memory and executive function	11
Theoretical links	17
Empirical Review	22
Group comparison research	25
Beyond two languages	29
Multilingualism as a continuous variable	31
Chapter Two: Methods	36
Research Design	36
Sample and Sampling	40
Instruments	41
Procedure	47
Ethical Considerations	47
Chapter Three: Results	49
Descriptive Statistics and Derived Variables	49
Language variables	50
Working memory variables	55
Demographic variables	57
Preliminary Analyses	58
Correlations	58
EFA	61
Bivariate Analysis	65
Multiple Regression Analysis	67
Chapter Four: Discussion	73
Key Findings and Interpretation	74
Methodological Concerns	86
General Reflection	92

Conclusion	97
References	101
Appendix A: Information Letter	110
Appendix B: Demographic Questionnaire	112
Appendix C: Humanities Ethics Clearance Certificate	115
Appendix D: Multiple regression analysis with AOA	116

List of Tables

Table 1: Language frequencies	50
Table 2: Descriptive statistics for language variables	52
Table 3: Descriptive statistics for working memory variables	55
Table 4: Descriptive statistics for demographic variables	57
Table 5: Correlations among key variables	59
Table 6: Correlations among AWMA tasks	62
Table 7: AWMA EFA results	64
Table 8: Bivariate regressions for English understanding	66
Table 9: Multiple regression results	69
Table D1: Multiple regression results with the addition of AOA	116

Chapter One

Literature Review

Introduction

Multilingualism is widespread in South Africa where the great majority of people grow up speaking multiple languages. Most children are educated in English even in cases where it is not their home language, despite constitutional provisions that schooling can take place in any of the country's 11 official languages (Rudwick, 2004). The mixture of languages spoken by South African multilinguals is unique to its particular context, but many aspects of the South African linguistic environment are relevant to an increasingly multilingual global context. Increasing migration has led to many children growing up with a language in their school and social environments that is not their home language (Extra & Vallen, 1997) and increasing globalisation has meant that children across the world are exposed to dominant global languages (Dor, 2004). It is now estimated that at least half of the world's population is multilingual and that growing up multilingual has become the norm rather than the exception (Grosjean, 2010).

The prevailing view was once that a child is incapable of acquiring multiple languages without incurring some sort of cost to their cognitive developmental (review in Hakuta, 1986). But in recent years this has been challenged by empirical research finding not only that the harmful effects of multilingualism are limited to tasks involving language-specific aspects such as lexical access or knowledge of syntax, but further that multilinguals actually seem to be advantaged in several aspects of executive functioning. A wide range of positive effects of multilingualism on cognition have now been established; research finding executive function (and especially inhibitory control) advantages in adults has been complemented by findings in infants (who maintain certain talents for longer and acquire others sooner), children (who reap executive function benefits from a young age) and older adults (who show exceptional cognitive reserve) (see Bialystok, 2017 for a review). A collection of non-results in young adults, meanwhile, has sparked an anti-advantage movement (Paap & Greenberg, 2013).

This wave of discovery has not yet reached every shore, however. Studies have for the most part grouped speakers into two groups, monolingual or bilingual, an approach that has been successful in answering broad questions about multilingual advantages and finding the range of effects listed above but has now started to come under question. What is a monolingual or a bilingual? How do we define them and how do we draw a line between them? What about the numerous ways in which two people identically labelled bilingual might differ from each other in their linguistic experience and proficiency? And what about the millions of people who speak more than two languages? Are we to assume their cognitive development is identical to those who speak two? Likewise, the effects of multilingualism on some aspects of cognition have been paid little to no attention and largely remain unknown. Effects on working memory, widely considered to be a central concept in executive function, and more generally in cognition, have only recently been examined with mixed results but some notable positive findings (Blom, Kuntay, Messer, Verhagen, & Leseman, 2014; Morales, Calvo, & Bialystok, 2013).

The focus of this study is on clarifying and illuminating these questions at the frontier of current knowledge. The central question asks whether people who are more proficient and experienced in the use of multiple languages have advantages in working memory. An exploratory design which examines the influence of a number of aspects of multilingualism simultaneously is adopted so that their effects are not conflated and they can be directly compared to one another. Each of 189 multilingual young adults are compared regarding the extent to which they are balanced in proficiency across their languages and across different kinds of proficiency within their languages, the age at which they started becoming multilingual (revealing how dependent advantages are on critical early periods of development), and the effect of speaking a third language, relative to speaking two.

The burst of research in this field has occurred for the most part in Europe and North America. This has meant that a very limited number of the world's languages has been studied, and that the particular effects of speaking languages from many parts of the world, South Africa included, remain a mystery. In one of the most unequal countries in the world

(in terms of wealth and income) (Leibbrandt, Finn, & Woolard, 2012) the effect of socioeconomic status, which is likely to be significantly correlated with both multilingualism and working memory, needs to be accounted for.

Another issue with multilingual research is that a range of different studies focus on different outcomes with little consideration for how their theoretical constructs relate to each other, resulting in a potpourri of different findings which are difficult to combine in any clear, meaningful way. Thus an aim of this study is to understand how the concepts and models encountered, especially those of multilingualism, working memory and executive function, relate to one another.

These questions are informed by a range of literature. Before performing an empirical study of multilingualism's effects on working memory, one needs to clarify and explore definitions of multilingualism, assess concepts and models of working memory and executive function, and explore the theoretical links between the phenomenon of multilingualism and these cognitive outcomes. After surveying literature interrogating these issues in the first section of this chapter, the second section reviews empirical studies that have investigated multilingual advantages in executive function. The focus of this empirical review is on advantages in working memory and on studies that explicitly consider more than two languages or model multilingualism as a continuous variable. Thereafter Chapter Two describes the research design, methods and procedure used to address these research questions, covering aspects such as the sample targeted, instruments used to collect data and procedure followed. Chapter Three describes the sample on a range of variables and presents the results of statistical analyses. Finally, Chapter Four discusses and interprets these results in relation to the existing literature and attempts to determine how the results observed have informed the research questions.

Theoretical Review

Defining multilingualism

The diversity and complexity of multilingual experiences, in South Africa and around the world, raises important questions about how multilingualism should be defined, conceived of and operationalised. This subsection introduces broad questions regarding the definition of multilingualism and why it is significant, leaving a discussion of its relation to working memory and executive function to later subsections.

Most people would agree that being multilingual involves speaking more than one language, but among researchers there is a persistent lack of consensus regarding the definition and operationalisation of multilingualism (Garcia & Li Wei, 2014). Our intuitive understanding of multilingualism is multi-faceted and can include notions of proficiency, experience, exposure and use across multiple languages. Some have defined multilingualism in terms of the community in which someone lives, but most definitions are based on some aspect of individual experience or proficiency (Sabourin & Vinerte, in press). Kohnert (2010) centres the definition of multilingualism around children receiving input in multiple languages in dynamic periods of language development between birth and adolescence. Many other exposure-based definitions similarly focus on early life as being critical, but recently even small amounts of later life exposure to other languages have been considered important (see Thompson, 2013). Because different components of cognition will be more or less affected by different aspects of multilingualism (Marian et al., 2007), researchers choose to focus on one or another of these aspects based on the theme of their research, and the pursuit of a universal definition may be futile or even counterproductive.

Most competing definitions of multilingualism assume that there are distinct languages and that someone can speak one or more of these languages. The recent wave of translanguaging theory critiques this assumption and claims that a person only has one dynamic language system with a range of “linguistic features” which can be applied to different societally-defined languages (Garcia & Li Wei, 2014, p. 14). According to this

theory it is incorrect to speak of people possessing multiple languages or language systems, and so research comparing bilinguals and monolinguals is misguided in simply expanding or multiplying conceptions of monolingualism (Garcia & Li Wei, 2014).

Notwithstanding the inconsistencies in the way multilingualism is defined, there is great interest in its possible effects on cognition. The integral and consistent involvement of language in all human activity distinguishes it from other forms of experience; beyond its use in communication, language is instrumental in thinking and the interpretation of sensory experience (Bialystok, 2017; Vygotsky, 1962). This has led many to speculate on how the presence of multiple languages might affect cognitive development, and whether the effect is quantitative (i.e. accelerating development in a dose-like manner) or qualitative (leading to conceptual jumps and qualitatively different forms of cognition). Effects of growing up multilingual on linguistic abilities are well established, with, for example, multilingual infants maintaining the ability of universal phonetic discrimination longer than monolingual infants, but many believe that there are more general cognitive effects beyond the domain of language (Bialystok, 2017).

Vygotsky (1962) thought that multilingualism stimulated and accelerated cognitive development due to its role in facilitating an understanding that the relation between words (or signs) and the things they signify is essentially arbitrary. Learning two (or more) completely different words for the same object helps a child to realise that nothing in the sign (such as its sound) represents the object it names and that its language system is one among many. The child is thus more likely to understand aspects of language in general terms in a way that facilitates the acquisition of higher mental operations. This is an example of multilingualism enabling a qualitative shift in cognition which then accelerates cognitive development. In a similar vein, research by Byers-Heinlein and Werker (2009) into the word-learning heuristics of 17 to 18 month old infants provides fascinating insight into the way multilinguals adapt to multiple languages and learn in different ways as a result. It is widely established that in learning words infants make use of the heuristic of disambiguation, meaning that they will tend to associate a novel word with an unfamiliar rather than a familiar object. Multilinguals, however, in being exposed to several basic labels for the same familiar

object may have learned to systematically violate the principle of disambiguation (Byers-Heinlein & Werker, 2009). The authors' findings are that while monolingual infants act in accordance with this tendency, bilinguals do so only marginally and trilinguals not at all. These results seem compatible with Vygotsky's (1962) belief that multilinguals are more aware of the arbitrary relation between labels and the things they designate.

This section has introduced questions of what multilingualism is and how it might affect cognitive development. The difficulty in defining multilingualism was raised and will be returned to later, but it is crucially important to come up with some precisely defined labels to avoid ambiguity, especially in this study which compares speakers of two and three languages. For the remainder of this report 'multilingualism' will be used as a general term to refer to the use of and proficiency in more than one language. Other more numerically specific terms will be used to refer to speakers of two languages (bilinguals) and three languages (trilinguals). Due to inconsistencies and vagaries in the use of these terms (especially the terms multilingual and bilingual) across research, the labels used in this report will sometimes differ from those used in the original papers. However, the labels will always accurately describe the number of languages spoken as per the definitions above.

Working memory and executive function

The majority of research investigating the effects of multilingualism on cognition has focused on some aspect of executive function and, most prominently, inhibitory control. Before discussing this research and relating it to the definitions of multilingualism and the South African context referred to above, it is necessary to describe and interrogate the concept of executive function, as well as the related concept of working memory which is the focal outcome of this study. A lack of consistency and understanding of the relations among psychological concepts can lead to misguided research questions (Bennett, Hacker, & Bennett, 2003), and so every effort will be made to explain how the concepts that are introduced relate to one another.

Executive function is a construct that has received an immense amount of interest (Alvarez & Emory, 2006). As a whole, executive functions are believed to have a domain-general, supervisory role (over other cognitive processes), and to be responsible for self-regulation and control of goal-directed behaviour (Miyake et al., 2000; Riggs, Shin, Unger, Spruijt-Metz, & Pentz, 2014). In performing a complex task, it is the role of executive functions to coordinate the cognitive processes required. executive functions are widely believed to depend on cortical regions and networks in the frontal lobe, and sometimes the terms executive function and frontal lobe are used interchangeably (e.g. Miyake et al., 2000). However, given that one is an anatomical concept and the other a functional one, several philosophical problems arise from identifying one with the other, and empirically they have been found to be partially dissociable (with non-frontal areas active during executive function tasks and frontal areas active during non-executive function tasks) (Miyake et al., 2000). A range of processes or functions have traditionally been grouped under the executive function banner, including inhibitory control, working memory, attentional control, switching or shifting between sets or rules, conflict management (Miyake et al., 2000; Riggs et al., 2014), as well as flexibility and monitoring (Bialystok, 2017). The most critical difference between different theories of executive function concerns whether executive function is a single entity or can be fractionated into distinct components. The unity of executive function components is challenged by some findings of low correlations among proposed components (which are comparable to correlations among other distinct aspects of cognition), and by neuroimaging, developmental and lesion studies suggesting some degree of fractionation (Riggs et al., 2014).

Miyake et al. (2000) made a much-referenced and widely accepted distinction between three components of executive function: inhibitory control (of prepotent responses), shifting between tasks or sets of rules (and managing interference from previously active ones), and working memory (information updating and monitoring). These components were chosen, rather than the alternatives listed above, for their potential to be measured by simple tasks (their tractability), their probable use in more complex tasks, and for what the authors perceived as their “circumscribed”, foundational nature (with other proposed executive functions likely to result from some combination of these) (Miyake et al., 2000, p.55).

Miyake et al. (2000) selected tasks which they anticipated would engage these functions and performed confirmatory factor analysis (CFA) to see whether a one, two or three factor model best explains task performance. A model with both unity and diversity was most supported, with the three components listed above being distinguishable but also sharing an underlying commonality.

The tasks which have been used to measure working memory are most important to this study and are discussed later, but the predominance of other aspects of executive function in existing multilingual advantage research and the links between working memory and executive function mean it is important to have some understanding of the tasks used to measure executive function as well. It is difficult (or impossible) to devise tasks which isolate aspects of executive function, because by definition executive functions work by acting on other capacities, and because different people may approach the same task with different strategies (Miyake et al., 2000). This introduces issues of task impurity. Miyake et al. (2000) addressed this issue by focusing on latent variables that can explain performance across tasks, but the fact that their tasks were selected with their three factors in mind means there are problems of circular reasoning in their method. This casts doubt on the extent to which their study can be taken as evidence that executive function can be (somewhat) fractionated into three components. Traditional tasks for measuring inhibitory control include the Stroop (1935) and Flanker tasks, where participants need to inhibit prepotent responses based on word reading and incongruent symbols, respectively. Switching between tasks or sets is generally measured by some task requiring multi-tasking. Other tests, such as the Attentional Network Test (ANT), measure executive function more broadly and yield information about several of its aspects, such as the efficiency of alerting, orienting and control in terms of executive attention (Fan, McCandliss, Sommer, Raz, & Posner, 2002).

Models of working memory, the focal outcome of this study, can now be understood within the broader context of theories of executive function. The term working memory was first used by Miller (1960, as cited in Cowan, 2008) to refer to the memory that is used in planning and executing actions. The incorporation of working memory into many of the models of executive function discussed above makes it clear that the processing aspects of

working memory are related to executive function. This processing aspect is the feature which many use to distinguish between working memory and short-term memory (STM). Baddeley, Eysenck and Anderson (2009) define STM as the ability to store small amounts of information over short periods, meaning it is defined not according to a theory of how this capacity functions in the mind but according to its experimental measurement. Working memory, on the other hand, is a theoretically defined system which is capable of both storing (using STM) and manipulating information in ways that are essential to cognition and learning (Baddeley et al., 2009).

Beyond the general distinction between working memory and STM (which is not uncontroversial), there are several differing theoretical approaches to describing working memory's functioning and structure, and its relation to long-term memory (LTM). Cowan (2008, p.1) distinguishes between models which see working memory as STM applied to cognitive tasks, as a multicomponent store holding and manipulating information in STM, and as "the use of attention to manage STM". Some examples of these different kinds of model are assessed below.

The most influential model of working memory was first presented by Baddeley and Hitch (1974) as a challenge to the prevailing conception of short-term memory as a static store of (exclusively) phonological information. The Baddeley and Hitch (1974) model contained three components: the central executive (responsible for coordinating the actions of the other two components and allocating attention based on the task at hand), the phonological loop (responsible for storing and processing auditory information), and the visuospatial sketchpad (responsible for storing and processing visual and spatial information). The separate existence of the latter two slave systems was justified based on evidence that performance on two simultaneous domain-specific tasks was almost as good as on each task individually. In many cases people who have poor working memory perform poorly on any task regardless of modality, a phenomenon that is said to derive from issues in the central executive, but sometimes people can have uneven performance with what appears to be a deficiency in one of the slave systems (Alloway, 2007). Evidence for the importance of subvocal repetition of items in short term memory led to the postulation of two subdivisions

of the phonological loop: the phonological store and the articulatory rehearsal system (Baddeley, 2000).

Later, Baddeley (2000) introduced an augmented working memory model with a fourth component, the episodic buffer, a hub which is responsible for combining information from long-term memory and the working memory slave-systems into coherent episodes, and which acts as the pathway through which short-term memories are transferred to long-term memory. This buffer could in theory process types of information (e.g. semantic) that the other two STM stores could not. This recent addition remains the least researched of the model's components and the most difficult to operationalise (Nobre et al., 2013).

Cowan (1988, 2008) presented an alternative, attention-based model in which STM is subsumed under working memory, as are other processing mechanisms which utilise STM. In this embedded model, the components of working memory are believed to be activated memory (a subset of LTM), and the focus of attention (a further subset of activated memory). Information held in the former is believed to decay over time unless it is refreshed by coming into the focus of attention, which has capacity limits (Cowan, 2008). Thus instead of being stored in separate components as in Baddeley's model, Cowan conceives of phonological and visuospatial information as two of the kinds of information that can be stored in activated memory, and the extent to which there is cross-interference between two chunks of information depends on how similar they are in modality. Although the phenomenon plays a role in Baddeley's (2000) model, the organisation of information (e.g. letters or numbers) in memory into larger meaningful chunks is especially important according to Cowan's (2008) model. 'BBC' is an example of a single chunk that is made up of three items (letters), and capacity limits are believed to exist for the number of chunks, rather than the total number of items, that can be held in memory. Another model that conceives of working memory as the intersection of attention and memory is that of Engle (2002). Engle's conception of executive attention (or working memory capacity) is similar to notions of executive function (as in Miyake et al., 2000) in that it is domain-general and limited in its capacity, but differs in that it is a continuous construct rather than one consisting of discrete components (Bialystok, 2017).

The tasks used to measure working memory generally include some STM component, with the participant needing to store some information in their mind, and some processing component where this information needs to be manipulated in some way (e.g. Alloway, 2007). Many of these tasks are differentiated as verbal or visuospatial based on the nature and modality of the content to be stored and processed. An example of the former is the Listening Recall task in which a participant hears a series of sentences and must judge if each is true or false while remembering the final word of each sentence in sequence (see Alloway, 2007). Visuospatial tasks generally involve some judgement based on visual information. For example, participants must select which shape is the odd one out while remembering the locations of the odd shapes out in the sequence in which they were presented. Measures of working memory based on these kinds of tasks have been found to correlate very well with measures of intellectual aptitude and fluid intelligence (better than measures of other specific psychological functions), which has led to increased interest in the construct (Alloway, Gathercole, & Pickering, 2006; Cowan, 2008). Recent speculation about what aspect of tasks is particularly good at predicting intelligence has suggested that the task needs to be sufficiently challenging to the control of attention (Engle, Tuholski, Laughlin, & Conway, 1999) or, a simpler suggestion, that it needs to prevent the participant from rehearsing (and thus from refreshing information stored in memory) (Cowan, 2008).

Due to the clear theoretical involvement of executive function processes in working memory, debates on the unity or diversity of executive functions have immediate implications for models of working memory. A unified conception of executive function precludes the possibility of working memory being a completely distinct component (and it would have to be understood as the outcome of general processing capacities acting together with short-term memory stores), while a diverse conception can explicitly include working memory as one of the components of executive function (as Miyake et al. (2000) do).

The understanding of working memory as one of the distinct components of the executive function system is complicated by the fact that some working memory models contain theoretical components which are similar (or even identical) to components of

executive function models. The similar descriptions and roles of the central executive in Baddeley's models (Baddeley & Hitch, 1974; Baddeley, 2000) and the executive control system in Miyake et al.'s (2000) model make clear the perils of simplistic hierarchical understandings. How can one component of a subsystem be the same as the supervisory component of the super system? Thus when understanding models of working memory it is important to determine whether the authors consider their model to be a subsystem of executive function, to be an independent system which engages executive function processes, or to actually be constitutive of EF. Neither of the working memory models discussed above are meant to be independent of EF. It is clear that in Cowan's (2008) theory working memory is inclusive of executive function processes that act on activated memory and direct the focus of attention, and in Baddeley's that the central executive involves executive functions (Baddeley & Hitch, 1974; Baddeley, 2000). Neither do these models conceive of working memory as some subsystem of a broader executive function system, but it is unclear whether they conceive of working memory as merely executive function processes as applied to STM, or whether there is some share of processing which is meant to be memory-specific. Engle's (2002) model, on the other hand, is clearly meant to be descriptive of executive function and so should be understood as an alternative to models of executive function.

Theoretical links between working memory, executive function and multilingualism

Multilingual advantage research aims to find advantages in domain-general cognitive functions, beyond the realm of language. In any assessment of multilingual advantages, the theory linking executive function to multilingualism is very important, and this is now discussed before the links between multilingualism and working memory are examined. Recent research has found that executive functions can be augmented by training and intensive and regular engagement (Yow & Li, 2015). With regard to language, two questions are particularly important: one, which aspect of multilingual experience involves the use of domain-general executive functions (and which functions are involved), and two, how can this use lead to improvements in the function over time (Bialystok, 2009; Hilchey & Klein,

2011)? A comprehensive theory of multilingualism and executive function should be able to shed light on these questions.

Green's (1998) model of how minds manage multiple languages states that the source of multilingual advantages lies in the experience of inhibiting words from languages that are not currently being spoken. According to Green's model, semantically linked units are activated in both languages in response to an experience or thought, and an inhibitory control mechanism is responsible for suppressing units that are irrelevant or inappropriate given the task at hand. The consistency of Green's account with the Miyake et al. (2000) model of executive function has contributed to its popularity, and recently Abutalebi and Green (2008) have expanded the model to give a more detailed account of (eight) control mechanisms which multilinguals recruit according to the interactional context. Heidlmayr et al. (2014) also give an inhibition-based account where inhibitory control is needed to avoid applying first language processing strategies to one's second language (and vice versa) to avoid impairing comprehension (rather than production as in Green (1998)).

These accounts of why cognitive control is necessary in managing multiple languages are based on behavioural evidence of joint activation of either lexemes (Green, 1998) or processing strategies (Heidlmayr et al., 2014) across languages. Another approach bases the need for cognitive control on neuroscientific evidence of shared pathways among multilinguals' languages (Buchweitz & Prat, 2013). Further, neuroimaging evidence of overlapping networks in linguistic and cognitive control tasks is routinely used to support the assumption that linguistic control mechanisms are (at least partially) domain-general (Bialystok, 2017). A belief in joint activation of languages and the subsequent need for inhibition does not necessarily imply a domain-general inhibitory control position, however. Dijkstra and van Heuven (2002) share Green's (1998) belief in co-activation of languages but posit language-specific (and automatic rather than active) control mechanisms (Dijkstra & van Heuven, 2002).

Support for inhibitory control accounts has come from several sources, with one being Heidlmayr et al.'s (2014) finding that the multilingual advantage in a Stroop task correlated

negatively with the duration of immersion in a foreign (second) language context. This implies that the less automatised and more interference-based processes which are prevalent when one is first becoming multilingual are central to executive function advantages. However in recent years Green's cognitive control account, which, by virtue of its focus on inhibition, stresses advantages based on speech production, has been criticised for being incapable of explaining multilingual advantages in infants who are not yet speaking (Bialystok, 2017; Bialystok, Craik, & Luk, 2012). Joint activation of a multilingual's languages is widely established (Kroll & Bialystok, 2013) and has been observed during comprehension as well as production (Marian & Spivey, 2003), but Green's (1998) account explicitly (and exclusively) revolves production as the mechanism by which advantages develop. Multilingual infants have been found to override habitual responses and generalise across memory cues better than their monolingual peers, and to have better encoding and recognition memory (Bialystok, 2017). This suggests that at least some of the multilingual advantages observed later in life begin to develop pre-verbally and before the kind of inhibition that Green proposes (Bialystok, 2017).

Bialystok (2017; Bialystok et al., 2012) moved away from Green's inhibition account by framing the multilingual advantage as an improved ability to deploy and manage attention. She believed that such an account, drawing on Engle's (2002) continuous working memory model, was better able to explain the findings in infants discussed above as well as the observed pattern of non-significant results for certain kinds of inhibitory tasks. Multilingual advantages have been found for inhibitory tasks which require suppressing distracting cues (like Stroop tasks) but not for those which require refraining from doing something, meaning there is an advantage in a specific kind of inhibition, while observation of facilitation and interference effects from non-active languages cast doubt on the idea that languages not in use are actively inhibited (Bialystok, 2017). Other kinds of experience that lead to cognitive advantages and do not seem predominantly based on inhibition have also been cited as a reason for the *insufficiency* of the inhibitory account in explaining the phenomenon of executive control advantages in general (Valian, 2015).

In other accounts Ransdell, Barbier & Niit (2006) focus on the improved metalinguistic awareness and working memory that come from heterogeneous linguistic experience involving the activation and inhibition of different codes, and Costa, Hernández, Costa-Faidella and Sebastián-Gallés (2009) and Pelham and Abrams (2014) focus on conflict management and the continuous monitoring of which language is appropriate for each communicative situation. Brito, Sebastián-Gallés and Barr (2014) speculate that the association of multiple words with the same object means that multilingual infants become more flexible in their application of memory, making use of a wider range of retrieval cues. Finally, some authors forsake the pursuit of a single mechanism and choose to focus on multiple mechanisms leading to advantages (see Valian, 2015).

The theoretical link between multilingualism and working memory has generally been seen as less clear than in the cases of inhibition and other executive functions (Bialystok, 2009), and this could be a big factor behind the paucity of empirical literature studying this relationship. However, the longitudinal nature of language means working memory is involved in holding chunks of linguistic information in mind and connecting them to preceding and following chunks (Riggs et al., 2014). The working memory system of multilinguals has to adapt to perform these linguistic processes with multiple streams of information simultaneously, as when translating between languages, and so might be more developed than that of monolinguals (Riggs et al., 2014). Besides, as the paragraphs above show, contemporary shifts in the explanation of which executive function processes multilinguals have exceptional experience in have made notions of working memory highly relevant. Many accounts now explicitly predict multilingual advantages in attention-based working memory systems (Bialystok, 2017; Blom et al., 2014; Costa et al., 2009), while other accounts focusing on conflict monitoring or effective selection of the target language also imply substantial engagement of working memory¹. Further, given a partially unified conception of executive functions which includes working memory (like that espoused by Miyake et al., 2000), and the significant body of positive findings for effects of multilingualism on executive functions, one would expect an advantage to exist for working memory.

¹ Neuroimaging studies have found the dorsolateral prefrontal cortex, an area associated with WM, to be active during language-switching tasks (Riggs et al., 2014).

While many of the theories above present hypotheses about which aspect of multilingual experience is responsible for executive function advantages, the proposed mechanisms through which executive functions are engaged are broadly and vaguely described. Further, the manner in which the use of a function augments that function over time often remains unexplained. It is possible that not every function is augmented through use, or that only particular types of use augment a function's capacities. Furthermore, the possibility of maximal levels in a function (beyond which it cannot be augmented by experience) must be considered. Many invoke positive findings for experience-based advantages in other domains such as music or video games as evidence for the fact that executive function can be trained (e.g. Heidlmayr et al., 2014), but such analogies are invalid substitutes for theoretical links between the engagement of a function and changes in that function over time. For example, the inhibitory control account (Green, 1998) goes to great lengths to explain how inhibitory control is instrumental in switching between languages, but gives no information as to how this inhibitory control mechanism may develop over time given this experience. Bialystok (2017) begins to address this question with her claim that it is more plausible that a continuous executive function construct will be augmented by use than some discrete construct. This shows the step from use to augmentation being directly addressed and Bialystok does well to explicitly look at implications for this relationship in the theory of executive function².

In conclusion, there are some clear theoretical links between multilingualism and aspects of executive function but the specific mechanism through which multilingual experience engages and enhances executive function over time needs further elaboration. The connection between multilingualism and working memory has been discussed and investigated relatively little, but there are good reasons to think that working memory plays a large role in linguistic and multilingual processes. Evidence surrounding this link will add to

² Another way of answering this question is to deny there is a distinction between use and augmentation. Some plasticity theorists conceive of all learning and all memory as neuroplasticity (e.g. Black, 1998). This already wide-ranging blanket conception of plasticity could perhaps be adapted to posit that all use of a given cognitive function will result in neurological change in the structures associated with that function (and therefore changes in ability and behaviour). The engagement with such a position lies outside the scope of this report.

the understanding of the mechanisms through which multilingualism affects cognition and, indirectly, contribute to the debate around the unity or diversity of executive functions.

Empirical Review

Having discussed various models of executive function and working memory and the theoretical links between multilingualism and these aspects of cognition, the remainder of this chapter will review pertinent empirical evidence surrounding the multilingual advantage. The first section will focus on studies that compare groups of monolinguals and bilinguals, the traditional form of study in the field, before further sections focus on studies that look at speakers of more than two languages and on studies that model multilingualism as a continuous variable. In each section the effects on a range of executive functions are examined before a comprehensive review of the effect on the focal outcome of working memory is presented.

Whether researchers model multilingualism discretely (comparing groups) or continuously (along a spectrum), they need to decide which aspects of multilingualism are crucial to its operationalisation so that they can determine the criteria for classification into one or another group (in the former kind of study) or select the language variables to be modelled (in the latter). The selection of these aspects depends on a combination of previous empirical findings and the author's theory of what multilingualism is and how it engages and affects executive function.

The previous section showed how empirical results have shaped theory about the relation between multilingualism and EFs. The current pattern of results, especially among infants, is not easily explained by the historically dominant inhibitory control account, and so alternative accounts have arisen that can better explain the pattern of results. Theory should also shape empirical research, however, by suggesting which aspects of multilingualism determine cognitive advantages (and when), and by generating hypotheses. As an example, the inhibitory control account would predict differences in age of acquisition (AOA) *after* the onset of speech to be significant, whereas other attention-based language-selection accounts

would also highlight differences in AOA in the period *before* speech production begins. AOA could be significant in determining advantages because of the developmental effects of either early critical periods (after which one cannot become multilingual in the same way) or prolonged exposure (Kaushanskaya & Marian, 2009).

Beyond AOA, many other aspects of multilingual experience could also be relevant in determining differences in EFs. The effects of additional languages would undoubtedly depend on one's proficiency in those languages. Proficiency would determine cognitive advantages in many of the theories discussed; for example, the inhibitory control account depends on the subject's inhibition of jointly activated lexemes in the non-target language (Green, 1998) whose level of activation should increase with the level of proficiency in that language³.

Linked to the proficiency focus are approaches which explicitly focus on balance between languages (recently advocated by Bialystok, 2017). If the mechanism which leads to multilingual advantages is particularly developed in those who speak their languages at equivalent levels of proficiency then one would expect differences between balanced and unbalanced multilinguals. Some researchers suggest that a mechanism of managing interference from competing languages implies that one should shift attention away from more long-term characteristics like balanced proficiency (and AOA) to those based on current language exposure, use and switching (Pelham & Abrams, 2014). Another potential determinant of multilingual advantages is the similarity or difference between a speaker's languages. Intuitively, more interference would be expected between more phonologically and orthographically similar languages, which would be relevant within an inhibitory control account. Similarly, language processing might be more challenging when it involves distinguishing between two languages that are rhythmically similar and share many cognates (Brito et al., 2014). On the other hand, if the mechanism leading to multilingual advantages works through switching between different grammatical and phonological sets it would seem

³ Some believe that even additional languages that a person has had little exposure to and speaks at a low proficiency can profoundly affect cognition and give rise to multilingual advantages (Thompson, 2013). In fact, some have speculated that it is those who are in the process of becoming multilingual (who are still of a low proficiency) who have the most exercised and augmented executive functions (especially inhibitory control) (Heidlmayr et al., 2014).

that the more different the languages are, the more difficult it is to switch between them, and thus the more pronounced the multilingual advantage. A pioneering study found that speakers of both highly similar and highly different language pairs had advantages over monolinguals but did not find any difference between these bilingual groups (Brito et al., 2014).

Recent evidence suggests that experiences need to be intensive and sustained (over a long period of time) to train or augment executive function capacities (Yow & Li, 2015). Language use is one of the most pervasive forms of experience, and the more balanced a multilingual is, the more time they spend switching between their languages, and the more intensive and regular this training will be. One may conceive of executive function or working memory either as something that can be altered or augmented only in certain developmental periods or as something whose nature and capacity is more sensitive to recent experience. This conception, one way or the other, would have significant effects on whether the researcher focuses on historical aspects (such as AOA and age at which fluency was achieved) or more recent experiential aspects (such as the degree of balanced exposure and switching between languages). A focus on proficiency can be accommodated within both approaches, as proficiency is inextricably tied into historical language use and experience but also directly affects the nature of current multilingual experience.

A range of factors which seem relevant in determining multilingual advantages have been presented with examples of how different theoretical positions favour a focus on certain of these aspects. The range of different effects observed seems to indicate that there is more than one process or mechanism leading to advantages (Bialystok et al., 2012). No attempt has been made to arrive at definitive conclusions about which aspects of multilingualism should be instrumental in its operationalisation; indeed, all of the aspects discussed above are compatible with one another and, if consistent with theoretical predictions, can be used in conjunction (see Cockcroft, Wigdorowitz, & Liversage, 2017; Yow & Li, 2015). The prevailing uncertainty surrounding the mechanism responsible for multilingual advantages makes an approach which models multiple aspects of multilingual experience in an exploratory study attractive. In this study three aspects are modelled and compared to one another: balance in proficiency, AOA and the number of languages spoken. These factors

together should account for the extent of engagement of many of the mechanisms discussed, but they prioritise longer term, more permanent characteristics over more contemporaneous variables focusing on recent experience.

Groups comparison research

This section will describe studies which compare people sorted into (apparently homogenous) groups of monolinguals and bilinguals (Yow & Li, 2015). As established in the previous section, the aspect of multilingualism that is essential in leading to the multilingual advantage should take centre stage in the classification of multilinguals.

Researchers who focus on the effects of AOA on multilingual advantages often sort bilinguals into early and late bilinguals to enable within-bilingual comparisons. Luk, De Sa and Bialystok (2011) compared English proficiency and flanker task performance in English monolinguals and early and late bilinguals who spoke a variety of mother-tongues and had become actively bilingual before or after the age of 10, respectively. They found that early (but not late) bilinguals had comparable English proficiency to monolinguals and that only early bilinguals exhibited a significantly smaller response time cost on incongruent trials relative to monolinguals. These results suggest that both AOA and proficiency are important for multilingual advantages but cannot be used to parse out their effects. Sabourin and Vinerte (in press) also compare early (from one to six years) and late (after six) bilinguals with monolinguals on their performance in cognitive control tasks, but they make a further distinction between simultaneous (from birth) and early sequential bilinguals, believing that different underlying processes may be involved in the development and use of language in each group. In performance on a Stroop test and an Attentional Network Test (ANT), they found significant differences between early and simultaneous bilinguals in mixed-language Stroop performance, and in the ANT task found only efficiency advantages between the simultaneous bilingual group and the monolingual group to be significant. These can be described as the most and the least bilingual groups and the results of this study suggest that studies that group together simultaneous and early bilinguals may be losing sensitivity to some of the unique developmental properties (and cognitive consequences) of simultaneous

language acquisition. Pelham and Abrams (2014) compared early and late Spanish-English bilinguals (who learned their second language before the age of seven and between seven and 13 respectively) to English monolinguals, and found performance advantages on ANT tasks for both bilingual groups relative to monolinguals, but no significant differences between the bilingual groups.

Very few studies have paid attention to the intrinsic properties of the languages spoken and the majority of studies have used European languages and have not considered the similarity of multilinguals' languages. The direction of the effect of similarity is unclear with some finding more pronounced effects for more similar languages, on the one hand, and others for more different languages, on the other (Bialystok, 2017).

The effect of participants' ages at testing, rather than at acquisition, has come under increasing scrutiny. Relative to other age groups, little research has looked at adult populations, and studies with young adults have resulted in a disproportionately high number of non-significant results (Valian, 2015; Bialystok, 2016). This research has been used by some (such as Paap, Johnson, & Sawi, 2015) as evidence that the bilingual advantage does not exist or is only present in very specific circumstances. Others, such as Bialystok (2016, 2017), have retorted that the reliable observation of advantages at other ages, and the observance of different structural and functional results in neuroimaging studies in the same age group⁴, must mean that there is some methodological issue that is driving these non-results. Young adults could be close to their peak performance in executive function (Yow & Li, 2015), meaning that there is little variation around a high mean and little chance of significant differences between groups, a phenomenon which could be compounded by ceiling effects in tests which are not challenging enough to reveal differences at the higher end of the distribution (Bialystok, 2016). Besides, the university students who make up most young adult studies are at a stage of their life where they are being challenged by a range of other types of experience which could have as much of an effect on their executive function as proficiency in multiple languages (Valian, 2015). The mystery surrounding multilingual

⁴ Relative to their monolingual peers, multilingual young adults have been found to have superior grey matter density and volume and white matter connectivity, and differences in functional activation which are associated with more efficient processing (Bialystok, 2017).

advantages in this age group was a significant reason for the use of a young adult sample in this study. Determining whether young adult multilinguals are advantaged in their working memory would help to establish whether there are advantages in some kinds of executive function or whether the lack of multilingual advantages extends to all kinds of executive function.

Non-linguistic factors that are predictive of executive function and correlated with multilingualism, such as SES, have recently become the focal point of both critiques of and new directions in multilingual advantage research (see Valian, 2015). Most research does not pay enough attention to the role of these covariates in determining executive function and influencing experimental results (Yow & Li, 2015), and most studies comparing cognition of bilinguals and monolinguals are conducted in similar developed, European language-speaking areas of the world with participants who are from an educated and high SES background. If there are systematic differences in nationality, SES or culture between the language groups in a study then it will be difficult for the researchers to claim that they have isolated a multilingual effect (Yow & Li, 2015).

SES is a broad term covering a range of variables such as income, wealth, education and asset ownership, and these factors and others they are correlated with (such as the amount of early childhood stimulation) are likely to account for a broad range of experiences which can affect cognitive behaviour and development (Valian, 2015; Bialystok, 2017). Recent studies of SES, and the interaction (or absence thereof) between this effect and the multilingualism effect, have indicated that the two factors are each significantly but independently correlated with executive function (e.g. Blom et al., 2014; see review in Bialystok, 2016). Because of these findings this study seeks to determine the effects of multilingualism while controlling for SES.

Another variable that is believed to be confounded with the multilingualism effect when not controlled for is that of culture. Multilinguals generally have to switch between different cultures, as well as languages, in their everyday life, and the cognitive challenge of managing the different expectations and norms of each could lead to changes in executive

function (Riggs et al., 2014). Some research has varied the country of origin of both monolinguals and multilinguals to control for the effect of culture, but has generally found significant effects for multilingualism after controlling for culture (which is itself significant in some cases) (Tran, Behseta, Ellis, Martinez-Cruz, & Contreras, 2015; Yang & Yang, 2016, as cited in Bialystok, 2017).

A multilingual advantage in inhibitory control tasks has been investigated more than other executive functions and is traditionally believed to be the best supported by the evidence (Valian, 2015). Multilingual advantages in working memory are relatively under-explored (Riggs et al., 2014), and the concluding part of this section will review some of the few studies that have investigated this issue. Morales et al. (2013) found advantages in working memory among a large sample of children, aged five to seven, which were especially pronounced when other executive functions, such as conflict resolution, were more highly taxed. Blom et al. (2014) found similar results among a sample of five- and seven-year-old children in the Netherlands, with Turkish-Dutch bilinguals possessing an advantage over their monolingual counterparts in verbal and visuospatial working memory tasks with a substantial processing component (when Dutch vocabulary and SES were controlled for). Riggs et al. (2014) found small but significant effects of bilingualism on executive function and on working memory in the 6th but not the 5th grade (as well as effects on growth in executive function from the 5th to the 6th grade) after controlling for gender, age and biculturalism. Ransdell et al. (2006) studied a sample of monolingual and bilingual university students and found that bilinguals were better at evaluating their own proficiency and had higher working memory as measured by a reading span task.

Others have failed to find a significant effect of multilingualism on working memory. Engel de Abreu (2011) found no differences in working memory (or fluid intelligence) between monolingual and simultaneous bilingual children between the ages of 6 and 8 that were matched on age, gender and SES. Kaushanskaya and Marian (2009) found no significant differences in working memory between English-Spanish bilinguals (evenly split into early and late bilinguals) and English monolinguals, but did find that only early bilinguals showed a correlation between phonological measures of working memory and

word learning performance. This could imply that these bilinguals are unique in applying this kind of working memory to the learning of new (phonologically unfamiliar) words, where in other groups phonological working memory abilities are more restricted to their native language's phonology. For any study the ability of the tasks used to represent an underlying construct needs to be carefully considered, and some of these studies (including those which failed to find an effect of multilingualism on working memory and some of the positive findings) used a limited range of tasks that might have been insufficient for detecting differences. In particular, Kaushanskaya and Marian (2009) measured working memory using only a non-word repetition task while Engel de Abreu (2011) used simple and complex span tasks. The relative lack of studies examining the effects of multilingualism on working memory, and the inconsistency of the instruments and findings in these studies, underlines the need for a methodologically robust investigation into the issue.

Multilingualism beyond two languages

None of the studies considered so far included speakers of more than two languages as a distinct group, and none of them consider (theoretically or empirically) what the effects of multilingualism on executive function beyond a second language might be. The expected effect of proficiency in an additional language (after the second) depends on one's theory for why bilingual advantages exist in the first place. The increased interference from additional languages would suggest a necessity for enhanced inhibitory control for speakers of more than two languages, while an additional option would mean that language selection would have to be performed with heightened attentional capacities. However, whether the (undoubtedly) greater demands on these mechanisms would lead to further advantages also depends on how one conceives of the manner in which the mechanism affects executive functions. If one holds a strictly dose-related, linear conception of how multilingual experience augments executive function then the presence of a third language should lead to additional advantages (J. Diamond, 2010). On the other hand, it could be the case that while speaking a second language generates advantages, speaking a third creates excessive conflict

and impairs general language development and cognitive performance (Brito et al., 2014)⁵. If there is a certain threshold of proficiency (or exposure) that is needed for multilingual advantages to manifest (see Sagasta Errasti, 2003 for more on the threshold hypothesis), and trilinguals generally have lower proficiency (or exposure) in their second and third languages⁶, then it would be expected that trilinguals will not have the advantages that bilinguals do (Brito et al., 2014).

Either of these scenarios would have worrying consequences for a great number of existing monolingual-bilingual studies that define bilinguals either as speakers of two languages (not bothering to find out or report whether participants speak additional languages) or speakers of two or more languages. If use of a third language confers advantages on speakers beyond those of speaking two then this research will have conflated these separable effects and overestimated the effect of speaking a second language. If speakers of three languages are at a disadvantage, on the other hand, this would imply that the majority of research has underestimated the beneficial effects of speaking two languages by unthinkingly including speakers of three (or more) languages.

The few studies to have explicitly researched speakers of more than two languages have been ambivalent in their results. Brito et al. (2014) found no significant difference between trilingual and monolingual infants in their ability to access memories flexibly in the face of changing contexts and retrieval cues, despite finding an advantage for bilinguals over monolinguals. They also found no differences between the groups for working memory, but acknowledge that a solitary working memory task may have been insufficient to capture differences. Poarch and Van Hell (2012) found an advantage for both bilingual and trilingual children (aged 5 to 8) over monolinguals in Simon and ANT tasks (but no difference between

⁵ This position, as applied to the developmental situation, might imply that there is a certain threshold of conflict among languages, below which the child is challenged in ultimately beneficial ways as they exercise their executive function in managing multiple languages, and above which the child is unable to learn to distinguish between languages and so suffers in their linguistic development and is not able to reap the executive function rewards of managing languages.

⁶ The assumption that trilinguals would necessarily have lower proficiency in their additional languages is challenged by research that finds that there is generally a strong correlation between proficiencies in additional languages, and that in becoming bilingual a person learns the general process of language learning and develops a metalinguistic awareness that can then be applied to the learning of further languages (Kahn-Horwitz, Schwartz, & Share, 2011).

the multilingual groups). Cockcroft et al. (2017) compared monolingual and trilingual young adults (categorised using both AOA and self-reported proficiency) and found working memory advantages for the trilingual group before and after controlling for SES. Kavé, Eyal, Shorek and Cohen-Mansfield (2008) found that general cognitive task performance in older adults improved with the number of languages spoken (up to four) after controlling for education.

In a study on French-German bilingual adults, Heidlmayr et al. (2014) found that the bilingual advantage in Stroop tasks was reinforced by the use of a third language. Chertkow et al. (2010) looked at the effect of bilingualism and trilingualism in augmenting cognitive reserve in older adults. They found no effect for speaking two languages but that speaking three had a protective effect, with later onset of symptoms and diagnosis of Alzheimer's disease. The authors also included both local (Canadian) and foreign multilinguals to address criticisms surrounding multilingual advantage studies which conflate culture effects with multilingual effects by using immigrant populations for multilingual groups. In immigrant but not Canadian multilinguals they found that speaking two languages significantly improves cognitive reserve.

Multilingualism as a continuous variable

The multi-faceted and heterogeneous nature of multilingual experience, and subsequent difficulty in modelling it, has been cited as a reason for the often inconsistent results in the literature (Yow & Li, 2015; Sabourin & Vinerte, in press). Some believe that this difficulty is compounded by a group comparison approach and can only be properly addressed by thinking of multilingualism as a continuous variable (along a spectrum) (Luk & Bialystok, 2013; Bialystok, 2016). Incera and McLennan (2017) argue that researchers commonly dichotomise (or otherwise group) language ability with little conceptual or statistical justification, and that treating groups as homogenous entities leads to a loss of information about individual differences, and can both decrease the chance of finding an effect when there is one and increase the chance of finding an effect when there isn't (Type I and Type II error, respectively). As well as reflecting the differences between multilinguals better than group

comparison can, the continuous approach is also uniquely able to test which aspect of multilingualism is necessary or particularly important for cognitive advantages. Using designs based on multiple regression analysis, many of the studies discussed below incorporate several aspects of multilingual experience.

Findings indicating significant developmental differences between simultaneous and early sequential bilinguals (Sabourin and Vinerte, in press) emphasise the importance of a fine-grained distinction between different AOAs. A continuous AOA variable may not capture the differences between simultaneous and early sequential bilinguals as well as treating each as a separate group, but will at least differentiate between these speakers due to the discrepancy in the ages when the acquisition of the language began. Luk et al. (2011) ran a model that included AOA as a continuous variable across their entire sample and found a significant correlation between a lower age and a lower response time cost on incongruent trials. Continuous modelling also enables a more fine-grained analysis of the effects of balance (in proficiency, exposure or use) among a multilingual's languages.

Tse and Altarriba (2014) used a standardised test of proficiency in (second language) English as a continuous measure of bilingualism in children aged five to nine. They looked at associations between these scores and measures of attentional control and found that second language proficiency scores were associated with superior conflict resolution and working memory capacity, after controlling for gender and SES. Yow and Li (2015) modelled multiple continuous measures of bilingualism capturing the AOA of second languages and the extent of balanced proficiency and usage across two languages. They tested all the components of executive function as specified by Miyake et al. (2000) to see whether the effects of these linguistic factors are function-specific (to one of the components of executive function) or executive function-general. The authors studied 72 English-Mandarin bilinguals who differed in the language variables described but who all had the same education level, culture, ethnicity and age group (18 to 25 years old) so that results could not be attributed to the confounding factors discussed above. The language indicators were created by subtracting the lower proficiency (and proportion of time using a language) from the higher, and by calculating an AOA in years variable, with all information coming from a self-report

questionnaire. Second language AOA was found to significantly predict the interference cost in Stroop tasks, while a latent balance variable based on both proficiency and use was associated with a smaller interference cost in Stroop tasks and a smaller mixing cost in a task-switching task. On the other hand, no significant effects on resistance to detractors and information updating and monitoring (working memory) were found.

In one case the two novel approaches described above have been combined and multilingualism in up to three languages has been modelled continuously. Incera and McLennan (2017) studied the extent to which balanced exposure among the languages of a multilingual (speaking two or three languages) is responsible for executive function advantages. Their continuous measure comes from self-reported information on the proportion of time that multilinguals are exposed to their less frequently spoken language (for bilinguals) or the sum of their lesser-spoken languages (for trilinguals). The score ranges from 0 to 50 so that speakers of more than two languages cannot have a higher score than bilinguals' maximum, and it was found that a higher degree of multilingualism and a lower age (also modelled along a continuum) were positively associated with Stroop (inhibitory control) but not Flanker (conflict management) performance. No interaction between age and multilingualism was found, meaning that there was no evidence of multilingualism contributing to cognitive reserve.

This study follows this novel approach and uses continuous measures of balance in proficiency in up to three languages. In so doing, the theoretical and practical reasons for modelling multilingualism continuously are heeded, and the near uncharted territory of third language effects is illuminated by distinguishing between speakers of two and three languages and by including third language proficiency information in the aforementioned balance variables. Additionally, the body of evidence for AOA effects, even at very early ages, led to the inclusion of a continuous measure of AOA in months.

Conclusion

Multilingual advantages in many executive functions have been found, and the theorised unity among these functions (Miyake et al., 2000) means that there should be accompanying multilingual advantages in working memory (Bialystok, 2009), an expectation that is also supported by recent theorising about the role of working memory in multilingual processes. However, findings for the relation between multilingualism and working memory have been erratic, with some positive significant findings (e.g. Tse & Altarriba, 2014; Blom et al., 2014; Morales et al., 2013) and some non-significant results (e.g. Engel de Abreu, 2011; Yow & Li, 2015; Brito et al., 2014). Inconsistencies in the tasks used and the ages at testing make these results difficult to compare, however, and could be driving differential results.

This literature review has revealed that while there has been a wealth of valuable and insightful research comparing bilinguals and monolinguals, the effect of additional languages (after two) has been studied very little and largely remains a mystery. The emphasis in the literature is increasingly on determining which aspect of multilinguals' heterogeneous experience is driving the development of cognitive advantages. Recent approaches have begun to argue for continuous measures of multilingualism, and a promising avenue of research within this approach has focused on balanced multilingualism. The age at which multilinguals are tested has been shown to influence results, and it is clear that more nuanced and precise multilingual advantage research needs to be performed with young adults. A need to factor in influential confounding variables such as SES and parental education has been established, as well as a need to perform research in cultural and linguistic contexts other than the typical American and European ones.

This study aims to contribute to the body of knowledge surrounding multilingual advantages by responding to the pressing questions and uncertainties identified above. The central research question of this study asks whether multilingualism leads to advantages in working memory in young adults. Based on the reviewed literature it was predicted that more multilingual participants would perform better in all working memory tasks, but especially visuospatial tasks (because of a history of multilingual disadvantages in linguistic tasks). The

lack of research comparing the effects of different aspects of multilingual experience makes a complementary question important: if there are advantages, which aspect of multilingual experience determines their development? The focus was on the following aspects of multilingualism: balance in proficiency among languages, balance in different kinds of proficiency within languages, the number of languages spoken and the age at which a person started becoming multilingual. No specific hypotheses about the importance of one factor relative to another were made due to there being insufficient research studying the relations among them on which to base these predictions. Another complementary question concerns whether SES, which is predicted to be correlated with working memory, has independent effects on working memory, and how this effect interacts with the effect of multilingualism.

Performing such a study in the context of South Africa addresses several of the shortcomings identified in the existing literature. The effect of multilingualism in South African (and, more broadly, non-Western) languages has been insufficiently studied, and the prevalence of multilingualism in many languages in the country means that both the effect of multilingualism in these particular languages and the effect of multilingualism in more than two languages can be studied simultaneously. Further, the particular effects of multilingualism in a developing country context where both SES and culture are highly variable addresses the historical lack of research in this kind of context and examines the influence of these factors. The restriction of the study to young adults ensured that the effects of prolonged experience in multiple languages could be studied, and meant that this study can address the lack of clarity concerning multilingual advantages in this age group.

Chapter Two

Methods

Research Design

The following research design was used to seek answers to the research questions described in the previous chapter. This design was non-experimental as there was no researcher manipulation of any independent variable among different participants or groups. Instead, a cross-sectional design was employed, with a range of data on pre-existing factors collected (at one time point only) from participants. The data collected was quantitative in nature and was analysed using a quantitative, correlational design. Some aspects of the research design were constrained by this study's use of previously collected data from an overarching project investigating multilingualism and working memory (see Sample and Sampling).

The dependent variables were four different components of working memory: verbal storage, verbal processing, visuospatial storage and visuospatial processing. The key independent variables modelled multilingualism by measuring balance in proficiency across and within a person's languages, whether a person speaks three languages or two, and the AOA of a person's second language, and these language variables were complemented by an SES score. The working memory variables were integer variables, the third language variable was a binary indicator variable, the SES score was an ordinal variable, and the balance variables were continuous. Ordinary Least Squares was used for all regressions.

This study was largely exploratory in nature for the following reasons. First, the broad nature of the variables involved in answering these questions, including working memory, multilingualism and SES, are all conceptually complex and any definition thereof will necessarily be contentious. The ways in which these relate to each other are unclear, all the more so for their inconsistent definitions, and often can only be explored in a holistic manner whereby several relationships are tested simultaneously. Second, so much of this research is entering new territory where there is neither sufficient empirical precedent nor sufficiently

comprehensive theory by which to determine specific directional hypotheses. While the domain-general nature of tasks in which previous multilingual advantages have been found supports the prediction that multilinguals would perform especially well in visuospatial working memory, there is insufficient evidence with which to predict which aspect of multilingualism would most determine these advantages. Therefore, a design was chosen which controlled for as many aspects as possible but which maintained broad research questions and looked for significant relationships among the key factors.

The ability of multiple regression analysis to model the effects of one variable while controlling for the effects of others was exploited so that this exploratory design could be implemented, with the effects of several aspects of multilingualism and that of SES estimated simultaneously. The specific effect of speaking a third language was incorporated while simultaneously modelling multilingualism continuously in the form of two balance variables capturing balance in proficiency between and within a person's languages, respectively. In so doing the lack of knowledge about the effects of speaking a third language was directly addressed while the benefits of modelling multilingualism continuously were harnessed. Other regressions further exploited this capacity to directly compare the effects of someone's proficiency (as captured in the balance variables) and the age at which they started becoming multilingual, thereby contributing to the debate over the relative importance of each factor. Despite findings indicating independent effects of SES and multilingualism on executive functions (Blom et al., 2014), the high correlation between linguistic background and economic outcomes in the South African context (Cornwell & Inder, 2008) makes it important to control for SES in order to isolate the effect of multilingualism.

Before multiple regressions were run descriptive statistics of the sample's demographic characteristics, languages spoken and other key factors were calculated. Thereafter preliminary analyses used correlation analysis, exploratory factor analysis (EFA) and bivariate regression analysis to examine pertinent relationships among the variables included in regressions as well as other variables of interest.

A series of philosophical positions underlie this research approach, concerning the object of study, the way this object is measured, and the way in which these measurements are analysed, respectively. The study of working memory, conceived of as a subset of executive function, assumes a philosophical position in which cognitive capacities can, at least partially, be conceived of as distinct modules or components. Further, the models of working memory that this study relies on (in terms of measures and interpretation) assume a conception of memory as comprising both storage and processing. The idea of things being *stored* in memory is itself critiqued by some who deny that something (or a representation thereof) needs to be (or even could be) stored in order to be remembered (Bennett, Hacker, 2003). Associated concepts such as memory encoding and retrieval, to the extent they depend on storage, are also implicated in this critique. Finally, the treatment of multilingualism as the extent of balance among different languages assumes a theoretical position which holds that people speak distinct language systems⁷.

Having established what the constructs to be analysed are (and the philosophical positions implied), the method of their measurement needs to be selected and justified. In our study, quantitative measures were used to represent these constructs in correlational analysis. The use of quantitative measures of psychological constructs is a traditional tool of the logical positivists, who believed not only that observations could be independent (of the researcher's beliefs and the research context) but also that *only* empirical observation could yield truth and determine universal facts (Schumacher & Gortner, 1992). The positivists stressed that nothing can exist in the mind without first existing in the senses, and so no claim can be valid unless it is based on sensory data (Zammito, 2004), a position firmly endorsed and entrenched by the behaviourists (Skinner, 1987). The claim of direct access to theory-neutral observation has long been subjected to serious critiques. While events may exist neutrally before perception, in the process of perception they are necessarily assimilated into existing conceptual structures and affected by the researcher's beliefs and frame of reference (Turner, 1968). What counts as an observation depends on the researcher's theories of mind and perception (Schumacher & Gortner, 1992). Post-positivism, in many ways a successor to

⁷ It is unclear whether the translanguaging position (Garcia & Li Wei, 2014) would undermine our entire research methodology, or just the rationale for and interpretation of it (with the reasons for the multilingual advantage resting not on the benefits of switching between language systems but on benefits from flexibly applying the sole dynamic language system to different socially-defined languages).

logical positivism, incorporated this criticism and posited that no observation can be independent of the theories and beliefs of the researcher, which determine both what counts as an observation and how it is interpreted (Zammito, 2004). Further, the concept of one objectively superior way of conducting science was rejected, meaning that many different methods of justifying a claim (including self-report and reason) are possible and that many socio-cultural and psychological aspects are important to justification (Zammito, 2004).

Consequently, this study was guided by the post-positivist paradigm, holding that no approach is a priori superior to another and that there are numerous ways of justifying a claim depending on the question at hand (Zammito, 2004). Subjective experience and intuition provide valuable insight into the processes of a multilingual mind, but are limited in the evaluation of differences in executive functions due to those functions' fast, automatic nature. Further, the multilingual advantage hypothesis rests on the existence of widespread cognitive advantages which can only be studied and isolated (after controlling for covariates) in a large sample quantitative design. On these grounds the use of quantitative measurements are justified for the investigation of the research question, but the specificity of the results must be considered. An awareness of the influence of the researcher's beliefs on every stage of the research process, and the implications for the generalisability of results, must be maintained (Schumacher & Gortner, 1992). Despite their more considered approach, post-positivists' use of quantitative designs is still criticised by qualitative and interpretative researchers as unjust, simplistic reduction of mental phenomena and properties to numbers and scores (Yanow & Schwartz-Shea, 2015). Some (e.g. Hyltenstam, 1988) choose to use qualitative measures of multilingualism, and the question of whether it is possible to measure multilingualism and its effects quantitatively, and what kind of information may be ignored by such measures, is returned to in Chapter Four.

Finally, taking the use of quantitative measures as a given, a position needs to be taken regarding the proper method of analysing quantitative information. The use of correlation analysis, specifically multiple regression, implies a frequentist approach to data analysis (as opposed to a Bayesian framework, for example).

Sample and Sampling

This study was based on a convenience sample of 189 young adult participants. For the current study 68 participants were interviewed in total, but of these one withdrew altogether while seven elected not to complete the AWMA (see instruments described below). A further five were dropped from the sample for having missing LEAP-Q proficiency information in their third language, resulting in the final pool of 55 participants. These participants were pooled with previously collected data from 106 participants collected in 2013 and 39 participants collected in 2014, resulting in a sample of 200. Of these, 11 participants were dropped for reporting only speaking one language so that the effect of within-multilingual differences (such as speaking a third language) could be clearly examined. This left the final sample of 189 participants.

Although this study did not use a between group design comparing speakers of different numbers of languages, different calls were made for less multilingual English mother-tongue speakers (the 39 participants from 2014) and speakers of two languages (the 55 participants from the current study) or three languages (the 106 participants from 2013) with South African language mother tongues. These calls were not completely rigid (and in the sampling process for the current study many participants who spoke more than two languages were interviewed) but were designed to ensure that a wide distribution of less and more multilingual participants were interviewed.

The sampling procedure for this study followed the previous data collection procedures (in 2013 and 2014) as closely as possible. As in previous years, this was a convenience sample of young adult university students from the University of the Witwatersrand (also known as Wits). All tests and questionnaires were administered in English, but the fact that all participants were studying in the medium of English (Wits' only medium of instruction) was grounds to believe that they were sufficiently proficient in the language to have understood the questions and tests administered, and their experience as students should have ensured they were test-wise and computer- and questionnaire-literate. To attract potential participants, announcements were placed on the Wits online portal, Sakai, and made in person

at first-year Psychology lectures. Calls focused on first-year Psychology students who were offered an incentive of a 1% course mark increase for participation in the study. The call described the subject of the study as ‘Language and Working Memory’ but gave little detail about the study’s hypotheses to avoid framing effects. Participants who showed interest in participating were sent documents with information about what participation and their consent would entail and were asked to confirm their participation (see Appendix A), after which an interview date and time were arranged.

As described above, a broad range of more and less multilingual participants were sought, and speakers of a number of South African languages including the Nguni and Sotho languages as well as English, Afrikaans, xiTsonga and tshiVenda were represented in the sample. The study also aimed for a sample with substantial variation in SES to reflect the reality of South Africa being one of the most unequal countries in the world (Leibbrandt et al., 2012). Descriptive statistics covering participants’ age, gender, preschool attendance, SES and the languages they speak are presented in Chapter Three.

Instruments

To implement the design of this study and adequately describe the sample, instruments were used to collect information about: 1) participants’ languages spoken and proficiency and AOA for each language; 2) participants’ demographic characteristics and SES background; 3) participants’ verbal and visuospatial intelligence; and 4) participants’ working memory capacities and abilities.

Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007)

The self-report Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007) was administered to determine the number of languages spoken by participants and their proficiency in reading, speaking and comprehension and AOA for each language. The LEAP-Q collects a range of information including a list of languages known in order of

dominance and of acquisition, the share of time a participant is exposed to each of their languages (relative to total language exposure) and preferred language choice for written and spoken contexts, along with information about hearing and visual impairments and learning and language disabilities. For each language further information is collected on the age at which acquisition began (in months), fluency was achieved, reading began, and reading fluency was achieved. On a scale from 0 (none) to 10 (perfect) respondents rate their level of proficiency in speaking, comprehension and reading for each language, before rating the contribution of a variety of factors to their learning of the language, and the proportion of time they are exposed to the language through a variety of media.

In multilingual research there is a general lack of “uniform assessment instruments” (Marian et al., 2007, p. 941), and in less studied languages such as those in this study this problem is especially acute. The LEAP-Q was created as an alternative to improvised self-report questionnaires of language proficiency which abound in the literature. It has been shown to have good internal and criterion-based validity (based on its ability to predict a range of standardised behavioural measures of language ability) (Marian et al., 2007) but has not been standardised in the South African context.

Demographic Questionnaire

To determine whether working memory differences among people are truly attributable to multilingualism it is important to control for other factors that are likely to affect working memory, especially where these may be correlated with language experience. A brief demographic questionnaire was administered (see Appendix B) which covered the educational history of the respondent (including whether they had attended preschool), and the educational attainment of their parents. Additionally, information was collected about their childhood home including the number of caregivers present and the presence of 26 different assets and factors which were used to create an asset index (created by SAARF, 2001) to proxy for SES. The items for this index included assets such as kitchen and household appliances, computers, and entertainment systems, as well as access to municipal services (such as water and flush toilets) and metropolitan or rural inhabitation. The assets a

participant reported as present were then weighted and summed to come up with a number which placed them on an ordinal scale from 1 (lowest) to 10.5 (highest) of SES groups. This scale has not been standardised or validated.

Wechsler Adult Intelligence Scale - 3rd Edition (WAIS-III; Wechsler, 1997)

Two verbal subtests from the Wechsler Adult Intelligence Scale - Third Edition (WAIS-III; Wechsler, 1997) acted both as measures of verbal intelligence and as objective measures of English proficiency.

The Vocabulary subtest examines knowledge of English words ordered from easiest to hardest, with word selection and difficulty determined by the frequency of words in the South African context. For each word the interviewer spoke the word out loud while showing the participant a card with the word written out. The participant then had to define the word (and choose between multiple definitions if they provided them) before the interviewer grades their response as 0, 1 or 2 based on how well the word has been defined.

The Similarities subtest examines respondents' ability to describe the conceptual similarities between words. A series of word pairs was presented with the conceptual similarity between the two words becoming progressively more abstract and difficult to perceive. For each item interviewers spoke the two words and asked in what way they were alike before grading participants' responses as 0, 1 or 2 based on how well and unambiguously the connection has been described.

Both tests are highly dependent on English lexical and conceptual knowledge, and the Similarities test in particular measures conceptual reasoning and processing without placing a heavy load on the storage component which is integral to working memory. These tests are useful in revealing relationships among aspects of verbal intelligence and other verbal tests and as more objective measures that can be correlated with variables based on self-reported data.

The version of the WAIS-III has been adapted and standardised for the South African context (Claasen, Krynauw, Paterson, & Wa ga Mathe, 2001).

Automated Working Memory Assessment (AWMA; Alloway, 2007)

Finally, the full Automated Working Memory Assessment (AWMA), a battery of 12 tests evenly divided between verbal storage, visuospatial storage, verbal processing and visuospatial processing (Alloway, 2007), was used to measure the working memory of participants, the outcome variable of this study. The simple storage tests involve recall of strings of data and do not involve a heavy processing component, whereas the processing tasks require the memorisation and recall of data while performing some sort of manipulation or processing. The ordering of the tasks is jumbled so that tasks from any given component are not performed in sequence and do not occur solely at the beginning or end of the test. The division of components is based on the Baddeley and Hitch (1974; Baddeley, 2000) model of working memory (discussed in Chapter One) and mimics its separation of working memory into verbal (phonological loop) and visuospatial (sketchpad) components, while the processing tasks correspond more with the central executive.

In the AWMA, each task contains a series of stages of ever longer spans and begins with a small span (generally between one and four based on the interviewer's judgement of what the interviewee can easily handle). Within a stage, a participant has to correctly recall four spans to progress to the next stage (where the span is longer by one) and cannot recall the span incorrectly more than twice or the task ends and their score is calculated. The audio or visual presentation of spans and score calculation is computerised while the interviewer is responsible for evaluating the correctness of participants' recall and processing responses.

Verbal storage is tested using the Digit Recall, Word Recall and Nonword Recall tasks which are all span tasks involving the recall of a sequence (of digits in the first instance, words in the second, and English-like non-words in the third) in the order in which it was heard. Performance in non-word recall tasks is believed to be affected by long-term lexical

and sub-lexical knowledge (Engel de Abreu, 2011) in the language the non-words are based on.

Visuospatial storage is measured using the Dot Matrix, Mazes Memory and Block Recall tasks, all of which involve replicating a pattern or shape in the order presented. The Dot Matrix task involves the recall of the position of a sequence of red dots in a 4x4 grid. The Mazes Memory task involves the recall of a path out of a maze as the maze gets progressively larger and the path progressively more convoluted. In the Block Recall task a number of cubes are scattered irregularly on a board and the participant needs to recall which cubes are pointed at (by an arrow) in sequence.

Verbal processing is measured using the Listening Recall, Counting Recall and Backwards Digit Recall tasks, all of which involve the recall of a string of figures or words after performing a calculation or manipulation. The Listening Recall task involves the respondent hearing a series of sentences and judging if each is true or false while remembering the final word of each sentence. The Counting Recall task involves the respondent counting the number of circles in a series of images consisting of circles and triangles before recalling the number of circles in sequence at the end. The Backwards Digit Recall task replicates the Digit Recall task but with the respondent tasked with recalling a progressively longer span of digits in reverse order.

Finally, visuospatial processing is measured by the Odd One Out, Mister X and Spatial Recall tasks in which a series of locations need to be recalled after answering questions that require visuospatial manipulation. In the Odd One Out task the respondent must say which shape (of three) is the odd one out in a sequence of trials before recalling the position of odd shapes in the sequence presented. The three shapes are always presented side by side from left to right and are two-dimensional. In the Mister X task the respondent must say whether two figures (rotated at different angles) are holding a ball in the same (or a different) hand before recalling the location of the balls in the sequence of trials. The figure on the left is always the right way up and is holding the ball in their left or right hand, while the figure on the right is rotated so that the ball can be in any one of six places. In the Spatial Recall task

two shapes are presented and the respondent must say whether the shape on the right which has a red dot above it and has been rotated is the opposite or the same as the one on the left, before recalling the locations of the dots in sequence. The dot in the figure on the right can be in any one of three places.

This battery of tests provided a measure of four components of working memory, each of which was used as the dependent variable in respective regressions. The presence of three tasks for each aspect meant that the measure of the component was not overly sensitive to the particularities of any one task, and provided a high enough number of items for the performance of factor analysis.

In a study with English children the AWMA was found to have good construct stability (with scores relatively stable over the course of a school year, and good diagnostic validity, with high convergence with other measures of working memory (Alloway, Gathercole, Kirkwood, & Elliott, 2008). The instrument has not been standardised or validated in the South African context, however.

The lack of standardisation or validation for the AWMA and LEAP-Q in South Africa would be a serious concern for their use in diagnostic or norm-setting purposes, but is not as grave an issue for the present research purposes. Nonetheless, this fact was borne in mind when interpreting results and some steps were incorporated into data analysis to examine the tests' suitability in the South African context (see Chapter Three).

Procedure

Interviews and tests were conducted by myself and 17 Honours students, who used the data collected (combined with the previously collected data) for research reports centred around working memory. All interviews were conducted one on one in a private, quiet room. The standardised (and in the case of the AWMA computer-administered) nature of the instruments means that the interview process would be similar among interviewers. Further, neither the cognitive tasks of the AWMA and WAIS-III nor the relatively neutral information about language experience collected by the LEAP-Q should yield different responses based on the interviewer. On the other hand, responses concerning assets and parental education levels collected in the demographic questionnaire may be more sensitive to interviewer characteristics. To guard against interviewer effects the ordering of the tasks and the instructions delivered were carefully standardised among interviewers during thorough training sessions. On arrival participants were given an information sheet and consent form, after which the LEAP-Q was administered. Then the WAIS-III tests were administered before participants were given a break to avoid pronounced effects of fatigue on responses. Thereafter participants performed the AWMA before ending with the demographic questionnaire. All tests were administered in English, and practice examples were used in the WAIS-III and AWMA to ensure the participant understood the demands and goals of each task. In total, the interview and test process took between 90 and 120 minutes.

Ethical Considerations

Before conducting any interviews ethical approval for this study and its data collection were obtained from the Humanities Research Ethics Committee (HREC Non-Medical; Protocol Number: MPSYC /17/003 IH, see Appendix C).

All participants were above the age of 18 and were therefore legally capable of giving consent, and were given an information sheet detailing what the data collection process would entail and how the collected data would be used. All participants signed a consent form which echoed this information and stressed the confidential and voluntary nature of the study,

with participants free to leave at any point without negative consequences. Although this never occurred, in case of participants experiencing distress, the interview would have been stopped and participants referred to the Emthonjeni Community Psychology Clinic or the Student Centre for Careers and Counselling Development Unit, both of which offer free counselling.

The information collected in questionnaires and the cognitive tests administered were not of a sensitive nature and so it was deemed unlikely that the interview procedure would evoke painful associations in participants. Nevertheless, the confidentiality of participants' responses in any publication or data set was considered to be highly important for ethical reasons and in making participants respond to questions and perform tests more comfortably. While complete anonymity is not possible in face-to-face interviews, information for each participant was anonymised before the data analysis phase to ensure confidentiality. In the data set, observations were identified by a number related to the interviewer and never by the interviewee's student number or name. Participants who indicated interest in staying informed about the study's results will be sent this research report, but no personalised results will be sent to anyone.

The data, without any identifying features for observations, will be kept after this study (in a locked cupboard and in password protected digital files) and potentially used in future research (with participants' consent).

Chapter Three

Results

Introduction

In Chapter One the merits of modelling multilingualism continuously rather than discretely, and in particular of modelling balance among languages, were discussed. The research design, outlined in Chapter Two, revolved around examining the effect of balanced proficiency among a multilingual's languages on the different components of working memory. The central form of analysis used in answering this question used a series of multiple regressions with working memory components acting as dependent variables and a control for socioeconomic status accompanying the language variables as independent variables. In this chapter the preliminary steps which were taken to enable the implementation and interpretation of these regressions are described before the multiple regressions themselves are presented. These preliminary steps are divided into two sections: the first contains descriptive statistics for all of the variables which were crucial in running and interpreting preliminary and multiple regression analysis, as well as a description of the process by which multilingualism and working memory variables were derived from the variables produced by the instruments; the second contains the results of preliminary analyses involving the analysis of correlations, EFA and bivariate regression analysis, which tested assumptions surrounding key variables and looked for simple relationships among them. At each step the results of statistical analyses are presented in tables and attention is drawn to significant findings.

Descriptive Statistics and Derived Variables

Before using variables in regression analysis it is important to understand how the various characteristics they measure are distributed in the sample. This section is divided into three subsections, with each describing the characteristics of a different group of variables used in this study. The first describes the language data which form the key independent variables; the second describes the cognitive measures used in the study, including the working memory variables which act as dependent variables; and the third describes the

demographic variables which act as controls in regressions or provide crucial background information. In each subsection the variables generated by the instruments and the process by which study variables were derived from these are described. Later, in the preliminary analysis section, the properties and validity of these derived variables are explored.

Language variables

The LEAP-Q captured a range of self-report data about participants' language experience and capacities. In this section three types of information are discussed: the languages which participants speak, their proficiency in these languages, and the age at which they began acquiring these languages. As described in Chapter Two, the sample of this study comprises 189 young adults, all of whom reported speaking two (n=54) or three (n=135) languages. The languages spoken by participants varied a great deal, as is common in South Africa (and especially Johannesburg) (Makalela, 2013), and Table 1 presents the eight most frequently reported first, second and third languages. The particular languages a person speaks do not play an explicit role in our research design but this information is useful in contextualising our study.

Table 1

Frequencies of first, second and third languages spoken

Languages by frequency	First language	Second language	Third language
1	English (36)	English (108)	isiZulu (34)
2	isiZulu (34)	Afrikaans (28)	English (33)
3	sePedi (34)	isiZulu (18)	sePedi (12)
4	tshiVenda (19)	seSotho (9)	seTswana (12)
5	seTswana (16)	seTswana (7)	seSotho (10)
6	xiTsonga (13)	isiXhosa (4)	isiXhosa (5)
7	siSwati (11)	tshiVenda (3)	xiTsonga (5)
8	seSotho (10)	Hebrew (3)	Afrikaans (4)
9	Other (16)	Other (9)	Other (20)
<i>N</i>	189	189	135

Note. Frequencies for each language are in parentheses.

The first column shows that more participants spoke English ($n = 36$) as their mother-tongue than any other language, followed by isiZulu ($n = 34$). Beyond these the Sotho languages were strongly represented with sePedi ($n = 34$), seTswana ($n = 16$) and seSotho ($n = 10$) the mother tongues of many participants. The similarities between these languages and the way in which they people switch between them in urban varieties has led many to claim that they can and should be treated as one language (Makalela, 2009). The second column reflects the unparalleled prominence of English as a second language ($n = 108$), followed by Afrikaans ($n = 28$) and isiZulu ($n = 18$). Finally, isiZulu is the most common third language ($n = 34$), followed by English ($n = 33$) and the three Sotho languages (collectively $n = 34$). While most of the languages reported were South African languages, a few reported speaking languages which are not common in (or indigenous to) the country, such as Mandarin or Hebrew.

For each of a participant's languages, the LEAP-Q collected self-report data on proficiency of three kinds (speaking, reading and understanding), rated on a scale from zero (none) to 10 (perfect), as well as the AOA for each language (in months). The means, standard deviations and ranges for each type of proficiency (and the average of the three) are presented for first, second and third languages in Table 2, alongside information for mean English proficiency and the balance variables (described below).

Table 2

Descriptive statistics for proficiency, age of acquisition (AOA) and balance variables

Language category	Variable	<i>M</i> (<i>SD</i>)	Range
First language	Reading	8.41 (1.81)	2 - 10
	Speaking	8.95 (1.28)	5 - 10
	Understanding	9.09 (1.09)	5 - 10
	Mean proficiency	8.82 (1.21)	4.67 - 10
	AOA (in months)	13.06 (15.55)	0 - 72
Second language	Reading	7.49 (2.41)	0 - 10
	Speaking	7.41 (2.23)	1 - 10
	Understanding	7.74 (2.15)	1 - 10
	Mean proficiency	7.55 (2.12)	1.33 - 10
	AOA (in months)	76.03 (34.32)	0 - 180
Third language	Reading	5.13 (3)	0 - 10
	Speaking	6.34 (2.18)	1 - 10
	Understanding	6.99 (2.14)	1 - 10
	Mean proficiency	6.16 (2.04)	1.67 - 10
	AOA (in months)	119.47 (63.92)	0 - 240
English	Mean proficiency	8.25 (1.44)	4.33 - 10
Balance	Within-language	-.74 (.45)	-2.6 - 0
	Between-language	-1.61 (1.08)	-5.4 - 0

Notes. $N = 189$ for balance variables and first and second language variables. $N = 185$ for English mean proficiency. $N = 135$ for third language variables.

It is clear from these statistics that participants' average proficiency was highest in their mother-tongue ($M = 8.82$, $SD = 1.21$), and that they were progressively less proficient in their second ($M = 7.55$, $SD = 2.12$) and third ($M = 6.16$, $SD = 2.04$) languages which were acquired later in life (with all differences significant at $p < 0.0001$ in paired t-tests). As expected, the mean AOA climbed steadily from first language ($M = 13.06$, $SD = 15.55$) to second ($M = 76.03$, $SD = 34.32$) and third ($M = 119.47$, $SD = 63.92$) (with all differences significant at $p < 0.0001$ in paired t-tests). There were inequalities among the kinds of proficiency as well, with reading substantially lower than other forms of proficiency in

participants' first language (speaking: $t(188) = -5.17, p < 0.0001$; understanding: $t(188) = -6.1, p < 0.001$) and, especially, third language (speaking: $t(134) = -4.93, p < 0.0001$; understanding: $t(134) = -7.75, p < 0.001$). The differences between mean speaking and understanding proficiency, on the other hand, appeared smaller across languages. Because of its importance for interpreting results of tasks carried out in English, the table gives the mean proficiency (across the three proficiency types) for English ($M = 8.25, SD = 1.44$). Four participants were missing English proficiency information, so this mean is calculated from a sample of 185.

Rather than using these proficiency variables as they are in absolute form, this data was used to create two balance variables: a within-balance variable which measures the balance among the kinds of proficiency within each of a speaker's languages, and a between-balance variable which measures balance in kinds of proficiency across a speaker's languages. Because this approach revolved around balance within and between languages, the ordering of participants' languages (by proficiency or by AOA) did not have an impact on these variables. Thus, for the entirety of the study, languages were sorted according to AOA.

Beyond its theoretical importance and novelty, the use of balance variables which are calculated relative to a participant's own maxima has the advantage of increasing the comparability of self-reported scores across participants. While the LEAP-Q has been shown to have good internal and criterion-based validity (Marian et al., 2007), the comparability of self-reported scores across individuals is always questionable. One person's '9' may be another's '8' and so comparing the absolute values provided by participants may not be so revealing a strategy. However, a single person's self-reported scores are likely to be more directly comparable and internally consistent. For example, a person's rating their proficiency in reading '8' and in speaking '9' seems likely to convey real information about their speaking being more proficient than their reading. This assumption will be tested in the Preliminary Analyses section.

For the within-language balance variable, the differences between a person's maximum self-reported proficiency and their other proficiencies was calculated and summed. So, if a

person rated their reading as '9', their speaking as '8' and their comprehension as '7', their within-language score for that language would be $(-1) + (-2) = -3$. The mean of the within-language balances from the languages a person speaks was then taken as the final within-language balance variable to be used in regressions.

For the between-language balance variable, a similar approach was used, except that the maximum for each proficiency type (e.g. reading) across a person's languages, was calculated. Then the distance of the other reading score(s) from the maximum was calculated (and, for speakers of three languages, summed). This yielded a measure of how balanced a person's speaking, reading and comprehension ability are across their languages. The mean of these three scores was then taken as the final between-language balance variable.

The descriptive statistics for these two balance variables are included in the last two rows of Table 2. Given the way these variables are calculated, a score of zero represents perfect balance and the magnitude of negative scores represents the extent of imbalance. Both the within-language ($M = -0.74$, $SD = 0.45$) and between-language ($M = -1.61$, $SD = 1.08$) balance variables indicate that there was considerable imbalance (and variation therein) in participants' proficiencies within and across languages. As would be expected, there was far greater imbalance among proficiencies across languages than within languages ($t(188) = -11.27$, $p < 0.0001$).

Working Memory Variables

The main aim of this study was to examine the effect of multilingualism on components of working memory, which were measured using the twelve tasks of the AWMA. The descriptive statistics of all 189 participants' performance in these tasks are presented in Table 3. In all of these calculations scores scaled to a reference population and normalised around 100 (with standard deviation = 15) were used.

Table 3

Descriptive statistics of Automated Working Memory Assessment (AWMA) components and task performance

Component	Tasks	<i>M</i> (<i>SD</i>)	Range
Verbal storage		90.54 (11.12)	69 - 127.3
	Digit recall	89.78 (11.36)	69 - 136
	Word recall	86.31 (13.16)	64 - 135
	Nonword recall	95.52 (18.18)	55 - 136
Verbal processing		93.08 (10.88)	68 - 125.7
	Listening recall	95.12 (13.78)	70 - 139
	Counting recall	94.96 (15.40)	57 - 148
	Backwards digit recall	89.17 (14.31)	65 - 132
Visuospatial storage		93.04 (12.59)	64.3 - 133
	Dot matrix	96.75 (16.76)	63 - 146
	Mazes memory	95.72 (15.74)	58 - 134
	Block recall	86.64 (15.18)	60 - 135
Visuospatial processing		97.38 (12.28)	75.7 - 133.3
	Odd one out	98.46 (14.65)	69 - 134
	Mister X	98.57 (15.37)	61 - 136
	Spatial recall	95.11 (14.80)	72 - 137
Verbal intelligence			
	Vocabulary	11.30 (2.69)	3 - 19
	Similarities	9.93 (2.87)	5 - 19

Notes. *N* = 189.

As can be seen in the table, each of the four stipulated components of working memory was measured by three different tasks. The mean scores for all of the tasks are below 100, implying that average performance in our sample was somewhat lower than, but within one standard deviation of, the average performance of the reference population. Mean scores were especially low on the Word Recall ($M = 86.31$, $SD = 13.16$), Block Recall ($M = 86.64$, $SD = 15.18$), Backwards Digit Recall ($M = 89.17$, $SD = 14.31$) and Digit Recall ($M = 89.78$, $SD = 11.36$) tasks. A look at the ranges and the standard deviations indicates that performance on some tasks, especially the visuospatial Dot Matrix ($M = 96.75$, $SD = 16.76$) and Mazes Memory ($M = 95.72$, $SD = 15.74$) tasks, was more variable than others.

For the dependent variables in multiple regression analysis, scores for the four working memory components were calculated as the mean of the scores for the three tasks measuring that component. These component variables, with their descriptive statistics, can be seen in bold before each set of tasks in Table 3. Mean performance was highest for visuospatial processing ($M = 97.38$, $SD = 12.28$), lowest for verbal storage ($M = 90.54$, $SD = 11.12$), and intermediate for verbal processing ($M = 93.08$, $SD = 10.88$) and visuospatial storage ($M = 93.04$, $SD = 12.59$). Mean scores for verbal processing and visuospatial storage were not significantly different from one another, while all other differences among components were significant (with p-values ranging from 0.006 to 0.0001 in paired t-tests).

WAIS-III scores were scaled to a reference population whose scores were normalised around a mean of 10 with a standard deviation of 3. This means that, like the AWMA, the scores from each test are on the same scale and can be compared to one another and more easily combined and interpreted. Participants performed better and less variably ($t(188) = 8.35$, $p < 0.0001$) in the Vocabulary test ($M = 11.30$, $SD = 2.69$) than in the Similarities test ($M = 9.93$, $SD = 2.87$).

Demographic variables

Even if they are not explicitly included as controls in multiple regression, it is important to understand the demographic characteristics of the sample, in order to know exactly who is being studied and to assist in the interpretation of regression results.

Table 4

Descriptive statistics for demographic variables

Variables	<i>M (SD)</i>	Range
Female	.67	
Preschool	.78	
Age	20.41 (1.82)	18.1 - 32.5
SES score	7.92 (2.18)	2 - 10.5

Notes. $N = 189$ for all variables except for Female ($N = 188$). ‘Female’ gives the proportion of the sample that is female. ‘Preschool’ gives the proportion of the sample that had attended preschool. SES score = socioeconomic status score.

The majority of participants were female (67%). The mean age ($M = 20.41$ years, $SD = 1.82$ years) just above 20 years old. Seventy-eight percent had attended preschool. The SES score ($M = 7.92$, $SD = 2.18$), obtained from the LSM questionnaire (SAARF, 2001) varies on an ordinal scale (in intervals of 1 or 0.5) from 1 (lowest category) to 10.5 (highest category). The mean score of 7.92 suggests that participants were on the middle to high end of the SES spectrum, but the considerable SD of 2.18 means that there was a spread of participants from substantially lower and higher SES groups.

Preliminary Analyses

Three stages of preliminary analysis were performed. In the first correlations among key variables were examined, in the second EFA was applied to the AWMA tasks to explore the underlying factors driving task performance, and in the third bivariate regression analysis was used to determine the extent to which variables based on self-report information were capable of predicting objective language and verbal intelligence tests.

Correlations among key variables

Table 5 contains pairwise correlations among key variables, including the cognitive scores (for the working memory components and the verbal intelligence tests), the language variables (capturing within- and between-language balance, whether a person speaks a third language, the AOA of a participant's second language and their mean English proficiency), and finally the demographic variables (capturing age, gender, whether the person had attended preschool and their SES score). This information provides an initial understanding of how these variables relate to one another, and provides evidence of simple relationships which can aid in the interpretation of the more complex relationships revealed by multiple regressions analysis.

Table 5

Correlations among key cognitive, multilingualism and demographic variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Verbal st.	—														
2. Verbal pr.	.41***	—													
3. Visuospatial st.	.35***	.48***	—												
4. Visuospatial pr.	.24**	.54***	.69***	—											
5. Vocabulary	.39***	.18*	.04	.08	—										
6. Similarities	.31***	.15*	.04	.07	.67***	—									
7. Within balance	.09	.11	-.08	-.01	.16*	.1	—								
8. Between balance	-.11	.16*	-.04	.01	-.16*	-.16*	.26***	—							
9. Third lang.	-.12	-.03	.09	.06	-.15*	-.07	-.26***	-.22**	—						
10. AOA	.03	-.1	.07	.04	-.13	-.12	.1	-.28***	-.01	—					
11. English prof. ^a	.14	.06	-.04	-.03	.27***	.3***	.31***	-.08	-.16*	.09	—				
12. Age	-.06	-.07	.11	.02	-.08	.00	-.05	.08	.08	-.03	-.15*	—			
13. Gender ^b	-.06	-.13	-.27***	-.18*	-.02	-.07	.02	-.10	.00	.07	.17* ^c	-.28***	—		
14. Preschool	.10	.05	.01	-.01	.22**	.29***	-.07	-.19**	-.08	-.09	.14	-.16*	.12	—	
15. SES score	.22**	.04	-.12	-.06	.45***	.42***	.11	-.28***	-.24***	-.05	.3***	-.19**	.15*	.34***	—

Notes. $N = 189$

Verbal st. = verbal storage, Verbal pr. = verbal processing, Visuospatial st. = visuospatial storage, Visuospatial pr. = visuospatial processing, Within balance = within-language balance in proficiency, Between balance = between-language balance in proficiency Third lang. = whether the participant speaks a third language, AOA = age of acquisition of second language, Preschool = whether the participant had attended preschool, SES score = socioeconomic status score

^a $N = 188$, ^b $N = 185$, ^c $N = 184$ * $p < .05$ ** $p < .01$ *** $p < .001$

Only significant pairwise correlations, of which there were many, are described here. As would be expected, all of the working memory components were significantly and positively correlated with one another (with r ranging from 0.24 to 0.69 and p -values ranging from 0.001 to 0.0001), as were the Vocabulary and Similarities scores ($r = 0.67, p < 0.0001$). The Vocabulary and Similarities scores were only positively correlated with the verbal working memory components (with r ranging from 0.15 to 0.39 and p -values ranging from 0.04 to 0.0001), and not with the visuospatial components.

Within-language balance was positively correlated with between-language balance ($r = 0.26, p < 0.05$), suggesting that one of these types of balance make the other more likely and that those who are more balanced among proficiencies within their languages are also more balanced across their languages, and vice versa. On the other hand, speaking a third language was negatively correlated with both within- and between-language balance ($r = -0.26, p < 0.05$; $r = -0.21, p < 0.05$), implying that speakers of three languages had significantly lower balance in their proficiency within and across their languages. Second language AOA was negatively correlated with between-language balance ($r = -0.28, p = 0.0001$), implying that those who started becoming multilingual later achieved a lower level of balance in proficiency among their languages. English proficiency was positively correlated with within-language balance ($r = 0.31, p < 0.0001$), implying that proficient English speakers were balanced among proficiency types across their languages, and negatively correlated with speaking a third language ($r = -0.16, p = 0.03$), implying that those who spoke more than two languages were less proficient in English.

Speaking a third language was also negatively correlated with the Vocabulary test ($r = -0.15, p = 0.04$), implying that those who spoke more languages had a smaller English vocabulary. Proficiency in English was strongly and positively correlated with performance in the Vocabulary ($r = 0.27, p < 0.0001$) and Similarities ($r = 0.3, p < 0.0001$) tests, but was not correlated with any of the working memory scores. Given the research question of this study, one of the most important findings is that between-language balance correlated positively with verbal working memory ($r = 0.16, p < 0.05$). On the other hand, between-language balance was negatively correlated with both the Vocabulary ($r = -0.16, p = 0.03$)

and Similarities ($r = -0.16, p = 0.03$) tests. Within-language balance correlated positively with the Vocabulary test ($r = 0.16, p = 0.02$)

A higher age was correlated with a lower SES score ($r = -0.19, p < 0.05$) and chance of preschool attendance ($r = -0.15, p < 0.05$). Females had significantly higher SES ($r = 0.15, p = 0.04$) and were significantly younger ($r = -0.28, p = 0.0001$). Females had significantly lower visuospatial storage ($r = -0.28, p = 0.0002$), visuospatial processing ($r = -0.18, p = 0.01$) and higher English proficiency ($r = 0.16, p = 0.02$) scores. Participants with a higher SES score had significantly less chance of speaking a third language ($r = -0.24, p = 0.0009$) and of being balanced among their languages ($r = -0.28, p = 0.0001$), were more likely to attend preschool ($r = 0.34, p < 0.0001$) and had significantly higher Vocabulary ($r = 0.45, p < 0.0001$), Similarities ($r = 0.42, p < 0.0001$), verbal storage ($r = 0.21, p < 0.05$) and English proficiency ($r = 0.3, p < 0.0001$) scores. Preschool attendance correlated positively with Vocabulary ($r = 0.22, p = 0.002$) and Similarities ($r = 0.29, p = 0.003$) scores, and negatively with between-language balance ($r = -0.18, p = 0.01$). Finally, older participants had lower proficiency in English ($r = -0.15, p = 0.04$).

Exploratory Factor Analysis of the AWMA

As mentioned in Chapter Two, the AWMA was not standardised nor validated in the South African context. To explore whether the constructs posited by the authors are similar to the underlying factors that explain performance in the twelve AWMA tasks in our sample and context, EFA was applied to the 12 AWMA tasks. Alloway, Gathercole & Pickering (2006) compared various models of working memory using CFA and determined that a three component model with one central processing component and two modality-specific stores best explained their sample's performance on the AWMA. This paper shows how factor analysis can be applied to the AWMA to test the underlying structure of WM, although the fact that the AWMA was designed with Baddeley's (2000; Baddeley & Hitch, 1974) model in mind (Alloway, 2007) means there will be some bias towards that model by design. Because of the focus on determining the underlying structure of latent variables driving results (rather

than a focus on data reduction), EFA was used rather than principal components analysis (Costello & Osborne, 2005).

Table 6
Correlations among Automated Working Memory Assessment (AWMA) tasks

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Digit recall	—											
2. Word recall	.48***	—										
3. Nonword recall	.27***	.47***	—									
4. Listening recall	.31**	.36***	.14***	—								
5. Counting recall	.28***	.20**	.06	.29***	—							
6. Backwards digit	.42***	.36***	.20**	.41***	.34***	—						
7. Dot matrix	.28***	.25***	.15*	.30***	.35***	.33***	—					
8. Mazes memory	.31***	.24***	.27***	.24***	.22**	.17*	.40***	—				
9. Block recall	.26***	.18*	.07	.24***	.35***	.32***	.59***	.33***	—			
10. Odd one out	.32***	.15*	.09	.27***	.32***	.34***	.55***	.32***	.44***	—		
11. Mister X	.24***	.25***	.03	.31***	.36***	.26***	.54***	.36***	.49***	.45***	—	
12. Spatial recall	.19**	.16*	.09	.30***	.44***	.36***	.52***	.38***	.44***	.48***	.60***	—

Notes. N = 189

Backwards digit = Backwards Digit Recall

* $p < .05$ ** $p < .01$ *** $p < .001$

The pairwise correlations among the AWMA tasks are presented in Table 6. These provide some indication as to how well tasks correlate within and across the working memory components they are designed to tap.

To satisfy the assumptions for the use of EFA there need to be a high number of strong correlations (above 0.3) among items. Almost all of the pairwise correlations among tasks were significant, with many correlations above 0.3, and some tasks, namely the Digit Recall, Word Recall, Backwards Digit Recall and Mazes Memory tasks, were significantly correlated with all other tasks. Further, a high number of observations is needed, with one rule of thumb being a ratio of at least ten observations per item (Costello & Osborne, 2005), which was satisfied with our sample size of 189. A more contested assumption is that each factor needs to have at least three items (Costello & Osborne, 2005).

After an initial EFA two factors were retained based on multiple criteria (as recommended by Costello & Osborne, 2005) including looking at eigenvalues (using a threshold of one) and scree plots (looking for notable breaks in the curve). Exploratory factor analysis was then rerun with only these two factors and, based on expectations that factors might be related to one another, an oblique rotation was applied (Costello & Osborne, 2005). The resultant factors and the loadings of each item on each factor can be seen in Table 6.

Table 7

Results from Exploratory Factor Analysis (EFA) applied to the 12 AWMA tasks

Items	Factor 1	Factor 2	Uniqueness
Digit recall	0.35*	0.54**	0.59
Word recall	0.24	0.68**	0.48
Nonword recall	0.08	0.54**	0.70
Listening recall	0.40*	0.35*	0.72
Counting recall	0.52**	0.13	0.72
Backwards digit	0.44*	0.40*	0.65
Dot matrix	0.75**	0.09	0.43
Mazes memory	0.47*	0.22	0.73
Block recall	0.67**	0.03	0.55
Odd one out	0.66**	0.06	0.56
Mister X	0.73**	0.01	0.47
Spatial recall	0.74**	-0.01	0.45

Notes. $N = 189$.

Backwards digit = Backwards Digit Recall

* factor loading $> .32$ (item considered to load) ** factor loading $> .5$ (item considered to strongly load)

An item was deemed to load onto a factor if it had a factor loading of 0.32 or more, and to strongly load with 0.5 or more (following guidelines from Costello & Osborne, 2005). The first factor, which explained 76.16% of the variance of the AWMA tasks, was loaded on by all items except for two of the verbal storage tasks (Word Recall and Nonword Recall). Further, it was strongly loaded on by all but one (Mazes Memory) of the visuospatial tasks and by one verbal task: Counting Recall. The second factor explained 30.09% of the total variance and was loaded on exclusively by verbal tasks: the three verbal storage tasks and two of the verbal processing tasks (Listening Recall and Backwards Digit Recall). It was strongly loaded on by the three verbal storage tasks. The first factor was proposed to tap some general processing ability which also incorporates visuospatial storage, and the second factor was proposed to tap verbal (or phonological) storage. All of the items loaded onto one of the two factors, but a few (Listening Recall, Backwards Digit Recall and Mazes Memory) did not strongly load onto either. There was some cross-loading, with the Digit Recall,

Listening Recall and Backwards Digit Recall tasks all loading onto both factors, but no item strongly loaded onto both factors.

Bivariate Analysis with Language Variables

A potential weakness of this study's methodology is its use of self-report data to model multilingualism. Bivariate regression analysis was used to see how well self-reported proficiency in understanding English predicts performance on objective English tests, following the precedent of similar tests by Ross (1998) and Marian et al. (2007). Participants' scores on the Vocabulary and Similarities tests from the WAIS-III were regressed on their self-reported proficiency in understanding English, in its original form as an absolute value. It was reasoned that self-reported proficiency in understanding would best capture participants' knowledge of words and the conceptual similarities between them, and so these simple linear regressions would provide some indication as to how much the self-report data follows and successfully proxies for objectively-measured proficiency. The results under model (1) in Table 8 show that self-reported English understanding proficiency was a significant determinant of both Vocabulary ($B = .70, p < .001$) and Similarities ($B = .68, p = .001$) scores.

Table 8

Bivariate regressions of verbal intelligence scores on self-reported understanding in English

Dependent Variable	Statistic	(1) Absolute value	(2) Less max. English prof.	(3) Less max. understand. prof.
Vocabulary	<i>B</i> (se)	.70 (.14)	1.14 (.26)	.80 (.15)
	<i>p</i>	< 0.001	< 0.001	< 0.001
	<i>R</i> ²	0.1274	0.0972	0.1314
	<i>F</i> (1, 183)	26.73	19.70	27.69
	<i>AIC</i>	866.27	872.58	865.42
Similarities	<i>B</i> (se)	.68 (.15)	1.01 (.28)	.75 (.17)
	<i>p</i>	< 0.001	< 0.001	< 0.001
	<i>R</i> ²	0.1033	0.0650	0.1002
	<i>F</i> (1, 183)	21.08	12.72	20.37
	<i>AIC</i>	898.02	905.76	898.67

Notes. *N* = 185

Less max. English prof. = Less maximum English proficiency (the maximum proficiency value in English is subtracted from English understanding), Less max. understand. prof. = Less maximum understanding proficiency (the maximum understanding proficiency across languages is subtracted from English understanding), *AIC* = Akaike Information Criterion

The next step tested whether using relative proficiency scores is as or more informative than using absolute scores. This was an important assumption underlying this study's methodology, with balance variables which weight proficiency scores according to an individual's maximum proficiency levels supposed to counteract the inconsistency of interpersonal standards. Two further sets of regressions were run, with the independent variable being the difference between English understanding and the maximum English proficiency (of speaking, reading and understanding), in the first set, and between English understanding and the maximum understanding proficiency across languages, in the second. These two variables capture the essence of the approaches used to create the within-language and between-language variables, respectively. The R-squared's (*R*²'s) and Akaike Information Criteria (*AIC*; Akaike, 1974) of the regressions on these different independent variables were then compared to see which approach explains most variance in the objectively measured WAIS-III scores. The additional comparison of the *AIC*s (with a lower number indicating less

loss of information and greater goodness-of-fit) is included because of the limitations of using the R^2 statistic to compare models (Akaike, 1974).

From Table 8 it can be seen that the first model which used absolute values (Vocabulary: $R^2 = 0.1274$, $AIC = 866.27$; Similarities: $R^2 = 0.1033$, $AIC = 898.02$) and the third model where proficiency in understanding English is measured as the distance of that score from the maximum understanding proficiency across languages (Vocabulary: $R^2 = 0.1314$, $AIC = 865.42$; Similarities: $R^2 = 0.1002$, $AIC = 898.67$) have very similar values for both statistics. This implies that both are equally good at explaining the variation in the objective verbal intelligence measures, and that both are superior to the second model which used scores relative to the English maximum (Vocabulary: $R^2 = 0.0972$, $AIC = 872.58$; Similarities: $R^2 = 0.0650$, $AIC = 905.76$).

Multiple Regression Analysis

The central questions of this study were addressed using multiple regression analysis. Each component of working memory measured by the AWMA (verbal storage, verbal processing, visuospatial storage and visuospatial processing) served as the dependent variable in respective regressions so that the effects of multilingualism on each of these aspects of working memory could be investigated. The use of multiple regression analysis allowed us to control for a covariate which is very likely to affect working memory and be correlated with multilingualism: socioeconomic-status (SES) (Blom et al., 2014). The inclusion of a measure of SES was necessitated by the high correlation between multilingualism and SES in the sample (see Table 4) and a desire to avoid conflating their effects. Preschool attendance is seen as an important factor in affecting later educational outcomes, especially in South Africa where school preparedness can be a big problem (Richter & Samuels, 2018), but was not included as a covariate because of a lack of data concerning the quality of preschool education, which has been found to be crucial in determining cognitive control outcomes (A Diamond, Barnett, Thomas, & Munro 2007) and is highly variable in South Africa (Richter & Samuels, 2018). Age was not included as a covariate based on the belief that working

memory capacity and functioning are stable in the age range we are studying (Alloway & Alloway, 2013).

The language variables described heretofore have incorporated a continuous measure of balance in proficiency among multiple languages. As well as stressing the importance of modelling multilingualism continuously, the review of literature in Chapter One revealed a dearth of studies looking at the effects of speaking more than two languages on executive function and working memory. Along with the within- and between-language balance variables described above, this study included a binary variable indicating whether a participant speaks a third language or not, thereby explicitly modelling the difference between speakers of two and three languages and contributing to this new and important question. This variable was interacted with the between-language balance variable, allowing us to isolate the effect of speaking a third language, holding the balance of proficiency across languages constant, and inversely to isolate the effect of increased balance holding the number of languages constant. This is important as those who speak more than two languages may be more imbalanced across them, and so a regression which does not include this interaction term will conflate this decreased balance effect with the effect of speaking an additional language.

Table 9

Multiple regression results

	(1)	(2)	(3)	(4)
Independent variables	Verbal storage	Verbal processing	Visuospatial storage	Visuospatial processing
Within balance	1.73 (1.92)	1.95 (1.89)	-1.03 (2.22)	.49 (2.17)
Between balance	-1.47 (1.42)	3* (1.4)	.13 (1.64)	2.22 (1.61)
Third language	-1.45 (2.99)	-1.15 (2.95)	-.21 (3.46)	-2.97 (3.38)
Between balance x Third language	.51 (1.67)	-1.71 (1.65)	-1.02 (1.92)	-3.1 (1.88)
SES	.80* (.41)	.46 (.4)	-.68 (.47)	-.25 (.46)
Constant	84.81* (4.18)	94.4* (4.13)	96.72* (4.83)	101.55* (4.73)
<i>F</i> (5, 183)	2.49	1.72	0.91	0.73
<i>R</i> ²	0.0636	0.0450	0.0242	0.0195

Notes. $N = 189$. For all independent variables coefficients are shown with standard errors in parentheses. Models (1) to (4) were four separate regressions with the name of the model describing the dependent variable of each.

Within balance = within-language balance in proficiency, Between balance = between-language balance in proficiency, Third language = whether the participant speaks a third language, Between balance x Third language = the interaction between the two variables, SES = socioeconomic status.

* $p < .05$

The results from the four regressions are described in Table 9. In the regression for verbal storage ($R^2 = 0.0636$; $F(5, 183) = 2.49$, $p = 0.03$) the SES score had a significant positive coefficient ($B = .8$, $p = .05$). No other variable was a significant predictor of verbal storage. In the regression for verbal processing ($R^2 = 0.0450$; $F(5, 183) = 1.72$, $p = 0.13$) between-language balance was a positive and significant predictor of verbal processing working memory ($B = 3$, $p = .03$). In the visuospatial storage ($R^2 = 0.0242$; $F(5, 183) = 0.91$, $p = 0.48$) and processing ($R^2 = 0.0195$; $F(5, 183) = 0.73$, $p = 0.60$) regressions there were no significant predictors among the independent variables, implying poor goodness-of-fit and explanatory power.

Post-regression diagnostics were applied to regressions results and residuals to examine whether the multiple regression assumptions for accurate coefficient estimation and hypothesis testing were met. For the sake of conciseness, the results of these checks will not be shown visually in tables or figures and will be described in general terms. To look for particular observations which may be excessively influencing regression results, each observation's square of (studentised) residuals and leverage (the extent to which they affect coefficients) was plotted on a graph. Across the four regressions no extreme outliers were found but some observations could be seen to have both a high residual and high leverage. These were identified as the most likely to be disproportionately influencing regression results and coefficients, and so regressions were rerun without them. However, their omission had only very small and insignificant effects on regression results, perhaps due to the relatively large sample size of the regression. Therefore, after checking that there were no apparent errors in their data, these observations were preserved in regressions.

The assumptions for coefficient estimation in multiple regression are that there are linear relationships between independent and dependent variables, that there is no high multicollinearity, and that errors are independently and identically distributed and homoskedastic. Scatter plots revealed no evidence of non-linear relationships between the dependent and independent variables which would have violated the first assumption, and all relationships appeared either linear or non-existent. Variance inflation factors were calculated after each regression to test for multicollinearity but no factor above seven was observed, indicating that multicollinearity was not a serious problem. Because our research design sampled each participant only once there was no issue with repeated measures from the same individual which might have cast doubt on the independence of errors. Given our staggered data collection (over a few years), one way in which our data could have violated this assumption would have been if errors from the same time point were more highly correlated within themselves than between time points. However, working memory is a relatively stable construct (Alloway & Alloway, 2013) which depends on many aspects of lifetime experience (see Chapter Two) and should not vary in any systematic way between years or times of year. Among the groups of participants from different waves of data collection in our study the only factors which might have varied between groups and affected working memory were

already controlled for in regressions (and thus were not in the error). Another way in which the independence of errors might not hold is if the errors differ systematically among interviewers, but a scatter plot of interviewer code against studentised residuals revealed no consistent pattern.

To look for heteroskedasticity, graphs with residuals plotted against fitted values were generated. After observing that the variance of residuals did not appear to change with the predicted value, the Breusch-Pagan test for heteroskedasticity (Breusch & Pagan, 1979) was used to see whether the residuals significantly varied as a function of predicted values (or independent variables). For all of the regressions the null hypothesis of homoskedasticity was not rejected. Therefore heteroskedasticity is not an issue, and given that the assumption of identically and independently distributed errors appears sound, the coefficients generated by these regressions are valid.

To test whether residuals were normally distributed - the key assumption for the validity of hypothesis testing (and thus reported p-values) - the Shapiro-Wilk test for normality (Shapiro & Wilk, 1965) was run. The null hypothesis of normality was not rejected for the verbal processing regression but was rejected for the other three regressions. Kernel density plots of the studentised residuals were plotted against normal distributions and showed residuals to be approximately normal but generally skewed to the right and, in the case of the verbal processing regression, to have an unusual shape around the mean.

In the course of reviewing the theoretical and empirical evidence surrounding multilingual advantages in Chapter One it became clear that the age at which a person becomes multilingual can heavily influence the degree and nature of subsequent cognitive development. Further, AOA and proficiency are highly correlated with one another, and so any analysis which models only one of these factors is likely to be conflating their effects to some extent. Therefore, after running regressions focusing on balance in proficiency, another set of regressions were run which added an independent variable for the age at which participants started becoming multilingual: second language AOA. Appendix D contains a tabulation of the regression results. Second language AOA was not found to be a significant

determinant of any of the working memory components, and its inclusion rendered all previously significant coefficients non-significant. In particular, the coefficient on between-language balance in the verbal working memory regression fell in magnitude and was no longer significant at the 5% level ($B = 2.70, p = .07$).

The AOA of participants' third language was not included in any regressions as its inclusion would have limited the sample to only those who spoke three languages and would have excessively narrowed the sample beyond the scope of this study's questions. The inclusion of second language AOA did not meaningfully affect any of the post-regression diagnostics and so they are not described for this set of regressions.

Chapter Four

Discussion

Introduction

Does multilingualism lead to advantages in working memory? If so, what kind of working memory is affected, and which aspects of multilingual experience are pivotal? The previous chapter described the results of the data analyses which were performed in the hope of answering these questions. This Discussion chapter is an interpretation of these results drawing on the range of theoretical and empirical literature covered in Chapter One. The results of this study provide unique insight into a number of phenomena that have not been extensively studied. Balanced proficiency among languages was found to be a significant determinant of the verbal processing aspect of working memory, but none of the variables measuring multilingualism significantly predicted any aspect of visuospatial working memory. This pattern of results, combined with other significant correlations among variables, suggested that balanced multilingualism may simultaneously lead to advantages in the storage and processing of verbal information and disadvantages in other linguistic tasks. In addition, the use of self-report proficiency information weighted against individual maxima emerged as a promising alternative to its use in absolute form, and SES was a significant determinant of verbal storage.

The first part of this discussion highlights and interprets key findings, taking the pattern of significant and non-significant results at face value and discussing how they inform the central research questions of this report. This interpretation draws on the results from the preliminary statistical analyses and is informed by the literature reviewed. The next part questions the extent to which this pattern of results can be taken at face value, reflecting on the methodological and theoretical issues which could have biased or distorted this study's results, and evaluating the risk of each. Thereafter the final section reflects on the study as a whole, highlighting the unique contributions the study makes, its theoretical and practical significance, and suggesting directions for future research in the area.

Key Findings and Interpretation

The strongest evidence for the nature of the link between multilingualism and working memory came from multiple regression analysis, by virtue of its ability to model numerous variables simultaneously and examine the relationships among them. In the first set of regressions, the only multilingualism variable to significantly predict one of the working memory components was between-language balance in proficiency, which was found to significantly predict the verbal processing component of working memory. None of the other variables, including within-language balance, whether or not the participant spoke a third language and the interaction term between this and the between-language balance variable, showed significant coefficients in any of the regressions. This suggests that speakers need to be highly balanced across languages in various kinds of proficiency (reading, speaking and understanding) in order to engage the mechanism leading to multilingual advantages. The significant negative correlation between speaking a third language and the balance variables meant that the effect of speaking a third language could have been cancelled out by the effect of decreased balance and underlined the importance of including and interacting both in the same regression in order to accurately estimate their effects. The non-significance of the interaction term suggests that the significant effect of between-language balance remains the same regardless of the number of languages spoken. The balance variables were significantly correlated with one another, but no significant relationship between the within-balance variable and any of the working memory variables was found in either regression or correlation analysis. This implies that balance in proficiency across languages is more important in determining (verbal) working memory advantages than balance across different kinds of proficiency (speaking, reading and understanding) within languages. The importance of balanced multilingualism in determining executive function is consistent with previous studies (Incera & McLennan, 2017), but the comparison of the effect of balance with the effect of the number of languages spoken is unprecedented.

This insight into the importance of between-language balance relative to other aspects of multilingualism was complemented by a set of further regressions which incorporated the AOA of participants' second languages. The significant negative correlation between this

AOA variable, which can be interpreted as the age at which participants started becoming multilingual, and the between-language balance variable implies that later multilinguals do not attain the same level of proficiency in their additional languages as do early multilinguals. This is consistent with previous findings (Luk et al., 2011) and hypotheses for critical periods in language development (Kaushanskaya & Marian, 2009). Because earlier multilinguals were generally more balanced among their languages, the change in the balance variable's coefficient (and significance level) with the inclusion of the AOA variable could be interpreted as the portion of its effect which was driven by differences in AOA. None of the regressions showed AOA to be a significant determinant of working memory, but the decrease in the magnitude and significance of the between-language balance variable caused by the addition of AOA to the verbal processing regression might mean that its previous significance was in part due to the effects of early life exposure to multiple languages among more balanced multilinguals.

The other variable which was significant in the first set of regressions was SES, which significantly predicted verbal storage scores. This finding is consistent with other research reviewed in Chapter One which found SES to significantly predict working memory independently of multilingualism. Socioeconomic status acts as a proxy for a range of factors including the content and amount of parental speech directed at children and the amount of time parents can spend with children, which have been found to predict vocabulary (Hoff, 2003; Rowe & Goldin-Meadow, 2009) and language processing (Fernald, Marchman, & Weisleder, 2013) development. To the extent that verbal short-term memory tasks depend on vocabulary (Verhagen & Leseman, 2016), these effects could be driving the effect found in this study and could explain why this advantage was restricted to verbal (and not visuospatial) storage. Socioeconomic status was significantly and negatively correlated with the between-language balance variable, raising concerns about how its effect on working memory might be conflated with the multilingualism effect. The inclusion of SES alongside the multilingualism variables in multiple regressions was therefore instrumental in estimating the independent effects of each factor.

All aspects of the interviews and tests in this study were carried out in English which raises questions about the effect this may have had on participants' comfort and confidence in performing tasks and their ability to understand and follow instructions. The fact that all participants were studying in English (all had met the university's English proficiency entrance requirements) was taken as sufficient evidence that they had enough experience in performing tests and filling out questionnaires in the language and therefore their scores should not have changed as a result of their not understanding English-medium instructions. The effect of English on scores becomes far more profound and complex for the tasks which were not only administered in English but also involved the description, memorisation or manipulation of English subject matter. The most important of these tasks, for their central use as dependent variables, were the verbal AWMA tasks, which for the most part involved the storage and processing of English content. The WAIS-III Vocabulary and Similarities tests tested participants' ability to define English words and describe the conceptual similarities between them.

Contrary to expectations, self-reported proficiency in English was not significantly correlated with performance in the verbal working memory tasks, but did correlate significantly and positively with the verbal intelligence (Vocabulary and Similarities) scores. In this sample self-reported English proficiency was not significantly correlated with between-language balance. This could be the result of a mutual cancelling out of a negative correlation in one group - English mother-tongue speakers who are highly proficient in English but speak other languages poorly (and are thus imbalanced among their languages) - and a positive correlation in another group - multilinguals for whom high additional language English proficiency will mean higher balance across languages. On the other hand, English proficiency was positively and significantly correlated with within-language balance, perhaps reflecting the fact that highly proficient English speakers are likely to have gone to better quality schools (Nel & Müller, 2010) and thus have smaller discrepancies between their reading, understanding and speaking proficiency.

While regression results showed that between-language balance determines verbal processing scores, correlations among this variable, the verbal working memory scores and

the two measures of verbal intelligence help to expand the interpretation of the balance variable's effects and provide evidence for the way in which this variable enhances task performance. Between-language balance is significantly and negatively correlated with both verbal intelligences scores. Combined with the lack of correlation between this balance variable and English proficiency, this suggests that those who are highly balanced across multiple languages perform worse on the verbal intelligence tests because of some *language-general* conceptual reasoning disadvantage rather than a lack of conceptual knowledge of English *in particular*. This conceptual disadvantage could be due to multilinguals having a lower vocabulary across their languages⁸ and not developing as deep a conceptual understanding due to a lack of sufficient exposure to any one language (Engel de Abreu, 2011). In less multilingual people, the ease of lexical access (Bialystok, 2009) and increased familiarity with one language may give speakers more time to reflect on conceptual meanings of words and their relations. This disadvantage in highly balanced multilinguals could coexist with advantages related to the storage and processing of linguistic material due to their experience in using multiple words across languages and translating among them. These speculations are useful in thinking about how multilingualism can have diverse and simultaneous effects on different cognitive capacities, and suggest that the language-general advantage driving advantages in verbal processing tasks is independent of other conceptual capacities. This means that a multilingual advantage in verbal working memory tasks is not inconsistent with established multilingual disadvantages in other linguistic tasks related to vocabulary, verbal fluency and lexical access (Bialystok, 2009; Engel de Abreu, 2011).

Because the four components of working memory proposed by the AWMA acted as dependent variables, the interpretation thus far depends on the assumption that the 12 tasks administered did indeed measure verbal storage, verbal processing, visuospatial storage and visuospatial processing. A previous study applied CFA to a sample of children performing the AWMA and found that a three component model with one modality-general central processing component and two domain-specific storage systems best explained participants' performance (Alloway et al., 2006). These components matched the three components (the

⁸ The belief that multilinguals have a lower vocabulary across languages is supported somewhat by a significant correlation indicating that speakers of three languages perform worse on the Vocabulary test than do speakers of two. This negative correlation could also be responsible for the lack of significance of the presence of a third language in the verbal working memory regressions.

central executive, the phonological loop and the visuospatial sketchpad) of the Baddeley and Hitch (1974) model of working memory, ignoring the episodic buffer (Baddeley, 2000). These results supported the AWMA's division of tasks, with performance on the storage tasks explained by the modality-specific stores and performance on the processing tasks explained by the conjunction of the domain-general processing component and the storage component related to the modality of the task at hand.

The EFA results of this study differ from these findings in substantial ways and raise questions about what is actually being tapped by the AWMA tasks in this sample. Two factors were found; one seemed to be specific to the storage of verbal information (something like a phonological loop) that explained performance in the verbal storage tasks and to a lesser extent in the verbal processing tasks. The other explained performance in all of the visuospatial tasks, with no distinction between storage and processing tasks, and to a lesser extent in the verbal processing tasks. The fact that all the processing tasks load onto this factor suggest that it represents some domain-general processing ability (such as the central executive) or general intelligence, which makes the strong loading of the visuospatial storage tasks surprising. This implies that there is no visuospatial-specific storage component and that visuospatial information, in contrast to verbal, is handled by domain-general capacities which both store and process information. This finding is in conflict with previous evidence for the existence of a passive visuospatial store (Bruyer & Scailquin, 1998)⁹.

Although this study's young adult sample is consistent with Baddeley's original intention to explain working memory functioning in adults (Baddeley & Hitch, 1974; Baddeley, 2000), the fact that Alloway et al. (2006) performed their CFA with a sample of children may reduce the extent to which those results can be compared with these EFA results. Still, the fact that EFA is exploratory by nature and that the factors found in this study differ so profoundly from those found by Alloway et al. (2006) limit the extent to which these

⁹ Further, research among American Sign Language speakers has found substantial evidence for a store and even an articulatory rehearsal system in the visuospatial modality (Wilson & Emmorey, 1997). This evidence suggests that such a system *can* exist but does not imply that it *does* exist in speakers of verbal languages.

results can be used to come to conclusions about the structure of working memory¹⁰. However, these findings are useful in determining that, in our sample, the visuospatial tasks appear to be failing in their efforts to tap a specific construct. If the effects of multilingualism can only be detected when looking at outcomes for more precisely defined aspects of executive function, the fact that visuospatial task performance can be explained by some general processing or intelligence factor may explain why it is not predicted by the multilingual variables in this study.

The overlap among executive functions and the diversity of abilities involved in performing any particular task mean that the task impurity problem has always plagued the measurement of specific executive functions (Miyake et al., 2000). In working memory tasks the extent to which this counts as ‘impurity’ varies depending on the extent to which the processing aspects of working memory are supposed to depend on (or even encompass) executive functions. However, the extent to which working memory tasks depend on general executive functions has previously been claimed to vary by modality, with visuospatial working memory tasks involving more substantial use of executive control, for example, than verbal working memory tasks (Alloway, Gathercole, Willis, & Adams, 2004). The loadings of the tasks in EFA also suggests that the distinction between verbal and visuospatial modalities cannot be made very clearly for some tasks, implying that scores for working memory in one modality will be influenced by working memory capacity in the other. For example, the strong loading of the Counting Recall task on the general visuospatial/processing factor echoes previous findings which have cast doubt on the purely verbal (as opposed to visuospatial) nature of the task (Absatova, 2015). The cross-loading on both factors of some tasks in the EFA raises further questions about this distinction. In some tasks a different distinction - that between storage and processing - is controversial; some claim that the Backwards Digit Recall task is a processing task in children but a storage task in adults (St Clair-Thompson, 2010).

¹⁰ In order to make claims about the structure of working memory or use the underlying factor structure to determine dependent variables for our analysis it would have been necessary to conduct CFA to explicitly compare competing models using inferential statistics (Costello & Osborne, 2005).

The interpretation thus far also depends on the multilingualism variables accurately reflecting the languages participants speak, their proficiencies in those languages and the ages at which they began acquiring them. Measures of linguistic proficiency can be based on an interviewer's perceptions and evaluations based on in-depth conversation, on performance in some standardised test, or, as in this study, on self-evaluation. There are certain aspects of one's own language abilities and experience which an individual has access to and which could never be captured as comprehensively or profoundly by some external measure or judgement. This is especially true when the objective measure is based on a very limited aspect of linguistic proficiency, with an example being Blom et al. (2014) using scores on vocabulary tests to represent multilingual balance. However, the use of self-report measures, which can be very simplistic and based on little background information, is also problematic (Bialystok, 2016). The extent to which a person has access to their own mind, and knowledge of their own experience, always needs to be balanced against the lack of awareness an individual may have of certain mental processes and capacities that function in an unconscious or very fast manner. Further, even when judgement concerns something that most people do have conscious access to, the standards and scales on which they base their self-evaluations, and the information that is emphasised, could be highly variable. Finally, self-report responses could always be influenced by the social desirability of certain responses (Riggs et al., 2014).

All of the language variables in this study, as well as the English proficiency variable discussed above, were based on the self-report information collected using the LEAP-Q. While there could be some variation in participants' evaluation of what counts as speaking a language, self-report information about the number of languages spoken and the age (in months) at which acquisition of each language began was assumed to be accurate and comparable across participants. The concerns about self-report proficiency information were allayed somewhat by bivariate regressions showing that self-reported understanding of English significantly predicted two objective measures of English (the Vocabulary and Similarities scores). Further, it was argued that proficiency scores weighted to individual maxima would increase the comparability of self-report information. The fact that both absolute scores and scores weighted against an individual's maximum understanding score

(in any of their languages) predicted a similar degree of variation in verbal intelligence scores (and led to a similar level of information loss) lent support to the notion that using internally weighted self-report information can be a promising alternative to using scores in absolute form. Although it was not found to perform better than absolute scores, the capacity of scores weighted to individual maxima to predict objective measures of language proficiency provides support for this study's methodology and reaffirms the potential of an approach focused on multilingual balance. The implications of these findings are limited, however, and showing that variables based on self-report significantly predict objective measures does not imply that all of the relevant aspects of language proficiency have been captured, or rule out the possibility that some objective measure of proficiency would have detected relationships between multilingual balance and working memory that self-report information could not measure.

Although the effect of gender was not included in regressions, its relation with working memory was explored using correlation analysis. Females had significantly lower visuospatial working memory than males, a correlation that is consistent with previous findings for gender differences in visuospatial tasks (Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003) but which may have to do with other factors such as participants' area of study. The low proportion of males in the study may therefore mean that participants in the higher range of visuospatial working memory who are needed to detect significant results were not sufficiently incorporated. Such reasoning is not supported by the fact that visuospatial scores were on average higher than verbal scores, meaning that significant differences in verbal working memory were detected despite poorer performance in tasks of that modality.

The lack of a significant relationship between multilingualism and (visuospatial) working memory is not unprecedented in samples of young adults (Kaushanskaya & Marian, 2009) nor in children (Engel de Abreu, 2011), but no previous non-significant finding was based on so comprehensive a measure of working memory. The particular pattern of significant effects for verbal working memory and a lack thereof for visuospatial working memory is unusual and is the inverse of that normally observed in multilingual advantage

research, where multilinguals are at a disadvantage in verbal tasks and at an advantage in non-verbal tasks (Bialystok et al., 2012; Valian, 2015). Furthermore, the absence of advantages on visuospatial tasks in this study means that multilinguals' advantages in verbal working memory cannot be attributed to some domain-general processing advantage outweighing a linguistic disadvantage (as can be done when working memory advantages in both modalities are found, e.g. Blom et al., 2014). The theory linking multilingualism to executive function and working memory advantages (reviewed in Chapter One) focused on domain-general advantages and so cannot explain this pattern of results. The demands of the working memory tasks used, with their substantial engagement of other executive functions and heavy processing aspect, seem to match the characteristics of the working memory tasks on which multilinguals outperformed monolinguals in previous studies (Blom et al., 2014; Morales et al., 2013). All of the processing tasks are challenging to the control of attention (with their dual requirements of remembering a sequence and performing some calculation using the information) and in many of them the rehearsal of the information to be maintained in memory is made difficult by the simultaneous processing requirements. Thus they meet the conditions necessary for a task to accurately reflect working memory (and predict intelligence scores) postulated by Cowan (2008) and Engle et al. (1999).

This kind of pattern is unusual but is not without precedent. Sabourin and Vinerte (in press) found bilingual adults to have greater advantages in cognitive control tasks which involved linguistic content (Stroop tasks) than in tasks involving non-linguistic content (ANT tasks). Both simultaneous and early bilinguals performed better in linguistic tasks but only simultaneous bilinguals performed better in tasks involving non-linguistic content. The particular nature of the Stroop task means that the authors largely restrict the interpretation of differential performance in that task to familiarity (or a lack thereof) in the particular language being tested and are hesitant to explore language-general (or metalinguistic) factors. The suggestion that only simultaneous multilinguals have non-linguistic advantages, and the implications thereof for this study, are returned to in the following section.

Can significant results in the verbal domain be used as evidence for a multilingual advantage in working memory? In the cognitive control literature it has been claimed that

significant differences need to be found in tasks that are not linguistically driven, and do not use linguistic content, to constitute evidence for a multilingual advantage (Hilchey & Klein, 2011). This is because the multilingual advantage is by definition domain-general and any advantage in a verbal task could be attributable to superior linguistic control (or in the case of this study language-specific memory or processing) rather than any general cognitive capacity (Sabourin & Vinerte, in press). Speculations along these lines could imply that multilinguals are especially good at storing and processing letters and words due to their extensive practice using this kind of content in multiple languages and in several complex multilingual processes, but that this enhanced efficiency or capacity cannot be applied to other kinds of information.

Although many researchers have embraced the assumption that linguistic control and cognitive control mechanisms are identical (e.g. Green, 1998), others believe that the co-activation of a multilingual's languages could be managed by a language-specific control mechanism (Engel de Abreu, 2011; Dijkstra & van Heuven, 2002) and some contrast observed advantages in linguistic tasks with non-results in non-linguistic tasks and use this as evidence for a distinction between the two types of control (Yow & Li, 2015). For the current study the key question is the extent to which multilingualism is likely to engage some language-specific aspect of working memory (which only becomes apparent in processing tasks), and the suggestions from the empirical analysis above need to be combined with other empirical evidence and theoretical models of working memory to arrive at an answer.

In Baddeley's (2000; Baddeley & Hitch, 1974) model the most straightforward way in which an advantage could be restricted to verbal working memory is through some sort of enhancement of the phonological loop. It is not clear how the phonological store could be augmented by experience, however, so if the advantage lies in the phonological loop it would have to derive from the development of the articulatory rehearsal system. A more advanced articulatory rehearsal system should lead to advantages in both storage and processing tasks, however, unless the system is somehow more efficient when processing demands are high, an interaction with processing demands which would imply some involvement of the central executive. Being domain-general, any augmentation of the central executive should be

observable in all tasks with a processing component (and not just verbal tasks). Is this model inconsistent with our results, therefore? One way of accommodating these results within the model might be to conceive of multilinguals as having more experience in applying their domain-general processing capacities to linguistic information in novel ways. Conceived thus, the domain-general central executive would not be augmented by multilingualism, but the ways in which this processing is applied to linguistic content would be. Several questions arise from such a position, however, and it is unclear whether these can be answered in a satisfying manner using Baddeley's (2000; Baddeley & Hitch, 1974) model. Are the ways in which the central executive interacts with the stores contained within the central executive? If so, then surely there are parts of the central executive that are not domain-general, as they relate specifically to one or another modality. But if knowledge or memory of how to apply central executive capacities to the storage components is not a part of the central executive then where is it in relation to the model? The central executive is described broadly and functionally as a supervisory component which coordinates the behaviour of the other components, but surely to be a meaningful component this would have to include information about how processing abilities are applied to stores. Unfortunately as one begins to enter this territory of asking specific questions related to the roles and extents of different components it is difficult to find answers due to the vagueness of the model and its components.

Cowan's (2008) model of working memory seems capable of providing a more plausible account of verbal processing specific multilingual advantages. Within the model some attentional processes, such as preferential processing of certain characteristics, are believed to be modality-specific (Gomes, Molholm, Christodoulou, Ritter, & Cowan, 2000). The lack of an observed advantage in the more passive short-term memory measured by the verbal storage tasks suggests that the advantage is not based on multilinguals having either a lower rate of decay for verbal information in activated memory or diminished interference between verbal chunks of information. That an advantage was found exclusively in tasks that involved more conscious and effortful processing, implying a more integral and nuanced role for the focus of attention, could suggest that multilinguals have an experience-induced advantage in using their focus of attention to refresh and maintain verbal content in activated memory while simultaneously devoting attention to some manipulation of this content. This

kind of ability, along with other strategies to manage interference among verbal chunks of information in activated memory, would intuitively form an integral part of many processes that multilinguals routinely perform (such as translation).

Further, the question of how use of some cognitive capacity leads to that capacity's enhancement over time, an issue which is so often ignored or opaque (see Chapter One), seems relatively straightforward in this interpretation. Halford, Cowan and Andrews (2007) speak about the manner in which working memory can improve over time, with chunk capacity limits remaining more or less fixed (based on empirical evidence) but the ability to form larger chunks with a certain kind of information increasing with experience. Besides, the allocation of attention is integral to learning and development (Gomes et al., 2000) and strategies (conscious or not) for deploying the focus of attention in challenging linguistic tasks could be learned and refined over time.

Besides these speculations about how models of working memory can accommodate domain-specific verbal effects, there are some empirical findings which seem consistent with this pattern. Kaushanskaya and Marian (2009) did not find any significant differences in the performance of a working memory storage task between monolingual and bilingual adults but found that early bilinguals were uniquely able to apply their working memory to phonologically unfamiliar information. This seems to echo the speculation above that novel ways of applying working memory capacities to phonological information may be responsible for verbal advantages rather than differences in working memory capacity itself. Similarly, Brito et al. (2014) found that multilingual infants were able to access memories more flexibly and made use of a greater range of retrieval cues than did monolinguals. These findings were interpreted by the authors as evidence for domain-general advantages, but it seems plausible that experience in multiple languages could lead to increased receptivity to retrieval cues and memory flexibility exclusively for phonological or verbal information.

Methodological concerns

In the previous section it was generally assumed that the pattern of effects, with balance among languages predicting verbal processing but having no effect on any aspect of visuospatial working memory, accurately captured the pattern of effects in the population under study. For a range of reasons, however, this may not be the case and the real effects in the population may differ from those found in this study. The nature of inferential statistics means that the conclusions that can be drawn from any empirical study are necessarily tentative. This is especially so in an exploratory study such as this where the statistical analyses are designed to explore the relations among a range of variables rather than testing one or two precisely defined hypotheses

On the whole the relationships between multilingual proficiency and experience and working memory outcomes were weaker than predicted based on previous theoretical reasoning and empirical findings. It is important to question the extent to which this lack of significant relationships is real. As Bialystok (2016) points out, a non-significant result is the absence of evidence for a relationship and is not, in itself, evidence for the lack of a relationship. Focusing on the example of a lack of effect of balanced multilingualism on visuospatial working memory, this section will consider a series of possible scenarios which could have given rise to the observance of this result. Within each scenario a range of determining factors, and the likelihood of each, will be discussed.

Scenario 1: effect exists in the sample but is not detected

In the first scenario there was an effect of multilingualism on visuospatial working memory in our sample, but for methodological reasons this effect was not detected. This could be due to factors related to the instruments or to factors related to the regression design.

Balance in proficiency among languages may determine differences in working memory but the instruments employed to measure either construct could be insufficiently valid, accurate or sensitive. Doubts about the constructs that are being tapped by the AWMA

were discussed in the previous section. Besides, even if working memory differences based on multilingualism do exist, ceiling effects in tasks may reduce the ability to detect these differences at higher ranges. This problem is believed to be especially acute among young adults (Bialystok, 2016) and the variety of tasks used to measure working memory in different studies means that very different constructs may be involved in performance of each (e.g. Engel de Abreu, 2011; Kaushanskaya & Marian, 2009). The fact that the AWMA is often used in samples of children (e.g. Alloway et al., 2006) raises concerns about it being sufficiently challenging for older age groups, but in the present sample the maximum span length was almost never reached and plots of participants' scores did not indicate any patterns resembling ceiling effects, diminishing the risk that this factor influenced results.

The possibility that the self-report measures may be insufficiently sensitive or insufficiently consistent among participants (Bialystok, 2016), and that a more objective measure of multilingualism may have better captured the differences among participants that are associated with differences in working memory, was discussed in the previous section. An important movement in multilingual advantage research, which was adopted in this study, is the modelling of multilingualism as a continuous variable. The virtues of such an approach (rather than a discrete one) were discussed in Chapter One, but by treating multilingualism as a spectrum of balance in proficiency and AOA we may have failed to detect particular thresholds which are highly deterministic of cognitive outcomes. One example of this would be the distinction between simultaneous and early bilinguals made by Sabourin and Vinerte (in press): if AOA effects are determined primarily by simultaneous language acquisition then directly comparing simultaneous bilinguals (who have zero months difference between the ages at which they began acquiring their two languages) with sequential bilinguals (who have some difference between them) will be more sensitive to multilingual advantages than comparing people along a spectrum of months. Likewise, if differences in a variable only determine advantages within a particular range, as would be the case if AOA only affects development in the period after the onset of speech production (Green, 1998) or in early critical periods (Kaushanskaya & Marian, 2009) then modelling this variable along a wider range would dilute and mask its effect. Such a phenomenon would have resulted in a detectable pattern in post-regression diagnostics however, and for the regressions no such

evidence was found in tests for non-linear relationships. Therefore the risk of having obscured any effect of multilingual balance or AOA by modelling them continuously does not seem substantial in this study.

A range of factors related to statistical design and methodology could result in an inability to detect significant relationships: the fit of the model, the comprehensiveness and suitability of controls, the regression strategy, and even the overarching (frequentist) statistical paradigm on which multiple regression analysis and hypothesis testing rests (Sivia & Skilling, 2006). With the exception of the regression for verbal processing, none of the multiple regressions exhibited normally distributed residuals, which means that the significance levels for coefficients need to be treated with caution. Another factor which jeopardises the extent to which regression results can be interpreted is the poor goodness of fit of the models, with the exception of the verbal storage regression. This implies that the independent variables were poor predictors of the variance in the working memory variables, and that other factors (or other measures of these factors) are needed to explain working memory differences. In Chapter One the importance of controlling for covariates (variables which are correlated with both multilingualism and working memory) was stressed as a necessary step in isolating multilingualism effects (Valian, 2015; Yow & Li, 2015). While our regressions controlled for SES, an important covariate which is often omitted in multilingual advantage research, other controls such as age, preschool attendance and parental education were omitted because of concerns around multicollinearity or because including them in regressions produced no significant effects. However, the poor goodness of fit of the models means that there could well be other variables that are significantly correlated with both multilingualism and working memory that are causing bias in the coefficients and their significance levels in regressions. The direction of this bias could be upwards or downwards depending on the signs of the correlations between the variables. One example of such an omitted variable was discussed in Chapter One, namely, the extent to which people need to navigate different cultures (Riggs et al., 2014). In the South African context it is plausible that this is significantly correlated with multilingualism so the two factors' effects may be conflated in this study. Unfortunately, investigating the issue was beyond the scope of the current study.

Scenario 2: effect exists in the population but not in the sample

In the second scenario the instruments and regression design faithfully reflected the lack of an effect in the sample, but the sample was not a good representation of the population for whom there *is* an effect. As Bialystok (2016) points out, no effect will be uniform and any relationship will differ in nature and magnitude in each individual. If the effect of multilingualism on working memory exists it will vary along some distribution, and every study will yield different results depending on where its participants fall on this spectrum. Thus it is to be expected that some studies will not find significant results, even if the effect exists in the population studied, the study's design is appropriate for the effect's measurement and the sampling procedure is sound. Thus, one has to look at the aggregate of significant results and non-significant results to draw reliable conclusions about the existence of any relationship. The chance of observing a non-significant result due to statistical variation when the effect exists in the population will diminish with the size of the sample and there is always the risk that the power of a given data set is insufficient to find an effect due to a small sample size. For studies of this nature a sample size of 189 is relatively large, however, and many of the studies reviewed in Chapter One found a significant relationship between multilingualism and aspects of executive function in substantially smaller samples (e.g. Blom et al., 2014 with a sample of 120), thus attenuating but not eliminating the risk that this sample is too small to detect effects or that it contains a disproportionate number of participants with effect sizes on the smaller end of the distribution.

Broadly, this study aimed to investigate effects among young adult South Africans and to sample all of the linguistic groups from less multilingual English mother-tongue speakers to more multilingual speakers of South African languages. However, the fact that this study used convenience sampling means there may be something particular about our sample which is driving our non-significant results. It cannot be claimed that the sample is representative of any population, as the sampling procedure did not ensure that each person in the target population had an equal chance of being included in the sample. Therefore there may be an excessively high proportion of certain kinds of people for whom the multilingual effect

differs from population parameters, or a concentration of some factor which we are not controlling for that is distorting or cancelling out the effect of multilingualism.

Scenario 3: effect does not exist in the population

In the final possible scenario the observed non-significance of the effect accurately reflects reality. In other words (balanced) multilingualism has no significant effect on visuospatial working memory, counter to speculation based on previous findings (Bialystok, 2017; Blom et al., 2014; Morales et al., 2013). Such a scenario could mean that multilingualism has no effect on visuospatial working memory in any way, but could also mean that this study investigated the wrong aspect of multilingualism or working memory. If young adults (and especially students) are indeed close to their peak in executive functioning (Yow & Li, 2015), and are at a stage where their working memory is thoroughly challenged by other activities (Valian, 2015), then multilingualism may indeed have no substantial effect in this population.

Evaluating the overall pattern of results

The factors discussed above create uncertainty about the extent to which non-significant results in visuospatial regressions can be taken as evidence for the lack of a relationship between multilingual balance and visuospatial working memory. Two important questions now arise. First, to what extent do these factors apply to an inverse scenario in which significant results were observed in the absence of any real effect (so-called ‘false positives’). The answer to this question will determine the extent to which the positive findings (for the between-language balance effect on working memory) can be believed to represent real relationships. Second, do these factors apply to the verbal working memory tests as much as they do to the visuospatial ones? Both of these questions are crucially important in determining the extent to which the overall *pattern* of results, with stronger effects for between-language balance and SES on verbal working memory than on visuospatial working memory, can be faithfully interpreted. Unless it can be reasoned that these factors affect some of the regressions or variables more than others then the pattern of

results observed should still convey something real about the way multilingualism affects working memory in our sample (and, more cautiously, in this context).

Many of the factors responsible for the misalignment of study observations and real effects discussed in scenarios one and two could just as likely have led to a scenario in which significant relationships were found despite no relationship existing in the population. Such factors include the effect of omitted covariates, the suitability of the regression design and statistical paradigm and the risk of bias due to a statistically anomalous or non-representative sample. The extent to which these positive results can be interpreted directionally is further complicated by the possibility of reverse causality; people who have high executive function or working memory may be more likely to learn multiple languages and achieve a high level of proficiency in them (Bak, 2016).

Some factors, such as the influence of English proficiency, logically apply more to verbal than visuospatial tests. If these factors serve to cloud true differences in verbal working memory, however, then the significant relationship observed is more, not less, consequential. One factor that affects the visuospatial storage and processing regressions more than the verbal ones is the consistent rejection of normality and goodness of fit in those regressions. Only the verbal storage regression exhibited significant goodness of fit and only the verbal storage regression had normally distributed residuals (according to a Shapiro-Wilk test). This suggests that the fit of the models and the suitability of the control variables was better in the verbal storage regression and that the coefficients and their significance levels were more accurate in the verbal processing regression. Further, the significant correlation between verbal processing and between-language balance, and between verbal storage and SES, replicates the pattern of results seen in the regressions and implies that their significance is not merely an artefact of the regressions' misspecification.

In summary, many of the factors that could be driving the non-significance of existent effects could also be driving the significance of non-existent effects in this study's models. However, almost no factor (except for reverse causality) can be argued to disproportionately affect the positive results. The verbal regressions seem to be more reliable than the

visuospatial ones, adding evidence for the ‘reality’ of the positive effect of multilingual balance on verbal working memory but also providing some reason to think that the visuospatial regressions may be relatively less trustworthy and therefore undermining the extent to which the pattern of results can be interpreted.

General Reflection

Previous sections of this chapter interpreted the results of this study and reflected on the extent to which these results could have been affected by a range of factors. While several limitations have been identified, the results have been found to be worthy of interpretation and to inform the research questions of this study in important ways. This section reflects on some of the merits of the research design, discusses the theoretical and practical implications of the results, and returns to some of the fundamental issues surrounding the link between multilingualism and working memory.

Sample and context

Beyond its importance for the study of young adult multilinguals, the sample was significant for the particular languages spoken by participants. The predominance of Bantu languages (and especially Nguni and Sotho languages) among the participants of this study means it is able to address under-studied questions about the nature of multilingual advantages in these languages, and in non-European languages more generally. Whether intrinsic properties of these languages could affect the nature of multilingualism and the effects on executive function and working memory were not directly explored in this study, but it is possible that language-specific properties were responsible for the unusual pattern of results observed.

Another issue which was not explicitly studied is the impact of the degree of difference among a multilingual’s languages. Both Nguni and Sotho languages are distinct from English in their agglutinative and tonal characteristics (Roux & Visagie, 2007; Zerbian & Barnard, 2010), and the fact that most participants in this study spoke English and one Nguni or Sotho

language means that highly different language pairs were common. On the other hand, with the consideration of (up to) three languages many participants would have also spoken language pairs from the same family (e.g. seSotho and seTswana) which are very similar in phonology, morphology and syntax (Makalela, 2009). This linguistic environment, with a rich spread of different and similar language pairs, makes South Africa an important place in which to study multilingual advantages, as they should manifest whether the mechanism leading to advantages is more engaged by switching between more distinct languages or by managing interference from more similar languages (Brito et al., 2014). The lack of visuospatial working memory advantages for multilinguals in this study is all the more surprising given this mix of language pairs, but because the similarity or difference between languages was not modelled no contribution can be made to hypotheses over which kind of multilingualism is more likely to lead to advantages. Future research in the country could explicitly compare bilinguals who speak two similar languages and bilinguals who speak two different languages with multilinguals who speak a combination of similar and different languages.

Comparing aspects of multilingualism

The importance of looking for significant differences among multilinguals based on a variety of characteristics and aspects of experience has been stressed in recent literature (Bialystok, 2017; Valian, 2015; Yow & Li, 2015). This study, which is one of the few to have focused on multilingual advantages in speakers of more than two languages, has important implications for the traditional approach to researching bilingual advantages. Finding that speaking a third (or fourth) language serves to augment the multilingual advantage (or affects cognition differently to speaking a second language) does not just complement classic bilingual advantage research; it *challenges* it. This is because many studies either define bilinguals as those who speak two *or more* languages (Bialystok, 2016), or fail to even consider whether the bilinguals in their sample speak additional languages other than the two studied (e.g. Tse & Altarriba, 2014). If speaking a third language leads to additional qualitatively similar effects to speaking a second then part of the effect observed in these studies will be due to the additional languages spoken by some of their participants, a factor

which should be modelled. If speaking a third language leads to disadvantages then the effects observed in these studies will be attenuated. Finally, if speaking a third language has qualitatively different effects to speaking a second then these studies will have conflated their effects and failed to understand the nuances of multilingual effects.

Because they do not simultaneously model information about proficiency, most studies that group speakers into broad categories (monolingual, bilingual etc.) have to adopt stringent inclusion criteria for a speaker's languages. For example, in classifying monolinguals and trilinguals Cockcroft et al. (2017) only count languages that a person has spoken since they were very young and is very proficient in. This is useful, and undoubtedly necessary, in narrowing the focus of studies to the high level multilingualism which is likely to lead to advantages, but it means that the phenomenon of more recent additional languages that are spoken at lower proficiency levels is completely ignored. Some research has found that languages spoken at low proficiency can determine multilingual advantages (Thompson, 2013), and even that advantages are most pronounced during initial exposure to a language (Heidlmayr et al., 2014), suggesting that the threshold hypothesis may not hold. This study's simultaneous modelling of the number of languages spoken and balance in proficiency meant that additional low-proficiency languages could be incorporated while factoring in their lack of equal standing with other languages.

The design of this study meant that several of the ambiguities surrounding the effect of speaking a third language could be clarified. Previously it has been difficult to separate the effect of speaking more languages from the (supposedly correlated) lower proficiency across additional languages (Brito et al., 2014). Indeed, in our sample those who reported speaking a third language were significantly less balanced across and within their languages, validating these concerns, but the multiple regression design meant that the effects of each variable could be understood independently of one another. This study does not directly support any of the hypotheses in the emerging trilingual advantage literature, as speaking a third language was not found to significantly predict any of the components of working memory. The bilingual advantage research pioneered by Bialystok and Majumder (1998), among others, challenged the previously dominant assumption that growing up multilingual had harmful

consequences for cognitive development. An important complement to the body of positive findings for multilingual advantages has been the lack (for the most part) of significant negative findings, except in naming and other linguistic tasks where multilinguals consistently perform more poorly (Bialystok, 2016). In this study multiple regression analysis extended this evidence to the effect of trilingualism and again no negative (nor positive) effect was found. On the other hand, the significance in the verbal processing regression of a variable measuring balanced proficiency among up to three languages is consistent with the notion that consideration of more than two of a speaker's languages can produce more sensitive measures of multilingualism.

Revisiting the link between multilingualism and working memory

The surprising non-significance of language variables in visuospatial working memory regressions means that it is important to reflect on alternative aspects of multilingual experience and working memory that can be studied in the future. Within quantitative studies of multilingualism there is a range of information that can be used to quantify multilingualism, and a range of ways of collecting that information. Measuring multilingualism through balance in proficiency may fail to detect the effects of multilingualism by focusing on an aspect of multilingual experience which is not key to determining cognitive advantages. As discussed in Chapter One, the basis for any measure of multilingualism should depend on the theory of how multilingualism leads to advantages in cognition. If the mechanism by which multilingualism leads to cognitive advantages is more dependent on recent experience of switching among languages (Pelham & Abrams, 2014), for example, than overall proficiency or AOA, then this study's balance-based operationalisation of multilingualism would have failed to detect multilingual advantages. Even if a continuous measure of multilingualism is indeed more suited to the study of multilingual advantages (Bialystok, 2017; Incera & McLennan, 2017), the particular variables (and interactions) used may not be optimal.

A central tenet of post-positivism is that the merits of different research methods, including different forms of data collection and the use of different kinds of data, must be

justified based on the details of the study at hand (Zammito, 2004). The question for this study then becomes whether the effects of multilingualism in the domain of cognition and working memory are more amenable to detection by qualitative investigation, with its more human and intuitive depth, or quantitative investigation, with its standardised tests and ability to compare large numbers of people relatively objectively.

Most of the discussion in this report has contrasted the informativeness and validity of different quantitative measures of multilingualism. At a more fundamental level, a study such as this which models the concepts involved as numbers can only capture certain aspects of the relation between the concepts (in this case multilingualism and working memory), and the decision about which aspect of multilingualism to focus on cannot justifiably be restricted to those aspects which are quantifiable. Though the information contained in quantitative measurements depends on what goes into the numbers, as discussed above, indubitably some aspects of a person's multilingual proficiency and experience, or at least the full depth and nuance thereof, cannot be captured numerically. Much of the effect of multilingualism could take the form of enhanced metalinguistic awareness (Ransdell et al., 2006), for example, which was not directly assessed in most of the studies reviewed here but could be a factor explaining some of the differential performance observed in quantitative tests. This is one example of a qualitative shift in thinking brought about by multilingualism which could not sensibly (or usefully) be conceived of as an increase of magnitude. Other changes may not concern superior awareness nor any improvement in existing capacities; multilinguals and monolinguals may utilise different processes to perform the same functions. It is interesting to think about whether Vygotsky's (1962) belief that multilingual children realise the arbitrary relation between word and reference earlier, a good example of an effect that is best understood qualitatively, might have an effect on performance in verbal working memory tasks. Would people (and specifically children) with a better understanding of the arbitrariness of a label be better (or possibly worse) at remembering and manipulating that label in certain ways, and could this awareness lead to them performing the task in different ways?

The difficulty of measuring the automatic, fast processes underlying working memory using qualitative or self assessment was used to justify its measurement using quantitative, psychometric tests in this study. The use of a scale of quantitative measures of working memory is closely related, conceptually and historically, to measures of fluid intelligence or IQ (Unsworth & Engle, 2007). Within the post-positivist paradigm guiding this study, these numbers are not meant to be equivalent to a person's cognitive capacity, nor are they meant to capture that capacity in its entirety. However, it is maintained that they are a valid way to represent certain aspects of cognition as numbers and that comparing people along a spectrum of numbers can be informative (in an intrinsically limited way). Even if this reasoning is valid, however, the particular aspect of working memory that is affected by multilingualism, if there is one, may still not be the aspect which is measured by the AWMA, and may not be quantifiable. The fact that working memory can, in some imperfect way, be represented by numbers, does not mean that it is always best to do so nor that every aspect of it can be captured in this way.

Conclusion

Partly due to the lack of research looking specifically at multilingualism's effects on working memory, much of this discussion and the literature review in Chapter One centred around studies looking at effects on other aspects of executive function. This means that several of the interpretations depend on assumptions about the relation between working memory and other executive functions. For example, the idea that our results add weight to Paap and Greenberg's (2013) claims against a multilingual advantage assumes that non-significant results from any study looking at multilingualism's effects on any executive function can be collected together in one coherent body of evidence. Studies like this one not only help to illuminate questions about the role of working memory in cognition and its relation to executive function, they are also important in building a body of working memory-specific multilingual advantage research on which future hypotheses and interpretations can be based.

An important contribution made by this study is the exceptional breadth of the data collected and the use of a comprehensive regression design. Whereas previous studies have used a limited number and type of tasks to measure working memory (often only one or two tasks, e.g. Ransdell et al., 2006), this study used a battery of tests which measured each of four different components of working memory with three tasks. This reduces the extent to which irrelevant idiosyncrasies of any task (or a participant's performance therein) will affect results and gives a more comprehensive measure of working memory. Likewise, the multilingualism variables and the SES score were based on several self-report responses, and a range of background information assisted in describing the sample and interpreting results. As discussed, a number of factors cast doubt on our pattern of significant and non-significant findings. However, the availability of a wide range of relevant variables and the use of comprehensive correlational and multiple regression analysis enabled an informed evaluation of the probability of each factor having affected results.

In this study a unique approach to measuring multilingualism was employed, with two scores capturing the extent of balance in proficiency within and across languages, respectively. Weighting self-reported proficiency scores to each individual's maxima (within each language or proficiency type) proved to be an original strategy that achieved the twin goals of explicitly measuring multilingual balance, which has emerged as a key determinant of cognitive advantages, and of creating a variable based on self-report data that has predictive power and is comparable across participants. English proficiency scores weighted against the maximum score across an individual's languages were shown to predict objective measures of English performance as well as absolute proficiency scores, and further evidence of the predictive power of these relative scores came from multiple regression results where between-language balance significantly predicted the verbal processing component of working memory.

That multilingual balance significantly predicted verbal working memory (in processing tasks) but not any aspect of visuospatial working memory appeared on first inspection to be in conflict with the predictions based on previous research where multilinguals were advantaged in domain-general executive function tasks but disadvantaged

in linguistic tasks. But a closer inspection of the relationships among the variables in this study and the pattern of empirical findings in the area suggested a more nuanced understanding of the influence of modality in working memory tasks is needed. A disadvantage in tasks which involve lexical access or verbal fluency (Bialystok, 2009) could coexist with advantages in memory processes applied to verbal content. If multilingual experience in storing and processing linguistic information leads to advantages that are specific to that kind of content then it would be expected that multilinguals' performance in verbal working memory tasks be superior to their performance in visuospatial working memory tasks. The plausibility of modality-specific working memory advantages for multilinguals coexisting with domain-general advantages in other aspects of executive function depends on the extent to which working memory and executive function are theoretically distinct. If most of the processing aspects of working memory come from other executive functions then it would be hard to explain why the strategies and improvements in processing can be specific to one kind of content in one aspect of executive function and not in another.

In an increasingly multilingual world it is vital that we discover more about the effects of speaking multiple languages. Multilingualism may affect several aspects of cognitive development and these need to be studied through a number of approaches before a more holistic understanding can be established. It is clear that the richness of multilingual experience cannot be understood only through the monolingual-bilingual dichotomy and that several aspects (qualitative and quantitative) need to be examined in incremental, nuanced ways. Knowledge of whether multilingualism leads to cognitive advantages, and which aspects are crucial to this, is helpful to millions of individuals, parents, teachers, and policy-makers in determining where, how and to what extent multilingualism should be encouraged and inculcated in homes and schools. But with so many people already multilingual another important implication of this research is in understanding the ways in which multilinguals are advantaged and incorporating this information into educational contexts. In contexts where multilinguals may be educationally disadvantaged due to socioeconomic factors, paying attention to the aspects of cognition in which they are advantaged by virtue of their linguistic experience could help to compensate for these disadvantages and provide students with

increased confidence. Though this study found multilingual advantages in verbal working memory, these results and others in the field are still at far too exploratory a stage to make explicit education or policy recommendations. However, the extension of the study of multilingual effects to working memory and the incorporation of additional languages beyond two are important steps towards a state of understanding from which more concrete recommendations can be made.

References

- Absatova, K. (2015). Applying automated working memory assessment and working memory rating scale in Russian population. *The Russian Journal of Cognitive Science*, 3(1-2), 21-33.
- Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, 23(4), 557-582.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.
- Alloway, T. P. (2007). *Automated Working Memory Assessment Manual*. London: Pearson Assessment.
- Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology*, 25(1), 84-93.
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the automated working memory assessment. *Educational Psychology*, 28(7), 725-734.
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, 77(6), 1698-1716.
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, 87(2), 85-106.
- Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, 16(1), 17-42.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. London: Psychology Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation* (pp. 47-89) Elsevier.
- Bak, T. H. (2016). The impact of bilingualism on cognitive ageing and dementia. *Linguistic Approaches to Bilingualism*, 6(1), 205-226.

- Bennett, M. R., Hacker, P. M. S., & Bennett, M. (2003). *Philosophical foundations of neuroscience* Blackwell Oxford.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(1), 3-11.
- Bialystok, E. (2016). The signal and the noise. *Linguistic Approaches to Bilingualism*, 6(5), 517-534.
- Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, 143(3), 233.
- Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240-250.
- Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics*, 19(1), 69-85.
- Black, J. E. (1998). How a child builds its brain: Some lessons from animal studies of neural plasticity. *Preventive Medicine*, 27(2), 168-171.
- Blom, E., Kuntay, 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.
- Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, , 1287-1294.
- Brito, N. H., Sebastian-Galles, N. & Barr, R. (2014). Differences in Language Exposure and its Effects on Memory Flexibility in Monolingual, Bilingual, and Trilingual Infants. *Bilingualism: Language and Cognition*. 18(4), 670-682.
- Bruyer, R., & Scailquin, J. (1998). The visuospatial sketchpad for mental images: Testing the multicomponent model of working memory. *Acta Psychologica*, 98(1), 17-36.
- Buchweitz, A., & Prat, C. (2013). The bilingual brain: Flexibility and control in the human cortex. *Physics of Life Reviews*, 10(4), 428-443.
- Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, 12(5), 815-823.

- 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 and Associated Disorders*, 24(2), doi: 118-125. 10.1097/WAD.0b013e3181ca1221
- Claassen, N. C. W., Krynauw, A. H., Paterson, H., & Wa ga Mathe, M. (2001). A standardisation of the WAIS-III for English-speaking South Africans. Pretoria: Human Sciences Research Council.
- Cockcroft, K., Wigdorowitz, M., & Liversage, L. (2017). A multilingual advantage in the components of working memory. *Bilingualism: Language and Cognition*, 1-15. doi:10.1017/S1366728917000475
- Cornwell, K., & Inder, B. (2008). Language and labour markets in south africa. *Journal of African Economies*, 17(3), 490-525.
- 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.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1-9.
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323-338.
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science (New York, N.Y.)*, 318(5855), 1387-1388. 318/5855/1387 [pii]
- Diamond, J. (2010). Social science. the benefits of multilingualism. *Science (New York, N.Y.)*, 330(6002), 332-333. 10.1126/science.1195067 [doi]
- 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.
- Dor, D. (2004). From englishization to imposed multilingualism: Globalization, the internet, and the political economy of the linguistic code. *Public Culture*, 16(1), 97-118.

- Engel de Abreu, P.M.J. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory, 19*(5), 529-537.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309.
- Extra, G., & Vallen, T. (1997). Migration and multilingualism in western europe: A case study of the netherlands. *Annual Review of Applied Linguistics, 17*, 151-169.
- 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.
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16*(2), 234-248.
- Garcia, O., & Wei, L. (2014). *Translanguaging: Language, Bilingualism and Education*. Basingstoke: Palgrave Macmillan.
- Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., & Cowan, N. (2000). The development of auditory attention in children. *Frontiers in Bioscience, 5*(1), d108-120.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(02), 67-81.
- Grosjean, F. (2010). *Bilingual*. Harvard University Press.
- Hakuta, K. (1986). *Mirror of language. the debate on bilingualism*. ERIC.
- Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences, 11*(6), 236-242.
- Heidlmayr, K., Moutier, S., Hemforth, B., Courtin, C., Tanzmeister, R., & Isel, F. (2014). Successive bilingualism and executive functions: The effect of second language use on inhibitory control in a behavioural stroop colour word task. *Bilingualism: Language and Cognition, 17*(3), 630-645.

- Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, *18*(4), 625-658.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368-1378.
- Hyltenstam, K. (1988). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual & Multicultural Development*, *9*(1-2), 67-84.
- Incera, S., & McLennan, C. T. (2017). Bilingualism and age are continuous variables that influence executive function. *Aging, Neuropsychology, and Cognition*, , 1-21.
- Kahn-Horwitz, J., Schwartz, M., & Share, D. (2011). Acquiring the complex English orthography: A trilingual advantage? *Journal of Research in Reading*, *34*(1), 136-156.
- Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705-710.
- Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging*, *23*(1), 70.
- Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, *43*(6), 456-473.
- Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514.
- Leibbrandt, M., Finn, A., & Woolard, I. (2012). Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa*, *29*(1), 19-34.
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605-621.
- Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, *14*(4), 588-595.
- Makalela, L. (2009). Harmonizing South African Sotho language varieties: Lessons from reading proficiency assessment. *International Multilingual Research Journal*, *3*(2), 120-133.
- Makalela, L. (2013). Translanguaging in Kasi-taal: Rethinking old language boundaries for new language planning. *Stellenbosch Papers in Linguistics Plus*, *42*(1), 111-125.

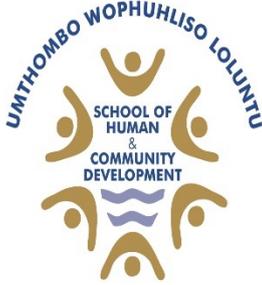
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition, 6*(2), 97-115.
- 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.
- Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology, 114*(2), 187-202.
- Nel, N., & Müller, H. (2010). The impact of teachers' limited english proficiency on english second language learners in south african schools. *South African Journal of Education, 30*(4), 635-650.
- Nobre, A. d. P., Rodrigues, J. d. C., Sbicigo, J. B., Piccolo, L. d. R., Zortea, M., Duarte Junior, S., & Salles, J. F. d. (2013). Tasks for assessment of the episodic buffer: A systematic review. *Psychology & Neuroscience, 6*(3), 331-343.
- 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.
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex, 69*, 265-278.
- Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 313.
- Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 113*(4), 535-551.

- Ransdell, S., Barbier, M., & Niit, T. (2006). Metacognitions about language skill and working memory among monolingual and bilingual college students: When does multilingualism matter? *International Journal of Bilingual Education and Bilingualism*, 9(6), 728-741.
- Richter, L., & Samuels, M. (2018). The south african universal preschool year: A case study of policy development and implementation. *Child: Care, Health and Development*, 44(1), 12-18.
- Riggs, N. R., Shin, H., Unger, J. B., Spruijt-Metz, D., & Pentz, M. A. (2014). Prospective associations between bilingualism and executive function in latino children: Sustained effects while controlling for biculturalism. *Journal of Immigrant and Minority Health*, 16(5), 914-921.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, 15(1), 1-20.
- Roux, J. C., & Visagie, A. S. (2007). (2007). Data-driven approach to rapid prototyping xhosa speech synthesis. Paper presented at the *Workshop Proceedings*, , 95
- Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science (New York, N.Y.)*, 323(5916), 951-953. 10.1126/science.1167025 [doi]
- Rudwick, S. (2004). 'Zulu, we need [it] for our culture': Umlazi adolescents in the post-apartheid state. *Southern African Linguistics and Applied Language Studies*, 22(3-4), 159-172.
- Sabourin, L., & Vinerte, S. (in press). Cognitive control among immersed bilinguals: considering differences in linguistic and non-linguistic processing. *Bilingualism: Language and Cognition*.
- Sagasta Errasti, M. P. (2003). Acquiring writing skills in a third language: The positive effects of bilingualism. *International Journal of Bilingualism*, 7(1), 27-42.
- Schumacher, K. L., & Gortner, S. R. (1992). (Mis)conceptions and reconceptions about traditional science. *ANS.Advances in Nursing Science*, 14(4), 1-11.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611.
- Sivia, D., & Skilling, J. (2006). *Data analysis: A bayesian tutorial* OUP Oxford.

- Skinner, B. F. (1987). Whatever happened to psychology as the science of behavior? *American Psychologist*, 42(8), 780.
- South African Audience Research Foundation (SAARF) (2001). *Living Standards Measure*. Retrieved on 4 March 2013 from: <http://www.saarf.co.za/LSM/lsm.asp>
- St Clair-Thompson, H. L. (2010). Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology*, 22(2), 286-296.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643.
- Thompson, A. S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal*, 97(3), 685-701.
- Tran, N. A., Behseta, S., Ellis, M., Martinez-Cruz, A., & Contreras, J. (2015). The effects of Spanish English dual language immersion on student achievement in science and mathematics. *E-Journal of Education Policy*,
- Tse, C., & Altarriba, J. (2014). The relationship between language proficiency and attentional control in cantonese-english bilingual children: Evidence from simon, simon switching, and working memory tasks. *Frontiers in Psychology*, 5, 954.
- Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, 133(6), 1038.
- Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, 18(01), 3-24.
- Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, 141, 65-82.
- Vygotsky, L. S. (1962). *Language and thought*. Massachusetts Institute of Technology Press, Ontario, Canada,
- Wechsler, D. (1997). *WAIS-III: Wechsler adult intelligence scale*. Psychological Corporation San Antonio, TX.

- Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences*, 35(4), 863-875.
- Wilson, M., & Emmorey, K. (1997). A visuospatial “phonological loop” in working memory: Evidence from american sign language. *Memory & Cognition*, 25(3), 313-320.
- Yanow, D., & Schwartz-Shea, P. (2015). *Interpretation and method: Empirical research methods and the interpretive turn*. Routledge.
- Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology*, 6, 164.
- Zammito, J. H. (2004). *A Nice Derangement of Epistemes: Post-Positivism in the Study of Science from Quine to Latour*. Chicago & London: The University of Chicago Press.
- Zerbian, S., & Barnard, E. (2010). Word-level prosody in Sotho-Tswana. Paper presented at the *Speech Prosody 2010-Fifth International Conference*.

Appendix A
Information Letter



Psychology
School of Human & Community
Development
University of the Witwatersrand
Private Bag 3, Wits, 2050
Tel: 011 717 4503 Fax: 011 717
4559

Dear Sir/Madam

As part of a partial fulfillment of a postgraduate Masters degree in Research Psychology at the University of the Witwatersrand, I am currently conducting research examining how language proficiency affects working memory. I understand that taking part in such research requires an investment of your time, but your participation will be greatly appreciated. I would therefore like to invite you to participate in this research. Participation is entirely voluntary and choosing to participate will not advantage or disadvantage you in any way. There are no foreseeable risks or benefits involved in your participation. Please note that you have the right to refuse to participate or to withdraw from the study at any time and this will not be held against you in any way.

The assessment procedure will be conducted individually at a time suitable to you, whereby a portion of the assessments will be computer-based and recorded. Participation in the study will involve signing consent forms, completing questionnaires, having some of your answers to questions audio-recorded, and completing a compilation of cognitive assessments. This will take approximately 90 minutes in total, and you would be required to stay for the duration of the assessments. After the completion of the assessments, you will be free to leave. By consenting to participate in this study, it will be declared that you have not suffered a previous head injury or concussion, are not currently taking drugs or alcohol, have not been formally diagnosed with a psychological disorder and/or are taking prescribed medication, have not been diagnosed with a learning or language disorder and have had no previous exposure to the assessments involved in this study.

Due to the face-to-face nature of this study, complete anonymity is not guaranteed. However, this information will not be linked to any results or traceable to any individual. Anonymity will be guaranteed in resulting reports, theses and/or publications. Complete confidentiality will be guaranteed since only the researcher and supervisor have access to the questionnaires and corresponding data, and anonymous participant codes will be used to identify both. The results of this research will not be used to examine individual performance, but rather group trends.

If you choose to participate, you will be consenting to the abovementioned information, and you will be giving permission for all data collected throughout the study to be used for analysis, reporting, and possible publication or conference presentation. You may receive general feedback on the outcomes of the study if requested, but given the confidential nature of the study, individual feedback is not possible.

Please see the contact details provided below if you have any further questions, concerns or you require feedback on the progress of the research. Thank you again for considering being part of this study. Please detach and keep this sheet for future reference.

Mr. Gabriel Espi-Sanchis
(Supervisor)
1778063@students.wits.ac.za

Prof. Kate Cockcroft
kate.cockcroft@wits.ac.za

Appendix B

Demographic Questionnaire

CODE

--	--	--	--

Demographic Questionnaire

Gender:

M	F
---	---

Date of Birth:

D	D	M	M	Y	Y	Y	Y
---	---	---	---	---	---	---	---

Home Language(s):

School Language(s):

Current Degree and Faculty:

Previous degrees or qualifications:

Current year of study (1st, 2nd, 3rd):

How many years have you been at university?

Did you ever fail a grade at school? If so, which one?

Did you ever require an intervention from a language specialist?

Did you attend pre-primary school?

Living Amenities & Caregiving

Educational and occupational status of your parents or primary caregivers:

Mother: Level of Education		Father: Level of Education	
No schooling		No schooling	
Less than primary school completed		Less than primary school completed	
Primary school completed		Primary school completed	
Secondary school not completed		Secondary school not completed	

Secondary school completed		Secondary school completed	
Tertiary education completed		Tertiary education completed	
Other		Other	
Current occupation:		Current occupation:	

Marital status of primary caregivers:

Married	
Living together as husband and wife	
Widow/widower	
Divorced/separated	
Other	

Number of caregivers in the household in which you spend the most time (please tick):

0	
1	
2	
>2	

Living Standards Measure:

Please answer the following questions according to your circumstances while growing up, and not in your current student accommodation if these are different.

Question	Answer	
1. I have the following in my household:		
TV set	TRUE	FALSE
VCR	TRUE	FALSE
DVD player	TRUE	FALSE
M-Net/DStv subscription	TRUE	FALSE
Hi-fi/music centre	TRUE	FALSE
Computer / Laptop	TRUE	FALSE
Vacuum cleaner/floor polisher	TRUE	FALSE
Dishwashing machine	TRUE	FALSE
Washing machine	TRUE	FALSE
Tumble dryer	TRUE	FALSE
Home telephone (excluding a cell)	TRUE	FALSE
Deep freezer	TRUE	FALSE
Fridge/freezer (combination)	TRUE	FALSE
Electric stove	TRUE	FALSE
Microwave oven	TRUE	FALSE

Built-in kitchen sink	TRUE	FALSE
Home security service	TRUE	FALSE
3 or more cell phones in household	TRUE	FALSE
2 cell phones in household	TRUE	FALSE
Home theatre system	TRUE	FALSE
2. I have the following amenities in my home or on the plot:		
Tap water in house/on plot	TRUE	FALSE
Hot running water from a geyser	TRUE	FALSE
Flush toilet in/outside house	TRUE	FALSE
3. There is a motor vehicle in our household	TRUE	FALSE
4. I am a city dweller	TRUE	FALSE
5. I live in a house, cluster or town house	TRUE	FALSE
6. I live in a rural area outside Gauteng and the Western Cape	TRUE	FALSE
7. There are no radios, or only one radio (excluding car radios) in my household	TRUE	FALSE
8. There is no domestic workers or household helpers in household (both live-in & part time)	TRUE	FALSE

Appendix C
Humanities Ethics Clearance Certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (SCHOOL OF HUMAN & COMMUNITY DEVELOPMENT)

CLEARANCE CERTIFICATE

PROTOCOL NUMBER: MPSYC/17/003 IH

PROJECT TITLE:

The relationship between proficiency in multiple languages and working memory: a study of multilingual advantages in South Africa.

INVESTIGATORS

Espi-Sanchis Gabriel

DEPARTMENT

Psychology

DATE CONSIDERED

12/07/17

DECISION OF COMMITTEE*

Approved

This ethical clearance is valid for 2 years and may be renewed upon application

DATE: 12 July 2017

CHAIRPERSON
(Professor Brett Bowman)

cc Supervisor:

Prof. Kate Cockcroft
Psychology

DECLARATION OF INVESTIGATOR (S)

To be completed in duplicate and **one copy** returned to the Secretary, Room 100015, 10th floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure be contemplated from the research procedure, as approved, I/we undertake to submit a revised protocol to the Committee.

This ethical clearance will expire on 31 December 2019

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix D

Table D1

Multiple regression results with the addition of AOA

	(1)	(2)	(3)	(4)
Independent variables	Verbal storage	Verbal processing	Visuospatial storage	Visuospatial processing
Within balance	1.74 (1.96)	2.15 (1.93)	-1.42 (2.26)	.11 (2.21)
Between balance	-1.49 (1.52)	2.70 (1.50)	.70 (1.75)	2.79 (1.71)
Third language	-1.44 (3.00)	-1.03 (2.97)	-.43 (3.47)	-3.18 (3.39)
Between balance x Third language	.52 (1.70)	-1.55 (1.67)	-1.32 (1.96)	-3.39 (1.91)
AOA	-.00 (.03)	-.01 (.03)	.03 (.03)	.03 (.03)
SES	.79 (.42)	.41 (.41)	-.58 (.48)	-.16 (.47)
Constant	84.89* (4.75)	95.66* (4.69)	94.30* (5.48)	99.18* (5.35)
<i>F</i> (5, 182)	2.06	1.49	0.90	0.75
<i>R</i> ²	0.0636	0.0467	0.0290	0.0243

Notes. $N = 189$. For all independent variables coefficients are shown with standard errors in parentheses. Models (1) to (4) were four separate regressions with the name of the model describing the dependent variable of each.

Within balance = within-language balance in proficiency, Between balance = between-language balance in proficiency, Third language = whether the participant speaks a third language, Between balance x Third language = the interaction between the two variables, AOA = the age of acquisition of a participant's second language, SES = socioeconomic status.

* $p < .05$