SECOND LANGUAGE LEARNING-INDUCED ENHANCEMENT OF EXECUTIVE FUNCTIONS IN AN INSTRUCTED LANGUAGE LEARNING CONTEXT

MEHMET AKINCI

BOĞAZİÇİ UNIVERSITY

2020

SECOND LANGUAGE LEARNING-INDUCED ENHANCEMENT OF EXECUTIVE FUNCTIONS IN AN INSTRUCTED LANGUAGE LEARNING CONTEXT

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Foreign Language Education

by

Mehmet Akıncı

Boğaziçi University

2020

DECLARATION OF ORIGINALITY

I, Mehmet Akıncı, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature.........MAK.........

Date ......August 18, 2020......

ABSTRACT

Second Language Learning-Induced Enhancement of Executive Functions in an Instructed Language Learning Context

This study probes the effects of L2 experience on the enhancement of executive functions (EFs) and the predictive validity of EFs for L2 success from a domain-general perspective in a longitudinal pre-test/post-test design. Participants in the experimental group were 165 Turkish high-school graduates receiving intensive L2 instruction in a university setting for six months (600 class contact hours), and the ones in the control group were 103 freshman students taking undergraduate courses in their first language. Two visual complex span tasks (symmetry and rotation) to measure working memory, antisaccade and flanker tasks to measure inhibition, and one standardized English proficiency task were administered before and after the instruction. Multivariate and univariate repeated measures ANOVA results indicated that both instruction types led to the enhancement of EFs except for flanker performances, yet no group differences were observed. Exploratory factor analysis revealed that complex span task and antisaccade performances had loadings on the same factor, named as executive attention, whereas flanker performance was independent from this dimension. Regression findings demonstrated that, albeit little, flanker but not executive attention could explain L2 success. Prior L2 experience was found to be the best predictor. The study concluded that L2 experience can contribute to the enhancement of EFs, yet executive attention might not play a significant role in L2 success at the end of a six-month intensive L2 instruction.

ÖZET

Yönlendirilen Dil Öğrenimi Bağlamında, İkinci Dil Öğreniminden Kaynaklı Yönetici İşlevlerin Gelişmesi

Bu çalışma, ikinci dil deneyiminin yönetici işlevlerin genişlemesi üzerindeki etkilerini ve yönetici işlevlerin ikinci dil başarısı ile ilgili yordama geçerliliğini, alan-genel perspektiften, ön-test/son-test boylamsal bir araştırma ile incelemektedir. Deney grubundaki katılımcılar, üniversite ortamında altı ay boyunca yoğun ikinci dil eğitimi alan (600 ders saati) 165 lise mezunu Türk öğrenciden oluşmaktadır. Kontrol grubundakiler ise, üniversite birinci sınıf derslerini ana dillerinde alan 103 Türk üniversite öğrencisidir. İşler belleği ölçen, iki adet görsel kompleks uzam testleri (simetri ve rotasyon); ketlemeyi ölçen antisakkad ve flanker testleri; ve bir adet standartlaştırılmış İngilizce dil yetisi testi öğretim öncesi ve sonrasında uygulanmıştır. Tekrarlı ölçümler için kullanılan çok değişkenli ve tek değişkenli varyans analizi sonuçları, her iki öğrenim şeklinin de flanker ketleme performansı hariç yönetici işlevlerde genişlemeye yol açtığını göstermiştir, ancak herhangi bir grup farkına rastlanmamıştır. Keşif faktörü analizi, kompleks uzam ölçekleri ile antisakkad ölçeğinin, yönetici dikkat adı verilen aynı faktör altında toplandığını, ancak flanker ölçeğinin bu faktörden bağımsız hareket ettiğini göstermiştir. Regresyon analiz sonuçları, yönetici işlevlerden sadece flanker ölçeğinin az da olsa ikinci dil öğrenim başarısını açıklayabildiğini, yönetici dikkatinin ise ikinci dil öğrenim başarısı ile ilişkisi olmadığını göstermiştir. Ayrıca, çalışma öncesi ikinci dil deneyimin, başarı konusunda en iyi açıklayıcı olduğu görülmüştür. Sonuç olarak, ikinci dil deneyiminin, yönetici işlevlerin genişlemesine katkısı vardır, denilebilir.

Ancak, yönetici dikkatin altı aylık yoğun bir dil eğitimi sonundaki başarıda önemli bir rol oynadığı söylenemeyebilir.

CURRICULUM VITAE

NAME: Mehmet Akıncı

DEGREES AWARDED

PhD in Foreign Language Education, 2020, Boğaziçi University
MA in Foreign Language Education, 2009, Boğaziçi University
BA in Foreign Language Education, 2006, Boğaziçi University

AREAS OF SPECIAL INTEREST

Bilingual memory, working memory, attention control, language assessment, research methodology, individual differences in language learning

PROFESSIONAL EXPERIENCE

Assessment Deputy Manager, İstanbul Şehir University, 2018-2020
Proficiency Test Item Writer, İstanbul Şehir University, 2016-2018
A2 Level Coordinator, İstanbul Şehir University, 2012-2016
English Language Instructor, İstanbul Şehir University, 2011-2012
English Language Instructor, Maltepe University, 2007-2010

AWARDS AND HONORS
Highest Honors List, Boğaziçi University, 2009
Honors List, Boğaziçi University, 2006

PUBLICATIONS

Journal Articles

Akıncı, M. (2020). Cognitive effects of third language experience and the role of language proficiency. *Boğaziçi University Journal of Education, 37*(1), 69-83.

Akıncı, M., & Yıldız, S. (2017). Effectiveness of corpus consultation in teaching verb+noun collocations to advanced ELT students. *Euroasian Journal of Applied Linguistics, 3*(1), 91-109.

Akıncı, M., Bektaş, S., Gülle, T., Kurt, S., & Kurt, Y. (2016). Ses temelli cümle yöntemi ile okuma-yazma eğitimi / The sound-based literacy instruction. *Boğaziçi Üniversitesi Eğitim Dergisi / Boğaziçi University Journal of Education, 33*(2), 97-115.

Conference Proceedings

Akıncı, M., & Ateş B. (2017). Assessment literacy to facilitate teacher intervention and improve student learning. In *Towards Higher Education (THE): Bridging the Gap. Boğaziçi University School of Foreign Languages ELT Conference*, İstanbul, Turkey.

Akıncı, M. (2015) Barriers encountered during the implementation of 1:1 computing projects and viable solutions. In *EDTECHIST Educational Technology Conference*, Istanbul, Turkey.

Akıncı, M., & Ateş B. (2014). Washback under control. In *IATEFL Testing, Evaluation and Assessment SIG Conference*, Granada, Spain.

ACKNOWLEDGEMENTS

There are many special people to whom I would like to thank for their invaluable contributions, both directly and indirectly, at different stages of my PhD journey. Yet, it is not possible to give a particular mention to every single of them here.

First and foremost, I am fully indebted to my supervisor, Dr. Gülcan Erçetin, for putting her valuable experience and unsurpassed knowledge at my disposal. It has been a greatly enriching experience to work under her guidance. I will remain grateful to her for taking time in many helpful discussions, reading my manuscripts, and giving such constructive and thoughtful feedback throughout my PhD journey. Without her effective supervision, understanding, and encouragement, this PhD would not have been achievable. It was her who pushed me further than I thought I could go. She has been and will be an inspiration and a role model to me in my academic career.

I would like to express my immense gratitude to the outstanding supervisory committee: Dr. Belma Haznedar, as an expert on bilingualism, and Dr. Ayşe Reyyan Bilge Yıldırım, as an expert on cognitive psychology. They have made a huge contribution to the thesis by supervising my progress from beginning to end. And to the other thesis committee: Dr. Şebnem Yalçın, and Dr. Duygu Özge for critically assessing and approving the work presented in this thesis. I thank them all for participating in my defense and helping me strengthen my thesis.

I gratefully acknowledge the academic and financial support of Boğaziçi University. I undoubtedly learned a lot through the PhD courses, and this research was supported by the funding received through a research grant (Grant No. 14621) from the Boğaziçi University Scientific Research Project Fund.

I would like to acknowledge the valuable input of Dr. Randall W. Engle who contributed to the discussion helping to shape the research design, took time to read my research proposal, and encouraged me to collaborate with his research assistants, Christopher Draheim and Jason S. Tsukahara, at Georgia Institute of Technology. I greatly appreciate their support and guidance during this journey by providing me with the measurement tools and answering my questions regarding this research.

These acknowledgements would be incomplete without mentioning my research assistants: Esra Güneri, Kerime Tamer, Afra Selma Taygar, Ahsen Kenez, Elif Altın, and Semanur Konuk. I want to thank them all for their excellent work during data collection. I would like to mention Dr. Yıldırım's help to recruit these assistants and make it possible to use the Cognitive Psychology Lab at İstanbul Şehir University.

My deepest appreciation belongs to my love, my dear wife, Tuğba Akıncı, who has always been by my side throughout this PhD journey and lived every single minute of it, and without whom I would not have had sufficient motivation and courage to embark on this journey in the first place and completed the process successfully. She has always believed in me and encouraged me to follow my dreams. And to my beloved daughters, Elif Mira and İpek Vera, for their understanding and patience, and their love.

I also want to express my deep gratitude to all my family, particularly my mum, Hatice Akıncı, my father, Sadık Akıncı, my grandfather, Mehmet Akıncı, and my grandmother, Makbule Akıncı, for their vigorous support and encouragement at all times.

I would like to thank all my colleagues at İstanbul Şehir University and closest friends for their faithful support and constant belief that I could do the best and finish the PhD.

Last but not least, I would like to thank all of those whom I have not mentioned here by name, yet supported me during this journey.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Executive functions (EFs; also called executive control or cognitive control) referring to the processes needed to regulate thoughts and behaviors (Miyake & Friedman, 2012) are essential for a huge number and variety of real word tasks (Diamond, 2013; Engle, 2018). Environmental factors and experiences in various activities have a positive influence on cognitive and brain plasticity, thereby enhancing EFs (Cuevas, Rajan, & Bryant, 2018). One of them is experience with a second language (L2) in that the joint activation of two languages in a bilingual mind requires intensive and frequent use of cognitive control processes for language management by maintaining attention to the target language and avoiding interference from the competing one (Bialystok, Craik, Green, & Gollan, 2009). Since EFs and language functions rely on the same integrated brain network, L2 experience occurring in the verbal domain is expected to lead to broad cross-domain effects (Li, Legault, & Litcotsky, 2014b).

Inhibition and working memory (WM) are the two major and most commonly studied components of EFs in relation to L2 experience. The view that long-term experiences with an L2, acquired at early ages, enable bilingual children and older adults to have better EFs as compared to monolinguals is a widely held perspective (Bialystok, Craik, & Luk, 2012). However, findings of the studies comparing young adults having life-long L2 experience with their monolingual counterparts have resulted in growing skepticism concerning the claims for L2 experience-induced positive effects on EFs. While a number of studies demonstrated that long-term L2 experience enhanced EFs as regards inhibition (Abutalebi et al., 2012; Bialystok, Craik, & Ryan, 2006; Blumenfeld & Marrian, 2011; Chung-Fat-Yim, Sorge, & Bialystok, 2017; Mechelli et al., 2004; Yang & Yang, 2016), and WM capacity (Antón, Carreiras, & Duñabeitia, 2019; Jiao, Liu, Wang, & Chen, 2019; Luo, Craik, Moreno, & Bialystok, 2013; Sullivan, Prescott, Goldberg, & Bialystok, 2016), others found null effects both on inhibition (Antón et al., 2019; Bialystok, Martin, & Viswanathan, 2005; Kousaie, Sheppard, Lemieux, Monetta, & Taler, 2014; Paap & Greenberg, 2013) and on WM capacity (Bonfieni, Branigan, Pickering, & Sorace, 2020; Morrison, Kamal, & Taler, 2019; Ratiu & Azuma, 2015; Smithson & Nicoladis, 2013). The possible contributing factors for the contradicting results can be listed as (1) insufficient interindividual variability due to the nature of the measurement tools and the calculation of dependent variable (DV) (Bialystok, 2017; Chung-Fat-Yim et al., 2017), (2) treating bilingualism as a categorical variable rather than a continuous one (Bialystok, 2016), (3) confounding variables (e.g., individual differences rather than L2 learning) (van den Noort et al., 2019), and (4) misconceptualization of EFs on the basis of the componential unity/diversity model of Miyake et al. (2000) (Bialystok, 2017; Paap, 2019).

Furthermore, challenging the critical period hypothesis (Lenneberg, 1967), recent scientific evidence has some indications for short-term L2 learning- or intensive L2 training-induced changes in brain areas, particularly the areas bound to both EFs and language learning, for the individuals whose L2 experience begins as of/after puberty (Gurunandan, Carreiras, & Paz-Alonso, 2019; Hosoda, Tanaka, Nariai, Honda, & Hanakawa, 2013; Mårtensson et al., 2012; Schlegel, Rudelson, & Tse, 2012; Sullivan, Janus, Moreno, Astheimer, & Bialystok, 2014; Qi et al., 2019). However, little research has been conducted to explore the behavioral changes in EFs induced by L2 learning as of/after puberty, and the extant one indicates inconclusive findings. Accordingly, while these L2 learners did not outperform their monolingual counterparts in their performances in inhibition (Kalia, Wilbourn, & Ghio, 2014; Sullivan et al., 2014; Xie, 2018), they showed greater WM, yet depending upon the task (Dong, Liu, & Cai, 2018; Kerrigan, Thomas, Bright, & Filippi, 2017), degree of proficiency (Yang, 2017), or time period of learning and novelty of the language (Huang, Loerts, & Steinkrauss, 2020). The contradiction is plausible given the fact that brain changes can emerge after three months of L2 experience whereas qualitative changes in behavior require more intensive and quality experience (Driemeyer, Boyke, Gaser, Büchel, & May, 2008).

In addition, the issue of EF-related individual differences in L2 learning has been the focus of neurological and behavioral studies. As for inhibition, while it is related to online language processes (Linck, Hoshino, & Kroll, 2008; Linck, Schwieter, & Sunderman, 2012), behavioral data suggest that inhibition is not a robust predictor of L2 learning outcomes (Ghaffarvand Mokari & Werner, 2019; Linck & Weiss, 2011, 2015). With respect to WM, both neurological (Yang & Li, 2012) and behavioral findings (Linck, Osthus, Koeth, & Bunting, 2014) suggest that individual differences in WM can predict L2 success. However, further research is needed to clarify the relationship between EFs and L2 learning outcomes. In other words, while individual differences in EFs can predict L2 success, L2 experience could also have facilitative effects on EFs (Grundy & Timmer, 2017).

Although numerous studies have been conducted to explore the predictive validity of WM in relation to L2 learning outcomes, the majority have been cross-sectional and few have adopted a longitudinal design. These longitudinal studies reported varying degrees of WM effects due to the use of different research designs and measurement tools, and different starting proficiency levels of participants.

Although these studies had a pre-test/post-test design, WM was tested either at pre-test (e.g., Chang, Wang, Cai, & Wang, 2019; Kormos & Sáfár, 2008; Sagarra, 2017; Serafini & Sanz, 2016) or at post-test (e.g. Linck & Weiss, 2011, 2015) but not at both times. Other methodological limitations are as follows: (1) low sample size, (2) data collection periods shorter than six months, (3) not administering both pre- and post- training L2 proficiency task, and (4) employing only one complex WM task, typically verbal in nature, which disregards the domain-general view of WM capacity.

Given the significant implications of EFs for our daily cognitive activities, the little amount of research examining the relationship between EFs and L2 learning experience particularly as of/after puberty, the inconclusive findings and methodological problems of the extant behavioral studies, and the need to clarify the relationship between EFs and L2 learning outcomes, it gains utmost importance to fill the gap in the literature. To achieve this goal, the current study aims to investigate the longitudinal effects of L2 learning experience on EFs, by employing a longitudinal pre-test/post-test design on which both EFs tasks and an L2 proficiency test were administered at the beginning and end of a six-month intensive language instruction to 165 high-school graduates as L2 learners who took an intensive language instruction and their 103 near-monolingual counterparts who studied freshman courses[1].

[1] The term 'near-monolinguals' is used by Yang (2017) to describe the participants who have been exposed to L2, yet their L2 proficiency is limited. In today's world, it is almost impossible to have young adult participants with pure monolingualism as the second language instruction, albeit at a low level, is provided in the school system prior to university education.

CHAPTER 2

LITERATURE REVIEW

In this chapter, theoretical background of the present study will be provided in detail. First, the conceptualization of executive functions (EFs) together with its two major components, namely working memory and inhibition, will be discussed. Then, the enhancement of EFs through second language learning (L2) will be examined. After that, the predictive power of EFs for L2 learning outcomes will be analyzed. In the end, the summary and the research questions as well as the hypotheses will be provided.

2.1 Executive functions

2.1.1 Conceptualization of EFs

Having burgeoning popularity due to its contribution to success in a variety of domains and fields, EFs have garnered a significant amount of attention and are one of the most widely cited, yet befuddling constructs due to a diversity of taxonomies to define and the heterogeneity of available tasks to measure them (Baggetta & Alexander, 2016; Karr et al., 2018; Packwood, Hodgetts, & Tremblay, 2011).

More specifically, Eslinger (1996) found that 33 definitions of EFs had been proposed by the researchers. More than 30 components of EFs were revealed in 11 papers published between 1974 and 2004 (Jurado & Rosselli, 2007), and in a relatively more recent review, 68 components were identified and reduced to 18 through removing semantic and psychometric overlap between terms (Packwood et al., 2011). In a systematic review of 106 contemporary empirical studies, Baggetta and Alexander (2016) identified 48 models of EFs employed to define the construct. The most frequently cited models were the ones offered by Miyake et al. (2000) and Diamond (2006, 2013) while the most frequently cited components were ‘inhibitory control or response inhibition’, ‘WM or updating’, ‘shifting or set-shifting’ (Baggetta & Alexander, 2016).

Baggetta and Alexander’s (2016) review also revealed that 109 tasks purported to measure the construct were administered (56 of them used only once), and the examination of these tasks germane to the components mentioned in the model of Miyake and colleagues indicated that WM was assessed 107 times by 33 different tasks, inhibition 92 times by 28 different tasks, and shifting 80 times by 33 different tasks. Many of the tasks were found to be employed to measure different components by different researchers. Another finding arising from the review was that one single task was used to assess one component (e.g., Patrick, Blair, & Maggs, 2008), which would call into question the reliability or validity of the research due to the task impurity problem since each task measures EF abilities besides task-specific non-EF abilities, and any single task is unlikely to measure the construct adequately (Bagetta & Alexander, 2016; Miyake & Friedman, 2012; Packwood et al., 2011). At this point, Baggetta and Alexander (2016) also highlight the similar role of employing tasks with a background of no rigorous psychometric evaluation. Presence and employment of a number of different measurement tools to assess EFs have contributed to an important contention about EFs, namely the issue of dimensionality. Extant research indicates that a factor analytic approach has been employed to explore the correlations among the tasks and a heterogeneous number of factors varying from one factor (e.g., Della Salla, Gray, Spinnler, & Trivelli, 1998) to as many as six factors (Testa, Bennett, & Ponsford, 2012) have been found.

Reevaluating previously supported factor solutions of confirmatory factor analyses (CFA) in seventeen articles published between 1998 and 2016, Karr et al. (2018) indicated a gradual differentiation of EFs from children to adults (i.e., unidimensionality of EFs for children and adolescents, and both unity and diversity and nested-factor models for adults). However, an issue of a publication bias towards Miyake et al. (2000) was highlighted, and interpretation of the inferences was suggested to be made with caution. In the following subsection, the most frequently cited model of EFs, the unity and diversity model of EFs by Miyake et al. (2000), will be discussed.

2.1.2 The unity and diversity model of EFs

Defining EFs as the processes controlling and regulating thought and action, Miyake et al. (2000) found three components of EFs (latent variables) in college students, namely inhibiting prepotent responses (inhibition), updating WM representations (updating), and shifting between tasks or mental sets (shifting). In their analysis through a parameterization of the correlated factors (three latent variables with three observed task data per each), Miyake et al. (2000) found that three latent models correlated with each other, thus sharing a common underlying ability (unity); on the other hand, since the confidence intervals for correlations did not include 1.0, they provided evidence for separability across the components with variance unique to each particular component (diversity).

To explore the unity in depth, Friedman and colleagues (Friedman et al., 2008, 2016; Friedman, Miyake, Robinson, & Hewitt, 2011) employed a bifactor parameterization and reached a model with a similar fit on which a common factor (Common EF) captured the unity predicting all nine tasks; and only updating and shifting exhibited unique variance. The explanation of the unity did not leave any unique variance for inhibition. In other words, the inhibition was argued to be virtually perfectly correlated with Common EF. Therefore, Miyake and Friedman (2012) contend that Common EF is related to one's ability germane to the active maintenance of task goals and goal-related information. It also plays a role in employing this information to bias lower-level processing in an effective way. Friedman and Miyake (2017) elaborate on the issue of Common EF by arguing that goal management and implementation is a requirement of all the EF tasks, yet it is particularly essential for the inhibition task which requires avoiding prepotent responses or conflicting information. Additionally, they provide an explanation to dispel the misconception about their proposal that Common EF is just about maintenance of goals, arguing that it also functions in retrieving right goals and implementing them at the right time (Gustavson, Miyake, Hewitt, & Friedman, 2015).

The commonality among the EF components is also indicated by neural data as well. A quantitative meta-analysis was conducted by Niendam et al. (2012) who analyzed 193 functional neuroimaging studies pertinent to EFs. The results revealed that, during the performance of functions across EF domains (e.g., inhibition, and WM), significant activations were observed on a common set of cortical and subcortical regions within the cognitive control circuit, which indicates that EFs are supported by a superordinate cognitive control network. The cortical regions were frontal (dorsolateral prefrontal cortex [DLPFC] and anterior cingulate cortex [ACC]), and parietal (inferior and superior parietal lobe) regions in addition to the regions related to verbal and auditory stimuli, namely occipital (Broadman area [BA], 19) and temporal regions (BAs, 13, 22, 37) respectively. Last, such subcortical regions as thalamus, caudate and putamen were also activated.

In short, the most commonly cited model of EFs, namely the unity and diversity model of Miyake et al. (2000), posits three latent variables or components that, though correlated (unity), are separable (diversity). The unity part or the common variance (i.e., Common EF) has a perfect correlation with inhibition (e.g., Friedman et al., 2008), and its function is thought to maintain goal-related information, and retrieve right goals and implement them (Friedman & Miyake, 2017). This is very similar to the mechanisms mentioned by Engle and colleagues in their framework of executive attention (e.g., Engle & Kane, 2004, Engle, Kane, & Tuholski, 1999a;), which will be discussed in detail in the following subsection.

2.1.3 The working memory model of EFs

Working memory (WM) has been a central concept to multiple theories of the control of thought and action in cognitive psychology (Cowan et al., 2005; Engle & Kane, 2004). The term, 'WM', was first utilized by computer scientists to refer to the structures set up in programs to hold information only temporary to be employed in carrying out certain operations (Newell & Simon, 1956). In human research, it was first used by Miller, Galanter, and Pribram (1960) who considered WM as a part of the mind allowing to store information about goals and sub-goals of immediate concern needed to perform ecologically appropriate actions.

In one of the most cited papers of the time, Atkinson and Shiffrin (1968) employed the term 'short-term memory' (STM) or 'immediate memory' in their modal model and described it as a unitary mechanism holding information temporarily to be used in processing. They also put an emphasis on the control processes receiving information into and taking it out of the short-term store. The memory span was measured with a simple span task presenting a list of verbal items and requiring a participant to repeat the list verbatim. The longest list repeated accurately was evaluated as her/his memory span.

Baddeley and Hitch (1974) consulted to the term, WM, to make their views distinguished from the modal model. In contrast to a single mechanism proposed by the modal model, they defined WM as a multi-component system to store information temporarily as it is processed in ongoing mental activities. The system has two modality-specific rehearsal buffers: one to store verbal information (phonological loop) and one to store visual and spatial information (visuo-spatial sketchpad); and one modality-independent central executive component that is an attentionally-based control system managing the flow of information in and out of its two slave storage components. In contrast to a simple span task, Baddeley and Hitch (1974) presented a task involving retention of a list in memory while at the same time carrying out another process, which caused interference to be dealt with to perform the task through a multicomponent storage and processing system restricted by time and capacity.

Working within the framework of Baddeley and Hitch (1974) to understand individual differences, Daneman and Carpenter (1980) devised such complex span tasks as reading span and listening span purported to adequately measure WM since both storage and processing mechanisms were to be engaged. Unlike simple span tests, a participant reads or listens to a set of sentences aloud (processing) and is required to recall the final word of each sentence (storage) in these complex tasks. The number of sentences increased from trial to trial; and the span of the subject is set according to the largest set size s/he recalls all the sentence-final words from.

Although Baddeley and Hitch (1974) incorporated earlier STM literature to their slave systems, neither Baddeley and Hitch nor Daneman and Carpenter made an explicit distinction between WM and STM in their studies. With respect to the distinction at a conceptual level, Cowan (1998), advocating the existence of a single storage system consisting of elements at different levels of activation, posited that the contents of the system can be regarded as long-term memory (LTM), most elements of which are in inactive state. Some of the elements, on the other hand, are above some threshold of activation yet outside the focus of attention. Additionally, there are some other elements in a higher state of activation and inside the focus of attention, and the maintenance of the elements in this state requires controlled, limited-capacity attention processes. Cowan considered STM as the subset of WM and claimed that STM is a simple storage component for LTM units above threshold but not within the focus of attention while WM includes the contents of the STM and controlled attention.

Endorsing Cowan's view of WM, Engle, Tuholski, Laughlin, and Conway (1999b) looked into the differences between STM and WM on an empirical level in a latent-variable study having both STM (forward-word span with dissimilar words, forward-word span with similar words, backward-word span with dissimilar words) and WM (complex span: reading span, operation span, counting span) tasks. Conceptually, their argument was that the shared variance between STM and WM tasks would be regarded as STM (domain-specific stores with associated rehearsal procedures), and the variance left over as the controlled attention or central executive component of WM[2]. The results revealed that STM and WM constructs are highly

---

[2] Engle et al. (1999a) made a difference between the concepts of WM and WMC, and named the controlled attention part of WM as WM capacity (WMC). Engle (2002) and Engle and Kane (2004) call this part 'executive attention' as well.

related but separable, and only executive attention or working memory capacity (WMC) could explain the higher-order functioning operationalized as the general fluid intelligence (Gf). Since the relationship between the attention capability and Gf was obtained through verbal WM tasks and non-verbal Gf tasks, the executive attention could be posited to be domain-general; that is, individual differences in executive attention could presumably reveal themselves in a wide variety of cognitive tasks.

One alternative explanation to the relation between the complex span tasks and Gf would be that the processing component of the WM tasks might play a role. However, Engle, Cantor, and Carullo (1992) controlling the processing speed, and Conway and Engle (1994) controlling the processing demands indicated that the correlation between WM and higher-order cognitive tasks is not a result of processing skills used in the WM tasks.

In their theory of executive attention, Engle and Kane (2004) propose a two-factor model to explain the nature of WMC: memory phenomenon and attention phenomenon. The former reflects the role of executive attention to maintain the task-related goals or information in active and accessible state, and the latter the role of inhibition to resolve response competition or conflict caused particularly by prepotent or habitual behaviors (attention control, AC).

Unsworth and Engle (2007) elaborate on the memory phenomenon with an explanation of two mechanisms: a limited-capacity yet dynamic system to maintain goal-related information in active state (primary memory, PM), and cue-dependent retrieval of relevant information from long-term storage (secondary memory, SM). PM is basically what Engle and Kane (2004) discuss in the memory factor above. As for the long-term storage, some goal-relevant information is displaced from PM due to its limited-capacity, so it should be retrieved from SM via a strategic search in the presence of irrelevant information. Unsworth and Engle (2007) contend that the individual differences in WMC occur due to the differences not only in active maintenance in PM but also in performing searches in SM.

Shipstead, Lindsey, Marshall, and Engle's (2014) study revealed that PM, SM, and AC are three dissociable mechanisms of WMC. They are all independently related to higher order cognition, and complex span tasks can reflect all of these mechanisms. Shipstead, Harrison, and Engle (2015) highlighted the strong relationship between attention control and WMC, since engaging attention in a controlled manner was found to be a critical aspect of WMC. In short, it could be argued that executive attention or WMC is needed in situations requiring active maintenance of task goals under the conditions of distraction or the retrieval of the goal-related information in the face of response competition, which is consistent with the explanations for the unitary nature of Common EF by Gustavson et al. (2015) and Friedman and Miyake (2017) in section 2.1.2.

With respect to these three mechanisms of WMC, the work in cognitive neuroscience has demonstrated that prefrontal cortex (PFC) plays a role in maintaining information in PM (Miller & Cohen, 2001) and in encoding and retrieval in SM (Simons & Spiers, 2003). Additionally, the anterior cingulate cortex (ACC) is found to be activated in monitoring the amount of AC (Botvinick, Braver, Barch, Carter, & Cohen, 2001). In the process of performing a complex span task, stronger activations were observed in the PFC and ACC (Kondo et al., 2004a) and stronger connectivity between the PFC and ACC for participants with higher WM span (Kondo, Osaka, & Osaka, 2004b).

2.1.4 Inhibition

In the last 30 years, inhibition, one of the core components of EFs, has become a central player in multiple research domains within psychology ranging from behavioral disorders to development of cognitive abilities. To define the construct, researchers have adopted different approaches, which has resulted in multiple definitions of inhibition, thereby leading to several conceptual distinctions in the extant literature. Despite having different interpretations, the majority of the scholars in the field speculate that inhibition is a multidimensional construct (e.g., Dempster, 1993; Harnishfeger, 1995; Nigg, 2000).

Dempster (1993) adopts the family of inhibition processes approach, conceptualizing the term as a general process that operates in various domains. He classifies inhibitory processes according to their psychological forms: motor, perceptual, or verbal; and argues that the classification can be further extended by including the location and temporal operating characteristics of interfering stimuli.

Harnishfeger (1995) classified inhibitory processes according to three dimensions: (1) intentional or unintentional, (2) behavioral or cognitive, (3) inhibition or resistance to interference. Unintentional inhibition refers to suppression prior to conscious awareness while intentional inhibition to suppressing irrelevant stimuli consciously. Behavioral inhibition is related to the control of overt behavior while cognitive inhibition involves the control of mental processes germane to suppressing unwanted or irrelevant thoughts and inappropriate meanings of ambiguous language, and gating irrelevant information from WM. Finally, inhibition as an active suppression process concerns, for example, the removal of irrelevant information from WM while resistance to interference is a gating mechanism that prevents irrelevant information or distractors from entering WM and is susceptible to performance decrements under the conditions of distracting stimuli or selective attention.

On the basis of Harnishfeger's (1995) suggestions, Nigg (2000) proposes four types of inhibition: (1) 'interference control' concerning the suppression of interference caused by resource or stimulus competition, (2) 'cognitive inhibition' dealing with suppression of irrelevant information from WM, (3) 'behavioral inhibition' involving suppression of prepotent responses, and (4) 'oculomotor inhibition' regarding suppression of reflexive saccades.

Working on Nigg's (2000) conceptualization, Friedman and Miyake (2004) proposed a trichotomy of inhibition functions: (1) 'prepotent-response inhibition' (PRI) as a combination of Nigg's behavioral and oculomotor inhibition, (2) 'resistance to distractor interference' (RDI) corresponding to Nigg's interference control, and (3) 'resistance to proactive interference' (RPI) which is similar to Nigg's cognitive inhibition. They define PRI as the ability to deliberately override dominant, automatic, or prepotent responses and measure it with the antisaccade task, the stop-signal task, and the Stroop task. They consider RDI as the ability to deal with interference from task-irrelevant information in the external environment and assess it with flanker, work meaning, and shape matching tasks. With respect to RPI, they define it as the ability to deal with the intrusions resulting from the previously relevant but currently obsolete information for the task at hand. They tested the distinctions among these functions, employing CFA. They first came up with a three-factor model, yet then with a more parsimonious model in which PRI and RDI fall along a single factor, and RPI on a different one. However, in their unity and diversity model of EFs, Miyake and colleagues operationalize the inhibition component of EFs only on the basis of PRI and do not include RDI in their definition.

Diamond (2013) proposes a dichotomy of inhibition functions: (1) self-control or behavioral inhibition, and (2) interference control. She further extends the interference control, dividing it into two sub-categories: selective attention and cognitive inhibition. Consistent with Friedman and Miyake's (2004) finding, she considers both PRI and RDI together in the same category, namely selective attention; and she relates cognitive inhibition with the suppression of prepotent mental representations.

As mentioned in the previous subsection, Engle and colleagues use the term 'attention control' rather than inhibition in their framework, and define it in a way that it combines PRI and RDI together. Therefore, in their studies, they employ the tasks measuring PRI (e.g., antisaccade) and the ones assessing RDI (e.g., flanker) as the measures of the same latent variable. Attention control is thought to be the ability to resolve response competition or conflict caused particularly by prepotent or habitual behaviors (Engle & Kane, 2004). They argue that attention control plays a critical role in dealing with the effects of proactive and retroactive interferences and avoiding the effects of distraction in the process of accomplishing goals.

Based on the multidimensional perspectives of inhibition as discussed above, it is apparently certain that the forms of inhibition are separate, yet there is research indicating that there are commonalities among the functions sharing substantially similar neural bases. Niendam et al.'s (2012) meta-analysis of 193 functional neuroimaging studies revealed that tasks (e.g., antisaccade, stroop, flanker, Go-No-Go) requiring inhibition elicited activation in frontal and parietal regions, including DLPC (BAs 9, 46), ACC (BA 32), and superior (BA 7) and inferior parietal lobe (BA

40). Such tasks are also associated with activation in other prefrontal (BAs 6, 10), occipital (BA 19), and temporal (BA 13) regions. Activation was also observed in such subcortical regions as caudate, thalamus, and putamen.

Furthermore, working on a unitary framework of supervisory attentional system (SAS) by Norman and Shallice (1986), Cieslik, Mueller, Eickhoff, Langner, and Eickhoff (2015) first classified inhibition into four categories and then investigated the core neural correlates of inhibition-related tasks requiring supervisory control needed for goal-directed, non-routine behavior. Their conjunction analysis revealed that a right-dominant midcingulo-insular-inferior frontal core network (anterior insula [aI] and inferior frontal junction [IFJ], dorsal anterior cingulate cortex [DACC] and pre-supplementary motor area [preSMA]) play a central role for the regulatory processes mediated by the SAS. The schema activation monitoring processes are supported by the right aI and the right IFJ. While the former is associated with monitoring and general implementation of general task set, the latter with the continuous reactivation of the relevant stimulus-response rule. The DACC/preSMA are related to mediating energization of the adequate (general) task schema. This finding opens the door to the possibility of a unitary perspective for inhibition, while it implies that other pre-frontal and parietal regions may be less domain-general in contrast to previous findings. They argue that left lateral prefrontal cortex (PFC) plays a more important role in the early stages of learning.

Consistent with the scholars mentioned above (Diamond, 2013; Engle & Kane, 2004; Friedman & Miyake, 2004), Cieslik et al. (2015) consider PRI and RDI tasks in the same category in their classification of inhibitory functions. In addition to the common activated SAS regions (the right aI and IFJ; and DACC/preSMA), increased activity in a dorsal attention network (the parietal lobe [around the intraparietal sulcus (IPS) and superior parietal lobe (SPL)] and the dorsal frontal cortex along the precentral sulcus) across the interference tasks was observed, which indicates commonalities among PRI and RDI tasks.

On the other hand, there are some other neurological studies indicating that PRI and RDI are slightly different (Groom & Cragg, 2015; Xie, Ren, Cao, & Li, 2017) in the sense that while the amplitude of N2 is enhanced by RDI, the P2 amplitude by PRI. However, the task employed in Groom and Cragg's (2015) study, namely go/no-go task, was not associated with the DACC/preSMA in Cieslik et al. (2015), while the other inhibition-related tasks were.

Another area of contention about inhibition is whether inhibition-related tasks capture the construct adequately or not, stemming from the low correlations obtained among inhibitory tasks across studies. Friedman and Miyake (2004) propose three factors that lead to the problem of low-correlation: (1) the use of tasks with no justification why they measure inhibition, (2) the low reliability due to the practice effect or the calculation of dependent variable, and (3) task-specific variance masking commonalities attributable to inhibition, namely task-impurity problem. In their review, Draheim, Mashburn, Martin, and Engle (2019) argue that the problems associated with measurement are mainly due to a reliance on reaction time (RT) scores, especially RT differences. One of the concerns about the use of RT is that it is sensitive to speed-accuracy interactions (Heitz, 2014). Another is the impurity of RT. More specifically, cognitive processes underlying RT is so complicated that it could result in faulty interpretations of RT-based scores (Miller & Ulrich, 2013).

The issue of low-reliability among different inhibition tasks due to their reliance on RT scores might weaken the unified concept of inhibition (see Rey-Mermet, Gade, & Oberauer, 2018 for a review). Paap, Anders-Jefferson, Zimiga, Mason, and Mikulinsky (2020a) argue that the Flanker, Simon, and Stroop tasks could not capture the construct adequately. Draheim et al. (2019) note Friedman and Miyake's (2004) suggestion that new tasks be needed to make new theoretical advancements. Therefore, Draheim, Tsukahara, Martin, and Engle (in review) administered a range of established, modified and new attention control tasks and found that new tasks such as threshold versions of the Stroop and flanker can improve the measurement of AC over RT-based ones, providing evidence for the unitary concept.

In this current study, the framework of executive attention or WMC by Engle and colleagues (Engle, 2002; Engle & Kane, 2004) is adopted since it is consistent with the explanations of Miyake and colleagues for Common EF (Friedman & Miyake, 2017). Thus, it is based on the individual differences in relation to the domain-general ability to control attention (Engle, 2002, 2018).

2.2 Enhancement of EFs

Since EFs play a significant role in a broad range of domains and fields, the development and enhancement of EFs has piqued the interest of the scientific community. Relatively more attention has been placed in the periods of early and mid-childhood; therefore, much is known about the development of EFs in these periods. Accordingly, it could be put forward that EFs undergo a protracted and gradual process, and some of its most profound changes are observed in early periods (Chevalier & Clark, 2018; Cuevas & Bell, 2014; Cuevas et al., 2018). The improvement continues into adolescence which is the optimal period for the development of EFs (Crone, Peters, & Steinbeis, 2018). As for the components of

EFs, inhibition and shifting stabilize during the period of early adolescence while WM can improve up until the age of 21 (Huizinga, Dolan, & van der Molen, 2006).

With respect to the factors contributing to the enhancement of EFs during childhood, Cuevas et al. (2018) highlight biological (e.g., brain development, genetics, temperament) and environmental factors (e.g., socioeconomic status [SES], caregiving, and education). Crone et al. (2018) highlight the effects of hormones on the improvement during the early adolescence or puberty, contributing to the maturity of brain connectivity patterns.

Although it is certain that the building blocks of EFs are acquired during the early childhood, and the change in EFs continues into adolescence, the individual differences in EFs are apparently stable according to the twin studies (e.g., Friedman et al., 2008, 2011) during these periods. However, the criticism about the twin studies should be noted. The adoption of equal environment assumption, homogeneity of the environmental factors, possible additive genetic effects, and the inadequate representativeness could cause an upward bias in heritability. Although the findings indicate that EFs are the most heritable psychological trait, high heritability does not necessitate that EFs are immutable since environmental factors can affect EFs (Friedman et al., 2008). In another twin study, Friedman et al. (2016) focus on the transition from adolescence to adulthood (a period of ages 17 to 23, in other words, emerging adulthood [Arnett, 2000]) since it is the period when performance matures thanks to the development in the neural circuitry in a context of experiencing a wide range of new environments and social roles. They found marked stability in Common EF, shifting and updating primarily due to genetic factors. On the other hand, a significant, though small, change was also noted only in Common EF due to environmental influences, which is consistent with the findings of Lyons et al. (2009) and Tucker-Drob and Briley (2014).

As Arnett (2000) maintains, during emerging adulthood, individuals experience considerable life changes in such different contexts as residence, education, and new social roles. However, Friedman et al. (2016) acknowledge that it is unknown what specific environmental factors lead to changes in Common EF during this transition period. One of the experiences that individuals go through and influence their cognition is the lifetime experience with L2, which is the topic of discussion in the next section.

2.2.1 Life-long L2 experience-induced effects on EFs

Weinreich (1953) introduced the idea that the presence or activation of more than one language results in interference, and bilingual individuals need to resolve the competition between these languages. Based on the Levelt's model (Levelt, Roelofs, & Meyer, 1999) and the SAS (Burgess & Shallice, 1996; Shallice & Burgess, 1996), Green (1998) proposes, in his inhibitory control (IC) model, that the ability or the goal to use two languages appropriately requires the regulation of SAS (through the guidance of top-down cues) mediating the process of contention scheduling on which language task schemas compete with each other to control output by altering their level of activation. Performing a language-specific task (e.g., translating from one language to another) requires the activation of goal-relevant, language specific schemas and the inhibition of the potential competitors for production as the irrelevant schemas in order to select between alternative responses. That is, for a bilingual speaker, the SAS regulates the process of activating the target-language and suppressing non-target one.

In the literature of bilingual advantage, Bialystok et al. (2009) suggest that experience with an L2 is considered to have a unique feature since the joint activation of two languages in a single mind requires the intensive and frequent use of EF-related cognitive processes for language management by maintaining attention to the target language and avoiding interference from the competing language. Since EFs and language functions rely on the same integrated brain network, L2 experience in the verbal domain is expected to lead to broad cross-domain effects (Li et al., 2014b). Additionally, in consistent with the adaptive control hypothesis (Green & Abutalebi, 2013), Yang, Ye, Wang, Zhou, and Wu (2018) indicate that language control processes are modulated by contextual interaction through adaptive changes in the regions and circuits that are related to specific control processes.

The view that lifetime experiences with an L2, acquired at early ages, enable bilingual children and older adults to have better EFs as compared to monolinguals is widely held (Bilaystok et al., 2012). However, the findings with young adults have resulted in growing skepticism concerning the claims for L2 experience-induced positive effects on EFs. In relation to L2 experience, the two major and most commonly studied components of EFs are inhibition and WM.

2.2.1.1 Effects of life-long L2 experience on inhibition

The most commonly studied component of EFs in relation to life-long L2 experience is inhibition which, despite conceptual differences (see section 2.1.4), generally concerns the deliberate inhibition of dominant, automatic or prepotent responses when needed (Miyake et al., 2000). Over the last decade, a wide range of behavioral and brain studies have been conducted to investigate the effects of L2 experience on inhibitory control abilities of young adults with life-long L2 experience. The behavioral data have revealed mixed findings. Positive effects of bilingualism on bilingual young adults' ability to use inhibition processes have been reported with measures of a flanker task (e.g., Yang & Yang, 2016), a Stroop task (e.g., Blumenfeld & Marrian, 2011), an antisaccade task (Bialystok et al., 2006), and an ambiguous figures task (Chung-Fat-Yim et al., 2017). However, null effects have also been observed. For instance, administering a Simon task to different groups including children, young adults, middle-aged adults, and older adults, Bialystok et al. (2005) found that young adults had equivalent performance with their monolingual counterparts. Likewise, Paap and Greenberg (2013) had monolingual and bilingual young adults complete four-standard EF tasks (antisacccade, Simon, flanker, and color-shape switching), and found nonsignificant language group contrasts. Similarly, Kousaie et al. (2014) indicated that young and older bilinguals and monolinguals had similar performances on a series of EF tasks including Stroop and Simon tasks. Additionally, Kalia et al. (2014) compared the performances of monolinguals, early bilinguals, and late bilinguals on an auditorily cued number numeral task, and revealed no clear advantages for one group. Late bilinguals, however, were less accurate as compared to the other two groups. Finally, controlling for the relevant demographic factors (i.e., age, SES, IQ, education, status of immigration), Antón et al. (2019) compared young bilinguals with their monolingual counterparts and found null-effects of bilingualism on their performances on the flanker, Simon, Stroop and numerical Stroop tasks.

The contradictory findings of the behavioral studies are obvious in the review papers as well. Van den Noort et al. (2019) found that the majority of the reviewed papers indicated bilingual advantage, and the studies with adult participants indicate bilingual advantage more as compared to the ones with children. Yet, van den Noort et al. (2019) noted the rarity of longitudinal studies. On the other hand, Paap's (2019) review revealed that studies indicating bilingual advantage are in a minority and he argues that the figures would be even smaller if the correction was made for publication bias and file-drawer problem. Additionally, Paap, Mason, Zimiga, Silva, and Frost (2020b) contend that van den Noort et al. (2019) could have had a tendency to give more weight to evidence confirming their hypothesis because their selection of papers was rather limited. They also questioned the issue of readily accepting the results of meta-analyses at face value.

As for the results of the brain studies exploring the relation between long-term L2 experience and inhibition, findings are relatively more conclusive compared to the behavioral ones. Mechelli et al. (2004), for instance, found that gray-matter density in the left inferior parietal cortex was greater in early or late bilinguals than monolinguals, and language proficiency was positively correlated with the density. However, behavioral measures did not show any significant group differences in the same study. In a combined functional and structural neuroimaging study by Abutalebi et al. (2012), it was revealed that dorsal anterior cingulate cortex (ACC) tightly bound to domain-general executive control functions was associated with language control and resolving non-verbal conflict. Young adults with life-long L2 experience were found to use this structure of the brain more efficiently in resolving cognitive conflicts and to have higher volume of gray-matter in this area. Likewise, Yang et al. (2018) found no flanker effect with bilinguals, indicating that dual-language context with two highly proficient languages enhances the neural efficiency of the inhibitory control network.

2.2.1.2 Effects of life-long L2 experience on WM

Since the components of EFs are intercorrelated with each other and life-long L2 experience has generally been found to have an effect on inhibitory control, WM has also been expected to be positively affected (Yang, 2017). Yet, a heavier cognitive load imposed on WM of L2 learners due to the dual language processing might have negative effects as well. Therefore, two possibilities arise. First, the high load on the WM might be a disadvantage (Tokowicz, Michael, & Kroll, 2004). Second, management of two simultaneously activated languages might lead to the development of an efficient mechanism, which would benefit the WM system (Bialystok, Craik, Klein, & Viswanathan, 2004; Miyake & Friedman, 2012). Calvo, Ibáñez, and Garcia (2016) argue that one possibility of more proficient bilinguals' better performance in conflict resolution and WM is that increased language processing demands improve the attentional skills that are essential in visual and verbal processing. Additionally, Huang et al. (2020) propose that learning an additional language requires not only processing a huge amount of new information, taxing WM but also manipulating the existing linguistic information and integrating it with the knowledge of the new language. Learners need to inhibit goal-irrelevant linguistic representations depending on the context, which depends on the function of WM.

A number of researchers have investigated the effect of life-long L2 experience on WM of young adults, by employing a variety of tools: simple verbal digit span tasks (Antón et al., 2019; Luo et al., 2013; Morrison et al., 2019; Ratiu & Azuma, 2015; Smithson & Nicoladis, 2013), simple nonverbal digit span tasks (Antón et al., 2019; Luo et al., 2013; Smithson & Nicoladis, 2013), complex verbal WM span tasks (Ratiu & Azuma, 2015; Smithson & Nicoladis, 2013), complex nonverbal WM span tasks (Ratiu & Azuma, 2015), or n-back (Morrison et al., 2019). Star-counting as a verbal WM measure and a modified flanker and probe tasks as nonverbal WM measures were employed as well (Sullivan et al., 2016). The modified flanker task was administered to increase the storage demand of WM while the conditional Go/No-go task to increase the processing demand of WM (Jiao et al., 2019).

These studies revealed contradictory findings as in the case of inhibitory control abilities. Bilinguals were found in some studies to outperform monolinguals on non-verbal tasks whereas they were at disadvantage on verbal tasks (Luo et al., 2013; Sullivan et al., 2016). On the other hand, Smithson and Nicoladis (2013) found no group differences on either verbal or non-verbal tasks. In addition, Ratiu and Azuma (2015) indicated that monolinguals had a better performance on a complex verbal task, yet there was no difference on simple verbal or complex visual tasks. The lower performance of bilinguals was explained by bilingual disadvantage in verbal tasks due to the dual language activation and processing cost in conflict resolution or weaker connections among lexical, semantic, and/or phonological/orthographic associations. On the other hand, Jiao et al. (2019) revealed cognitive benefits of bilingualism on more demanding WM tasks but not on simple tasks. Likewise, controlling the demographic factors effectively, Antón et al. (2019) found null-effects of bilingualism on the forward versions of the Corsi and digit span tasks while a potential bilingual advantage on their backward versions. They highlighted the issue of domain-specificity and -generality in line with the neurological findings (e.g., Li, Christ, & Cowan, 2014a), and argued that domain-specific networks play a role only during encoding while domain-general ones during encoding, maintenance and retrieval. For the bilingual advantage to emerge, the more demanding situations or tasks involving more complex processing and retrieval, in other words, requiring domain-general WM system, would be required. In the same vein, Morrison et al. (2019) found null-effects of bilingualism on both forward and backward digit span tasks. However, bilinguals exhibited larger P300 amplitude during their n-back performances, reflecting that bilinguals have more cognitive resources to complete the tasks with less effort and more efficiency, implicating that more challenging tasks would elicit group differences. Finally, highlighting the interaction of different dimensions of bilingual experience, Bonfieni et al. (2020) revealed that the positive cognitive effects of age of acquisition and proficiency were found only when these two dimensions were examined together. The exclusion of other dimensions of the individual variability (e.g., age and level of education) in their analyses still indicated the positive effect, yet the methodological concerns of this exclusion were noted as well.

The confounding findings were also clear in meta-analyses. Grundy and Timmer's (2017) meta-analysis of 27 independent studies with participants from different age groups (children, young adults, and older adults) revealed that L2 experience enhanced WM capacity. On the other hand, von Bastian, De Simoni, Kane, Carruth, and Miyake (2017) and Lehtonen et al. (2018) found no bilingual advantage on WM domain. Von Bastian et al.'s review indicated that neither age nor task mode was a moderator in the effect sizes. It should be noted that Lehtonen et al. claim that they reached the null effects after implementing bias correction strategies.

Mixed findings with the behavioral data of young adult bilinguals in their inhibitory control skills and WM can be attributed to a number of factors. First, interindividual variability might be insufficient for group differences in performance to emerge (Chung-Fat-Yim et al., 2017) possibly due to the ceiling levels in tasks measuring inhibition of young bilingual adults (Bialystok, 2017). Second, bilingualism is treated as a categorical variable in a majority of the studies. Yet, few people can be labeled as truly monolingual with no exposure to another language and few bilinguals can be found to have equivalent language experiences (Bialystok, 2016). Alternatively, researchers can use bilingual experience as a continuous variable, taking the L2 proficiency of participants into consideration (Bialystok, 2017). Third, the nature of the dependent variable (DV) based on response time (RT) cost score (latency switch costs) can be problematic from methodological and psychometric perspectives. To circumvent the problem, accuracy-based measurement tools and/or threshold versions of the existing ones could be employed (Draheim et al., 2019). Fourth, individual differences rather than L2 learning, (e.g., SES [Calvo & Bialystok, 2014]), can play confounding roles (van den Noort et al., 2019). Last, as Paap (2019) argues, the inconsistencies in the findings could be related to the conceptualization of EFs and the operational definition in the empirical studies. Bialystok (2017) criticizes the conceptualization of EFs generally based on the componential unity and diversity model of Miyake et al. (2000) that does not accurately capture the differences in cognitive processes of monolinguals and bilinguals. She offers Engle's executive attention model (Engle, 2002; Engle & Kane, 2004) as one alternative domain-general system that has two main differences: (1) WM capacity is a continuous construct, thus can show experience-dependent plasticity, and (2) WM capacity is rooted in the use of attention that is missing in the unity/diversity model. Yet, it should be noted that, as mentioned in section 2.1.3, the framework of executive attention of Engle and colleagues is based on the similar mechanisms proposed by Friedman and Miyake (2017) to account for the nature of unity part of their model (i.e., Common EF).

2.2.2 Short-term L2 experience- or intensive L2 training-induced changes on EFs

Challenging the critical period hypothesis (Lenneberg, 1967), recent scientific evidence has some indications for short-term L2 learning- or intensive L2 training-induced changes in brain areas including the ones bound to both EFs and language learning, for the individuals whose L2 learning experience begins as of/after puberty. Abutalebi and Green (2007) argue that even if it occurs after puberty, the representation and processing of L2 converges with that of L1 with growing proficiency. They contend that L2 learning will have functional and structural consequences on the brain due to various sources of difficulty at low levels of proficiency. Given vocabulary learning for example, the sources could be as follows: weaker neural connections between the concept, lemna and word form; interference from a prepotent concept name; and higher demand on control processes for resolving competition. Their review of relevant papers reveals that if not mastered to a higher degree of proficiency, L2 engages more extended portions of the left prefrontal cortex. Controlled processing for an L2 would entail cognitive control processes to resolve the competition and conflict between languages; therefore, activations were found outside the language-related areas in the brain, yet in areas related to cognitive control (e.g., left prefrontal cortex, and ACC) (Abutalebi, 2008; Abutalebi & Green, 2007).

At a comprehension level, particularly reading comprehension, Gurunandan et al. (2019) examining the functional differences between intermediate and advanced level L2 learners found that proficiency plays a significant role in modulating the similarity between L1 and L2 and the connectivity between language comprehension and language control regions (DLPFC and ACC). Although the similarity between L1 and L2 activation was found to be higher in the intermediate

L2 learners, functional connectivity with control areas was higher in the advanced L2 learners. These are indicative of the fact that L2 learning after puberty can display functional plasticity of language comprehension networks as well.

With respect to the changes after L2 training, Schlegel et al. (2012) analyzed the tensor imaging scans of 11 English undergraduate students taking a nine-month intensive Chinese language course, and found progressive changes in white matter tracts in relation to left hemisphere language areas together with changes in frontal lobe tracts crossing the genu of the corpus callosum. Additionally, Mårtensson et al. (2012) conducted a study with 14 university aged students enrolled at the Swedish Armed Forces Intelligence and Security Centre in order to be an interpreter. The participants studied a completely foreign language (Arabic, Dari, or Russian) from scratch. The findings revealed that taking a three-month intensive language training led to the changes in the structure of language-related brain regions, specifically increased cortical thickness in fronto-temporal cortex of the left hemisphere and gray matter volume in the left middle frontal gyrus. Moreover, in Hosoda et al.'s (2013) study, participants learned 240 vocabulary items in 16 weeks of time. The findings revealed positive plastic changes. More specifically, training intervention increased volume of gray matter in the inferior frontal gyrus (IFG). The researchers speculated that reinforcement learning could contribute to enhancing executive control of IFG over the mechanism employed to acquire L2 vocabulary. Last, Qi et al. (2019) investigated the changes in the brain induced by a Mandarin Chinese learning experience as L2 of 24 university students (mean age = 23.2) at Massachusetts Institute of Technology. The participants took 60 class hours of instruction at an introductory level in one month. Qie et al. found increased activation in the left IFG and left superior parietal lobe to the Mandarin speech in L2 at the end of a four-week period. They concluded that right IFG engagement is needed for adults to achieve L2 success during initial learning, yet right IFG disengagement for long-term retention of L2 skills.

With respect to the neurological changes in relation to inhibition, Sullivan et al. (2014) explored the effect of early stage L2 training on the executive control of participants who took introductory Spanish courses. It was found that experimental group showed electrophysiological changes in the post-test with larger P3 amplitude on the go-no go task, indicating the strengthening of the neural network involved in response inhibition. Yet, no behavioral differences were observed between the groups. The contradiction is plausible given the fact that brain changes can emerge after three months of L2 experience whereas qualitative changes in behavior require more intensive and quality experience (Driemeyer, et al., 2008).

Xie (2018) investigated the role of L2 proficiency in performance on a flanker task for English major undergraduate students in China with varying levels of L2 background (three proficiency levels) who received intensive L2 training for an academic year (16 hours per week and two semesters). He found that L2 proficiency positively influences the conflict monitoring (overall RTs in all three conditions: congruent, incongruent, and neutral) but has no effect on the flanker effect (RT differences between congruent and incongruent trials). Yet, the details about the L2 courses that participants took were not provided in the article.

As for WM, Kerrigan et al. (2017) conducted a study by administering both visuo-spatial (the change blindness task, and the forward and backward Corsi blocks tasks) and verbal (the word and alpha span tasks) measures. The participants were 30 undergraduate students with little L2 exposure (control group) and 30 other undergraduate ones who had started learning an L2 as of their puberty (experimental group) but reported a balanced use of both their L1 and L2 on a daily basis. They had different L1 backgrounds and ten of them reported experience in a third language (L3). The results revealed that the ones with more experience in L2 were faster and more accurate on visual tasks as compared to the ones with little L2 experience, yet no group differences were found on the verbal tasks. However, it should be noted that the verbal tasks were administered in L2 for the experimental group, in L1 for the control one.

In order to examine the role of L2 proficiency level in relation to EFs, Yang (2017) administered verbal and visual digit span tasks (both forward and backward) to one group of near-monolinguals and two groups (mid- and high-proficiency) of Korean learners of English whose length of residence in the USA varied. It was found that participants with mid-proficiency in L2 outperformed the others on both verbal (forward and backward) and visual (only forward) digit span tasks, yet no other group differences were observed. The advantage for the mid-proficiency L2 learners was explained through the intensive use of memorizing, replaying, and monitoring strategies during the early stages of L2 learning to deal with the extra load due to their comparatively lower L2 proficiency. The researchers argued that dual-language practices were natural for the proficient participants, reducing their dependence on the use of these strategies. Low sample size and a cross-sectional design rather than a longitudinal one have been noted as the limitations of the study.

Dong et al. (2018) compared the performances of two groups of Chinese university students taking either consecutive interpreting (CI) or English for general purposes course on n-back, L2 listening span, and letter running tasks before and after they took 32 hours of instruction in L2 for one academic semester. Both groups improved their performances on n-back accuracy and listening span tasks. Yet, a possible role of practice effect was noted. On the other hand, there was no improvement for either group on running letter task. Only CI group improved their updating efficiency (RTs on n-back) after the training, reflecting that updating efficiency and recalling process in the CI task share the same processes of attentional control.

Huang et al. (2020) investigated the impact of L2 and L3 learning on the enhancement of WM with first and second year Chinese undergraduate students learning English and/or a third-language (either Japanese or Russian). They performed an operation span task before and after taking a nine-month intensive language education (12 to 16 hours per week for L2; 4 to 10 hours per week for L3). Huang et al. found enhancement on the WM capacity of all groups yet more on the first-year students'. As for L3, the effect was only seen in the first-year students, indicating that WM is more actively engaged, thereby trained, at early stages of learning a new language, and learning a new language trains WM more than improving the existing one. However, such limitations were noted as having a limited sample size and the absence of control group with no experience of intensive L2 instruction. They ruled out the possibility of a practice effect by highlighting the varying degrees of effect sizes.

In short, EFs are among the most heritable psychological traits, yet they are not immutable thanks to environmental factors, one of which is L2 experience. Although bilingual advantage is clear in children with lifetime L2 experience, the findings germane to young adults are inconclusive, due to methodological limitations of existing studies. With respect to L2 experience after puberty, it is apparently possible for the changes in the brain and behavioral performances to emerge. In other words, lifetime or short-term experience in performing multilingual language tasks is assumed to offer benefits associated with domain-general executive control abilities, which may suggest that L2 experience leads to enhanced EFs.

2.3 Individual differences in EFs and L2 learning outcomes

Considering the literature discussed so far, one could argue that L2 experience contributes to the enhancement of EFs. However, another significant issue discussed in the literature is the role of EF-related individual differences in L2 success, which could be construed as claiming that EFs contribute to success in L2 learning. Here the assumption is that EFs are the causal mechanisms underlying performance in a variety of domains including L2 processing and proficiency outcomes (Linck et al., 2014). A number of behavioral and brain studies have been conducted to investigate the predictive validity of EFs for L2 learning outcomes. However, there is considerably less research investigating the role of inhibition in L2 learning as compared to that of WM, and very little research with a longitudinal research design.

2.3.1 Individual differences in inhibition

Few studies have been conducted to explore the predictive value of inhibition for L2 learning outcomes. The findings can be discussed in two main categories: online language processes and L2 success. While the studies investigating online language processing revealed a relationship, the behavioral ones did not.

With respect to online language processing, first, investigating the relationship between inhibitory control abilities of high-proficient L2 learners (32 Spanish-English, and 26 Japanese-English bilinguals) and the resolution of cross-language activation, Linck et al. (2008) revealed that bilinguals with higher inhibitory control abilities were better at suppressing the activation of lexical candidates during L2 picture naming. Second, Linck et al. (2012) examined the inhibitory control abilities of 56 native English speakers with high-proficiency level in French (L2) and mid-proficiency level in Spanish (L3) through a multilingual language switching task. They found greater reliance on inhibitory control when there was a great amount of cross-language competition; and thanks to more efficient inhibitory control system, less amount of inhibition was needed to support non-dominant language naming. It should be noted that only one task, the Simon task, was employed to measure the inhibition in both of these studies.

As for the relationship between inhibition and L2 success, Linck and Weiss (2011, 2015) conducted two studies with the same design but with different participants. These studies can be thought to be the first to explore the relationship by using a longitudinal (test-retest) design. In both studies, the Simon task was administered to measure inhibition at the end of the semester, and a proficiency test at the beginning and end of the semester. In the first study, the participants were 24 English speaking university students enrolled in a semester-long either introductory Spanish or German course, whereas in the second study, the participants were 25 English learners of Spanish taking introductory Spanish courses at university. In neither of the studies, inhibition emerged as a strong predictor. Yet, low sample size was reported as a limitation, and a question remained unclear if inhibition could emerge as a robust predictor at later stages of L2 learning or in more intensive learning contexts.

Ghaffarvand Mokari and Werner (2019) investigating the role of inhibition in L2 ability of 30 adult pre-intermediate learners of English to identify and discriminate English vowels after a 5 hour-phonetic training. They employed the Stroop and retrieval-induced forgetting task (RIF task measuring resistance to proactive interference; in other words, cognitive inhibition) as the measures of inhibition. They found a significant relationship between L2 success and RIF, yet no relationship between L2 learning and performance on the Stroop. They highlighted the role of suppression of irrelevant memory items in learning L2 phonology.

2.3.2 Individual differences in WM

WM is one of the central constructs in many L2 theories in proficiency development and processing. Over the last two decades, numerous studies have been conducted to explore the predictive validity of WM in relation to L2 learning outcomes. Linck et al.'s (2014) meta-analysis of 79 studies revealed that WM is positively associated with L2 success. Larger effect sizes were observed with executive control component of WM (complex span tasks) as compared to the storage one (simple span tasks) and with verbal measures as compared to non-verbal ones. However, the majority of these studies are cross-sectional, comparing the WM of learners at different levels of L2 proficiency. Although Linck et al. highlight that they considered publication bias and file-drawer effect into consideration and made analyses accordingly, Wen and Li (2019) argue that the review did not include some relevant studies.

Regarding the role of WM in vocabulary and grammar learning, Kempe, Brooks, and Kharkhurin (2010) and Martin and Ellis (2012) found a predictive relationship when learners had no L2 background. However, Engel de Abreu and Gathercole (2012), and Jean and Geva (2009) showed that the prediction was weak and unstable in learning L2 vocabulary. Martin and Ellis (2012) argue that WM is needed when learners are involved in the tasks requiring heavy processing load and active participation, which was supported not only by Suzuki and DeKeyser (2017) comparing mass practice with spaced or distributed practice, but also by Sanz, Lin, Lado, Stafford, and Bowden (2016) and Li, Ellis, and Zhu (2019) exploring the role of grammar instruction given prior to task performance. Last, while Goo (2012) and Li (2013) found the role of WM in learning explicit grammar knowledge, Kim, Payant, and Pearson (2015) and Li et al. (2019) in implicit knowledge.

In one of rare fMRI studies, Yang and Li (2012) investigated the neural correlates of grammar learning, by training 43 university students to learn artificial grammar sequences in two conditions: explicit or implicit. Both behavioral and neurological data showed that performance on a verbal WM task was correlated with success in sequence learning. The deployment of WM varied across two learning conditions. More specifically, the left dorsolateral prefrontal cortex was seen to be an important neural marker to be successful in explicit learning condition.

With regard to reading, WM's predictive power is well established both for L1 (Daneman, & Merikle, 1996) and L2 reading. Complex WM span tasks are found to be more indicative of L2 reading outcomes as compared to simple span tasks (Harrington & Sawyer, 1992), and in more challenging reading tasks: inferential reading rather than literal reading (Alptekin & Erçetin, 2009), and story comprehension rather than dialogue comprehension (Andersson, 2010). Prior L2 knowledge plays a mediating role in WM (Alptekin & Erçetin, 2010), and the predictive power is observed at initial stages rather than advanced stages of learning (Walter, 2004).

Expanding the framework of maintenance and disengagement (proposed by Shipstead, Harrison, & Engle, 2016), Martin et al. (2019) explored the role of WM and fluid intelligence in L1 reading comprehension and L2 vocabulary learning. They revealed that individual differences in learning are related to maintaining active information and disengaging from no longer relevant information. Yet, they noted that measures tapping disengagement (i.e., Gf and updating WM tasks) increase the prediction of performance beyond complex span tasks alone, which is consistent with Christopher et al. (2012) and Was, Rawson, Bailey, and Dunlosky (2011) highlighting the role of LTM in language comprehension.

As to the role of WM in L2 speaking, Fehringer and Fry (2007) indicated its role in fluency, Ahmadian (2012) in speaking through within-task planning rather than speaking through strategic planning prior to task performance, and Kormos and Trebits (2011) in storytelling on the basis of pre-sequenced pictures rather than in unstructured task.

Albeit quite limited, there is research exploring the role of WM in L2 listening and writing. Miki (2012) found that WM is indicative of both literal and inferential comprehension in L2 listening. However, the WM task was listening span task in L2 with a potential of being a confounding variable (Linck et al., 2014). In writing, while L2 writing is associated with verbal WM (Abu-Rabia, 2003; Olive, Kellogg, & Piolat, 2008), yet it is not related to spatial WM (Olive et al., 2008).

So far, the cross-sectional studies reviewed above have primarily demonstrated that WM capacity is indicative of L2 success. Yet, apparently, the nature of both learning and WM tasks can play a role in contradictory results. Few studies with a longitudinal research design have been conducted.

Kormos and Sáfár (2008) conducted the first longitudinal study to explore the predictive validity of WM for L2 learning outcomes, collecting data in two consecutive years to have sufficient number of participants. The participants were 121 Hungarian learners of English enrolled in a one-year long intensive L2 program at a secondary school (100 low proficiency and 21 high proficiency learners). In the first year, a non-word span task was employed to measure short-term memory (STM) at the end of the program, and a backward digit span to assess WM in the second year. However, high-proficiency learners did not perform the backward digit span. As a post-test, one proficiency test was administered in both years. The findings indicated that while the STM measure was significantly correlated with the L2 proficiency of high proficiency learners, it played no role in L2 success of low proficiency ones. On the other hand, WM was found to be highly correlated with overall English language competence of low proficiency learners.

Apart from exploring the role of inhibitory control in L2 learning outcomes as mentioned in the previous subsection, Linck and Weiss (2011, 2015) also investigated the predictive validity of WM capacity with low proficiency learners in a classroom context. The participants completed an operation span task and L2 proficiency test at the end of a semester-long L2 instruction. The results revealed that WM was an important predictor of L2 success.

In another longitudinal study, Serafini and Sanz (2016) investigated whether the role of WM in L2 learning outcomes varied according to initial proficiency level of L2 learners. The participants were 87 native English learners of Spanish in an instructed university setting. However, due to participant attrition, the number was reduced to 33. They took Spanish courses at different levels (23 low proficiency, 33 mid-proficiency, 31 high proficiency) for one academic semester. They completed operation span and digit span tasks at the beginning of instruction. Additionally, they completed the same L2 proficiency tasks in three sessions (at the beginning, at the end, and after the instruction). The findings revealed that WM was significantly related to grammatical development of low and mid-proficiency learners while it played a minimal role in high proficiency learners' L2 development.

Investigating the role of WM in L2 grammar and reading development of low-proficiency L2 learners, Sagarra (2017) conducted two experiments. In the first one, the participants were 82 English learners of Spanish enrolled in introductory Spanish courses at an American university for two semesters (second and fourth). They completed a reading span task in the middle of the first semester, and grammar and reading pretests at the end of the second semester and grammar and reading posttests at the end of the fourth semester. The results indicated that WM neither modulated the improvement in L2 proficiency nor had an effect on any of the linguistic test. The researcher argued that WM was not a significant predictor because the RST did not involve a taxing processing component. In the second experiment, with 330 low proficiency learners of Spanish, a reading span task with a more taxing processing component was employed. In contrast to the first experiment, WM capacity was found to be a strong predictor of L2 grammar and reading development over the course of one semester.

Finally, Chang et al. (2019) conducted two experiments to look into the role of WM in L2 learning outcomes. In the first one, 150 eighth graders learning English in China took the researcher-developed reading span task in L2 and midterm and final exams. The results revealed a significant relationship between WM and L2 skills (grammar, writing, and reading), and WM was found to be a key predictor of reading abilities. In the second experiment aiming to explore the effects of different components of WM, a battery of a modified reading span (in L2), Stroop, n-back, and number-letter transfer tasks was given to 80 out of 150 participants from the first study. The findings indicated that verbal WM and n-back tasks significantly accounted for reading comprehension. However, information regarding the intensity of the instruction and the proficiency level of the participants was not provided in the article.

In short, despite a substantial number of studies exploring the predictive validity of WM in relation to L2 learning outcomes, few longitudinal studies exist in the literature. These longitudinal studies reported varying degrees of WM effects due to the use of different research designs and measurement tools, and different starting proficiency levels of participants. Although these studies had a pre-test/post-test design, WM capacity was tested either at pre-test or at post-test, but not at both times. Other methodological limitations of these longitudinal studies can be listed as low sample size and data collection periods shorter than the optimum length of six months or over (Linck & Weiss, 2011, 2015; Sagarra, 2017 [Experiement I]; Serafini & Sanz, 2016), and not measuring L2 proficiency at the beginning and end of the study (Chang et al., 2019; Kormos & Sáfár, 2008). Additionally, in all of these studies WM was measured with only one complex WM task, typically verbal in nature, adopting the domain-specific view of WM yet disregarding the domain-general view. It should be noted that the operation span task, the most commonly administered complex span task across the fields, for example, was not found to be suitable for discriminating university level participants (Draheim, Harrison, Embretson, & Engle, 2017). Ideally, multiple tasks should be used to best measure a cognitive construct like WM capacity, so that the construct-irrelevant and task-specific variance can be minimized (Draheim, Hicks, & Engle, 2016).

2.4 Summary and the goal of the current study

EFs having important implications for everyday cognitive operations differ among people and the constant use of these functions in dual language management can strengthen domain-general EFs processes. The research conducted to examine the role of life-long or short-term L2 experience in enhancing two major components of EFs of young adults, namely inhibition and WM, has yielded mixed results. To explain the contradictory results, such factors have been listed as insufficient interindividual variability in relation to the calculation of dependent variable and nature of the measurement tools, treating bilingualism as a categorical variable rather than a continuous one, disregarding individual differences rather L2 learning, and misconceptualization of EFs. Moreover, the issue of EF-related individual differences in L2 learning has been examined in numerous studies, yet results remain inconclusive. Some limitations concerning the extant research have been listed as low sample size, shorter data collection period than six months, not administering both pre- and post- training L2 proficiency task, adoption of domain-specific view only, and lack of multiple measures to assess EFs. Finally, there remains a need to clarify the relationship between EF components and L2 learning outcomes.

Based on these considerations above, the following questions are investigated in this current study:

1. Does intensive L2 education have an effect on the development of EFs of L2 learners?
    a. Does L2 learners' WM capacity significantly differ between the beginning and end of six months of intensive L2 instruction?
    b. Does L2 learners' inhibitory control ability significantly differ between the beginning and end of six months of intensive L2 instruction?

2. Is there a significant difference between the changes in EFs of L2 learners and those of near-monolingual freshman students over six months of education?
    a. Are there significant differences between L2 learners and near monolingual freshman students in terms of WM capacity at the beginning and the end of six months of education?
    b. Are there significant differences between L2 learners and near monolingual freshman students in terms of inhibition at the beginning and the end of six months of education?
3. Do the baseline EF components form coherent dimensions that are relatively independent of one another? If yes, do the baseline EF dimensions explain significant amounts of variance in L2 proficiency after six months of L2 instruction beyond initial L2 proficiency?

In the light of the findings of extant research on inhibition (e.g. Sullivan et al., 2014; Xie, 2018) and WM (Huang et al., 2020; Kerrigan et al., 2017; Yang, 2017) showing that short-term and/or intensive L2 training can lead to enhanced EFs, intensive L2 education for six months is hypothesized to enhance both inhibitory control abilities (Hypothesis 1) and WM of L2 learners (Hypothesis 2). Next, the participants exposed to intensive L2 instruction are expected to have significantly higher gain scores in terms of both inhibition (Hypothesis 3) and WM capacity (Hypothesis 4) compared to near-monolingual freshman students since experience in L2 requires the utilization of more EF-related cognitive processes for the management of dual language activation (Abutalebi, 2008; Abutalebi & Green, 2007; Bialystok et al., 2009; Green, 1998; Li et al., 2014b; Yang et al., 2018). With respect to the EF components, tasks measuring inhibition are expected to form a dimension independent from the one composed by WM tasks in line with the unity and diversity model of Miyake et al. (2000) (Hypothesis 5). Although the framework of executive attention framework of Engle and colleagues encompasses inhibition (attention control in their terms) as a mechanism of WMC (Shipstead et al., 2014), complex span tasks and inhibition load on different latent variables in their studies. In this study, administering an accuracy-based task (antisaccade) and a threshold version of flanker is expected to increase the correlation between the inhibition tasks and the likelihood of having them load in the same factor (Draheim et al., in review). Additionally, L2 learners' baseline measure of WM is hypothesized to be a robust predictor of L2 success at the end of a six-month intensive L2 instruction (Hypothesis 6), given the findings of longitudinal research concerning the predictive validity of WM for L2 learning outcomes (e.g., Kormos & Sáfár, 2008; Linck & Weiss, 2011, 2015; Sagarra, 2017; Serafini & Sanz, 2016), and the proposal of the executive attention framework of Engle and colleagues (Engle & Kane, 2004) that complex span tasks predict higher order abilities since they tap general attention capabilities. A similar prediction is also made for inhibition (Hypothesis 7) although this hypothesis would be inconsistent with the findings of Linck and Weiss (2011, 2015) regarding inhibition. However, the limitations of these studies will be dealt with successfully in the current study.

CHAPTER 3

METHODOLOGY

3.1 Participants and the context of research

This study was conducted at a foundation university in Istanbul, Turkey, in the 2018-19 academic year. This university offers bachelor's degree programs either in Turkish or English. In order to gain admission, high-school graduates in Turkey have to take a university entrance exam that is administered by Measurement, Selection and Placement Center (ÖSYM) once a year in three rounds: a Basic Proficiency Test (TYT), an Area Qualification Test (AYT), and a Foreign Language Test (YDT). Approximately 1.5 million high school graduates take these multiple-choice format examinations. TYT consists of 40 Turkish, 20 social science, 40 basic mathematics, and 20 science questions. AYT consists of 40 Turkish, 40 social science, 40 mathematics, and 40 science questions. YDT consists of 80 second language questions, yet only the ones who would like to study a foreign language teaching or literature take YDT together with TYT and AYT. All students take the TYT exam, and in the AYT exam they answer the questions from the fields according to the program that they plan to apply for. Five types of scores are calculated for each student, on the basis of their performances on these exams and their high-school GPA: TYT, AYT verbal, AYT quantitative, AYT equally weighted, and YDT. Accordingly, the students are ranked among each other for each type of score, which is used during their admission to a university program. Since no participants in the current study were admitted to the university with their TYT or YDT scores, the weightings of the fields in the AYT exam and the related programs are provided in Table 1. Across the programs, 40% of a total exam score comes from TYT with the weightings of the fields: Turkish: 13%; Social Sciences: 7%; Basic Mathematics: 13%; and Science: 7%.

Table 1. Types of AYT Scores And Weightings of Fields, And Examples for Related Programs

Types of Scores	Fields in the Exam	Weightings (%)	Programs
Quantitative Scores (QS)	Mathematics	30	Industrial Engineering, Computer Science and Engineering, Architecture, Medicine, Statistics
	Physics	10	
	Chemistry	10	
	Biology	10	
Verbal Scores (VS)	Turkish Language and Literature	18	Cinema and Television, Public Relations and Advertising, History, Turkish Language and Literature
	History - 1, 2	7, 8	
	Geography - 1, 2	5, 8	
	Philosophy	8	
	Religion	8	
Equally Weighted Scores (EWS)	Turkish Language and Literature	18	Political Science and International Relations, Psychology, Management, Economics, Law, Interior Architecture and Environmental Design
	History - 1	7	
	Geography - 1	5	
	Mathematics	30	

Upon admission to the university where the present study was conducted, those students who would study in an English-medium program need to take and pass the university's English proficiency test in order to be able to enroll in freshman undergraduate courses. The minimum score to pass the proficiency exam is equivalent to 78 on the TOEFL IBT. Those students who failed the proficiency exam take the university's placement exam in English and were placed into one of the four levels in the intensive language education program: very low proficiency level (elementary), low proficiency level (pre-intermediate), mid-proficiency level (intermediate), or high proficiency level (upper-intermediate) classes. These students were exposed to either one semester or one academic year of second language (L2) education until they demonstrated that they attained the necessary proficiency in

English to proceed to their undergraduate studies. One group of participants of the present study (experimental group) were elementary level learners of English, and all of them needed to follow the intensive L2 education program for one academic year. Detailed information about the program is provided in Table 2.

Table 2. Overview of The One Academic Year Intensive L2 Education Program And Hours of Instructions

Semester	Level	Proficiency	Length of Instruction	
Fall				
	Elementary	Very Low	9 weeks	216 hours
	Pre-Intermediate	Low	8 weeks	192 hours
Spring				
	Intermediate	Mid	8 weeks	192 hours
	Upper-intermediate	High	8 weeks	192 hours
Total			33 weeks	792 hours

The participants in the experimental group were 165 elementary level L2 learners (88 females, aged 18.7 years on average) registered in the intensive language education program in the 2018-19 academic year.

L2 learning took place at an instructional context where non-native teachers of English followed a modular curriculum based on the Common European Framework for Reference (CEFR). Instructions were delivered through multiple procedures of focus on form and focus on forms, and skill- and/or structure-based thematic course books and in-house materials were used. Particular assessment tools were administered to decide if learners could proceed to an upper proficiency level to study or not.

Since the participants completed intermediate level between pre-test and post-test times, they were exposed to a total of 600 hours of language instruction in six months (25 weeks). Table 3 presents their number, programs, departments, and ratings on the university enterance exam. As can be seen, they were from various fields, and had heterogeneous academic success levels.

Table 3. Programs, Departments And Admission Scores with Rankings of The Participants in The Experimental Group

Program	Types of Admission Scores	Ranking*	Department	N	Percentage
Engineering and Natural Sciences	Quantitative	$25K^{th}$ - $280K^{th}$	Civil Engineering	4	20.60
			Computer Science and Engineering	12	
			Electrical and Electronics Engineering	5	
			Industrial Engineering	12	
			Mechanical Engineering	1	
Architecture and Design	Quantitative	$50K^{th}$ - $400K^{th}$	Architecture	5	7.20
			Industrial Design	3	
	Equally Weighted	$80K^{th}$ - $260K^{th}$	Interior Architecture and Environmental Design	4	
Humanities and Social Sciences	Verbal	550^{th} - $700K^{th}$	History	2	27.3
			Turkish Language and Literature	5	
	Equally Weighted	$5K^{th}$ - $750K^{th}$	Philosophy	2	
			Political Science and International Relations	13	
			Psychology	19	
			Sociology	4	
Communica-tions	Verbal	$2K^{th}$ - $750K^{th}$	Public Relations and Advertising	2	6.0
			Cinema and Television E	7	
			New Media and Communication	1	
Law	Equally Weighted	400^{th} - $180K^{th}$	Law	42	25.5
Islamic Studies	Verbal	114^{th} - $200K^{th}$	Islamic Studies	7	4.2
Management and Administrative Sciences	Equally Weighted	$10K^{th}$ - $700K^{th}$	International Trade and Management	6	9.1
			Management	5	
			Management Information Systems	4	

*Ranking indicates the range of students' ranking out of 1.5 million high-school graduates in the specific score type.

Only 26.7% of them reported to have received music education before. None reported to have received English language instruction above elementary level, taken language instruction in a different foreign language than English, or spent time abroad. Their parents had 11 years of formal education on average.

They were not given any compensation for their participation in the study. They were told that if they participated, they would be able to learn their performances in inhibition and WM tasks at the end of the study. The ones who attended the sessions during class time were not marked absent.

Upon admission, those students who would study in a Turkish-medium program are able to enroll in their freshman courses. They usually have a limited proficiency in L2 and do not have to take any intensive L2 instruction to proceed to their faculty studies. Instead, during the first semester of their undergraduate study, they take an English language course at an elementary level as one of their must-courses. During the data collection time, they received 30 hours of English language instruction. One hundred three near-monolingual freshman students (70 females, aged 18.6 years on average) with no or minimal exposure to another language were recruited as the comparison group in the current study. Table 4 presents their number, programs, departments, and ratings of these participants registered in Turkish-medium program in the 2018-19 academic year. As can be seen, they were from various fields, and had heterogeneous academic success levels. In comparison to the experimental group, although the range of departments is lower in the control group, heterogeneity and rankings of the participants are similar.

Only 23.5% of them had taken music education. None reported to have received English language instruction above an elementary level, taken language instruction in a different foreign language than English, or spent time abroad. Their parents had 10.5 years of formal education on average.

Table 4. Programs, Departments And Admission Scores with Rankings of The Participants in The Control Group

Program	Types of Admission Scores	Ranking*	Department	N	Percentage
Engineering and Natural Sciences	Quantitative	120Kth - 230Kth	Industrial Engineering	5	4.9
Architecture and Design	Quantitative	18Kth - 240Kth	Architecture	19	38.8
	Equally Weighted	60Kth - 500Kth	Interior Architecture and Environmental Design	21	
Humanities and Social Sciences	Equally Weighted	50Kth - 700Kth	Political Science and International Relations	2	39.8
			Psychology	39	
Communications	Verbal	8Kth - 880Kth	Cinema and Television	2	11.7
			Public Relations and Advertising	10	
Management and Administrative Sciences	Equally Weighted	290Kth -690Kth	Management Information Systems	5	4.9

*Ranking indicates the range of students' ranking out of 1.5 million high-school graduates in the specific score type.

The same compensation procedure was followed for the participants in the control group.

3.2 Materials

In the present study, one standardized placement test was administered to measure the proficiency level of the participants; and two WM capacity tasks and two inhibitory control tasks were used to measure EFs of the participants. In addition, one background questionnaire was employed.

3.2.1 Background questionnaire

The researcher developed a background questionnaire for the participants (see Appendix A [see Appendix B for the Turkish version]). It included parts inquiring gender, academic program, ranking in the university entrance exam, language and education background, parents' level of education, music education background, and the length of staying abroad. The average years of mother and father education are considered as the indicator of SES. As for the music education background, a period of six months or longer is thought to be noted as a participant has taken music education.

3.2.2 Placement exam

To measure the proficiency level of participants in English, the English Placement Test (EPT) (2006) by University of Michigan was employed. This test is a standardized one with 100 multiple choice questions. It has four sections with varying number of questions: grammar with 30, vocabulary with 30, listening with 20, and reading with 20 questions. Participants have 50 minutes to answer the questions. In the grammar section, they choose the most appropriate option to fill in a blank in a dialog given for each question. In the vocabulary section, they choose the best lexical item that completes a sentence for each question. In the listening section, they hear either a question or a statement. If it is a question, they choose the best option that answers the question. If it is a statement, they choose the best option that corresponds to the statement. In the reading section, they answer the questions related to the sentences or paragraphs that they read. The participants receive one point for each correct answer and zero points for incorrect or unanswered ones. The maximum score is 100.

3.2.3 The tasks of EFs

With respect to the tasks to measure EFs, visual tasks are chosen for both WM and inhibition since the present research has adopted a domain-general perspective of WM. Visual complex span tasks, namely symmetry and rotation span tasks, were employed to assess WM (Foster et al., 2015) while antisaccade (Hutchison, 2007; Kane, Bleckley, Conway, & Engle, 2001) as a benchmark test of inhibition (Hutton & Ettinger, 2006; Rey-Mermet et al., 2018), and a threshold version of flanker (Draheim et al., in review) for inhibition. These accuracy-based and psychometrically rigorous measures were thought to be the ones that can help to reach more valid conclusions about enhanced cognitive functioning in bilinguals as stated by Lehtonen et al. (2018). E-prime 2.0 was employed to perform all the EF tasks.

3.2.3.1 Symmetry span task

Participants recall sequences of red squares within a 4x4 grid matrix after performing a symmetry-judgment task as a simple processing (distractor) task. They first perform three practice sessions. In the storage alone practice session, the participants see sequences of red squares that appear in the matrix, and at recall they click the correct locations of the red squares in the matrix in the order they appear. In the processing alone practice session, the participants perform a symmetry-judgment task including an 8x8 matrix with some squared filled in black. They decide whether the pattern is symmetrical about its vertical axis. It is symmetrical approximately half of the time. After this practice section, the program calculates each participant's mean time (plus 2.5 standard deviations [SDs]) needed to complete the processing practice for each task, which is employed as a time limit for the processing portion of the experimental session for that specific individual. Each answer period that is longer than this time limit automatically counts that trial as an error. This practice session includes 15 symmetry-judgment tasks.

In the final practice session combining the matrix recall task with the symmetry-judgment task, immediately after deciding whether the current matrix is symmetrical, participants are presented with a 4x4 matrix with one of the cells filed in red for 650 milliseconds (ms). At recall, they recall the sequence of red-square locations by clicking on the cells of an empty matrix in the correct order in which the locations are displayed in the current set (see Figure 1 for a sample trial).

The real trial consists of three sets of each set size, with the set sizes ranging from two to five. The order of set sizes is random for each participant, and +/- 3 SD accuracy criterion is imposed to make sure that participants perform the symmetry judgment tasks accurately. Percentage of the accuracy is displayed in red in the upper right-hand corner of the screen. Partial-credit unit scoring procedure is applied to calculate the DV (Conway et al., 2005).

3.2.3.2 Rotation span task

Participants recall a series of serially presented items, the presentation of which is interrupted by a simple processing (distractor) task. The processing task is to mentally rotate a normal or mirror-reversed G, F, or R as a recall cue (approximately two cm tall, rotated at 0^0, 45^0, 90^0, 135^0, 180^0, 235^0, 270^0, or 315^0) and indicate if it is normal or mirror-reversed; it is normal half of the time (see Figure 1 for a sample trial). After the response to the recall cue, the screen is blanked for 500 ms, and a large or small arrow radiating out from the center of the screen in one of eight directions is displayed for one second. In total, there are 16 arrow size X arrow direction combinations each of which is used approximately equally often in the task and none of which is repeated within a set. When the arrow disappears, another recall cue appears. When the last recall cue is presented, the participants recall all of the arrows from the preceding displays in the order they appear. Set sizes range from two to five letter-arrow displays per trial. The scoring procedure is the same as the symmetry span.

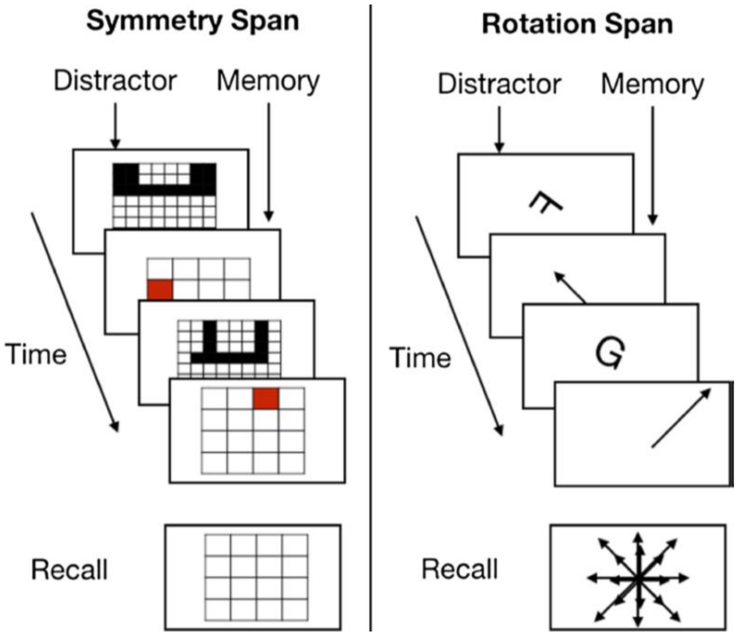

Figure 1. Sample trials on symmetry and rotation span tasks

3.2.3.3 Antisaccade

In this task, participants divert their gaze from a peripheral flash occurring one side of the screen and identify and report a letter briefly presented on the opposite side of the screen. They first see a variable duration fixation screen with a fixation cross in the middle lasting a random amount of time between 2000-3000 ms followed by an alerting tone for 300 ms. Then, an asterisk (distractor) appears for 300 ms on either the right or left hand side of the screen followed immediately by a target "Q" or an "O" for 100 ms on the opposite side of the screen than the asterisk. The location of the asterisk and target letter are both masked for 500 ms by "##". Participants are expected to ignore the asterisk and instead look away to the other side of the screen to catch the target letter and press the associate key on the keyboard.

There are three practice sessions. First, participants perform the 30 response mapping trials in which the cue and letter are presented in the center of the screen. In the second one, they complete 15 prosaccade trials in which the cue and letter are displayed on either the left or right side of the screen, yet they appear on the same side.

The last practice session includes the 10 antisaccade trials in which the cue and the letter appear on either the left or right hand side of the screen, yet they occur on the opposite sides of the screen. After completing the practice sessions, participants perform the experimental block including 60 antisaccade trials (see Figure 2 for a sample trial). Here are left and right letter locations and two letter responses equally divided across the trials. The proportion of the correctly identified target letters across all the trials is the dependent variable.

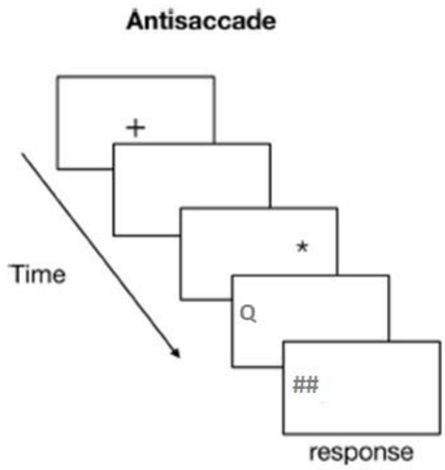

Figure 2. A sample trial on antisaccade task

3.2.3.4 Flanker deadline

This task is a modified version of the arrow flanker that employs an adaptive procedure to estimate the subject’s threshold (Draheim et al., in review). Eighteen blocks of 18 trials each (total 324 trials) are administered. Each trial has a response deadline that limits how long a participant has to respond before hearing a loud beep and losing the opportunity to respond on that trial. This deadline either decreases (less time to respond) if the participant is accurate on at least 15 trials within each block or increases (more time to respond) if their accuracy rate is below that. The first block has a response deadline of 1050 ms. For the first six blocks, the response deadline either decreases by

90 ms or increases by 270 ms for the next block, depending upon the accuracy of the participant on at least 15 of the 18 trials. For subsequent blocks, the response deadline either decreases by 30 ms or increases by 90 ms. If the response deadline besets below 150 ms after any block, it is automatically set to exactly 150 ms. The stimuli remain on the screen up until the response deadline. Each block has 12 congruent (e.g., → → → → →) and six incongruent trials (e.g., → → ← → →) in random order with a randomized 400-700 ms. The response deadline is the same for both trial types. Further, both congruent and incongruent trials are treated equally in determining if the deadline increases or decreases for the next block (i.e., regardless of the congruency, participants need to respond accurately before the response deadline on 15 of 18 total trials in order for the response deadline to decrease). The dependent variable is the response deadline after the final block. Participants respond by using the keyboard on which arrow stickers are placed ('z' for left pointing arrows and '.' for right pointing arrows).

3.3 Design and procedure

In February 2018, an ethics approval for the research was received from the Ethics Committee at the university where the data were collected (see Appendix C). In July 2018, a pilot study was conducted with 30 L2 learners who had started as elementary in the intensive language program in September 2017 and were upper intermediate level learners during the piloting. Two research assistants from the Department of Psychology were recruited to collect the data for two weeks in the Cognitive Psychology Lab of the same university. The instructions for the tasks had been translated into Turkish and were found to be clear. Significant correlations between inhibitory tasks (r = .37, p < .05) and between WM tasks (r = .55, p < .01) were found.

3.3.1 Real trial I

The participants in the experimental group were first given the EPT placement test during the second week of their study in fifteen elementary level classes on September 25, 2018. As of that time, they were invited to the laboratory at their convenience. Upon arrival at the lab, they filled in an informed consent form first, and completed a background questionnaire, and then carried out four EFs tasks in a counterbalanced order, the completion of which took 45-50 minutes. Two participants at a time were able to perform the tasks in two different quiet lab rooms. In addition to the assistants having worked during the piloting, four more research assistants from the same department were recruited to collect data for the real trials.

The participants in the control group were first given the EPT placement test during the fourth week of their study in the first semester (October 15-19) in their classrooms where they were taking ENG 101 English Language Introductory Course. As of that time, they were invited to the laboratory at their convenience. Upon arrival, they followed the same procedure as the experimental group did.

The data collection procedure took six weeks in total. The first three weeks were mainly allocated to the experimental group and the other three to the control group.

3.3.2 Real trial II

The second phase of the real trial was conducted in March 25 - May 3, 2019. The data collection in this second phase took as long as the first one did. First, the same participants were invited to laboratory at their convenience and their appointments were organized in a way that each participant did the tasks after a six-month period. They completed the EF tasks in the same order as they did during the first phase in which a counterbalance order was followed.

One hundred forty-three out of 165 participants of the experimental group completed the EF tasks in the second session of the real trial while 86 out of 103 participants of the control group did so. One hundred fifty-four out of 165 participants in the experimental group were given the EPT placement test on May 2 when the majority of them (N = 119) had just completed their intermediate level study and were taking upper-intermediate level language courses while the others were studying intermediate level courses. In six months of time, the participants in the experimental group took approximately 600 class hours of English language instruction while the ones in control group took approximately 30 class hours of English language instruction. Since the participants in the control group did not take any intensive language education, they were not given the EPT in the second phase of the experiment.

Upon their arrival at the laboratory, the participants in the control group were asked if they had taken any other English language courses or got involved in English-related activities apart from their elementary level introductory course. None reported to have done so.

None of the participants from either group reported to have taken any music education between the first and second phases of the experiment.

3.4 Data trimming and outlier analyses

First, the accuracy in the processing stage of WM tasks was taken into consideration, and the observations for each participant in WM tasks were excluded if their accuracy deviated from the mean by more than 3 SDs. To ensure normal distributions for the EF and proficiency variables, for each participant, observations that deviated from the mean by more than 3 SDs were excluded. Outliers were checked in a boxplot and the observations that were outside 2.2 times the interquartile range above the upper quartile and below the lower quartile were excluded as well. No data transformation was implemented since some analyses to answer the research questions would be on the basis of the performance differences on the same tasks between two stages of the experiment.

CHAPTER 4

FINDINGS

This chapter presents the analyses and results in relation to enhancement of executive functions (EFs), dimensions of EFs, and the predictive validity of EFs for second language (L2) learning outcomes.

4.1 Enhancement of EFs

The first two research questions were concerned with the role of receiving intensive L2 instruction on the enhancement of EFs in comparison to receiving undergraduate freshman education in mother tongue. It should be noted that while antisaccade and flanker tasks were employed to measure inhibition component of EFs, rotation and symmetry span tasks tapped WM capacity.

Table 5 presents the descriptive statistics for the performances of all participants on proficiency and EF measures both at Time I and Time II. There was a six-month time difference between Time I and Time II. Overall, the participants demonstrated higher WM and inhibition performances in Time II, as compared to their performances in Time I.

Table 6 shows the descriptive statistics separately for the participants who took an intensive L2 instruction (experimental group) and the near-monolingual freshman students in a Turkish program (control group). It is clear that both groups had higher inhibition and WM task performances in Time II compared to Time I.

Table 5. Descriptive Statistics for All The Participants at Time I And Time II

	Time	N	Mean	SD	Min	Max	Skewness	Kurtosis
Antisaccade								
	I	258	.77	.15	.41	1	-.73	-.42
	II	206	.88	.08	.64	1	-1.05	.52
Flanker								
	I	247	638.74	150.84	390	1110	1.03	.92
	II	207	610.29	114.76	390	870	.67	-.02
Rotation								
	I	253	24.79	7.61	5	42	-.47	-.27
	II	225	26.77	7.04	6	41	-.46	.03
Symmetry								
	I	257	27.39	6.86	8	40	-.38	-.16
	II	225	29.51	6.55	10	42	-.48	-.18
Proficiency								
	I	253	16.52	7.28	0	34	0.14	-.68
	II*	154	48.10	10.97	25	78	0.26	.70

Note. Higher Flanker scores indicate lower inhibition performances.
*Statistics for proficiency in Time II depict the performances of the experimental group only.

Table 6. Descriptive Statistics for Both Groups at Time I And Time II

Tasks	Time	Group	N	Mean	SD	Min	Max	Skewness	Kurtosis
Antisaccade									
	I	Experimental	161	.77	.15	.41	1	-.73	-.42
		Control	97	.76	.14	.45	1	-.48	-.73
	II	Experimental	131	.89	.08	.68	1	-1.24	1.20
		Control	75	.87	.09	.64	1	-.74	-.24
Flanker									
	I	Experimental	149	632.42	153.74	390	1110	1.00	0.80
		Control	98	648.37	146.57	390	1110	1.12	1.26
	II	Experimental	131	601.60	112.21	390	870	0.84	0.19
		Control	76	625.26	118.29	390	870	0.42	-0.11
Rotation									
	I	Experimental	154	24.69	7.82	6	40	-0.47	-0.47
		Control	99	24.95	7.31	5	42	-0.47	0.15
	II	Experimental	139	27.29	6.97	6	41	-0.42	-0.08
		Control	86	25.94	7.11	6	40	-0.54	0.20
Symmetry									
	I	Experimental	157	27.62	6.90	11	40	-0.30	-0.38
		Control	100	27.02	6.82	8	40	-0.52	0.20
	II	Experimental	139	29.75	6.70	10	42	-0.51	-0.13
		Control	86	29.13	6.32	13	40	-0.46	-0.22
Proficiency									
	I	Experimental	159	16.74	7.14	3	34	0.23	-0.58
		Control	94	16.15	7.53	0	32	0.02	-0.85
	II	Experimental	154	48.10	10.97	25	78	0.26	0.70
		Control	-	-	-	-	-	-	-

Although the experimental group obtained higher scores across the tasks in Time I than the control group did, independent-samples *t*-tests indicated that there were no significant differences between the groups in their performances across the tasks (antisaccade: $t(256) = 1.5$, $p > .05$; inhibition: $t(245) = -.81$, $p > .05$; rotation: $t(251) = -.27$, $p > .05$; symmetry: $t(255) = .69$, $p > .05$; proficiency $t(251) = .69$, $p > .05$). A paired-samples t-test comparing the L2 proficiency of the participants in experimental group between Time I and Time II revealed a significant difference, $t(148) = 38.54$, $p < .001$. The mean scores of proficiency indicated that these participants were elementary or beginner level learners of English at Time I and reached intermediate level of proficiency at Time II (Walter & Hentschel, 2013).

Table 7 presents the correlation matrix for the EF measures of all participants at both Time I and Time II. Although the highest magnitude of correlations was observed between WM tasks at both Time I and Time II (almost .50), the correlation between inhibition tasks was even below.20 at Time II.

Table 7. Correlation Matrix for The EF Measures Administered at Time I And Time

	1	2	3	4	5	6	7
1. Antisaccade_T1	-						
2. Flanker_T1	$-.253^{**}$	1					
3. Rotation_T1	$.282^{**}$	$-.261^{**}$	1				
4. Symmetry_T1	$.349^{**}$	$-.221^{**}$	$.492^{**}$	1			
5. Antisaccade_T2	$.673^{**}$	$-.252^{**}$	$.259^{**}$	$.379^{**}$	1		
6. Flanker_T2	$-.193^{**}$	$.482^{**}$	$-.152^{*}$	$-.234^{**}$	$-.176^{*}$	1	
7. Rotation_T2	$.305^{**}$	$-.188^{**}$	$.495^{**}$	$.439^{**}$	$.317^{**}$	$-.269^{**}$	1
8. Symmetry_T2	$.332^{**}$	$-.262^{**}$	$.423^{**}$	$.543^{**}$	$.324^{**}$	$-.194^{**}$	$.488^{**}$

Note. T1 (measured in Time 1 as pre-test); T2 (measured in Time 2 as post-test).
$^{*}p < .05$, $^{**}p < .01$

These correlations make it possible to conduct a repeated-measures of multivariate analyses of variance (MANOVA) to explore the joint effect of group and time germane to the enhancement of WM capacity, yet impossible to do so for the inhibition tasks because they are not within desired range (.20-.60) at Time II (Meyers, Gamst, & Guariono, 2006). Therefore, to look into the enhancement in inhibition, one repeated-measures analysis of variance (ANOVA) for each inhibition task was performed separately. These analyses are reported in the following subsections.

It is also indicated that the magnitude of correlations between antisaccade and WM at Time I got higher or remained similar at Time II, yet the one between inhibition tasks slightly went down.

4.1.1 Enhancement of WM

In order to examine whether receiving six months of intensive L2 instruction enhances the WM capacity of L2 learners and whether there is a significant difference between the changes in the WM capacity of L2 learners and those of near-monolingual freshman students, a 2 X 2 repeated-measures MANOVA was performed to investigate the joint effects of group and time on WM capacity of the participants. Two dependent variables were used: rotation span and symmetry span. The between-subjects independent variable was group (two levels: experimental, control) and within-subjects independent variable was time (two levels: pre-test, post-test). Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity, with no violations noted. The results yielded no significant interaction between group and time or no main effect of group. Yet, there was a statistically significant main effect of time on the combined dependent variables, F (2, 210) = 14.46, p < .001; Wilks' Lambda = .89; η^2 = .114; observed power = .998. When the results for the dependent variables were considered separately, using a Bonferroni adjusted alpha level of .025, the main effect of time was observed on both rotation, F (1, 211) = 9.44, p = .002; η^2 = .04; observed power = .864, and symmetry, F (1, 221) = 19.88, p < .001; η^2 = .08; observed power = .993.

4.1.2 Enhancement of inhibition

To check the joint effects of group and time on inhibition performances of participants as measured by antisaccade, a 2 X 2 (Group: experimental; control; Time: pretest; posttest) repeated-measures ANOVA was conducted with group as a between-participants factor and time as a repeated factor. Preliminary assumption testing was conducted to check for normality, univariate outliers, homogeneity of variance-covariance matrices, with no violations noted. Although there was no significant interaction between group and time or no main effect of group, the findings revealed a significant main effect of time on antisaccade performances of the participants, F (1, 199) = 146.57, $p < .001$; Wilks' Lambda = .58; $\eta^2 = .424$; observed power = 1.

To check the joint effects of group and time on inhibition performances of the participants as measured by flanker, a 2 X 2 (Group: experimental; control; Time: pretest; posttest) repeated-measures analysis of variance (ANOVA) was conducted with group as a between-participants factor and time as a repeated factor. Preliminary assumption testing was conducted to check for normality, univariate outliers, homogeneity of variance-covariance matrices, with no violations noted. The findings revealed no interaction between group and time or no main effects of group or time.

Overall, the findings germane to the enhancement of EFs of the participants revealed that the performances of both groups on WM tasks and antisaccade as an inhibition task improved within a six-month of time, yet their performances on flanker as an inhibition task did not. No group differences were observed across the tasks.

4.2 EFs and L2 performance

This part presents the analyses and findings related to the third research question concerning the dimensions of EFs components in Time I and predictive validity of the EF measures in Time I for L2 learning outcomes in Time II beyond initial L2 proficiency (i.e., proficiency measured at Time I).

Table 8 presents the correlation matrix for the EF measures of the participants in the experimental group at Time I, and the proficiency measures both at Time I and Time II.

Table 8. Correlational Matrix for Performances of Experimental Group on EF Measures at Time I And Proficiency Measures at Time I And Time II

	1	2	3	4	5
1. Proficiency_T2	1				
2. Proficiency_T1	.452***	1			
3. Flanker_T1	-.242**	-.236**	1		
4. Antisaccade_T1	.087	.135	-.203*	1	
5. Symmetry_T1	.148	.178*	-.256**	.308***	1
6. Rotation_T1	.096	.178*	-.213*	.262***	.511***

Note. T1 (measured in Time I); T2 (measured in Time II)
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

The magnitude of correlations between antisaccade and complex span tasks is higher than the one between antisaccade and flanker. Although complex span tasks were significantly correlated with proficiency at Time I, they were not at Time II. Initial proficiency had the highest correlation with L2 learning outcomes (i.e., proficiency measured at Time II).

4.2.1 Coherent dimensions of EFs

In order to look into whether the baseline EF components formed coherent dimensions that are relatively independent of one another, a principal component analysis (PCA) was conducted on the four measures of EFs at Time I with orthogonal rotation (varimax). The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .65, exceeding the recommended value of .6 (Kaiser 1960), and all KMO values for individual measures were > .61, which is above the acceptable limit of .5 (Field, 2009). Bartlett's test of sphericity (Bartlett, 1954), $\chi 2$ (6) = 66.362, $p < .001$, indicated that correlations between items were sufficiently large for the PCA.

The PCA yielded one component with eigenvalue exceeding 1, explaining 47.21 % of the variance. Upon having the second component with an eigenvalue of .84 and examining the Cattell's (1966) scree test, it was decided to include a two-component solution which explained a total of 68.13% variance, exceeding the limit of 60%. Table 9 shows the factor loadings after rotation. The items that cluster on the same components suggest that Component 1, including rotation span, symmetry span, and antisaccade, represents 'executive attention', and Component 2, including only flanker, represents'inhibition'.

Table 9. Rotated Factor Loading

	Factor 1	Factor 2
Rotation Span	.836	-.026
Symmetry Span	.831	-.142
Antisaccade	.524	-.294
Flanker	-.130	.968

4.2.2 Predictive validity of EFs for L2 success

Upon checking the dimensions of EFs, a hierarchical regression was conducted in order to check the predictive validity of the baseline measures of EFs for the L2 success of the participants in the experimental group at the end of the six months of L2 instruction beyond initial L2 proficiency.

Prior to conducting the analysis, the relevant assumptions of this statistical analysis were tested. A sample size of 139 participants was deemed adequate given three independent variables to be included in the analysis (Hair, Black, Babin, Anderson, & Tatham, 2006; Tabachnick & Fidell, 2007). The independent variables were proficiency at Time I (initial L2 proficiency), flanker, and executive attention. Executive attention was the factor score obtained through the PCA in the previous subsection. Even though this factor score was significantly correlated with proficiency at Time I ($r = .18$, $p < .05$), it was not with proficiency at Time II ($r = .11$, $p > .05$). Therefore, it was removed from the equation.

Residuals and scatterplots indicated the assumptions of normality, linearity, and homoscedasticity were all satisfied (Mertler & Vannatta, 2002). An examination of correlations revealed that no independent variables were highly correlated with each other (Field, 2009). Also, the collinearity statistics, tolerance and VIF, were within accepted limits. Therefore, the assumption of multicollinearity was met (Field, 2009; Mertler & Vannatta, 2002). The assumption of independent errors was also sustained because Durbin-Watson value was within accepted limits (Field, 2009).

A two-stage hierarchical multiple regression was conducted with Proficiency in Time II (L2 success) as the dependent variable. Proficiency in Time I as initial proficiency was entered at stage one of the regression to control for initial proficiency. The Flanker scores were entered at stage two. The regression statistics are presented in Table 10.

Table 10. Summary of Hierarchical Regression Analysis for Predictors of Learning Outcomes

Variable	β	T	sr^2	R	R^2	ΔR^2
Step 1				.43	.18	.18
Initial Proficiency	.43	5.48***	.18			
Step 2				.46	.21	.03
Initial Proficiency	.39	4.94***	.00			
Flanker	-.16	-2.03*	.03			

Note. $N = 139$; *$p < .05$, ***$p < .001$

The hierarchical multiple regression revealed that at stage one, initial proficiency contributed significantly to the regression model, F (1,133) = 30.01, $p < .001$, and accounted for 18.4% of the variation in L2 learning outcomes. Introducing Flanker scores explained an additional 2.5% of variation in L2 learning outcomes and this change in R^2 was significant, F (2,132) = 17.43, $p < .001$.

4.3 Summary of the findings

The current research looked into the enhancement of EFs through intensive L2 instruction and the role of EF-related individual differences in L2 success, thereby clarifying the relationship between EFs and L2 experience. The findings revealed that both intensive L2 education and undergraduate freshman education led to enhancement in WM capacity and antisaccade performances, yet no enhancement on flanker performances was observed. It was also found that WM capacity and antisaccade performances had loadings on the same factor, namely executive attention, while flanker performance was independent from executive attention. Finally, as for the predictive validity of EFs for L2 learning outcomes at the end of a six-month intensive L2 instruction, only flanker performances were found to be a significant predictor, explaining 2.5% of the total variance. Initial L2 proficiency, on the other hand, accounted for 18.4% of the total variance.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The primary objective of this study was to clarify the relationship between executive functions (EFs) and second language (L2) proficiency in a university classroom setting by investigating the effects of L2 experience on the enhancement of EFs and the predictive validity of EFs for L2 success. Adopting a domain-general perspective of Engle and colleagues (Engle 2002, Engle & Kane, 2004), EFs were measured through multiple WM (rotation span and symmetry span) and inhibition (antisaccade and flanker) tasks, all in visual nature.

The data were collected from a large sample with relatively homogenous background through the measures of EFs along with a standardized English language proficiency measure. In a pretest-posttest design experiment, the experimental group was exposed to a six-month intensive L2 instruction leading to intermediate level of proficiency in English from elementary level, whereas the comparison group was exposed to undergraduate freshman education in their first language (L1). Confounding variables such as SES, ethnicity, education, and age were controlled, and psychometrically more robust EF measures were employed. To the best of the researcher, the present study is the first to investigate both bilingual advantage and predictive power of EFs for L2 success on the basis of these theoretical and methodological foundations.

5.1 Findings for the enhancement of EFs

The first two research questions concern the effect of intensive L2 instruction on the enhancement of EFs as compared to the undergraduate freshman education in L1.

We hypothesized that a six-month intensive L2 instruction would enhance both inhibition (measured by antisaccade and flanker tasks) and WM capacity (measured by symmetry and rotation span tasks) of the experimental group, and more enhancement would be observed in EFs of the experimental group as compared to those of the control group.

Findings indicated that although both groups improved their WM capacity after a six-month education, the experimental group did not outperform the control one at the end of this period. As regards inhibition, similar results were observed based on the antisaccade performances. Yet, neither improvement nor group differences were observed on the performances of either group on flanker. Although these findings confirm our predictions concerning the positive role of intensive L2 instruction on the enhancement of WM capacity (first hypothesis) and partially on inhibition (second hypothesis), they run contrary to the predictions regarding the cognitive advantage of L2 experience over undergraduate education on WM capacity (third hypothesis) and on inhibition (fourth hypothesis).

The enhancement on WM capacity and response inhibition through intensive L2 experience, as predicted by the first and second hypotheses respectively, could be explained by the cognitive processes that L2 learners go through. L2 learners in the experimental group might have benefited from L2 experience entailing a joint activation of two languages in a single mind and employed EF-related cognitive processes to maintain attention to the target language and avoid interference from the competing language (Abutalebi, 2008; Abutalebi & Green, 2007; Bialystok et al., 2009). This could have resulted in activation outside the language-related areas in the brain as well, yet in areas related to cognitive control (e.g., left prefrontal cortex and anterior cingulate cortex [ACC]). For instance, on the basis of their finding that young adults with life-long L2 experience used this brain structure of the brain more efficiently in resolving cognitive conflicts and had higher volume of gray-matter in this area, Abutalebi et al. (2012) highlight that dorsal ACC is tightly bound to domain-general executive control functions and associated with language control and resolutions of non-verbal conflicts. Li et al. (2014b) argue that EFs and language functions are dependent upon the same integrated brain network and L2 experience in the verbal domain leads to cross-domain effects. In the same vein, Yang et al. (2018) indicate that language control processes are modulated by contextual interaction through adaptive changes in the regions and circuits that are related to specific control processes.

The findings of the current study regarding the enhancement on both antisaccade and complex span tasks seem to accord with the inhibitory control (IC) model by Green (1998) positing that the ability or goal to use two languages appropriately requires the regulation of supervisory attention system (SAS) through top-down cues to activate the target-language and suppressing non-target one. The control mechanisms used in performing antisaccade would be relevant to the mechanisms employed during the regulation of SAS and bilingual language control to manage the potential conflict among competing representation. Hence, it would be argued that enhancement on EFs relies heavily on response inhibition since L2 experience trains the use of inhibitory control ability. This explanation would sound more plausible given the mediating role of attention control on the relationship between complex span tasks and primary memory tasks (Shipstead et al., 2014) and given the large effect size for the enhancement on response inhibition, yet a small effect size for the enhancement of WM capacity in the current study. In other words, the enhancement on antisaccade as one of the mechanisms of WM capacity could have triggered the enhancement on complex span tasks. However, it should be noted that no group differences were observed in the current study. Enhancement was also found in the EFs of those in control group.

The findings indicating a nonsignificant group difference on the enhancement of EFs against the predictions of the third and fourth hypotheses of the present study could be discussed in light of the existing research exploring the bilingual advantage of lifelong L2 experience. Although the amount of L2 experience and age of acquisition would not be comparable to the participants in those studies, the present research has such methodological advantages as controlling confounding variables like SES, age, and ethnicity, and administering psychometrically more robust EF measures.

With regard to inhibition, while our findings of nonsignificant group differences between the L2 learners in the experimental group and the near-monolingual control group corroborate previous research outcomes (Antón et al., 2019; Bialystok et al., 2005; Kousaie et al., 2014; Paap & Greenberg, 2013), they contrast with the findings of others (Bialystok et al., 2006; Blumenfeld & Marrian, 2011; Chung-Fat-Yim et al., 2017; Yang & Yang, 2016). For instance, Paap and Greenberg (2013) who administered antisaccade and flanker tasks similar to the ones used in the current study did not observe any differences between bilinguals and monolinguals, whereas Bialystok et al. (2006) presented evidence for the bilingual advantage based on antisaccade. Such contradictory findings could be explained by the demographic factors as confounding variables in the studies of lifelong L2 experience. Controlling many demographic factors, Antón et al. (2019) found null-effects comparing bilingual and monolingual groups, yet, instead of the antisaccade task, they used the flanker, Simon, and Stroop tasks which are subjected to the criticism germane to reliability and validity (Draheim et al., 2019; Paap et al., 2020a; Rey-Mermet et al., 2019). The findings of the current study, based on the threshold version of flanker (Draheim et al., in review) that is supposed to overcome the methodological weaknesses of the flanker task, corroborate those of Antón et al. (2019).

With regard to WM, our findings echo the results of the studies showing null-effects of bilingual advantage (Jiao et al., 2019; Morrison et al., 2019; Ratiu & Azuma, 2015; Smithson & Nicoladis, 2013). Yet, this comparison should be read with caution since WM tasks used in these studies show a great deal of variety ranging from forward digit span to verbal complex span tasks. Few studies employed non-verbal tasks (Ratiu & Azuma, 2015; Smithson & Nicoladis, 2013), yet they were not the same in nature as the ones we employed. Our findings are not supportive of the findings of Antón et al. (2019), Jiao et al. (2019), and Morrison et al. (2019) indicating bilingual advantage on more demanding WM tasks such as a backward version of a visual WM task, and a more demanding n-back task. Thus, our findings do not support the argument that, for the bilingual advantage to emerge, WM tasks should be more demanding and involve complex processing and retrieval that require domain-general WM system. The visual complex span tasks (symmetry and rotation span) employed in the current study revealed null-effects of bilingual advantage. The results could have been different had we employed verbal complex span tasks. However, it should be noted that the language of the verbal tasks would be a confounding variable with verbal measures, especially when the tasks are constructed in a non-dominant language for the bilinguals. Additionally, several studies have shown bilingual disadvantage on complex verbal tasks (e.g., Luo et al., 2013; Ratiu & Azuma, 2015; Sullivan et al., 2016).

The lack of differences between the participants that received intensive language education in L2 and those who received undergraduate education in L1 does not support theoretical predictions by a number of scholars in the field (e.g., Abutalebi, 2008, 2012; Abutalebi & Green, 2007; Green, 1998; Li et al., 2014b). This could be explained by referring to Ullman's (2001, 2004, 2006) declarative/procedural (DP) model of language learning. Specifically, the model posits that learners employ two distinct memory systems specialized for different types of learning and knowledge: declarative memory (DM) and procedural memory (PM). While the DM is involved in the acquisition and storage of semantic and episodic knowledge, the PM is responsible for learning and performing routinized skills. Whereas DM is implicated in explicit learning, PM in implicit learning. As regards L1, aspects of lexicon depend upon DM while the learning and use of grammar upon PM. As for L2 learning, particularly after puberty, both lexicon and grammar are dependent upon DM. This is probably because instructed L2 learners memorize linguistic forms and employ the associative lexical memory to generalize patterns, and most of them go through explicit learning process. Aspects of grammatical processing in late L2 learning are less dependent upon left frontal and basal ganglia structures in L2 than L1 (i.e., the brain regions associated with PM). As such, adult L2 learners in the present study might have relied primarily on their DM to learn the L2 just like undergraduate students learning academic content in their L1.

In order to reach a solid intermediate level of knowledge as in the case of present study, most language learning takes place through explicit instruction of grammatical structures and vocabulary supported by practice activities in various formats (Kormos & Sáfár, 2008), which bears resemblance to acquiring knowledge in other domains (e.g., mathematics). Gupta and Tisdale (2009) argue that involvement of both DM and PM systems are not restricted to word learning, they are involved in all types of learning. For example, Menon (2016) indicates that DP anchored in the medial temporal cortex (MTL) (the hippocampus) plays a significant role in learning mathematics. Considering the results of the present study demonstrating enhancement on EFs of both groups yet no group differences, it could be argued that undergraduate freshman courses offered to the control group such as Politics, Economy, Humanity and Art cover semantic knowledge that consists of concepts, facts, ideas, and beliefs (Tulving, 1983) through explicit classroom instruction, thereby requiring the participants in the control group to employ DM systems as the ones in experimental group did in learning L2.

This explanation could be supported by the findings arising from the review of 46 papers on functional neuroanatomy of adult L2 learning by Tagarelli, Shattuck, Turkeltaub, and Ullman (2019). They found a network of structures that seem to be involved in both lexical and grammatical L2 learning and reflect the involvement of DM. Similar to the findings of Menon (2016) germane to learning mathematics, the review also indicated that prefrontal control circuits anchored in the anterior insula, ventrolateral prefrontal cortex, and DLPFC serving flexible hubs for integrating information across attentional and memory systems were also activated during adult L2 learning. Additionally, the neural data from Niendam et al.'s (2012) review indicated that the commonality among the EFs components is associated with the ventral system, which plays a significant role in the DM systems as well (Tagarelli et al., 2019). Therefore, it would be plausible to claim that the acquisition of semantic knowledge through both attentional and DM systems could have enhanced the EFs of both groups.

The DP model (Ullman, 2004, 2006) also posits that late L2 learners tend to become native like through extensive practice and experience, which reflects a shift from the DM to PM system, and PM is associated with basal ganglia including such subcortical regions as caudate and putamen. These regions also reflect the commonality among the components of EFs (Niendam et al., 2012) and are found to be activated during grammar learned through implicit learning or later stages of language learning (Tagarelli et al., 2019). The participants in the current study were late L2 learners with nonnative levels of proficiency who rely primarily on their lexical/semantic resources. As such, it can be conjectured that more exposure that is longer than a 600 class-hour instruction for six months would lead to sufficient L2 proficiency (apparently higher than intermediate), is needed for the group differences to emerge. Maybe, then the control areas (e.g. basal ganglia) would be employed and trained more and the bilingual advantage would appear as argued by Abutalebi and colleagues. It should be noted that gaining sufficient L2 proficiency results in the disappearance of neural differences between L1 and L2 representation in the brains of adult language learners (Perani & Abutalebi, 2005). This proficiency-dependent explanation would be consistent with the findings of Kerrigan et al. (2017) indicating the cognitive advantage of higher L2 proficiency level on visual-spatial WM task, Bonfieni et al. (2020) showing the additive role of higher proficiency on the control of proactive control mechanisms, and Gurunandan et al. (2019) revealing higher functional connectivity with control areas such as DLPFC and ACC in the advanced L2 learners at a comprehension level.

The PM related explanation for null-effects of bilingual advantage could also be a resolution to the discussion about the plasticity after critical periods (Birdsong, 2018). Some existing research exploring bilingual advantage as of or after puberty (e.g., Hosoda et al., 2013; Mårtensson et al., 2012; Qi et al., 2019; Schlegel et al., 2012) revealed group differences on the basis of neurological findings, indicating the possibility of brain changes after critical periods, yet no group differences on behavioral tasks. In order for behavioral differences to appear, more extensive exposure and practice would be needed to enhance PM related brain areas.

The enhancement on WM capacity and response inhibition of both groups could also be related to experiences that they went through as emerging adults. As argued by Arnett (2000), emerging adulthood, a period of ages 17 to 23, encompasses experiences of a wide range of new environments and social roles. The participants in the present study were the first-year university students and their school setting changed considerably from a high school environment to the university one. They experienced substantial life changes in such different contexts as residence and new social roles, entailing taking on and fulfilling new responsibilities. These experiences could also be an answer to the question why Huang et al. (2020) found enhancement on WM capacity more on the first-year university students as compared to the second-year ones despite having similar amount of foreign language exposure. The explanation about the role of experiences on the enhancement of EFs implicates the DP model (Ullman, 2004, 2006) as well. The experiences the participants of the current research underwent could be construed as episodic knowledge whose storage and acquisition require DM systems as semantic knowledge does.

Although the enhancement of EFs on complex span and antisaccade tasks was found for both groups in the present study, neither group improved their performances on a threshold version of flanker. This result partially runs contrary to the second and fourth hypotheses of the present study, predicting enhancement on inhibitory control abilities of the experimental group and group differences on these abilities, respectively. There are two possibilities to account for this result. First, given the findings that Common EF is perfectly correlated with prepotent response inhibition (Friedman et al., 2008, 2011, 2012), and that a significant, though small, change is only possible in Common EF during emerging adulthood (Friedman et al., 2016), it would make sense to observe an enhancement on antisaccade measuring prepotent response inhibition rather than on flanker measuring resistance to distractor interference. The second possible explanation is that the use of threshold flanker tasks to measure inhibition could have been insufficient to measure the construct adequately. More elaborate discussion about the relationship between prepotent response inhibition and WM capacity together with resistance to distractor interference and the use of a threshold version of flanker will be provided while discussing the findings concerning the third research question and the fifth hypothesis.

5.2 Dimensions of EFs and predictive power of EFs for L2 success

The third question concerning the dimensions of EFs and their predictive value for L2 learning outcomes beyond initial L2 proficiency yielded interesting results.

5.2.1 Dimensions of EFs

With regard to the dimensions of EFs, it was predicted by our fifth hypothesis that antisaccade and flanker performances would form a single component as indicators of inhibition, while visual complex span tasks would constitute the other component in line with the previous research outcomes (e.g., Cieslik et al., 2015; Friedman & Miyake, 2004; Shipstead et al., 2014). However, the PCA results identify that WM capacity and antisaccade performances had a significant loading on the same factor component named as executive attention, while flanker performance was independent from executive attention. This finding was unexpected but not necessarily out of line given two important issues in the field: the differences in the conceptualization of inhibition, and methodological concerns about the inhibition tasks.

First, the low correlation between antisaccade and flanker could be supportive of the conceptualizations of Dempster (1993), Harnishfeger (1995) and Nigg (2000) who argue that inhibition is a multidimensional rather than a unitary construct. More specifically, Nigg (2000) proposes that interference control germane to suppression of interference caused by resource or stimulus competition is conceptually different from behavioral inhibition concerning suppression of prepotent responses. While Friedman & Miyake (2004) named the former as resistance to distractor interference (RDI), they regarded the latter as prepotent response inhibition (PRI). They found that PRI (assessed by such measures as antisaccade) and RDI (assessed by such measures as flanker) fall along a single factor. Arguing against the findings of Miyake and Friedman, the findings of the present study are compatible with the conceptualization of Nigg (2000), and some extant neurological data indicating that PRI and RDI are different (Groom & Cragg, 2015; Xie et al., 2017) in the sense that while the amplitude of N2 is enhanced by RDI, the P2 amplitude by PRI. These findings might call into question the unitary conceptualization of inhibition, specifically the approaches (e.g., selective attention by Diamond, 2013, and attention control by Engle and colleagues) treating PRI and RDI as the components of the same latent variable.

Second, inhibition-related tasks, particularly the flanker task, could not have captured the construct adequately. Friedman and Miyake (2004) attribute the low correlations between inhibition tasks to task-related factors one of which is the calculation of dependent variable. Draheim et al. (2019) argue that the problems associated with measurement are mainly due to a reliance on reaction time (RT) scores, especially RT differences; and they suggest the use of new accuracy-based tasks. Draheim et al. (in review) employing a threshold version of flanker reached improvement in the correlation and defended the unitary concept in relation to inhibition. In the current study, the flanker task that they developed was employed together with antisaccade, namely the benchmark test of inhibition (Hutton & Ettinger, 2006; Rey-Mermet et al., 2018) and one of the most reliable tasks in Engle's lab (Engle, 2018). Since antisaccade has proved itself as psychometrically the most robust tool to measure inhibition, this finding of the current study, then, would be interpreted as evidence calling into question the validity of the use of a threshold version of flanker as well.

It could be suggested that binning procedure revealing 10 to 20 different bin values (Draheim et al., 2016) could be incorporated into the calculation of dependent variable of the threshold version of flanker to reach more interindividual differences since a limited number of thresholds were obtained in the current study. It should be noted, though, that the results would have been different if more than one task with a threshold version (e.g., Stroop) had been employed to measure inhibition, and confirmatory factor analysis (CFA) had been conducted rather than PCA. This would help to deal with the task-impurity problem as stated by Friedman and Miyake (2004).

Additionally, manipulating the nature of flanker tasks (a line task as in the present study or no-line task), Shipstead, Harrison, and Engle (2012) indicated that distractor effects are minimized in the line task on which visual context can guide individuals to devote less attentional resources to processing the distractor. In other words, bottom-up guidance is provided to focus on the target. However, the no-line flanker task eliminating the bottom-up guidance requires top-down control, thereby revealing attention-related differences. This finding is compatible with Dempster (1993) proposing that the classification of inhibition can be further extended by including the location and temporal operating characteristics of interfering stimuli. In short, in order for the individual differences to emerge, the nature of the flanker tasks could be reconsidered together with the calculation of dependent variable.

The finding of a strong correlation between complex span tasks and antisaccade in the present study corroborates the previous research outcomes by Kane et al. (2001) and Unsworth, Schrock, and Engle (2004). They indicated that high-WM participants were found to have significantly better performances in maintaining the goal to look away from the flashing stimulus, in other words, suppressing the automatic orienting response. Since individual differences in complex span tasks can be explained by antisaccade performances, the factor component with significant loadings from complex span tasks and antisaccade was named as executive attention in this present study.

This PCA finding could be argued to be compatible with the argument of Miyake and Friedman (2012) that the functioning of Common EF represents the ability to maintain goal-related information, and the proposal of Friedman and Miyake (2017) that Common EF is similar to executive attention proposed by Engle and colleagues (Engle, 2002; Engle & Kane, 2004). Given the perfect correlation between prepotent response inhibition with Common EF in Friedman and colleagues' studies (e.g., Friedman et al., 2008) and the operationalization of executive attention by Engle and colleague through the use of complex span tasks, the finding of the current study could be asserted to provide evidence for the similarity between Common EF and executive attention. It should be read with caution, though, since antisaccade had a loading of .52, while symmetry and rotation span tasks a loading of .83 on the executive attention component in the current study.

In the model of EFs of Miyake and colleagues, Common EF has a perfect correlation with inhibition, yet updating and switching exhibit unique variance. In parallel with this explanation, within the framework of maintenance and disengagement as a revised view of executive attention, Shipstead et al. (2016) propose that despite having shared variance with complex span tasks, updating measures tap disengagement more. According to this proposal, the dominant functioning of complex span tasks reflects the ability to control attention in a goal-directed manner in the face of contextually inappropriate prepotent actions, but at the same time, it also represents the ability to do cue-dependent search from secondary memory. This could be the reason why antisaccade was not found to be perfectly correlated with complex span tasks in the present study as it was with Common EF. However, it could be argued that the findings would be more supportive of the proposal that Common EF could be similar to executive attention and operationalized by complex span tasks if more than two complex span and prepotent response inhibition tasks had been employed, and the analyses had been conducted through CFA.

The findings germane to the dimensions of EFs could also shed light into the findings for the second research question related to the enhancement of EFs discussed in section 5.1. Given the stronger effect size in the enhancement on antisaccade performance as compared to the one on complex span tasks in the current study, and the possibility of change only on Common EF and the perfect correlation between PRI and Common EF (Friedman et al., 2016), it could be proposed that the enhancement on complex span tasks could have been triggered by the enhancement on attention control. However, this interpretation requires further research with a larger pool of measurement tools in a different research design.

5.2.2 Predictive validity of the dimensions of EFs for L2 success

With respect to the predictive value of the dimensions of EFs for L2 learning outcomes at the end of a six-month intensive L2 instruction beyond the initial L2 proficiency, a hierarchical regression was conducted. The findings revealed that the initial proficiency and baseline measure of flanker significantly accounted for 18.4 and 2.5% of L2 success respectively. However, visual WM tasks and antisaccade performances both individually and as one factor component named executive attention had no correlation with L2 success. It should be noted that executive attention had a significant, though weak, correlation with the initial proficiency level.

5.2.2.1 Predictive value of WM for L2 learning outcomes

These findings of the third research question pertaining to the predictive power of WM for L2 success run contrary to the sixth hypothesis predicting that baseline WM capacity would be a robust predictor of L2 success at the end of a six-month intensive L2 instruction. This finding might be thought not to reconcile with the literature implicating that greater domain-general WM resources could lead to better higher-order cognitive abilities (Engle & Kane, 2004), more specifically better L2 performance (Martin et al., 2019). As such, it seems to cast doubt upon the domain-general predictive power of executive attention. However, there are several possibilities to account for this result. First, individual differences in WM capacity could have been compensated during L2 instruction that took place in a natural university classroom context for six months. In order for the individual differences in WM to appear, learners should be in conditions where they need to maintain task goals in primary memory (PM), and retrieve goal-related information from the secondary memory (SM). This information is displaced from PM to SM due to irrelevant stimuli. (Unsworth & Engle, 2007). This could be more possible in situations where laboratory learning tasks and/or manipulation of instruction in a single session are employed since they would place different demands on learning and memory functions, thereby increasing the load as compared to the context of the present study. In these laboratory conditions, the buildup of PM could easily be manipulated, thus retrieval from SM would be required as well, and high WM learners can outperform low WM ones. As in the case of intensive L2 instruction for a longer period, however, learners could benefit from repetitions provided in the classroom, online materials available outside the classrooms, homework assigned in relation to the skills covered in the classroom. In other words, learners would have the opportunity to compensate for their low WM abilities. The findings of the current research indicating that complex span tasks had a significant correlation with initial L2 proficiency, yet were not significantly correlated with L2 learning outcomes sound to be in tune with this explanation.

This instruction-related explanation would corroborate the findings of Sanz et al. (2016) indicating that exercising traditional pedagogical approach in which grammar instruction was explicitly given prior to practice decreases the chances for high WM spans to outperform their counterparts; and Suziki and DeKeyser (2017) demonstrating that WM effects appear in mass instruction rather than spaced or distributed instruction. Suziki and DeKeyser contend that presenting similar vocabulary and grammar rules in a short period of time would increase the load in WM, thereby causing more interference that low WM spans could not deal with. Finally, given the significant, albeit low, correlation between executive attention and L2 proficiency at the beginning of the education in the present study, one is tempted to say that instruction with an extended period of time neutralizes the effect of WM on L2 success.

The explanation germane to the nature of instruction would be in tune with extant research implicating the role of task demands on WM effects across language skills. With respect to vocabulary and grammar, for instance, WM effects appear when learners are involved in the tasks requiring active participation (Martin & Ellis, 2012), and the load was increased by giving within-task feedback (Li et al., 2019). Additionally, with regard to L2 reading skills, inferential reading rather than literal reading (Alptekin & Erçetin, 2009), story comprehension rather than dialogue comprehension (Andersson, 2010) could reveal WM effects. Finally, as for speaking skills, speaking through within-task planning rather than speaking through strategic planning prior to task performance (Ahmadian, 2012) and storytelling on the basis of pre-sequenced pictures rather than in unstructured task (Kormos & Trebits, 2012) are the situations where high WM spans could outperform low WM spans. In sum, the effects of WM constraints could become more manifest when participants perform tasks imposing heavy demands on WM.

The second explanation for the nonsignificant effect of WM capacity on L2 learning outcomes is that complex span tasks would not able to capture disengagement adequately (Shipstead et al., 2016), and fluid measures including tasks of fluid intelligence or updating WM would be needed to explain robust explanation of individual differences in higher-order cognition abilities including language (Christopher et al., 2012; Martin et al., 2019; Was et al., 2011). Consistent with the definition of updating WM by Miyake et al. (2000), disengagement functions to reduce inappropriate information by ignoring irrelevant information and/or suppressing no longer relevant information. Likewise, Engle (2018) suggests that complex span tasks reflect the ability to maintain information in the face of divergent thought, while fluid intelligence tasks concern the ability to think of something when it is important, yet disengage or unbind it when it is not (i.e., functionally forget it). In sum, complex span tasks would not tap disengagement or retrieval processes sufficiently. As argued by Roediger (2000), despite receiving inadequate attention, retrieval processes could be the most critical elements to understand how memory works, echoing the proposition of Serafini and Sanz (2016) calling attention to the role of retrieval functions in L2 learning. Finally, the finding of Cheng et al. (2019) indicating a significant effect of n-back, an updating WM task, on longitudinal L2 success is in accordance with the explanation concerning the role of disengagement.

The third explanation of the non-significant WM effect on L2 learning outcomes has to do with the code-specific or modality specific hypothesis. It posits that entities belonging to such global categories as verbal and spatial representations are stored in and retrieved from separate memory areas in the brain through functionally and structurally distinct storage and processing resources specific to verbal and visual ones (Rösler & Heil, 2003). This could also explain why the findings of the present study adopting domain-general perspective and employing visual complex span tasks are incongruent with those of existing longitudinal L2 research (e.g., Chang et al., 2019; Kormos & Sáfár, 2008; Linck & Weiss, 2011, 2015; Sagarra, 2017, Serafini & Sanz, 2016) adopting a domain-specific one and using verbal STM or verbal complex span tasks.

The verbal STM tasks were non-word repetition and backward digit span (Kormos & Sáfár, 2008), operation span (Linck & Weiss, 2011, 2015; Serafini & Sanz, 2016), and reading span tasks (Chang et al., 2019; Sagarra, 2017). Engle, Tuholski et al. (1999b) argue that the shared variance between verbal STM measures and verbal complex span tasks encompasses domain-specific storage for verbal memory representations and associated rehearsal procedures and strategies, which could explain the reason why the verbal nature of the WM measures employed in longitudinal L2 research account for the differences in L2 learning outcomes. The role of strategies was proved by Turley-Ames and Whitfield (2003) indicating that improving encoding strategies can facilitate low span learners and obscure the correlation between operation span task and reading performance. The explanations based on the domain-specific perspective are also congruent with the findings of Martinez and Singleton (2019), indicating that intensive L2 experience puts bilinguals at an advantage in terms of phonological STM, and gain them better elaborative rehearsal strategies especially learning experience in a more explicit fashion. Finally, consistent with the domain-specific perspective, the review of Linck et al. (2014) indicates that effects of verbal WM tasks yielded higher effect-size.

It should be noted that higher WM capacity can be a cause or result of L2 proficiency (Martinez & Singleton, 2019), implicating that initial proficiency can be a confounding variable. In none-of the longitudinal L2 research mentioned above, however, proficiency tasks were not employed as pre-test measures. To obtain more rigorous results about the predictive validity of the domain-specific measures of

WM, it would be better to employ both proficiency and WM tasks as both pre-tests and post-tests measures.

With respect to the role of initial L2 proficiency in L2 success, the results of the present study indicated initial L2 proficiency accounted for 18.4% of L2 learning outcomes, which could be explained through the role of LTM. Preexisting knowledge already stored in LTM influences the perception of the incoming information and provides memory structures and representations with which this information can be integrated, thereby enhancing encoding - that is, one of the main processes of LTM and a prerequisite for subsequent memory processing (Dehn, 2015). The learners with less or no L2 background would need to form memory representations through new synaptic connections among relevant neurons. Yet, for the ones with better L2 background, input would activate the existing synaptic connections for the already stored engrams, thereby shortening the time needed to consolidate the representations. Retrieval and association over extended period of time would strengthen the connections between the components of declarative memory (McClelland, McNaughton, & O'Reilly, 1995; Paller, 2003).

This perspective opens the door to the possibility that long-term working memory (LT-WM) by Ericsson and Kintsch (1995) would play a significant role in L2 processing and comprehension. Accordingly, LTM functions as an efficient extension of WM, grouping items into chunks and associate the chunks with familiar patterns already stored in LTM. This is facilitated by expertise and strategy use in the domain. As for comprehension, individuals construct a representation in LTM first and expand it by integrating new information with relevant parts remaining accessible through retrieval structures (Kintsch, 1998, 2000).

Additionally, the predictive role of initial L2 proficiency level echoes the arguments of Ryskin, Levy, and Fedorenko (2020) and the findings of Shain, Blank, van Schijndel, Schuler, and Fedorenko (2020) about linguistic prediction as well, and can contribute to the debate concerning the effects of language experience versus executive resources on linguistic prediction. Ryskin et al. highlight that prior linguistic experience has a significant effect on the ability to make linguistic predictions about the incoming input, thereby affecting processing and comprehension. Additionally, Shain et al. revealed that linguistic predictions about upcoming words using cognitive processes are primarily carried out by the language network. Nevertheless, this should be read with caution given the overlap between the language network and domain-general circuits and the role of intensive instruction possibly compensating the low EF abilities of L2 learners. As argued by Fedorenko (2014), domain-general circuits are employed, yet it remains a question how often they are engaged and how theoretically significant it is.

5.2.2.2 Predictive power of inhibition for L2 learning outcomes

It was predicted by the seventh hypothesis that the more efficient inhibitory control system learners have, the higher the proficiency level they reach. However, this is not the case found in the present study revealing null effects of antisaccade performances on L2 learning outcomes, yet a significant, albeit low, effect of flanker.

The difference between the effects of inhibition tasks on L2 success draws attention to the conceptual differences between PRI and RDI and their functions as mentioned in section 5.2.1. While PRI deals with prepotent response inhibition, RDI the suppression of interference caused by resource or stimulus competition. One possible explanation is that that the prepotent interference resulting from L1 is related to the interference within learners' mind, yet flanker could be more related to the interference caused by stimulus outside. Apparently, learning an L2 in a classroom context might require learners to use attention to suppress interference in learning conditions not prepotent responses. In other words, in L2 learning, dealing with external interference would be argued to be more effective than internal interference. In this sense, it could also be argued that the flanker task could tap the ability to suppress irrelevant information as disengagement does while antisaccade is more related to goal-maintenance. However, it is a noteworthy possibility that each distractor in the classroom or outside the class would yield prepotent responses for the learners as well. For this reason, task-specific non-EF abilities would be argued to be related to L2 learning. These speculations call for further research with a larger pool of inhibition tasks.

This result would be inconsistent with the literature implicating that suppression of L1 activations might have a significant exploratory role in online L2 processing (e.g., Linck et al., 2008). One potential explanation for the null results is that the role of PRI would be significant in fine-grained online L2 processing outcomes but not in coarser measures of L2 proficiency or more importantly not during a long and intensive L2 instruction. Learners may have employed other mechanisms or strategies to compensate for a deficiency in cognitive processing. To support multifaceted, complex task of language learning, learners with lower PRI abilities rely more heavily on other strategies or cognitive resources (Linck & Weiss, 2015). As a result of this, facilitative effect of inhibition disappears. Another potential explanation of null PRI results is that obtaining proficiency from elementary to intermediate level might rely on L1 to aid L2 processing so heavily that L1 inhibition was not needed much (Linck & Weiss, 2015). Having better PRI abilities may predict L2 learning outcomes at more advanced proficiency levels.

It is worthwhile to note that a behavioral approach is adopted in the present study and the finding as to the predictive value of PRI measured by antisaccade would provide strong evidence for existing research outcomes by Linck and Weiss (2011, 2015), revealing no correlation between the Simon task (measuring PRI) and L2 learning outcomes, and answer their questions as to whether inhibition would emerge as a robust predictor in more intensive learning contexts. This finding is also consistent with Ghaffarvand Mokari and Werner (2019) revealing no relationship between the Stroop and phonological L2 performance after a 5 hour-phonetic training. Obtaining the null effects of response inhibition through the use of antisaccade which can deal with the methodological concerns related to the Simon and Stroop tasks (Paap et al., 2019) contributes to the validity of the findings of null-effects as well.

5.3 Conclusion

The current research aims to fill the gap in the literature of bilingual advantage and individual differences in EFs, stemming from the little amount of research examining the relationship between EFs and L2 experience particularly as of/after puberty, the inconclusive findings and methodological constraints of the extant behavioral studies, and the need to clarify the relationship between EFs and L2 learning outcomes. To this end, 165 Turkish high-school graduates learning English as an L2 in the experimental group in an instructed L2 context and 103 Turkish freshman students in the control group were recruited, and the longitudinal data were collected through the administration of symmetry and rotation span tasks to measure WM capacity, antisaccade and a threshold version of flanker to measure inhibition, and a standardized L2 proficiency test in a pre-test/post-test design.

The findings indicated an enhancement on EFs of both groups on WM and antisaccade performances, yet no enhancement on flanker performances was observed for either group. It was also found that WM capacity and antisaccade performances had loadings on the same factor, named as executive attention, while flanker performance was independent from executive attention. Finally, as for the predictive validity of EFs for L2 learning outcomes, EFs had no significant relationship with L2 success, except for flanker performances. Prior L2 experience was found to be the best to predict L2 success significantly, though. The role of inhibitory control (IC) theory by Green (1998) and declarative/procedural model by Ullman (2001, 2004, 2006) were consulted to explain the results pertaining to enhancement of EFs. The domain-specific rather than domain-general perspective for the role of EF-related individual differences in L2 success, and multidimensional rather than unitary perspective of inhibition seem to be robust to explain the other findings. As regards the relationship between EFs and L2 experience, from the point of domain-general perspective, it could be concluded that L2 experience can contribute to the enhancement of EFs, yet EFs might not play a role in L2 success at the end of a six-month intensive L2 instruction.

5.4 Implications

The findings of the present study have both conceptual and methodological, and pedagogical implications.

With respect to conceptual and methodological implications, first, in order for bilingual advantage to appear, L2 learners should take more intensive and/or longer instruction, given the current investigation period (600 class contact hours in six months), so that they can reach advanced levels of proficiency and train inhibitory control abilities more (Driemeyer et al., 2008). In the same vein, DP model by Ullman (2001, 2004, 2006) should be considered in order to understand the processes that language learners go through. Second, conceptual differences should be taken into consideration before employing the inhibition tasks. The low correlation among these tasks (Miyake & Friedman, 2004) could also be related to conceptual differences as well as psychometric issues. Further, as regards the inhibition tasks criticized by Paap et al. (2020a), the manipulation of the nature should be integrated into the research exploring a more effective way to calculate dependent variable (e.g., Draheim et al., in review). Third, in light of the findings arising from the majority of L2 and those of the current research, domain-specific rather than domain-general perspective seems to function better to explain the role of individual differences in employing cognitive resources in L2 learning outcomes. Nevertheless, it is worthwhile to note that adopting the framework of maintenance and disengagement (Shipstead et al., 2016) would provide a more sophisticated picture of the role of domain general perspective in L2 success. Fourth, since L2 experience can lead to enhancement on EFs, the predictive validity of the EFs measures in a post-test only design could be misinterpreted. Last, given the significant role of prior L2 experience in L2 success, pre-test measures of proficiency should be administered. If possible, learners with no L2 background should be recruited as participants.

The findings of the current research have practical implications for applied linguistics and language pedagogy. First, the findings suggest that the DP model has the potential to elucidate the L2 learning processes. As such, techniques enhancing learning and memory in the declarative and/or procedural memory systems could be integrated into the L2 learning curricula. As argued by Ullman and Lovelett (2018), item-level approaches targeting specific items or skills through such techniques as spaced repetition, retrieval practice, deep encoding, the enactment effect, and the method of loci, and learner-level approaches stressing the role of sleep, exercise, diet, and mindfulness can be taken into consideration. Second, in line with DP model, the length of instruction and the amount of practice provided for learners should be redesigned to facilitate automatic processes in the brain. Third, the findings of the present study indicate that prior L2 experience can explain the L2 success considerably more significantly than EFs can. Hence, it stands to reason that the similarity in the way to treat learners with little prior exposure (i.e., absolute beginners) and those with relatively higher proficiency (i.e., elementary learners) can result in numerous challenges for both groups of learners. Therefore, syllabus design, pedagogical sequencing, development of materials, classroom practice and/or assessment should be evaluated and designed accordingly, considering the presence of learners with mixed-English proficiency. For example, types of roles that low proficiency learners take on can influence their engagement and interactions in carrying out communicative tasks (Dao & McDonough, 2017) or language learning strategies they apply can influence their learning processes (Fewell, 2010).

5.5 Limitations and further research

Limitations of the present study together with suggestions for further research are worth mentioning.

First, the current research was based on the framework of the executive attention and the results pertaining to WM capacity are limited to the findings arising from complex span tasks. Adopting the framework of maintenance and disengagement (Shipstead et al., 2016) and the use of fluid measures such as Gf and updating WM tasks together with complex span tasks would yield a more sophisticated picture of the extent to which the relationship between the cognitive capacity and individual differences in higher-order cognitive abilities is (Engle, 2018; Martin et al., 2019). Therefore, further research should endeavor to explore the individual differences in L2 learning outcomes in a classroom environment in light of this framework. Furthermore, employing more than two measurement tools for each construct and conducting CFA to explore the dimensions of EFs could deal with the task-impurity problem (Miyake & Friedman, 2004). Accordingly, structural equation model could be performed to look into the predictive power of EFs for L2 success.

Second, the adoption of a domain-general perspective and the employment of visuals tasks in the current study make it highly difficult to compare the findings of this present research with those of previous L2 research exploring the role of EFs in L2 success since the majority of the research was based on the domain-specific perspective and the use of tasks, verbal in nature. Futhermore, given the generally held view that WM is not limited to the executive attention or central executive and it also encompasses domain specific storage with associated rehearsal procedures and strategies (Baddeley & Hitch, 1974, Cowan, 1998), further research can include a larger pool of verbal and visual, complex and simple span tasks to capture the construct more adequately. Hence, it would be possible in due course to test the hypotheses of the current research from a domain-specific perspective as well.

Third, our results are not universally applicable rather specific to L1 Turkish learners of L2 English studying in an instructed setting in one of the foundation universities (73 in total) in Turkey. Therefore, caution should be exercised in generalizing the results to young adult learners of English in other settings and with different degrees of L2 proficiency since each language school has its own teaching philosophy, curriculum and institutional goals. It would thus be interesting to replicate the study with L2 learners starting as elementary and reaching advanced levels of L2 proficiency at different institutions. Thus, the condition of providing an intensive instruction over six months or a year could be fulfilled as well (Serafini & Sanz, 2016). Such studies could shed light on whether bilingual advantages emerge when L2 is proceduralized or automatized. They could further show whether the degrees of proficiency affect the predictive power of EFs in L2 success if the proficiency measure was administered at the end of each level.

Fourth, it was very unlikely for the improvement on the post-test for WM and inhibition to be due to practice effect of implementing the same tasks twice since there was a six-month time lapse between two administrations in the present study. Nevertheless, the hypotheses of the current research could be tested in a further study through Solomon four-group design so as to rule out the possibility of the practice effect (Salthouse & Tucker-Drob, 2008).

Last but not least, this study focused exclusively on cognitive variables to explain the individual differences in L2 success. Yet, to construct a solid framework for L2 learning process, both cognitive and affective variables and their interactions could be taken into consideration as posited by the dynamic systems theory (Serafini, 2017). Thus, further research could be suggested to determine the extent to which affective and cognitive variables interact with each other in predicting L2 success.

APPENDIX A

BACKGROUND QUESTIONNAIRE

Dear participants,

This questionnaire consists of questions concerning your personal information, educational and language background. Your responses will be employed only for the purposes of the present study and kept confidential.

Thank you for your contributions.

Mehmet AKINCI
Assessment Coordinator
Preparatory School İstanbul Şehir University

Contact Information

Name: ________________________ Date: ____________________

Please answer the questions below accurately.

PART-A PERSONAL INFORMATION

1. Birth year: ___________
2. Gender: Female / Male
3. Nationality: _____________
4. Mother's level of education: __________________
5. Father's level of education: __________________
6. Mother tongue: _____________________
   Were you raised in a setting where a different language rather than Turkish was used?
   ______NO
   ______YES, if yes, it is impossible for you to participate in this study. Please inform the researcher and DO NOT CONTINUE.

PART B EDUCATIONAL BACKGROUND

7. Your department/academic program at the university_____________
   Your ranking on the university entrance exam _____________

8. Have you ever taken musical education?
   ____ NO
   ____ YES, if yes, the initial age of taking the education: _______;
   the length: ________

PART-C FOREIGN LANGUAGE BACKGROUND

9. Did you take any foreign language education before the university?
   _____ NO
   _____ YES, if yes, which language(s) ______________________
   the initial age(s) of starting the education ________

10. Have you ever been abroad where the foreign language(s) is (are) spoken?
    ____ NO
    ____ YES, if yes, which country/countries __________
    the length of stay ____________

APPENDIX B

BACKGROUND QUESTIONNAIRE (TURKISH)

Değerli katılımcı,

Bu anket, kişisel bilgileriniz, genel eğitim geçmişiniz ve yabancı dil öğrenim geçmişiniz ile ilgili sorular içermektedir. Vereceğiniz bilgiler sadece araştırmanın kapsamında kullanılacak olup, gizli tutulacaktır.

Katkılarınızdan dolayı teşekkür ederim.

Mehmet AKINCI
Ölçme ve Değerlendirme Koordinatörü
Hazırlık Okulu İstanbul Şehir Üniversitesi

İletişim Bilgileri

Ad ve Soyad: ________________________ Tarih: _____________________

Aşağıdaki soruları en doğru şekilde cevaplayınız.

PART-A KİŞİSEL BİLGİLER

1. Doğum yılınız: ___________
2. Cinsiyetiniz: Kadın / Erkek
3. Uyruğunuz: _____________
4. Annenizin eğitim durumu: ___________________
5. Babanınızın eğitim durumu: ___________________
6. Ana diliniz: ______________________
   Türkçe'den başka bir dil konuşulan bir ortamda büyüdünüz mü?
   ______HAYIR
   ______EVET ise; bu araştırmaya katılımınız mümkün olmamaktadır. Lütfen araştırmacıyı bilgilendiriniz ve ankete DEVAM ETMEYİNİZ.

PART B GENEL EĞİTİM GEÇMİŞİ

7. Üniversitedeki bölümünüz: _____________
   Üniversite sınavındaki sıralamanız: _____________

8. Hiç müzik eğitimi aldınız mı?
   ____ HAYIR
   ____ EVET ise, eğitime başladığınız yaş: _______;
   eğitim aldığınız toplam süre: ________

PART-C YABANCI DİL ÖĞRENİMİ GEÇMİŞİ

9. Üniversiteden önce hiç yabancı dil eğitimi aldınız mı?

   _____ HAYIR
   _____ EVET ise, hangi dil ya da diller ________________
   başladığınız yaş ya da yaşlar: __________

10. Yukarıda belirttiğiniz dil ya da dillerin konuşulduğu ülkelerde bulundunuz mu?
    ____ HAYIR
    ____ EVET ise, kaldığınız yer ya da yerler __________
    kaldığınız süre ____________

APPENDIX C

ETHICS COMMITTEE APPROVAL

İSTANBUL
ŞEHİR
ÜNİVERSİTESİ

ARAŞTIRMA ETİK KURUL KARARLARI

Toplantı Tarihi : 23.02.2018
Toplantı Sayısı : 04/2018
Toplantı Saati : 11:00
Toplantıya Katılanlar : Prof. Dr. Hatice AYNUR (Başkan)
Prof. Dr. Nihat BULUT
Prof. Dr. Cem BEHAR
Doç. Dr. Eda YÜCESOY
Yrd. Doç. Dr. Sinem ELKATİP HATİPOĞLU
Yrd. Doç. Dr. Betül NİZAM
Yrd. Doç. Dr. Eyyüp Said KAYA
Yrd. Doç. Dr. Hızır Murat KÖSE

Karar No : 1

İstanbul Şehir Üniversitesi Araştırma Etik Kurulu, proje yürütücüsü Prof. Dr. Gülcan Erçetin ve proje araştırmacısı Mehmet Akıncı tarafından sunulan 'İkinci Dil Eğitiminin Yönetici İşlevlerin Gelişmesine Etkisi'' İsimli proje taslağını değerlendirmiştir.

Verilen bilgilere göre; Bu proje, üniversite düzeyinde alınan yoğun ikinci dil eğitiminin öğrencilerin zihinsel yönetici işlevleri üzerindeki etkisini araştırmayı hedeflemektedir. İkinci dil öğrenimi ve deneyimi ile yönetici işlevlerin ilişkisini araştıran çalışmalar genellikle yönetici işlevlerin iki ana parçası olan ketleme denetimi ve işler bellek üzerinde odaklanmıştır. Bu çalışmalar, ikinci dili erken yaşlarda öğrenmiş ve hayat boyu ikinci dil deneyimine sahip çocuk ve yaşlılarda yönetici işlevlerin tek dilli akranlarına göre daha iyi olduğunu gösterirken, genç yetişkinler ile yapılan çalışmalar kesin olmayan bulgular ortaya koymuştur. Genç yetişkinlerdeki çelişkili bulgular, kullanılan ölçme araçlarına ve bağımlı değişken hesaplamasına bağlı faktörlerden kaynaklanabilir. Örneğin, bireyler arası yeterli varyasyon olmaması, iki dilliliğin kategorik bir değişken olarak ele alınması, ve yönetici işlevlerin hatalı kavramsallaştırılması bu faktörler arasında sayılabilir. İkinci dili geç yaşta öğrenen bireyler üzerinde yapılan nörolojik çalışmaların, kritik dönem varsayımının aksine, beyinde dil öğrenimi deneyimine dayalı değişiklikler tespit etmesine rağmen çok az çalışmada yönetici işlevlerdeki davranış değişikliği incelenmiştir. Az sayıdaki mevcut çalışmalarda ise ketleme ve işler bellek işlevlerini ölçmek için tek bir araç Aşağıda isimleri ve imzaları bulunan İstanbul Şehir Üniversitesi Araştırma Etik Kurulu üyeleri, araştırmacı tarafından kurula sunulan yukarıdaki bilgiler ışığında, belirtilen araştırmanın yürütülmesinde etik açıdan bir sakınca görmemektedir.

H.A.

kullanılmış, az sayıda katılımcıdan veri toplanmış ve yoğun olmayan kısa süreli eğitimler sunulmuştur. Bu projede, bu tür yöntemsel sorunları giderecek bir desen uygulanacaktır. Buna göre, çalışmanın deney ve kontrol grubunda en az 100 kişiden veri toplanacak, deney grubu en az 6 aylık bir yoğun dil eğitimi alacak, dil yetisi standart bir dil testi ile ölçülecek, ketleme ve işler bellek işlevlerinin her biri için iki ayrı ölçme aracı kullanılacaktır.

Projenin gerekçesi evrensel ve yerel katkıları göz önüne alınarak aşağıda sıralanmıştır.

Projenin evrensel katkıları:

a) Yoğun ikinci dil öğreniminin ve deneyiminin yönetici işlevler üzerindeki etkisini araştıran çalışmalar yok denecek kadar azdır. Çalışma bu boşluğun doldurulmasına katkıda bulunacaktır.

b) Mevcut araştırmalar ya ketleme denetimi veya işler bellek üzerine odaklanmış ve yöntemlerindeki sorunlar nedeniyle geçerli ve güvenilir bulgular sunma konusunda zayıf kalmışlardır. Bu çalışma, yoğun ikinci dil öğreniminin ve deneyiminin yönetici işlevlerin iki ana parçası olan hem ketleme denetimi hem de işler bellek üzerindeki etkisini araştıran; değişkenleri birden fazla araç kullanarak ölçen ve örneklem sayısını en az 100 kişi olarak hedefleyen güçlü bir yönteme sahip ilk çalışma olacaktır.

Projenin yerel katkıları:

a) Bu çalışmayla, ketleme denetimi ve işler bellek için geçerliliği ve güvenilirliği test edilmiş ölçme araçlarının Türkçe versiyonları hazırlanarak benzer çalışmaların Türkiye'de yapılmasının önü açılmış olacaktır.

b) Bildiğimiz kadarıyla, Türkiye'de ikinci dil öğretiminin yönetici işlevler üzerindeki etkisini inceleyen bir çalışma yapılmamıştır. Bu çalışma, üniversitede alınan yoğun İngilizce eğitiminin bilişsel süreçlerin gelişimine katkısını inceleyen ilk çalışma olacaktır.

Bu çalışma üç aşamada gerçekleşecektir. Birinci aşamada, testlerin Türkçe versiyonları hazırlanarak pilot çalışmaları yapılacaktır. İkinci aşamada, eğitim başlamadan önce katılımcılara ikinci dilde yeti testi, ketleme ve işler bellek ön-testleri verilecektir. Üçüncü aşamada, birinci yılın sonunda eğitim öncesi verilen testler son-test olarak tekrar uygulanacaktır.

I. Pilot Çalışma

Ketleme ve işler bellek testleri, Georgia Teknoloji Enstitüsü'ndeki Dikkat ve İşler Bellek Laboratuvarı'nda Prof. Randall Engle'ın başkanlığında geliştirilmiştir. Bu çalışmada 4 ayrı test kullanılacaktır (http://englelab.gatech.edu/index.html): Antisakkad Testi, Flanker Testi, Simetri Testi, Rotasyon Testi. Bu testlerin yönergeleri İngilizce olduğundan yönergeler Türkçeye çevirilecek ve E-prime yazılımı ile Türkçe versiyonları pilot çalışma için hazır hale getirilecektir. Testlerin pilot çalışması, İstanbul Şehir Üniversitesi Hazırlık Okulu'nda yoğun İngilizce eğitimi alan ve üniversitenin Türkçe programlarında birinci sınıfa kayıtlı olan 30 kişilik öğrenci grubuna bilgisayarda uygulanacaktır. Pilot çalışmada yönergelerin anlaşılırlığı ve maddelerin güvenilirliği incelenecektir.

H.A.

II. Eğitim öncesi testlerin uygulanması
Deney grubu olarak, İstanbul Şehir Üniversitesi'nin eğitim dili İngilizce olan programlarına katılmaya hak kazanmış öğrencilerinden İngilizce yeterlilik sınavını geçememiş ve başlangıç seviyesinde eğitim almak üzere Hazırlık okuluna kaydolmuş ve de pilot çalışmada yer almamış en az 100 kişilik bir gruba, akademik yılın başında bir adet standart İngilizce yeterlilik testi, iki adet ketleme testi ve iki adet işler bellek testi uygulanacaktır. Karşılaştırma grubu olarak, aynı üniversitenin eğitim dili Türkçe olan programlarına katılmaya hak kazanmış, yoğun İngilizce eğitimi almayacak olan, birinci sınıfa kaydolmuş ve pilot çalışmada yer almamış olan en az 100 kişilik bir gruba da eş zamanlı olarak aynı testler uygulanacaktır.

III. Birinci yıl eğitimi sonrası testlerin uygulanması
Aynı katılımcılara, birinci akademik eğitim yılının sonunda, eğitim öncesi verilen testlerin aynısı uygulanacaktır. Bu uygulama sürecinde, deney grubundaki katılımcılar, yoğun İngilizce eğitimlerini tamamlamış, akademik programlarında eğitim almaya hak kazanmış olacaklardır. Karşılaştırma grubundaki öğrenciler Türkçe eğitim gördükleri birinci sınıf derslerini tamamlamış olacaklardır.

Çalışmaya katılacak insanlarda aranacak özellikler:
Çalışma, ana dili Türkçe ve İngilizce yeterliliği başlangıç seviyesinde olan, başka bir dil öğrenme geçmişi olmayan, üniversitenin eğitim dili İngilizce ya da Türkçe olan programlarına kabul alan katılımcılarla yapılacaktır.

Aşağıda isimleri ve imzaları bulunan İstanbul Şehir Üniversitesi Araştırma Etik Kurulu üyeleri, araştırmacı tarafından kurula sunulan yukarıdaki bilgiler ışığında, belirtilen araştırmanın yürütülmesinde etik açıdan bir sakınca görmemektedir.

Prof. Dr. Hatice AYNUR
President

Prof. Dr. Nihat BULUT
Member

Prof. Dr. Cem BEHAR
Member

Doç. Dr. Eda YÜCESOY
Member

Yrd. Doç. Dr. Hızır Murat KÖSE
Member

Yrd. Doç. Dr. Betül NİZAM
Member

Yrd. Doç. Dr. Sinem ELKATİP HATİPOĞLU
Member

Yrd. Doç. Dr. Eyyüp Said KAYA
Member

REFERENCES

Abu-Rabia, S. (2003). The influence of working memory on reading and creative writing processes in a second language. *Educational Psychology*, *23*(2), 209-222.

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466-478.

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R, Cappa S. F., & Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex, 22*(9), 2076-2086.

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242-275.

Ahmadian, M. J. (2012). The relationship between working memory capacity and L2 oral performance under task-based careful online planning condition. *TESOL Quarterly*, *46*(1), 165-175.

Alptekin, C., & Erçetin, G. (2009). Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and reading span task make a difference? *System*, *37*(4), 627-639.

Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, *33*(2), 206-219.

Andersson, U. (2010). The contribution of working memory capacity to foreign language comprehension in children. *Memory*, *18*(4), 458-472.

Antón, E., Carreiras, M., & Duñabeitia, J. A. (2019). The impact of bilingualism on executive functions and working memory in young adults. *PloS ONE*, *14*(2), 1-30.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469-480.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York, NY: Academic Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47-89.

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind, Brain, and Education*, *10*(1), 10-33.

Bartlett, M. S. (1954). A note on the multiplying factors for various χ2 approximations. *Journal of the Royal Statistical Society*, *16*(2), 296-298.

Bialystok, E. (2016). The signal and the noise: Finding the pattern in human behavior. *Linguistic Approaches to Bilingualism, 6*(5), 517-534.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin, 143*(3), 233-262.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*(3), 89-129.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*(2), 290-303.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240-250.

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(6), 1341-1354.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *The International Journal of Bilingualism, 9*(1), 103-119.

Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology*, *9*, 1-17.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition, 118*(2), 245-257.

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2020). Cognitive control in bilinguals: Effects of language experience and individual variability. *Bilingualism: Language and Cognition*, *23*(1), 219-230.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34*(4), 263-272.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition, 130*(3), 278-288.

Calvo, N., Ibáñez, A., & García, A. M. (2016). The impact of bilingualism on working memory: A null effect on the whole may not be so on the parts. *Frontiers in Psychology*, *7*, 1-4.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245-276.

Chang, X., Wang, P., Cai, M. M., & Wang, M. (2019). The predictive power of working memory on Chinese middle school students' English reading comprehension. *Reading & Writing Quarterly*, *35*(5), 458-472.

Chevalier, N., & Clark, C. A. (2018). Executive function in early and middle childhood. In S. A. Wiebe & J. Karbach (Eds.), *Executive function: Development across the life span* (pp. 43-57). New York, NY: Routledge.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., … Olson, R. K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology: General*, *141*(3), 470-488.

Chung-Fat-Yim, A., Sorge, G. B., & Bialystok, E. (2017). The relationship between bilingualism and selective attention in young adults: Evidence from an ambiguous figures task. *The Quarterly Journal of Experimental Psychology*, *70*(3), 366-372.

Cieslik, E. C., Mueller, V. I., Eickhoff, C. R., Langner, R., & Eickhoff, S. B. (2015). Three key regions for supervisory attentional control: Evidence from neuroimaging meta-analyses. *Neuroscience & Biobehavioral Reviews*, *48*, 22-34.

Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354-373.

Cowan, N. (1998). *Attention and memory: An integrated framework*. New York, NY: Oxford University Press.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42-100.

Crone, E. A., Peters, S., & Steinbeis, N. (2018). Executive function development in adolescence. In S. A. Wiebe & J. Karbach (Eds.), *Executive function: Development across the life span* (pp. 58-72). New York, NY: Routledge.

Cuevas, K., & Bell, M. A. (2014). Infant attention and early childhood executive function. *Child Development*, *85*(2), 397-404.

Cuevas, K., Rajan, V., & Bryant, L. J. (2018). Emergence of executive function in infancy. In S. A. Wiebe & J. Karbach (Eds.), *Executive function: Development across the life span* (pp. 25-42). New York, NY: Routledge.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433.

Dao, P., & McDonough, K. (2017). The effect of task role on Vietnamese EFL learners' collaboration in mixed proficiency dyads. *System*, *65*, 15-24.

Dehn, M. J. (2015). *Essentials of working memory assessment and intervention.* Hoboken, NJ: John Wiley & Sons, Inc.

Della Salla, S., Gray, C., Spinnler, H., & Trivelli, C. (1998). Frontal lobe functioning in man: The riddle revisited. *Archives of Clinical Neuropsychology*, *13*(8), 663-682.

Dempster, F. N. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. In M. L. Howe & R. Pasnak (Eds.), *Emerging themes in cognitive development* (pp. 3-27). New York, NY: Springer.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. Craig (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70-95). New York, NY: Oxford University Press.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168.

Dong, Y., Liu, Y., & Cai, R. (2018). How does consecutive interpreting training influence working memory: A longitudinal study of potential links between the two. *Frontiers in Psychology*, *9*, 1-12.

Draheim, C., Harrison, T. L., Embretson, S. E., & Engle, R. W. (2017). What item response theory can tell us about the complex span tasks? *Psychological Assessment*, *30*(1), 116-129.

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example. *Perspectives on Psychological Science*, *11*(1), 133-155.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508-535.

Draheim, C., Tsukahara, J. S., Martin, J. D., & Engle, R. W. (in review). Attention control: Its measurement and nature. Manuscript submitted for publication in *Journal of Experimental Psychology: General*.

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., & May, A. (2008). Changes in gray matter induced by learning-revisited. *PloS ONE*, *3*(7), 1-5.

Engel de Abreu, P. M., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, *104*(4), 974-986.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19-23.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*(2), 190-193.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 972-992.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation*, *44*, 145-200.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999a). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 102-134). Cambridge, England: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999b). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309-331.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211-245.

Eslinger, P. J. (1996). Conceptualizing, describing, and measuring components of executive function: A summary. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 367-395). Baltimore, MD: Paul H Brookes Publishing.

Fedorenko, E. (2014). The role of domain-general cognitive control in language comp. *Frontiers in Psychology*, *5*, 1-17.

Fehringer, C., & Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica*, *41*(1-2), 37-72.

Fewell, N. (2010). Language learning strategies and English language proficiency: An investigation of Japanese EFL university students. *TESOL Journal*, *2*(1), 159-174.

Field, A. (2009). *Discovering statistics using SPSS.* Los Angeles, LA: Sage Publications.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, *43*(2), 226-236.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101-135.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186-204.

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, *52*(2), 326-340.

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*(5), 1410-1430.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201-225.

Ghaffarvand Mokari, P., & Werner, S. (2019). On the role of cognitive abilities in second language vowel learning. *Language and Speech*, *62*(2), 260-280.

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, *34*, 445-474.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67-81.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515-530.

Groom, M. J., & Cragg, L. (2015). Differential modulation of the N2 and P3 event-related potentials by response conflict and inhibition. *Brain and Cognition*, *97*, 1-9.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: Comprehensive meta-analysis. *Second Language Research, 33*(3), 325-340.

Gupta, P., & Tisdale, J. (2009). Word learning, phonological short-term memory, phonotactic probability and long-term memory: towards an integrated framework. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3755-3771.

Gurunandan, K., Carreiras, M., & Paz-Alonso, P. M. (2019). Functional plasticity associated with language learning in adults. *NeuroImage*, *201*, 1-9.

Gustavson, D. E., Miyake, A., Hewitt, J. K., & Friedman, N. P. (2015). Understanding the cognitive and genetic underpinnings of procrastination: Evidence for shared genetic influences with goal management and executive function abilities. *Journal of Experimental Psychology: General*, *144*(6), 1063-1079.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis (Vol. 6)*. Upper Saddle River, NJ: Pearson Prentice Hall.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175-204). San Diego, CA: Academic Press.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25-38.

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*, 1-19.

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., & Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: A multimodal imaging study. *The Journal of Neuroscience, 33,* 13663-13672.

Huang, T., Loerts, H., & Steinkrauss, R. (2020). The impact of second-and third-language learning on language aptitude and working memory. *International Journal of Bilingual Education and Bilingualism*, 1-17.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017-2036.

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 645-662.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: A critical review. *Psychophysiology*, *43*(3), 302-313.

Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language and its role in predicting word recognition ability. *Applied Psycholinguistics*, *30*(1), 153-185.

Jiao, L., Liu, C., Wang, R., & Chen, B. (2019). Working memory demand of a task modulates bilingual advantage in executive functions. *International Journal of Bilingualism*, *23*(1), 102-117.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, *102*(2), 199-225.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213-233.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151.

Kalia, V., Wilbourn, M. P., & Ghio, K. (2014). Better early or late? Examining the influence of age of exposure and language proficiency on executive function in early and late bilinguals. *Journal of Cognitive Psychology*, *26*(7), 699-713.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169-183.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147-1185.

Kempe, V., Brooks, P. J., & Kharkhurin, A. (2010). Cognitive predictors of generalization of Russian grammatical gender categories. *Language Learning*, *60*(1), 127-153.

Kerrigan, L., Thomas, M. S., Bright, P., & Filippi, R. (2017). Evidence of an advantage in visuo-spatial memory for bilingual compared to monolingual speakers. *Bilingualism: Language and Cognition*, *20*(3), 602-612.

Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory. *Studies in Second Language Acquisition*, *37*, 549-581.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Kintsch, W. (2000). The control of knowledge activation in discourse comprehension. In W. J. Perry & A. Grob (Eds.), *Control of human behavior, mental processes, and consciousness: Essays in honor of the 60th birthday of August Flammer* (pp. 137-170). Mahwah, NJ: Lawrence Erlhaum Associates, Inc.

Kondo, H., Morishita, M., Osaka, N., Osaka, M., Fukuyama, H., & Shibasaki, H. (2004a). Functional roles of the cingulo-frontal network in performance on working memory. *NeuroImage, 21*, 2-14.

Kondo, H., Osaka, N., & Osaka, M. (2004b). Cooperation of the anterior cingulate cortex and dorsolateral prefrontal cortex for attention shifting. *NeuroImage, 23*, 670-679.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*(2), 261-271.

Kormos, J., & Trebits, A. (2011). Working memory capacity and narrative task performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 267-285). Amsterdam, the Netherlands: John Benjamins.

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience, 8*, 250-262.

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394-425.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice, 2*(12), 59-67.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-38.

Li, D., Christ, S. E., & Cowan, N. (2014a). Domain-general and domain-specific functional networks in working memory. *NeuroImage*, *102*, 646-656.

Li, P., Legault, J., & Litcofsky, K. A. (2014b). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex, 58*, 301-324.

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *The Modern Language Journal*, *97*(3), 634-654.

Li, S., Ellis, R., & Zhu, Y. (2019). The associations between cognitive ability and L2 development under five different instructional conditions. *Applied Psycholinguistics*, *40*(3), 693-722.

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon, 3*(3), 349-374.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861-883.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, *15*(3), 651-662.

Linck, J. A., & Weiss, D. J. (2011). Working memory predicts the acquisition of explicit L2 knowledge. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 101-114). Washington, DC: Georgetown University Press.

Linck, J. A., & Weiss, D. J. (2015). Can working memory and inhibitory control predict second language learning in the classroom? *SAGE Open*, *5*(4), 1-11.

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, *28*(1), 28-34.

Lyons, M. J., York, T. P., Franz, C. E., Grant, M. D., Eaves, L. J., Jacobson, K. C., … Kremen, W. S. (2009). Genes determine stability and the environment determines change in cognitive ability during 35 years of adulthood. *Psychological Science*, *20*(9), 1146-1152.

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, *63*(1), 240-244.

Martin, J. D., Shipstead, Z., Harrison, T. L., Redick, T. S., Bunting, M., & Engle, R. W. (2019). The role of maintenance and disengagement in predicting reading comprehension and vocabulary learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(1), 140-154.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34,* 179-413.

Martinez, D., & Singleton, J. L. (2019). The effect of bilingualism on lexical learning and memory across two language modalities: Some evidence for a domain-specific, but not general, advantage. *Journal of Cognitive Psychology*, *31*(5-6), 559-581.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419-457.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain: Proficiency in a second language and age at acquisition affect grey-matter density, *Nature, 431,* 757.

Menon, V. (2016). Memory and cognitive control circuits in mathematical cognition and learning. In M. Cappelletti & W. Fias (Eds.), *The mathematical brain across the lifespan* (pp. 159-186). Elsevier.

Mertler C., & Vannatta, R. (2002) *Advanced and multivariate statistical methods*. Los Angeles, LA: Pyrczak Publishing.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: Sage.

Miki, S. (2012). Working memory as a factor affecting L2 listening comprehension subskills. *Kumamoto University Departmental Bulletin Paper, 10,* 119-128.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex functioning. *Annual Review of Neuroscience, 24,* 167-202.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Holt, Rinehart & Winston.

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review, 20*(5), 819-858.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8-14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Morrison, C., Kamal, F., & Taler, V. (2019). The influence of bilingualism on working memory event-related potentials. *Bilingualism: Language and Cognition*, *22*(1), 191-199.

Newell, A., & Simon, H. (1956). The logic theory machine: A complex information processing system. *IRE Transactions on Information Theory, 2*(3), 61-79.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241-268.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220-246.

Norman, D., & Shallice, T., 1986. Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self regulation* (pp. 1-18). New York, NY: Plenum Press.

Olive, T., Kellogg, R. T., & Piolat, A. (2008). Verbal, visual, and spatial working memory demands during text composition. *Applied Psycholinguistics, 29*(4), 669-687.

Paap, K. R. (2019). The Bilingual advantage debate. In J. W. Schwieter (Ed.), *The handbook of the neuroscience of multilingualism* (pp. 701-735). Hoboken, NJ: Wiley Blackwell.

Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., & Mikulinsky, R. (2020a). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cognitive Research: Principles and Implications, 5*(1), 1-27.

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., Mason, L., Zimiga, B., Silva, Y., & Frost, M. (2020b). The alchemy of confirmation bias transmutes expectations into bilingual advantages: A tale of two new meta-analyses. *Quarterly Journal of Experimental Psychology*, *73*(8), 1290-1299.

Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *Journal of Clinical and Experimental Neuropsychology, 33*(4), 456-470.

Paller, K. A. (2003). The principle of cross-cortical consolidation of declarative memories. In R. H. Kluwe, G. Lüer, & F. Rösler (Eds.), *Principles of learning and memory* (pp. 155-169). Basel, Switzerland: Birkhäuser.

Paradis, M. (2013). Late-L2 increased reliance on L1 neurocognitive substrates: A comment on Babcock, Stowe, Maloof, Brovetto & Ullman (2012). *Bilingualism: Language and Cognition*, *16*(3), 704-707.

Patrick, M. E., Blair, C., & Maggs, J. L. (2008). Executive function, approach sensitivity, and emotional decision making as influences on risk behaviors in young adults. *Journal of Clinical and Experimental Neuropsychology, 30*(4), 449-462.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202-206.

Qi, Z., Han, M., Wang, Y., de los Angeles, C., Liu, Q., Garel, K., … Perrachione, T. K. (2019). Speech processing and plasticity in the right hemisphere predict variation in adult foreign language learning. *NeuroImage, 192,* 76-87.

Ratiu, I., & Azuma, T. (2015). Working memory capacity: Is there a bilingual advantage? *Journal of Cognitive Psychology, 27(1),* 1-11.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 501-526.

Roediger, H. L. (2000). Why retrieval is the key process in understanding human memory. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn Conference* (pp. 52-75). Philadephia, PA: Psychology Press.

Rösler, F., & Heil, M. (2003). The principle of code-specific memory representations. In R. H. Kluwe, G. Lüer, & F. Rösler (Eds.), *Principles of learning and memory* (pp. 71-91). Basel, Switzerland: Birkhäuser

Ryskin, R., Levy, R. P., & Fedorenko, E. (2020). Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. *Neuropsychologia*, *136*, 1-12.

Sagarra, N. (2017). Longitudinal effects of working memory on L2 grammar and reading abilities. *Second Language Research*, *33*(3), 341-363.

Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for the interpretation of longitudinal change. *Neuropsychology*, *22*(6), 800.

Sanz, C., Lin, H., Lado, B., Stafford, C., & Bowden, H. (2016). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied Linguistics, 37*(5), 669-692.

Schlegel, A. A., Rudelson, J. J., & Tse, P. U. (2012). White matter structure changes as adults learn a second language. *Journal of Cognitive Neuroscience*, *24*(8), 1664-1670.

Serafini, E. J. (2017). Exploring the dynamic long-term interaction between cognitive and psychosocial resources in adult second language development at varying proficiency. *The Modern Language Journal*, *101*(2), 369-390.

Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, *38*(4), 607-646.

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). FMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, *138*, 107307.

Shallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*, 1405-1412.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2012). Working memory capacity and visual attention: Top-down and bottom-up guidance. *Quarterly Journal of Experimental Psychology*, *65*(3), 401-407.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception, & Psychophysics*, *77*(6), 1863-1880.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*(6), 771-799.

Shipstead, Z., Lindsey, D. R., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116-141.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*, 637-648.

Smithson, L., & Nicoladis, E. (2013). Verbal memory resources predict iconic gesture use among monolinguals and bilinguals. *Bilingualism: Language and Cognition*, *16*(4), 934-944.

Sullivan, M. D., Janus, M., Moreno, S., Astheimer, L., & Bialystok, E. (2014). Early stage second-language learning improves executive control: Evidence from ERP. *Brain and Language*, *139*, 84-98.

Sullivan, M. D., Prescott, Y., Goldberg, D., & Bialystok, E. (2016). Executive control processes in verbal and nonverbal working memory. *Linguistic Approaches to Bilingualism*, *6*(1), 147-170.

Suzuki, Y., & DeKeyser, R. (2017). Exploratory research on second language practice distribution: An aptitude x treatment interaction. *Applied Psycholinguistics*, *38*, 27-56.

Tabachnick, G. B. & Fidell, S. L. (2007). *Using multivariate statistics*. Boston, MA: Pearson Education, Inc.

Tagarelli, K. M., Shattuck, K. F., Turkeltaub, P. E., & Ullman, M. T. (2019). Language learning in the adult brain: A neuroanatomical meta-analysis of lexical and grammatical learning. *NeuroImage*, *193*, 178-200.

Testa, R., Bennett, P., & Ponsford, J. (2012). Factor analysis of nineteen executive function tests in a healthy adult population. *Archives of Clinical Neuropsychology, 27*, 213-224.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, *7*(3), 255-272.

Tucker-Drob, E. M., & Briley, D. A. (2014). Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychological Bulletin*, *140*(4), 949-979.

Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, *49*(4), 446-468.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(2), 105-122.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1-2), 231-270.

Ullman, M. T. (2006). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex*, *42*(4), 480-485.

Ullman, M. T., & Lovelett, J. T. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, *34*(1), 39-65.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104-132.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1302-1321.

Van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., … Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, *9*(3), 1-30.

Von Bastian, C. C., De Simoni, C., Kane, M., Carruth, N., & Miyake, A. (2017, November). Does being bilingual entail advantages in working memory? A meta analysis. In *58th Annual Meeting of Psychonomic Society, Vancouver*.

Walter, C. (2004). Transfer of reading comprehension skills to L2 linked to mental representation of text and to L2 working memory. *Applied Linguistics*, *25*, 315-339.

Walter, D., & Hentschel J. (2013, September). *CAMLA English Placement Test Forms D-F: Development Report.* Retrieved from https://michiganassessment.org/wp-content/uploads/2020/02/20.02.pdf.Res_.CaMLAEnglishPlacementTestEPTFormsD-F-DevelopmentReport.pdf.

Was, C. A., Rawson, K. A., Bailey, H., & Dunlosky, J. (2011). Content-embedded tasks beat complex span for predicting comprehension. *Behavior Research Methods*, *43*(4), 910-915.

Weinreich, U. (1953). *Languages in contact: Findings and problems.* New York, NY: Linguistic Circle.

Wen, Z., & Li, S. (2019). Working Memory in L2 Learning and Processing. In J. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 365-389). Cambridge, England: Cambridge University Press.

Xie, L., Ren, M., Cao, B., & Li, F. (2017). Distinct brain responses to different inhibitions: Evidence from a modified Flanker task. *Scientific Reports*, *7*(1), 1-10.

Xie, Z. (2018). The influence of second language (L2) proficiency on cognitive control among young adult unbalanced Chinese-English bilinguals. *Frontiers in Psychology*, *9*, 1-9.

Yang, E. (2017). Bilinguals' working memory (WM) advantage and their dual language practices. *Brain Sciences*, *7*(7), 86-107.

Yang, J., & Li, P. (2012). Brain networks of explicit and implicit learning. *PloS ONE*, *7*(8), 1-9.

Yang, J., Ye, J., Wang, R., Zhou, K., & Wu, Y. J. (2018). Bilingual contexts modulate the inhibitory control network. *Frontiers in Psychology*, *9*, 1-14.

Yang, S., & Yang, H. (2016). Bilingual effects on deployment of the attention system in linguistically and culturally homogeneous children and adults. *Journal of Experimental Child Psychology*, *146*, 121-136.