Doctoral thesis submitted to
the Faculty of Behavioural and Cultural Studies
Heidelberg University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy (Dr. phil.)
in Psychology

Title of the thesis

Exploring the potential of foreign language learning for dementia prevention: Effects on executive attention and Executive Functions in healthy older adults

presented by
Judith Alina Großmann

year of submission
2023

Dean: Prof. Dr. Guido Sprenger
Advisors: Prof. Dr. Patric Meyer
Prof. Dr. Dr. h.c. Andreas Kruse

Meinem Großvater

in liebevoller Erinnerung und Dankbarkeit

Acknowledgements

I would like to dedicate my sincere gratitude to everyone who supported me along the way to my doctoral thesis.

First and foremost, I want to thank my direct advisors Prof. Dr. Patric Meyer and Dr. Birgit Teichmann for their professional and organizational guidance in all aspects of my dissertation. Thank you for your scientific advice, your trustworthy collaboration, and your constant assistance with creative ideas in the face of questions, doubts, and obstacles. Only because of your dedication have I been able to complete my dissertation over the last few years. Thank you so much for your confidence. I also want to appreciate Prof. Dr. Dr. h.c. Andreas Kruse for preparing the second opinion for my doctoral thesis.

Institutional thanks go to the Network Aging Research (NAR), particularly to Prof. Dr. Dr. h.c. Konrad Beyreuther and Prof. Dr. Hans-Werner Wahl for providing the academic setting for my dissertation. Thank you for the helpful advice and valuable interdisciplinary exchange during the weekly meetings. Without the provision of all necessary facilities, such as office and experimental rooms, as well as computers, this dissertation would not have been feasible. In this respect, I would also like to highlight the SRH University Heidelberg, which also offered a room for testing.

A warm thank also goes to my colleagues at the NAR, who have been an essential source of professional, encouraging, and amicable support. In this context, I would especially like to name the “Mensa Group”, which met regularly for lunch breaks and enjoyable conversations, including Simge Celik, Merve Degirmenci, Dr. Patrick Heldmann, Dr. Julia Kirch, Dr. Eesha Kokje, Dr. Eva-Luisa Schnabel, Dr. Julia Schneider, and Dr. Anton Schönstein.

I am also grateful to all administrative members of the NAR, who have always been supportive by proofreading, designing flyers, and resolving technical issues: Taisiya Baysalova, Anna Tzanetoulakou, and Andreas Sokoll.

With regard to financial support, I would like to start by acknowledging the Klaus Tschira Foundation, without which this dissertation would not have been possible. I am also highly grateful for the support and advice of Dr. Steffen Aschenbrenner. Thank you for your unhesitating readiness to apply for funding for the language courses of my doctoral project at the SRH Clinic Karlsbad-Langensteinbach, which was gratefully granted.

In this context, I would also like to appreciate the F+U Academy of Languages Heidelberg for granting discount on the language lessons and SCHUHFRIED GmbH, Mödling, for supporting my doctoral project by providing free access to cognitive tests from the Vienna Test System. Finally, I am also thankful for all the funders who made open access publishing possible, including the German Research Foundation within the funding program "Open Access Publikationskosten", the Baden-Wuerttemberg Ministry of Science, Research and the Arts, and the Ruprecht Karl University of Heidelberg.

Sincere thanks are also sent to the student assistants and interns for their invaluable help and tireless efforts. In this regard, I would like to thank Ursula Consiglio, Jasmin Dressler, Julian Feher, Vanessa Guenther, Tom Schuster, Karina Senftner, Franziska Vetter, and Lea Wendler for their support in participant recruitment and assessment, as well as Verena Magdalena Koelsch for her assistance in developing this study. I cannot leave out to acknowledge Elena Dillmann and Britta Arenas Vallejo, who organized, and Elena De Angelis and Inge Stegmann, who conducted the language courses. Also, without the remarkable and reliable commitment of the study participants, this endeavor would not have been feasible.

Last but not least, I would like to express my heartfelt thanks to my family, especially my parents and my life partner Philipp, as well as my friends for their loving support and for sharing and celebrating accomplished milestones with me, as well as bringing me to different thoughts at times. Thank you also, Philipp and Nadine, for critically proofreading my dissertation. My deepest appreciation, I would like to dedicate to my grandfather Prof. Dr. Karl Fuchs. He was my additional silent advisor. He supported my academic career from the start by teaching me how to write scientific papers, how to plan and carry out a research project, and how to cope with uncertainties. He also critically read my manuscripts, offered suggestions, and always had an open ear whenever I had questions. Without his help, I probably would not have gotten this far and wouldn't have acquired so many skills and such great enthusiasm for science.

Parts of this dissertation have already been published in the following articles:

Degirmenci, M. G., **Grossmann, J. A.**, Meyer*, P. & Teichmann*, B. (2022). The role of bilingualism in executive functions in healthy older adults: A systematic review. *International Journal of Bilingualism*. *26*(4), 426–449. The original publication is available at https://doi.org/10.1177/13670069211051291.

Grossmann, J. A., Aschenbrenner, S., Teichmann*, B. & Meyer*, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial. *Frontiers in Aging Neuroscience*. *15*, 1–16. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185.

Grossmann, J. A., Koelsch, V. M., Degirmenci, M. G., Aschenbrenner, S., Teichmann*, B. & Meyer*, P. (2021). Effects of foreign language learning on executive functions in healthy older adults: Study protocol for a randomised controlled trial. *BMC Geriatrics. 21*(1), Article 122, 1-14. The original publication is available at http://dx.doi.org/10.1186/s12877-021-02051-x.

*Patric Meyer and Birgit Teichmann share the last authorship.

Extracted sections are indicated at the corresponding parts in this dissertation. Copyright of the original article is retained by the authors (*BioMed Central*, *Frontiers*). The accepted manuscript may be used in the authors' own dissertation or thesis (*Sage Journals*).

Table of contents

Zusammenfassung

Hintergrund: Die Weltbevölkerung altert, wodurch die Häufigkeit von Demenzerkrankungen zunimmt. Seit einiger Zeit wird vermutet, dass Fremdsprachenlernen im höheren Lebensalter die kognitiven Funktionen verbessern und somit ein gesundes kognitives Altern unterstützen könnte. Bislang gibt es jedoch nur wenige Studien, welche darüber hinaus widersprüchliche Ergebnisse lieferten. Ziel [dieser Dissertation und folglich] dieser Studie war es daher zu untersuchen, ob das Erlernen einer Fremdsprache die exekutive Aufmerksamkeit und die Exekutiven Funktionen bei gesunden älteren Erwachsenen verbessern kann. Darüber hinaus wollten wir Faktoren ermitteln, die kognitive Veränderungen bei Fremdsprachenlernenden beeinflussen könnten, wie z. B. die Kognitive Reserve, vorherige Fremdsprachenkenntnisse und -verwendung sowie die globale Grundkognition zu Beginn der Studie.

Methoden: In einer randomisierten kontrollierten Studie wurden 34 einsprachige Personen im Alter zwischen 65 und 80 Jahren entweder einer Sprachlern- oder einer Wartelisten-Kontrollgruppe zugeteilt. Die Teilnehmenden besuchten für täglich 1,5 Stunden an fünf Tagen pro Woche über insgesamt drei Wochen einen Spanischkurs für Anfänger. Die Wartelistenkontrollgruppe erhielt keine Intervention, hatte aber die Möglichkeit, am Ende der Studie am Sprachkurs teilzunehmen. Bei allen Teilnehmenden wurden die exekutive Aufmerksamkeit (Primäres Outcome), die Exekutiven Funktionen, die Wortflüssigkeit und die Aufmerksamkeit (Sekundäre Outcomes) vor, unmittelbar nach dem Kurs oder nach einer Wartezeit von drei Wochen für die Kontrollgruppe und drei Monate nach dem Kurs beziehungsweise der Wartezeit untersucht.

Ergebnisse: Das Lernen einer Fremdsprache führte weder unmittelbar, noch drei Monate nach dem Kurs zu einer signifikanten Verbesserung der primären oder sekundären Outcomes. Moderationsanalysen ergaben jedoch, dass Teilnehmende mit niedrigerer globaler Grundkognition sich tendenziell stärker in der Reaktionsinhibition verbesserten als Personen mit einer höheren globalen Kognition. Dieser Zusammenhang war in der Wartelistenkontrollgruppe nicht sichtbar.

Diskussion: Unsere Ergebnisse deuten darauf hin, dass das Erlernen einer Fremdsprache nicht generell die exekutive Aufmerksamkeit oder die Exekutiven Funktionen verbessert. Personen mit einer geringeren Grundkognition könnten dennoch vom Lernen einer Fremdsprache im Bereich der Reaktionsinhibition profitieren, einem Bereich, der besonders vom kognitiven Alterungsprozess betroffen ist. Unsere Ergebnisse verdeutlichen die Notwendigkeit, den Schwerpunkt von Demenzpräventionsmaßnahmen auf Bevölkerungsgruppen zu richten, die besonders vulnerabel gegenüber kognitivem Abbau sind. Darüber hinaus könnte ein verstärkter Einsatz individualisierter Lernmethoden, einschließlich der Anwendung von technologiegestütztem Lernen, den Teilnehmenden ermöglichen, auf ihrem entsprechenden Leistungsniveau zu üben. Dies könnte die Wahrscheinlichkeit erkennbarer kognitiver Verbesserungen erhöhen.

Abstract

Background: The world's population is aging, increasing the prevalence of dementia. Recently, foreign language learning in later life has been suggested to improve cognition and thus support healthy cognitive aging. To date, however, there are only a few studies with conflicting findings. Therefore, the purpose of [this dissertation and consequently of] this study was to examine whether learning a foreign language can improve executive attention and executive functions in healthy older adults. Additionally, we sought to identify factors affecting cognitive change in foreign language learners, such as cognitive reserve, previous foreign knowledge and usage, and global cognition at baseline.

Methods: In a randomized-controlled trial, we assigned 34 monolinguals between the ages of 65 and 80 to a language learning or a waiting list control group. The participants enrolled in a Spanish course for beginners that met five days a week for 1.5 h for a total of 3 weeks. The waiting list control group received no intervention but had the opportunity to join the language training at the end of the study. All participants underwent an assessment of executive attention (primary outcome), executive functions, verbal fluency, and attention (secondary outcomes) before, immediately after the course, or after a waiting period of 3 weeks for the control group and 3 months after the course or the waiting period.

Results: Foreign language learning did not significantly improve primary or secondary outcomes, neither immediately nor 3 months after the course. However, moderation analyses revealed that participants with lower global baseline cognition tended to improve more on response inhibition than individuals with higher baseline cognition. This relationship was not evident in the waiting list control group.

Conclusions: Our results suggest that studying a foreign language does not generally improve executive attention or executive functioning. Nevertheless, individuals with poorer baseline cognition may benefit cognitively from foreign language learning in response inhibition, a domain particularly affected by cognitive aging. Our findings highlight the need of focusing dementia prevention efforts on groups that are more vulnerable to cognitive decline. Additionally, more individualized approaches, including utilizing technology-assisted learning, might enable participants to practice at their performance level, increasing the likelihood of discernible cognitive gains.

The abstract of this study and concurrently of this dissertation was recently published in *Frontiers in Aging Neuroscience* in the scientific article “Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial, pp. 1–2”. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185.

Overview

The object of this dissertation was to investigate the effects of foreign language learning on executive attention and Executive Functions in healthy older adults. Thematically, this dissertation is to be assigned to the research area of dementia prevention. It is important to note that dementia prevention does not imply the absence of cognitive decline and dementia. Individuals who engage in dementia prevention may still develop dementia. However, dementia may occur later in life than it would have if no dementia risk reduction measures had been taken.

The study presented in the following chapters of this dissertation was conducted at the Network Aging Research (NAR) at Heidelberg University and the SRH University Heidelberg, Germany. Chapter 1 provides the theoretical and empirical background of the research field investigated in this dissertation. A comprehensive review of the role of bilingualism on executive functioning in healthy older individuals was part of this dissertation and preliminary work for the completed study. The manuscript was published in the *International Journal of Bilingualism*. Sections of the publication are included in chapter 1.4.3 and are referenced accordingly. The original publication is available at https://doi.org/10.1177/13670069211051291. The manuscript was mainly written by my co-author Merve Gul Degirmenci. As the second author, I critically reviewed and corrected the manuscript for relevant intellectual content. Together with Merve Gul Degirmenci, I also screened all relevant titles, abstracts, and full texts and performed a quality assessment for all studies finally included in the review. I also checked all extracted data for completeness and correctness.

Chapter 1.5 outlines the research questions and hypotheses of the study presented in this thesis. The methods are described in Chapter 3 and have previously been published in a study protocol in *BMC Geriatrics*. The original publication is available at http://dx.doi.org/10.1186/s12877-021-02051-x. A study protocol comprises details on how a study was conducted, including hypotheses, sample characteristics and selection, intervention design, and data analysis. The importance of published study protocols before recruitment is completed for the scientific integrity and rigor of conducted studies has been stressed by some researchers (see Chan et al., 2013 for a review). It enables authors of systematic reviews, researchers, and funders to evaluate the scientific rigor of a study design and published results (Li et al., 2016). A published study protocol also allows a comparison between the intended aims and investigations at the beginning of a study and the final report in the publication (Al-Marzouki et al., 2008). This increases the transparency of scientific work (Li et al., 2016) and facilitates the replication of key elements of a published study (Chan et al., 2013).

Under the supervision of my advisors, Patric Meyer and Birgit Teichmann, I developed the study protocol including the conception of the study. My other co-authors contributed to the study design. I mainly drafted the manuscript with the assistance of my co-author Verena Magdalena Koelsch. All co-authors have read and approved the ready version of the article, which was accepted for publication in 2021.

Regarding the conduct of the study, I was responsible for its implementation. I prepared material for assessment, recruited participants, organized a school and teachers to give the language courses for the study, and ensured accurate timing of the recruitment phase, course dates, and appointments with participants. Together with my co-author Steffen Aschenbrenner, I applied for funding for the language courses at the SRH Clinic Karlsbad-Langensteinbach, Germany, which we were successfully granted. I conducted the assessments and recruited, instructed, and supervised students, research assistants, and interns in recruitment and data collection procedures. I also processed, analyzed, and interpreted the data.

The findings of this study are presented and discussed in chapters 4 and 5, respectively. Chapter 5 also offers an outlook on suggestions for further research and concludes with an overall summary of the thesis. Chapters 4 and 5 have been published in *Frontiers in Aging Neuroscience* in the scientific article “Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial”. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185. I wrote the manuscript, which was supervised and approved by my co-authors Steffen Aschenbrenner, Birgit Teichmann, and Patric Meyer. Patric Meyer and Birgit Teichmann share the last authorship.

1. Theoretical background

1.1 Introduction

People aged 65 years and older are the world's fastest-growing age group. According to current estimates, their number will increase from the current 727 million to about 1.5 trillion by 2050 (United Nations, 2020). As a result, age-related health problems and diseases will become more prevalent as the population ages. Among these diseases, dementia is one of the main reasons for dependency and need for care in elderly people and the seventh main cause of mortality (World Health Organisation, 2021). Presently, dementia affects almost 60 million individuals around the world. According to estimates, this number will more than double by 2050 due to growing populations on the one hand and an increasingly aging population on the other (Nichols et al., 2022). Consequently, dementia is one of the most serious global health issues faced nowadays. Due to the dearth of curative or persuasive course-modifying pharmaceutical treatment approaches, research into non-pharmacological therapies for dementia prevention has emerged as a promising target of scientific activity.

A seminal paper by Livingston et al. (2020) suggests that 40% of the overall risk of developing dementia is modifiable. The authors identified twelve risk factors across the lifespan and determined their respective contributions to the overall risk. Their life course model demonstrates that it is never too early or too late to start dementia prevention. Even in early adulthood, factors such as low education have a significant impact. In middle and older adulthood, risk factors that negatively affect the development of neuropathological alterations in the brain include low social contact, depression, hearing disorders, hypertension, diabetes, smoking, obesity, excessive alcohol consumption, physical inactivity, traumatic head injury, and air pollution. Notably, the prevalence of the risk factors mentioned depends on socio-economic status. Therefore, it is not surprising that countries with lower socioeconomic status will experience a considerable increase in dementia diagnoses compared to those with higher socioeconomic status. For example, the projected number of people with dementia in Germany will probably not even double by 2050, whereas the numbers in countries in Northern Africa or the Middle East will more than triple (Nichols et al., 2022).

Besides intensive efforts to advance pharmacological treatments and prevention approaches and political endeavors to increase educational opportunities, many efforts are also being dedicated to identifying further protective factors for dementia. One potential protective factor – bilingualism – has proven particularly promising (Bialystok, 2017). However, despite intensive research, convincing reviews demonstrating a noticeable delay in dementia and an improvement in cognitive performance are lacking. In the absence of a clear cause-effect relationship, another line of research has emerged in the last decade: the study of the cognitive effects of foreign language learning in older adulthood (Antoniou et al., 2013). A further interest of the authors was the identification of interventions that can be applied in older adulthood to guard against cognitive decline and dementia. However, to date, there have been only a few studies in this field. To help fill this research gap, this thesis explored the cognitive consequences of learning a foreign language in healthy older adults.

1.2 Normal and pathological cognitive aging

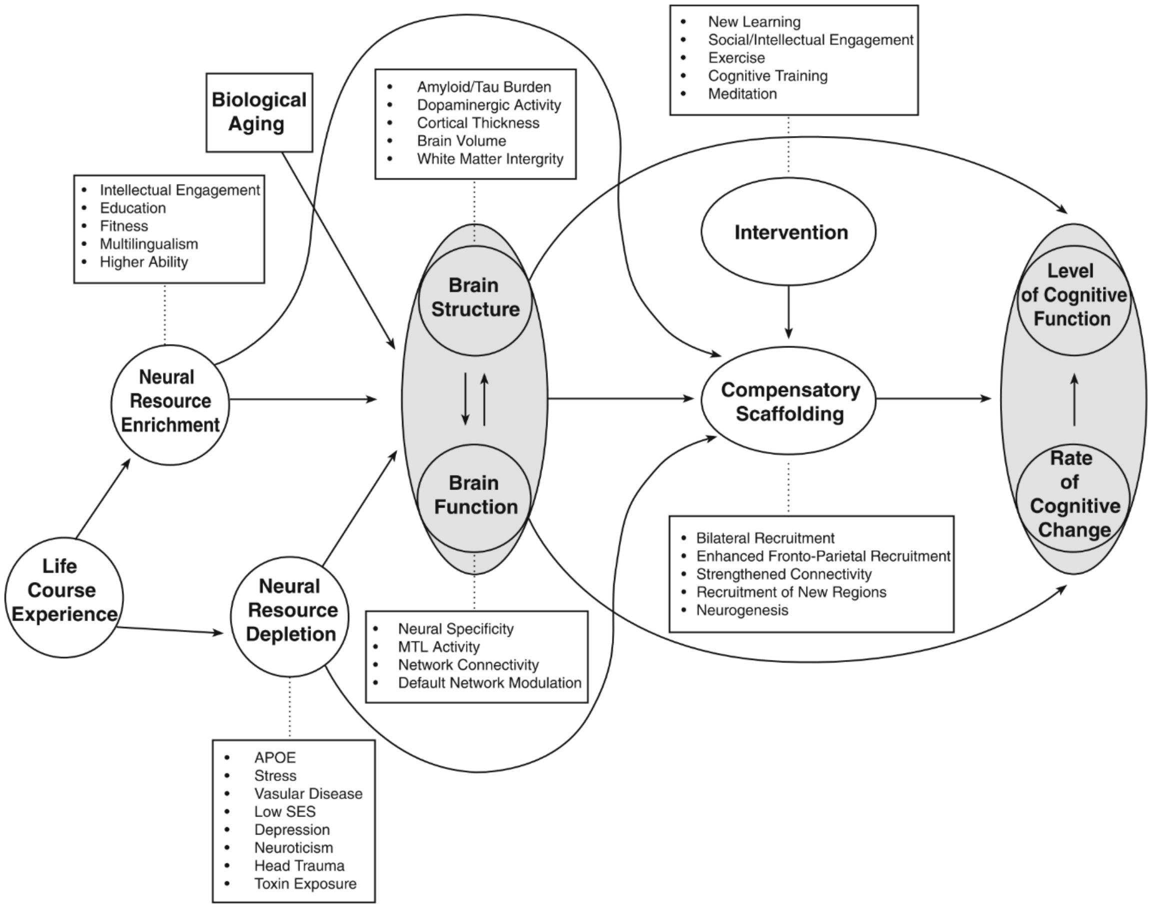
The term "cognitive aging" refers to a pattern of mild, age-related impairments of cognitive functions (Whalley et al., 2004). At the neuronal level, cognitive aging involves an array of structural and functional changes, including, but not limited to, a decrease in grey matter, particularly in the caudate, hippocampus, cerebellum, and frontal areas (Raz et al., 2005), and a frontally emphasized reduction of white matter integrity (Head et al., 2004). At the behavioral level, cognitive aging involves the gradual decline of a variety of cognitive functions, including processing speed, memory, visuospatial abilities, and Executive Functions. Verbal functions such as knowledge storage, in contrast, remain largely intact into older adulthood and even increase over the lifespan (Park et al., 2002).

There are debates about whether there is dedifferentiation or a general comprehensive decline of cognitive functions with increasing age. Dedifferentiation in this context means that younger adults engage a specific brain area to perform a given task, whereas older adults use more areas to complete the same activity (Baltes & Lindenberger, 1997). According to the scaffolding theory of aging and cognition, this decline in specialization may reflect a compensatory mechanism. With age, additional circuits are recruited, especially in the prefrontal cortex, to substitute for other structures whose function has decreased or become ineffective (Park & Reuter-Lorenz, 2009). It may also be possible that age-associated impairments of cognitive processes in some domains mediate declining functions in others. Domains particularly suspected here are decreasing inhibitory functions (see also chapter 1.4.1; Hasher & Zacks, 1988) and slowing information processing speed (Salthouse, 1996). Given the various structural and cognitive alterations related to aging, it is doubtful that a single process is responsible for cognitive decline. According to Park and Reuter-Lorenz (2009), cognitive aging is best understood as the combined influence of age-related neuronal and cognitive decline as well as compensatory processes. The revised life course model of the scaffolding theory of aging and cognition (STAC-r; Reuter-Lorenz & Park, 2014) provides a broad overview of mechanisms involved in neuronal and cognitive aging (see **Figure 1**). It also shows that the scaffolding mechanisms mentioned above, such as enhanced fronto-parietal recruitment and employment of new structural areas, neurogenesis, new learning, and intellectual engagement, can help compensate for brain structural and functional decline.

The disease stage of cognitive decline is dementia (Antoniou & Wright, 2017). At the behavioral level, dementia is characterized by increasing difficulties with learning and a progressive decline in cognition and the ability to cope with everyday life activities (Antoniou & Wright, 2017). The most well-known cause of dementia in individuals over 64 years is Alzheimer's disease (Gauthier et al., 2021). Pathophysiologically, it is characterized by amyloid beta peptide accumulation, tangle formation, synaptic dysfunction, and neuronal death in the brain. The underlying pathophysiological process of Alzheimer's Dementia (AD) and the clinical symptomatology can best be conceptualized as a continuum. It progresses from the preclinical stage through mild cognitive impairment (MCI) to mild and severe dementia (Sperling et al., 2011).

AD begins with a long preclinical or asymptomatic phase. In this phase, biomarkers indicative of AD, such as amyloid burden, are present but without any cognitive impairment, e.g., memory loss. Individuals at this stage are more susceptible to dementia. However, most people with biomarkers suggestive of AD will never develop dementia (Livingston et al., 2020).

Figure 1. Conceptual model of the scaffolding theory of aging and cognition (STAC-r) CC BY Reuter-Lorenz & Park (2014)



In people with MCI due to AD, biomarker evidence of AD has increased and can no longer be fully compensated by scaffolding mechanisms. Thus, individuals with MCI exhibit clinical symptoms of memory loss and language problems, such as difficulty finding words or decreased abstract thinking. Additionally, alterations in mood, behavior, and emotions can occur (Gauthier et al., 2021). These problems can be subtle but already noticeable by the affected person, close relatives, and friends. Everyday activities can still be performed independently. After two years, about 15% of people with MCI will eventually be diagnosed with dementia (R. C. Petersen et al., 2018).

Mild cognitive deficits associated with AD may precede the onset of the disease by up to ten years (Morris, 2005). This understanding has fueled the call for proactive treatment of the disease, ideally long before the first cognitive symptoms manifest (Sperling et al., 2011).

As the STAC-r model indicates, the cognitive aging process is not linear (Whalley et al., 2004), but can be actively modified. This knowledge can be harnessed to prevent dementia. However, pharmacological attempts to prevent the onset of dementia or to slow its progression have had little success (Salloway et al., 2014). Therefore, research has primarily focused on proactive therapy of behavioral symptoms such as cognitive dysfunction in early stages of the disease or the preclinical stage in cognitively healthy older adults.

The notion of cognitive reserve (CR) is among the most auspicious non-pharmacological avenues to preventing dementia. It exhibits some overlap with the STAC-r-model, which is more dynamic and emphasizes the role of plastic reorganization in mitigating neuronal degeneration (Reuter-Lorenz & Park, 2014). The notion of CR particularly underscores the significance of lifetime experiences for cognitive functioning and will be presented in the next chapter.

1.3 Cognitive reserve

CR can be conceptualized as the ability to endure a higher degree of age-associated or pathological neuronal decline without experiencing cognitive, functional, or clinical symptoms (Stern et al., 2020). The reserve heuristic was initially established to explain the observation that the degree of histopathological alterations at autopsy not always corresponded to the extent of clinical impairment, as was seen, for instance, in cases of AD (Katzman et al., 1988). Thus, CR moderates neuropathology and clinical symptoms (Stern, 2012).

Reserve is divided into two independent systems that synergistically contribute to resilience to neuropathology: neuronal reserve and CR. The neuronal reserve is a quantitative concept and refers to the brain's ability to resist pathological changes through greater brain volume, higher synaptic density, and a more significant number of healthy neurons (Stern, 2009). With a high neural reserve, sufficient neural substrate remains to ensure cognitive functionality despite pathology (Stern, 2012).

Conversely, CR is a qualitative concept defined as the capability of cognitive and functional processes to adapt to demands in terms of better efficiency, capacity, or flexibility. Functional processes are understood as networks of regions in the brain associated and interconnected with the execution of a task. At the same time, these networks also interact with each other (Stern et al., 2020). Therefore, it is assumed that the brain of a person whose cognitive networks work more efficiently or possess greater capacity or flexibility may be better able to cope with neuropathology-related alterations (Stern, 2009).

Notwithstanding their independence from one another, there are no clear boundaries between the neural reserve and CR. On the contrary, both contribute synergistically to reserve (Stern, 2012). For example, life experiences can also influence brain anatomy through neurogenesis, angiogenesis, or promotion of apoptosis resistance (Stern, 2009).

Furthermore, CR is not rigid or invariable but is influenced by life experiences (Stern et al., 2020). These experiences are not limited to the early years of life, such as years of education. A cognitively demanding occupation and engaging in cognitively stimulating activities – even at an older age – are also recognized as reserve-building factors (Opdebeeck et al., 2016). These factors are usually approximately determined indirectly based on autobiographical information rather than directly (Valenzuela & Sachdev, 2006).

From this theoretical framework, two central hypotheses emerged. According to the CR hypothesis, engaging in cognitively stimulating activities or life experiences contributes to CR and thus reduces the risk of dementia. The CR hypothesis also predicts that individuals with higher levels of cognitive function will have a reduced dementia risk than those showing poorer cognitive performance (Whalley et al., 2004).

Nowadays, there is ample empirical evidence for both hypotheses. For example, a meta-analysis including over 29.000 subjects demonstrated that individuals with high CR have a nearly 50% lower risk of developing dementia than individuals with low CR (Valenzuela & Sachdev, 2006). This finding can be explained by the fact that those with higher CR can tolerate more pathology. They thus experience cognitive impairment later in the course of the disease than people with lower CR. However, cognitive decline can accelerate from a turning point when neuropathology becomes so marked that compensatory mechanisms begin to fail (Stern, 2012).

Another meta-analysis examined 126 studies with 128.328 participants (Opdebeeck et al., 2016). The researchers found a positive relationship between the three largest, generally accepted proxy measures of CR – educational attainment, occupational activity, and cognitively stimulating leisure activities – and cognitive function: The higher the CR, the greater the overall degree of cognitive functioning. However, the strength of the relationship varied depending on the particular proxy measures and cognitive functions considered. For example, the correlation was moderate for the level of education, whereas it was weak for occupational activity and participation in cognitively stimulating leisure activities.

It is crucial to highlight that the observed associations between CR and cognitive functioning could also be of reverse causality. For example, subtle cognitive impairments could reduce interest in and ability to participate in activities that require social and cognitive skills (Sajeev et al., 2016). Indicative of the causality of the relationship are intervention studies that use a pre-post design to examine the effects of cognitively stimulating activities on cognitive functioning levels. This knowledge has prompted numerous intervention studies investigating the cognitively stimulating effects of intellectually challenging activities such as crossword solving, reading, or cooking (Iizuka et al., 2019) and musical activity (Bugos et al., 2007). A large-scale meta-analysis also showed positive effects of cognitive training on cognition that were long-lasting beyond the training period (Nguyen et al., 2019). Cognitive improvements resulting from cognitively stimulating activities were found in healthy older adults as well as in individuals with MCI (Pillai et al., 2011) and diagnosed with dementia (Zhang et al., 2017).

However, what experiences and activities contribute to CR remains a subject of research. One lifestyle factor commonly attributed to CR and which has been debated extensively is bilingualism (Bialystok, 2017). The next chapter presents the underlying mechanisms of bilingual cognitive language control and the empirical state of research.

1.4 Bilingualism

In the past two decades, numerous studies have been conducted to determine how bilingualism affects cognitive functions in older adults. Bilingualism is a highly heterogeneously defined construct. All characteristics including language proficiency, current usage in different contexts (e.g., home, school), history of acquisition, or socio-linguistic context in which both languages are used may contribute to the bilingual experience (Surrain & Luk, 2019). Another challenge for definition is that bilingualism is not a categorical variable but a continuous one. Bilingualism has, therefore, experienced different operationalizations in literature, ranging from very general to highly specific definitions. However, most studies define bilingualism based on proficiency, either self-reported or objectively measured (Surrain & Luk, 2019). As I refer to a broad range of studies in the following, I also define bilingualism based on proficiency as the capability to communicate in two languages at a fluent level, following Bak (2016).

In its nature, bilingualism is hardly different from many other cognitively stimulating experiences and activities. Nevertheless, its omnipresence in everyday life and across the lifespan, as well as the broad neural network bilingualism addresses (see chapter 1.4.1), makes it unique as an environmental factor to potentially protect against cognitive decline and dementia (Gallo et al., 2022). Indeed, a recent systematic review and meta-analysis suggests that the onset of symptoms of AD is delayed by almost 4.7 years, and the age of diagnosis is delayed by 4.2 years compared to monolinguals (Brini et al., 2020). However, not all systematic reviews support these findings. For example, in their systematic review, Mukadam et al. (2017) reported insufficient proof to imply that bilingualism would reduce the incidence of dementia. Yet, the authors included studies with each type of dementia, which may partly explain the conflicting results, as different causes of dementia involve different brain pathologies. The next chapter expands on the mechanisms and models to elucidate why bilingualism might delay dementia.

1.4.1 Mechanisms

Neuronal networks involved in bilingual language processing overlap with neurocognitive networks affected by cognitive aging (Abutalebi & Green, 2016; Raz et al., 2005) and those related to MCI and dementia (Gallo et al., 2022). Bilingualism could therefore strengthen brain regions and their connections to areas (e.g., fronto-striatial and fronto-parietal), which subserve cognitive control and are susceptible to cognitive decline in the presence of neuropathology such as seen in AD (Antoniou & Wright, 2017).

However, the exact mechanisms for the protective effects of bilingualism are still poorly understood. According to the theory of CR, bilinguals employ more alternative brain structures, which may allow them to tolerate greater brain pathology before cognitive symptoms manifest (Antoniou & Wright, 2017). This assumption is supported by Duncan et al. (2018). In monolingual and multilingual patients with MCI and AD, the authors measured cortical thickness and grey matter density. Both groups were matched for episodic memory, global cognitive functioning, age, and education. The authors found increased cortical thickness and higher grey matter density in cortical regions linked to language regulation and cognitive control in multilingual patients with MCI and AD. These areas comprised the left and right inferior frontal gyrus (IFG), left medial superior frontal gyrus, right ventromedial prefrontal cortex, left and right anterior temporal gyri, left parietal lobule, left and right cerebellum, and right cerebellar tonsil. In addition, they detected greater cortical thinning and reduced tissue density in the posterior parahippocampal gyri and the rhinal sulci – areas associated with memory functioning – in multilingual patients with AD compared to monolinguals. Thus, multilinguals had comparable memory dysfunction to monolinguals with AD despite having more neuropathology. This result supports the notion that bilingualism may enhance CR. Similarly, Schweizer et al. (2012) also reported significantly greater amounts of brain atrophy in areas relevant to the diagnosis of AD in cognitively and educationally matched bilinguals compared to monolinguals with AD.

In line with the CR hypothesis, bilinguals have also been reported to possess more substantial integrity of white matter fibers in the corpus callosum, stronger anterior-posterior connectivity (Luk, Bialystok, et al., 2011), and increased connectivity in the executive control network (Perani et al., 2017). The results of these studies suggest that bilingualism is associated with more extensive neural networks, which could increase neural reserve. A greater neural reserve may partially compensate for early cognitive dysfunction and thus delay the onset of clinical symptoms of AD in bilinguals.

According to the studies presented above, multilingual language regulation is accomplished through a broad and domain-general executive network that includes parts of the prefrontal cortex, inferior parietal lobules, and left caudate nucleus (Green & Abutalebi, 2013). Thus, the protective effect of bilingualism might also be mediated by its impact on executive control since many of the structures that are better preserved in bilinguals are directly or indirectly linked through associations with the so-called Executive Functions (EF; Abutalebi et al., 2015). Indeed, some systematic reviews suggest that bilinguals perform better in EF than monolinguals (Donnelly et al., 2019; Monnier et al., 2021).

1.4.2 Executive Functions

EF is a generic term for various higher-order cognitive processes necessary for goal-directed thoughts and behavior as well as for emotional control (Friedman & Miyake, 2017). They enable us to plan actions, adapt quickly and flexibly to new situations, suppress impulses, stay focused, or solve problems. Consequently, they are essential for efficient functioning in many aspects of life, e.g., mental and physical health, quality of life, or professional success (Diamond, 2013). The prefrontal cortex is the key structure responsible for the control of EF (Stuss & Benson, 1984). Therefore, the development of EF is directly connected to the maturation process of the prefrontal cortex. The trajectory generally follows a bell-shaped curve (Ferguson et al., 2021). EF develop gradually during childhood and deteriorate during older adulthood (V. Anderson et al., 2014).

Over time, different models of EF have been established (e.g., Baddeley's influential model of working memory [WM; Baddeley, 1986]). However, citing the most commonly used and widely accepted model of EF, there are three core domains of EF: inhibition of pre-potent responses, updating of WM representations, and shifting between tasks or mental sets (Miyake et al., 2000). These domains are distinct from each other while still interdependent and form the basis for further higher-level components of EF, such as logical thinking, problem-solving, and planning (Diamond, 2013; Miyake et al., 2000). It should be noted that an alternative model was later proposed that combined the three domains under a Common EF factor (Friedman et al., 2008). This Common EF factor can be defined as “the ability to actively maintain task goals and goal-related information and use this information to effectively bias lower-level processing” (Miyake & Friedman, 2012, p. 5). It explains some unique variance but also captures the variance of the shifting and the updating factor. The inhibition factor is absent because once the common EF factor is added to the latent factor model, the inhibition factor no longer explains any unique variance. As the Common EF factor is critical for efficient suppression of responses (Munakata et al., 2011), some authors propose that the Common EF factor might more easily be defined as inhibition. This simplified definition gives more insight into the underlying or shared mechanism of the two other domains of EF (Valian, 2015). However, from the perspective of Friedman and Miyake (2017), there is nothing special about inhibition. Moreover, the term Common EF is more general and flexible to be seen as an underlying factor of the other two domains (Valian, 2015). Still, as both models have been supported in the literature and show a similar fit to the data (Friedman & Miyake, 2017), we refer here to the original model, as this is the best known.

Inhibition is regarded as the ability to deliberately suppress automatic and dominant responses (Miyake et al., 2000) and comprises functions such as response inhibition, interference control, and self-regulation (Diamond, 2013). Inhibition enables one to react not according to a solid inner disposition or external cues but to what is appropriate or required in a particular situation (e.g., ignoring nonessential stimuli, postponing reward). Inhibitory control has been shown to be particularly affected by aging (Hasher et al., 1991). For example, older adults have more difficulty suppressing distractions. This difficulty points to a deficit in inhibitory control in aging, potentially partly mediating the relationship between age and performance in other domains of EF (Darowski et al., 2008). For example, WM deficits can be explained by the selection of irrelevant information and by the inefficient deletion of WM contents that are no longer relevant for task processing (Park & Reuter-

Lorenz, 2009). Bilingualism, in turn, has been associated with a better ability to ignore irrelevant stimuli (Costa et al., 2009) and to inhibit prepotent responses (Bialystok et al., 2004).

Updating and monitoring of WM representations is known as the ability to monitor and code inbound data for significance for a given task. Additionally, updating involves actively and adequately manipulating outdated and irrelevant data stored in WM (Miyake et al., 2000). Updating is closely linked to the notion of WM, which is described as the ability to retain and process information in mind (Baddeley & Hitch, 1994). WM is fundamentally necessary for understanding written or spoken language, doing mathematics, or mentally organizing information (e.g., making a daily plan or schedule). Updating is the ability to actively incorporate new information into thoughts or plans (Diamond, 2013). Like inhibitory control, WM declines early in aging (Fournet et al., 2012). Bilingualism has been associated with slightly higher WM capacity (Monnier et al., 2021).

Shifting – also called set-shifting, task switching, and cognitive or mental flexibility – is commonly recognized as the ability to switch between exercises, operations, and mental sets (Miyake et al., 2000). In everyday life, shifting is necessary, for example, for changing perspectives both spatially and interpersonally. If a problem is not solvable with one strategy, shifting enables us to change proceedings or take advantage of unexpected opportunities. Shifting between mental sets or tasks also involves aspects of the other two domains of EF. Inhibition is required to deactivate the previous mental set or task, while updating is needed to load the new task set into WM (Diamond, 2013). The literature suggests that aging is less associated with a problem in switching between mental sets but instead with greater difficulty in retaining multiple competing mental sets in WM (Kray & Lindenberger, 2000). Bilinguals, again, have been found to be better capable to shift between different tasks (Garbin et al., 2010).

It is hypothesized that the so-called bilingual advantage in EF results from the continuous load on cognitive language control during the processing of two languages. This assumption stems from the observation that representations of both languages are functionally active in the cortex irrespective of the one employed (Kroll et al., 2015). Thus, to suppress interference from one language while speaking in the other, cognitive language management is required. According to the Adaptive

Control Hypothesis – one of the most commonly cited models of bilingual language control (Green & Abutalebi, 2013) – the pattern of EF involvement depends on the interactional context of bilingual language usage: single language, dual language, and dense code-switching. Different cognitive control processes are required in these kinds of scenarios.

In a single language context – in contrast to a dual language environment – one language is solely employed in one context, putting less demand on cognitive control processes. A dual language context is regarded as a scenario in which both languages are used. The two languages are utilized interchangeably with various conversational partners. Thus, switching between languages may occur within an utterance. Moreover, in the dense-code-switching context, individuals routinely transition between languages within a phrase or word, linking vocabulary from one language to elements of the other. For example, the term “gedownloaded” origins from the German prefix “ge-“ and the English suffix “-ed” and the English verb “to download”. Green and Abutalebi (2013) distinguish eight cognitive control processes differentially recruited depending on the language context: goal maintenance, conflict monitoring, interference suppression, salient cue detection, selective response inhibition, task disengagement, task engagement, and opportunistic planning. Goal maintenance and conflict monitoring are rather recruited in the single language context. In contrast, the dual language context practically engages all cognitive control processes except opportunistic planning, which is more explicitly involved in the dense code-switching context.

The Adaptive Control Hypothesis still needs further testing. It is only one among several models attempting to define EF involvement in bilingual language processing (see Gallo et al. [2022] for an overview). As the next chapter will demonstrate, the empirical state of research on the effects of bilingualism on EF in healthy older adults is highly inconsistent. This discrepancy has several reasons, one of which is that the various contexts of bilingual language usage have received little attention in the past.

1.4.3 Empirical state of research in healthy older adults

Parts of this chapter have previously been published by my co-authors and me in the scientific article "Degirmenci, M. G., Grossmann, J. A., Meyer, P. & Teichmann, B. (2022). The role of bilingualism in executive functions in healthy older adults: A systematic review. *International Journal of Bilingualism*. *26*(4), 426–449. Copyright © 2022 (Copyright Holder)". The extracted parts are marked by indentations. Changes from the original publication are indicated in brackets. The original publication is available at https://doi.org/10.1177/13670069211051291.

Previous systematic reviews and meta-analyses have both supported and contradicted the bilingual advantage on EF (Adesope et al., 2010; de Bruin et al., 2016; Donnelly, 2016; Donnelly et al., 2019; Grundy & Timmer, 2017; Hilchey et al., 2015; Hilchey & Klein, 2011; Lehtonen et al., 2018; van den Noort et al., 2019; Zhou & Krott, 2016). These contradictory findings demonstrate either the absence of such an advantage as has been suggested as conceivable by some authors (Paap et al., 2015) or the lack of an in-depth and comprehensive understanding of the factors underlying a possible advantage. There are numerous factors whose influences we have yet to fully explore, such as the role of language typology, i.e., whether the two languages spoken are typologically similar or not (Antoniou & Wright, 2017). In addition, age of acquisition (Luk, de Sa, & Bialystok, 2011), socio-demographic factors, socioeconomic status, or migration background might also influence or mask the effects of bilingualism, to mention only a few of the influential aspects discussed in the literature (Celik et al., 2020).

Another factor that could have a significant influence is the age of the population under consideration. Effects of bilingualism might be more pronounced in older adults, as EF reach the peak of their functionality in early adulthood, from which point they gradually decline (Park et al., 2002). Empirically, therefore, we see a greater variability of EF with increasing age in adulthood (Bialystok et al., 2008). In addition, older adults often engage in less intellectually stimulating activities than younger adults, such as a cognitively stimulating occupation or computer games. Therefore, the effects of bilingualism might remain masked in younger adults, as bilingualism competes with the positive effects of other cognitively stimulating activities. The age of the participants could, therefore, partly explain the discrepant results in the literature.

> To our knowledge, there is no published systematic review or meta-analysis that reported only the results of older adults. Although almost all of the former reviews [listed above] included aging studies, only two studies reported separate results related to older adult participants (Donnelly, 2016; Lehtonen et al., 2018). Donnelly (2016) reported significantly larger effect sizes for interference costs than global RT [reaction time] amongst older adults. Lehtonen et al. (2018) reported a larger difference in a task-switching paradigm favouring older bilinguals over younger ones. (p. 429)

As there was no summary of the state of research in healthy older adults, we conducted a systematic review. We included 24 studies from 22 peer-reviewed published articles involving 1130 monolingual and bilingual older adults from ten different countries. The age of participants ranged from 64.5 to 80.9 years. We considered an extensive number of tasks associated with the three most common EF-domains: inhibition, shifting, and updating (Miyake et al., 2000). We also considered the influence of possible moderator variables, namely language proficiency, testing language, age of acquisition, country, socioeconomic status, and immigration background.

> Our systematic review provides no evidence for a general bilingual advantage in EFs [*sic*] in healthy older adults. However, when focusing on single EF domains, bilingualism seems to be reliably associated with an advantage in inhibition, most pronounced for the Stroop test and Simon task. For the shifting domain, less clear results were obtained, but this domain was also least studied. For the updating domain, no evidence of a bilingual advantage was found at all. Individual participant characteristics and methodological problems in studies may have caused the heterogeneity of the results and might explain some of the inconsistencies between the studies. (p. 444)

> Since all the studies did not control the same confounding variables, it was quite challenging to investigate the dynamics behind the studies that conducted the same EF tasks and demonstrated heterogeneous results. Critically, it is still not clear whether these effects were moderated by individual characteristics or not. The Stroop test, which is the most conducted task [in the studies included in the systematic review], is a robust example to explain the inconsistency in this dynamic. For example, studies that reported different results included participants from the same immigration background (Ansaldo et al., 2015; Antón et al., 2016; Clare et al., 2016; Kousaie & Phillips, 2012) differed in AoA [age of acquisition] or vice versa (Antón et al., 2016; Bialystok et al., 2008; Kousaie & Phillips, 2012). Yet again, it was difficult to sort studies regarding controlled variables and making an assumption based on them. (…) Therefore, well-designed studies that measure all core domains of EFs with valid and reliable tasks and that consider confounding variables are urgently needed. (pp. 443-444)

1.5 Foreign language learning in older age

The state of research presented above suggests that bilingualism may improve performance in EF. Better EF performance may contribute to the formation of CR, which in turn could lead to a later onset of dementia. However, there is still debate concerning the precise mechanisms involved (Antoniou & Wright, 2017). Furthermore, and perhaps because the exact mechanisms still need to be understood, several studies contradict the assumption of a so-called bilingual advantage calling its generalizability into question.

The diverse studies available to date, particularly in the literature on behavioral effects of bilingualism on EF, have sparked academic debate about the degree to which bilingualism leads to observable performance benefits in cognitive domains. Out of this debate and to elucidate a causal link between foreign language proficiency and usage and cognitive function, the idea was proposed to investigate whether foreign language training in older age could promote cognitive functions (Antoniou et al., 2013; Antoniou & Wright, 2017). Not only early (foreign language acquired in childhood) but also later life (adulthood) bilingualism is associated with relevant cognitive advantages, e.g., in selective attention (Vega-Mendoza et al., 2015). While challenging, learning a new language is nevertheless feasible for older adults (Kliesch et al., 2018) and might, therefore, be an effective method to augment cognitive functioning and thus promote healthy cognitive aging.

Foreign language acquisition may have cognitive advantages that are similar to those of bilingualism, despite obvious differences between the two, such as in terms of language proficiency. As with bilingualism in general, acquiring a foreign language is a complex mental activity that possibly involves several cognitive functions. Foreign language learners will, in most cases, not become bilingual, but the very process of acquisition could lead to measurable cognitive improvements (Meltzer et al., 2021).

Another argument in favor of the claim that foreign language learning can already show cognitively protective effects is again derivable from the theory of CR (see chapter 1.3). The theory posits that cognitive training can strengthen cognitive functions. In systematic reviews and meta-analyses, cognitive training has repeatedly shown positive effects on memory performance, EF, attention, processing speed, fluid intelligence, and subjective cognitive performance in healthy older adults (Kelly et al., 2014) and people with MCI (Reijnders et al., 2013). Not surprisingly, frequent and continuous cognitive training is also recommended for people with acquired brain injury, e.g., after stroke (Maurer-Karattup et al., 2022). In addition, including a social component in cognitive training can lead to more significant and long-lasting cognitive enhancements (Kelly et al., 2017; R. C. Petersen, 2011). Moreover, not only cognitive training but also purposeful, intervention-based engagement in intellectually stimulating cognitive activities promotes cognitive functionality across multiple cognitive domains. In their systematic review, Iizuka et al. (2019) detected that cognitive improvements were particularly pronounced when activities involved acquiring new abilities, provided substantial intellectual stimulation, and contained components of active communication, as is the case with foreign language learning.

1.5.1 Mechanisms

Parts of this and the following chapter have previously been published in *Frontiers in Aging Neuroscience* in the scientific article “Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial”. The extracted parts are marked by indentations. Changes from the original publication are indicated in brackets. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185.

> Which cognitive domains foreign language learning is addressing remains debated. More recent theories suggest that bilingual language control is organized hierarchically (Branzi et al., 2016). Correspondingly, language control may be fundamentally rooted in the attentional domain, more specifically in executive attention (Bialystok, 2017), which can be seen as the underlying performance on a broader range of EF tasks (Hilchey & Klein, 2011). According to Posner's and Peterson's influential model of attention, executive attention is an aspect of attentional control that is particularly necessary in high-conflict situations, e.g., when two languages are competing for processing (Posner & Petersen, 1990). The executive attention network comprises two sub-networks: a cingulo-opercular system, responsible for monitoring behavior, and a frontoparietal/dorsolateral prefrontal system, enabling switching between tasks or mental sets within a task (S. E. Petersen & Posner, 2012). Thus, executive attention is regulated by a network involving the anterior cingulate cortex (ACC) and lateral prefrontal areas (Fan et al., 2002), which are also of great importance in the early stage of foreign language acquisition (Pliatsikas, 2020). (p. 2)

The ACC is crucial for monitoring linguistic and cognitive conflicts (Luk, Green, et al., 2011) and is suggested to serve as a supervisory attentional system in bilingual language control (Branzi et al., 2016). “Therefore, executive attention might be one of the domains most directly addressed by foreign language learning” (Grossmann et al., 2023, p. 2).

Indeed, the Dynamic Restructuring Modell suggests that several anterior regions, such as the IFG, the medio-frontal gyrus, and the ACC, are linked to EF. These regions are recruited in the initial phase of learning a foreign language (Pliatsikas, 2020) and are also structures that are strongly affected by cognitive aging (Raz & Rodrigue, 2006). In support of the Dynamic Restructuring Modell, Bubbico et al. (2019) detected a significant augmentation in connectivity in the right IFG, right superior frontal gyrus, and superior parietal lobule in the only study to date that also assessed functional changes following language learning in older adults. Furthermore, a strong positive link was discovered between the magnitude of functional connection in the right superior frontal gyrus and global cognitive ability. Thus, changes in functional connection scores correlated positively with alterations in the MMSE score. This result was regarded by the authors as supporting the idea that the language course rather than the 4-month time interval between the two assessments caused the noticed gains in cognition in the language learning group.

In a nutshell, language learning may activate a broader neural network than other cognitively stimulating activities. It may increase not only language-related functions but also cognitive functions, even in older persons, as their brains are still plastic (Antoniou et al., 2013).

1.5.2 Empirical state of research in healthy older adults

> “Research on training-related changes in cognition after foreign language acquisition in healthy older adults is still limited. So far, only eleven studies have addressed this subject. [See **Table 1** for an overview of the studies available in this research field up to the submission of this thesis.] Six of these studies reported significant cognitive improvements after learning a foreign language (Bak et al., 2016; Bubbico et al., 2019; Long et al., 2020; Meltzer et al., 2021; Pfenninger & Polz, 2018; Wong et al., 2019).” (Grossmann et al., 2023, pp. 2–3)

It is worth emphasizing that the methodology varied considerably across the studies, with varying quality in study design. The age groups were heterogeneous and most studies' sample size was modest. Both previous skills in the foreign language to be learned and in other foreign languages were controlled unevenly or not at all. Additionally, the interventions varied in duration and frequency. Previous studies also measured a variety of and sometimes only general cognitive domains (e.g., global cognition). These study characteristics, as well as the strengths and shortcomings of earlier studies, are outlined in the following paragraphs.

Study design

Only five out of the eleven available studies were randomized controlled trials. Two of these studies failed to find a discernible cognitive benefit of studying a foreign language (Berggren et al., 2020; Valis et al., 2019). Contrarily, two studies (Bubbico et al., 2019; Wong et al., 2019) discovered improvements in global cognition after learning a foreign language, while a third study (Meltzer et al., 2021) demonstrated enhancements in updating and interference suppression (Stroop task). "However, in Bubbico et al. (2019), this supposed effect of foreign language learning was due to a decline in performance in the control group. In Wong et al. (2019), the intervention group improved their performance in working memory besides global cognition, but not compared to an active or passive control group" (Grossmann et al., 2023, p. 3). Relative to a passive control group, the improvement in Meltzer et al.'s study was seen in both the foreign language learning group that learned Spanish using an app and the active control group. The active control group underwent an app-based computerized cognitive training program. Therefore, it was not unexpected that this group showed measurable cognitive improvements. The main interest of the authors was to examine whether the group that used an app to study a foreign language would improve more than the passive control group. They were also interested in comparing the cognitive improvement of the language learning group with that of the active control group, which had received direct instruction in the cognitive abilities that the training was meant to enhance.

Three studies applied a quasi-experimental design with non-randomized group assignment (Bak et al., 2016; Kliesch et al., 2022; Ramos et al., 2017), while two others were pilot studies with no control group (Pfenninger & Polz, 2018; Ware et al., 2017). Long et al. (2020) also compared the results of studying foreign languages at three proficiency levels rather than using a passive control group. Except for two studies (Bak et al., 2016; Wong et al., 2019), most previous studies had no follow-up period (Berggren et al., 2020; Bubbico et al., 2019; Kliesch et al., 2022; Long et al., 2020; Meltzer et al., 2021; Pfenninger & Polz, 2018; Ramos et al., 2017; Valis et al., 2019; Ware et al., 2017).

Age of participants

Most studies exclusively involved elderly persons whose ages ranged from 55 to 90 years overall. Only Bak et al. (2016) and Long et al. (2020) included younger adults resulting in a wide age range from 18 to 85 years across both studies. According to Bak et al. (2016), age had no impact on improvements in attentional shifting. Long et al. (2020) also found no proof that language learning had age-dependent benefits for attentional shifting and selective attention. Nevertheless, the applicability of these findings to older persons needs to be approached with care.

Sample size

The studies also varied considerably in sample size. The two pilot studies by Pfenninger and Polz (2018) and Ware et al. (2017) were the smallest, with 12 and 14 participants, respectively. In Bubbico et al. (2019), only 14 subjects participated in the language learning arm and 12 in the passive control condition. Most studies with a control group included approximately 20 subjects in each group (Bak et al., 2014; Kliesch et al., 2022; Meltzer et al., 2021; Ramos et al., 2017; Valis et al., 2019). The most meaningful trials in terms of sample size were those of Berggren et al. (2020) and Wong et al. (2019), with a total of 160 individuals divided into two groups and 163 people divided into three groups, respectively.

Foreign languages taught

Concerning languages taught, most studies chose English (Bubbico et al., 2019; Pfenninger & Polz, 2018; Valis et al., 2019; Ware et al., 2017; Wong et al., 2019), and two Scottish Gaelic (Bak et al., 2016; Long et al., 2020). Others opted for Romanic languages such as Spanish (Kliesch et al., 2022; Meltzer et al., 2021), or Italian (Berggren et al., 2020). The choice was often based on the ease of learning, e.g., English, Spanish, or Italian, but also the personal relevance of the language to study participants. For example, in the study by Ramos et al. (2017), participants spoke Spanish but lived in the Basque Country. Therefore, learning Basque was beneficial for them.

Previous foreign language skills

Studies varied widely regarding the level of prior experience allowed in the foreign language taught. Some deliberately included different levels of previous knowledge (Long et al., 2020) or considered various courses according to prior experience (Bak et al., 2016; Pfenninger & Polz, 2018; Valis et al., 2019). For example, in Ware et al. (2017), subjects with beginner, intermediate, and advanced English skills were taught in one class. Some permitted prior experience while requiring that participants could still be classified as beginners (Berggren et al., 2020; Bubbico et al., 2019). Only in the studies by Ramos et al. (2017), Wong et al. (2019), and Meltzer et al. (2021) subjects were complete novices.

Only one study employed a structured diagnostic instrument to measure previous foreign language competency in the language being taught (Valis et al., 2019). In other studies, the language level was determined by a researcher (Ware et al., 2017) or a teacher (Bubbico et al., 2019). Berggren et al. (2020) permitted self-evaluation of knowledge as non-existent or extremely poor.

In most studies, no information on foreign language skills other than the language of tuition was provided (Berggren et al., 2020; Valis et al., 2019; Ware et al., 2017; Wong et al., 2019). Two studies allowed no more than school knowledge in any other language (Kliesch et al., 2022; Pfenninger & Polz, 2018). In Bak et al. (2016) and Long et al. (2020), previous foreign language proficiency in other languages was self-determined using a standardized questionnaire. However, no information was provided on the degree of monolingualism or bilingualism within the cohorts. Only two studies included only monolinguals. In Ramos et al. (2017), no prior experience in learning a second language was allowed. In the study by Kliesch et al. (2022), foreign language skills were self-rated based on a standardized language questionnaire. Additionally, no previous foreign language study in the last ten years was permitted.

Duration and frequency of foreign language learning interventions

Foreign language interventions ranged in duration from 1 week (Bak et al., 2016; Long et al., 2020) to 30 weeks (Kliesch et al., 2022) and 8 months (Ramos et al., 2017). Low-frequency courses were only offered once a week for around 2 h (Bubbico et al., 2019; Valis et al., 2019; Ware et al., 2017). Medium-frequency courses were held three times per week with a total of approximately 5 h of teaching per week (Berggren et al., 2020; Pfenninger & Polz, 2018; Ramos et al., 2017; Wong et al., 2019) or five times per week with 30 min each (Meltzer et al., 2021). High-frequency courses comprised 14 h of formal foreign language instruction spread across multiple weekdays, as well as evening leisure events in the target language (Bak et al., 2016; Long et al., 2020).

Longer courses were not necessarily related to cognitive improvements (Kliesch et al., 2022; Ramos et al., 2017). On the contrary, even 1 week classes of high frequency (Bak et al., 2016; Long et al., 2020) or a 4 week course of medium frequency (Pfenninger & Polz, 2018) could prove cognitively beneficial. In contrast to studies with high-frequency courses (Bak et al., 2016; Long et al., 2020), neither of the studies with low-frequency classes showed significant cognitive benefits (Bubbico et al., 2019; Valis et al., 2019; Ware et al., 2021). Thus, it appears that more frequently held, albeit shorter, interventions, e.g., 5 h of training per week spread across several days per week (Bak et al., 2016; Long et al., 2020; Meltzer et al., 2021; Pfenninger & Polz, 2018) are more likely to be cognitively advantageous than longer but less frequently held ones.

Cognitive domains and tasks

Prior research on cognitive outcomes of foreign language learning has explored various cognitive functions, including attention, EF, intelligence, memory, and global cognition. Most studies that reported significant improvements in cognition employed tasks that measured aspects of executive attention (Melrose et al., 2017), e.g., the Stroop task (Meltzer et al., 2021; Pfenninger & Polz, 2018), selective attention (Long et al., 2020), and attentional shifting (Bak et al., 2016; Long et al., 2020). In addition, Meltzer et al. (2021) found a significant effect on updating by utilizing an n-back task paradigm.

However, it should be emphasized that one out of the four studies named in the paragraph above implemented merely a quasi-experimental design (Bak et al., 2016). Another study only piloted data and had no control condition, making it impossible to rule out practice effects (Pfenninger & Polz, 2018). Also, Long et al. (2020) did not include a passive control group. Another study that demonstrated cognitive advantages of learning a foreign language was the study of Wong et al. (2019). As mentioned above, foreign language learners improved in global cognition and updating. However, compared to an active and passive control group, which exhibited no significant enhancements, there was still no significant interaction effect.

Also, the researchers could not detect cognitive improvements in spatial and verbal intelligence, (associative) memory (Berggren et al., 2020; Bubbico et al., 2019), (sustained) attention (Bak et al., 2016; Long et al., 2020; Wong et al., 2019), alertness (Kliesch et al., 2022), information processing speed (Bubbico et al., 2019), selective attention (Bak et al., 2016), and divided attention (Kliesch et al., 2022). Additionally, some studies found also no improvements in domains of EF, including inhibition using the Simon task (Meltzer et al., 2021; Wong et al., 2019), task switching (Bubbico et al., 2019; Ramos et al., 2017), updating (Berggren et al., 2020; Kliesch et al., 2022; Wong et al., 2019) and verbal fluency (Bubbico et al., 2019; Kliesch et al., 2022), or in global cognition (Bubbico et al., 2019; Valis et al., 2019; Ware et al., 2017).

To summarize, the available evidence to date is rather equivocal. There is, nevertheless, modest evidence of an advantage of foreign language acquisition in executive attention and updating. However, most studies revealed little to no cognitive gains, particularly in intelligence, memory, and attention, without executive components such as sustained attention, alertness, and information processing speed.

Control and investigation of confounding factors

As cognitive decline is a fundamental part of the normal cognitive aging process but can transition into pathological aging (see chapter 1.2), global cognitive status at baseline should be carefully evaluated. However, most studies did not ensure that only older persons without signs of cognitive impairment were included or that participants with suspected cognitive impairment were excluded (Bak et al., 2016; Long et al., 2020; Pfenninger & Polz, 2018; Ramos et al., 2017; Ware et al., 2017). In contrast, Wong et al. (2019) only included participants who scored 0 or 0.5 on the Clinical Dementia Rating Scale (CDR; Morris, 1993). Other studies accounted for cognitive decline by defining cut-off scores in cognitive screening instruments such as the Mini-Mental Status Examination (MMSE; Berggren et al., 2020; Folstein et al., 1975) or the Montreal Cognitive Assessment (MoCA; Kliesch et al., 2022; Meltzer et al., 2021; Nasreddine et al., 2005; Valis et al., 2019). However, every study established different cut-off scores. While Kliesch et al. (2022) and Valis et al. (2019) only considered subjects with a MoCA score of 26 or higher, Meltzer et al. (2021) set a meager minimum score of 21, which is far below the original author's threshold of 26 (Nasreddine et al., 2005). Bubbico et al. (2019) also excluded individuals with suspected cognitive decline following the screening, however, without specifying a cut-off score.

Educational background was another factor taken into account, as in Long et al. (2020), who only included subjects with university degrees. They also found that individuals with lower Gaelic skills improved more in selective attention and attentional switching than individuals with more advanced proficiency. Despite observing no general increase in cognition, Kliesch et al. (2022) found language learners with poorer cognition at baseline to outperform those in the active and passive control group on most cognitive measures. However, this effect was only visible in the first 20 weeks of the language course.

"To conclude, by now, there is, at best, weak evidence that foreign language learning may improve cognitive functioning in healthy older adults" (Grossmann et al., 2023, p. 3). To explain conflicting results, one must consider that all previous studies differed considerably in the aforementioned variables. A direct comparison of the available studies is, therefore, hardly possible. Thus, the evidence to date is insufficient to draw significant conclusions about the appropriate dose of the intervention, variables to be explored, acceptable prior foreign language knowledge, or other affecting factors to be accounted for. For example, low-frequency studies (Bubbico et al., 2019; Valis et al., 2019; Ware et al., 2017) solely used cognitive screening measures of global cognition, such as the MMSE or the MoCA, as their only or primary outcome. These tests are less sensitive for discrete cognitive gains, especially in older adults without cognitive impairment, as even older adults often already perform at near-to-ceiling levels in these tests. It is, therefore, difficult to state if the choice of outcome, the low intensity of the course, or both contributed to these studies' failure to find cognitive benefits following foreign language training.

Based on the findings of previous studies and the theoretical framework, executive attention may be a promising outcome affected by foreign language learning. Moreover, consistent with findings from related research disciplines, training frequency may be more decisive than course length. Since research remains far from determining cognitive characteristics altered by learning a foreign language and the circumstances under which these alterations occur, further controlled explorative studies are needed before large-scale studies can be performed.

Table 1. Characteristics and main cognitive outcomes of foreign language learning studies including healthy older adults

Reference	Design	Mean age (range) in years	Intervention	Duration	Cognitive domains and tasks	Results
Bak et al. (2016)	Q	49.14 (18–78)	FLL (Scottish Gaelic) (*n* = 33) Different activities (active control) (*n* = 16) Passive control (*n* = 18)	1 week: 14 h in total (+ Gaelic entertainment in the evening)	Attention: TEA (sustained attention, selective attention, attentional switching)	Significant age-independent improvement of the FLL in the TEA attentional switching ($p < .001$). No significant improvement in the control conditions ($p = .13$). In the FLL an improved performance compared to baseline was evident at 9-month follow-up for all participants who kept practicing no less than 5 h a week after the classes ended.
Berggren et al. (2020)	RCT	69.35 (65–75)	FLL (Italian) (*n* = 90) Relaxation training (active control) (*n* = 70)	11 weeks: 2 times/week 2.5 h each	Spatial intelligence: Ravens progressive matrices, WASI-II Matrices Verbal intelligence: Analogies, Syllogisms, and Verbal Inference Updating: Numerical updating, n-back Memory: Item-associative memory tasks (word-word, face-name, picture-picture)	No proof that acquiring a foreign language enhances cognition in older individuals relative to relaxation training. In each group, the mean change in task performance between pre and post assessment was not significant.

Table 1. continued, page 2/4

Reference	**Design**	**Mean age (range) in years**	**Intervention**	**Duration**	**Cognitive domains and tasks**	**Results**
Bubbico et al. (2019)	RCT	67.75 (59–79)	FLL (English) (*n* = 14) Passive control (*n* = 12)	16 weeks: 1 time/week 2 h each	Global cognition: MMSE Executive Functions: Frontal Assessment Battery Shifting: TMT-B – A Verbal fluency: F-A-S Test Attention: TMT-A Memory: Babcock Memory test	Significant between-groups difference over time (p = .01) in the MMSE. However, this difference was due to decreased performance in the control group (M_{Pre} = 29.35, M_{Post} = 28.28), while the intervention group remained stable over time (M_{Pre} = 27.23, M_{Post} = 27.81).
Kliesch et al. (2022)	Q	68.40 (64–75)	Combined computerized and face-to-face FLL (Spanish) (*n* = 28) Strategy game training (active control) (*n* = 17) Passive control (*n* = 16)	30 weeks: 2-5 h/week	Updating: 2-back task, operation span task Attention: Divided attention Verbal fluency: Letter fluency	No evidence of a cognitive improvement in the FLL in comparison to the active or the passive control group. However, individuals in the FLL with lower baseline cognition increased their performance more than individuals from either control group in most cognitive measures. The effect was strongest for WM capacity.

Table 1. continued, page 3/4

Reference	Design	Mean age (range) in years	Intervention	Duration	Cognitive domains and tasks	Results
Long et al. (2020)	Q	? (21–85)	FFL (Scottish Gaelic) (*n* = 105)	1 week: 14 h in total (+ Gaelic entertainment in the evening)	Attention: TEA (sustained attention, selective attention, attentional switching)	Significant improvement in selective attention and attentional switching ($p < .001$). Those with lower levels of Gaelic improved the most ($p < .01$). No hint that the effects of studying of foreign language on attention depended on age.
Meltzer et al. (2021)	Pseudo RCT	69.87 (65–75)	Computer-based FFL (Spanish) (*n* = 28) BrainHQ (active control) (*n* = 24) Passive control (*n* = 24)	16 weeks: 5 times/week 30 minutes each	Updating: N-back Inhibition: Simon task, Stroop task	Significant improvement in the 2-back condition of the n-back task ($p < .02$) and in the color-word condition of the Stroop task ($p = .01$) for both the FFL and the BrainHQ group relative to the passive control. The BrainHQ group experienced a more pronounced improvement in overall processing speed than the FLL group in comparison to the passive control.
Pfenninger and Polz (2018)	Q (pilot)	? (63–89)	FLL (English) (*n* = 12)	4 weeks: 3 times/week 2 h each	Inhibition: Stroop Task Attention: A-K-T	Significant improvement in the Stroop task ($p < .05$).
Ramos et al. (2017)	Q	68.3 (60–80)	FLL (Basque) (*n* = 26) Passive control (*n* = 17)	8 months: 3 times/week 2x2 h, 1x1.5 h/week	Shifting: Color-shape switching task	No significant improvement ($p > .05$).

Table 1. continued, page 4/4

Reference	Design	Mean age (range) in years	Intervention	Duration	Cognitive domains and tasks	Results
Valis et al. (2019)	RCT	70.9 (55+)	FLL (English) (*n* = 20) Passive control (*n* = 22)	12 weeks: 1 time/week 135 min each	Global cognition: MoCA	No significant improvement (p = .46).
Ware et al. (2017)	Q (pilot)	75.42 (63–90)	Computer-assisted FLL (English) (*n* = 14)	17 weeks: 1 time/week 2 h each	Global cognition: MoCA	No significant improvement (p = .67).
Wong et al. (2019)	RCT	70.97 (60–85)	Computer-based FLL (English) (*n* = 53) Computer games (*n* = 51) Music appreciation (passive control) (*n* = 49)	26 weeks: 2 or 3 times/ week 1.5–2.5 h each (5 h/week)	Global cognition: ADAS-Cog Updating: Auditory Reading Span Task, Digit Span Backwards Inhibition: Simon task Attention: ANT, Digit Span Forwards	Significant improvement in the ADAS-Cog (FLL: p < .001; Games: p < .001). No significant improvement for the passive control group. However, no significant interaction between groups. Improvements remained stable at a 3-month follow-up. The FLL group also showed significant improvement in the Auditory Reading Span Task and the Digit Span Backwards.

Note. ADAS-Cog = Alzheimer's Disease Assessment Scale Cognitive Subscale; A-K-T = Geriatric Concentration Test; ANT = Attention Network Test; FLL = Foreign language learning; MMSE = Mini-Mental Status Examination; MoCA = Montreal Cognitive Assessment; Q = Quasi-experimental; RCT = Randomized controlled trial; TEA = Test of Everyday Attention; TMT-B/-A = Trail-Making Test part B/part A.

2. Aims and research questions

The overall aim of this doctoral thesis was to investigate whether and to what degree learning a foreign language affects cognition in healthy older adults, given the research gaps discussed in the previous chapter. Considering the theoretical background and results of some of the previous studies, cognitive effects of foreign language learning in older age might arise most readily in tasks that measure aspects of executive attention. Thus, in this dissertation, a randomized controlled trial (RCT) was designed defining measures of executive attention as the primary outcomes. Additionally, a larger variety of EF tasks was included as secondary outcomes given the ongoing lack of consensus on the specific cognitive processes underlying foreign language learning.

This chapter – apart from the exact hypotheses statements – has been published in *Frontiers in Aging Neuroscience* in the scientific article “ Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial”. The extracted parts are marked by indentations. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185.

The hypotheses were established as follows:

- **Hypothesis 1:** Foreign language learners will improve their cognitive performance in executive attention and EF compared to a passive control group immediately after taking a foreign language class.
- **Hypothesis 2:** Foreign language learners will improve their cognitive performance in executive attention and EF compared to a passive control group 3 months after taking a foreign language class.

As this study is exploratory, we were also interested in factors influencing cognitive progress among foreign language learners. (…) As higher levels of CR are associated with better cognitive performance, the intervention might more likely introduce a significant improvement in cognition in those with lower levels of CR. The same might apply to previous foreign language proficiency and

> usage. Even though our study included only monolinguals, almost every older adult in Germany learned at least one language in school or later in life through courses. Therefore, we considered it important to analyze the possible role of prior foreign language knowledge skills and usage. We assumed that individuals who are less familiar with foreign languages might benefit more from foreign language learning because their brains are less adapted to cognitive control mechanisms of foreign language acquisition and usage than those of individuals with more previous foreign language experience. (p. 3)

Hypotheses 3 and 4 were defined as follows:

- **Hypothesis 3:** The level of CR will predict cognitive change in executive attention and EF in foreign language learners.
- **Hypothesis 4:** Previous foreign language experience will predict cognitive change in executive attention and EF in foreign language learners.

> In deviation from the original study protocol published (Grossmann et al., 2021), we added the baseline level of general cognition as a predictor, following more recent findings by Kliesch et al. (2022) (…). In their study, the authors found that language learners with lower baseline cognition improved more on cognitive tasks than those in an active or passive control condition during the first 20 weeks of the intervention. Their finding is in line with the CR hypothesis. As individuals with lower cognitive performance are at higher risk for dementia (Valenzuela & Sachdev, 2006), these individuals may also have more room for improvement when engaging in cognitively stimulating activities. (p. 3)

We formulated hypothesis 5 as follows:

- **Hypothesis 5:** The level of global baseline cognition will predict cognitive change in executive attention and EF in foreign language learners.

3. Methods

Parts of the following chapters, as well as figures and tables on methods, have previously been published by my co-authors and me in the scientific article "Grossmann, J. A., Koelsch, V. M., Degirmenci, M. G., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2021). Effects of foreign language learning on executive functions in healthy older adults: Study protocol for a randomised controlled trial. *BMC Geriatrics. 21*(1), Article 122, 1–14". These parts comprise chapters 3.2, 3.3, 3.4, 3.5, 3.6, 3.8, 3.9, and 3.10 (Grossmann et al., 2021, pp. 3–10). To use tenses consistently in the methods section, I have changed the future tense to the past tense (as shown in brackets) from the original publication. I also adapted the numbering and labeling of headlines, figures, and tables, as well as the citation style to the format of this dissertation. It should be noted that the referenced chapters, tables, and figures are in British English, as in the original publication, while the rest of this dissertation is in American English. The original publication is available at http://dx.doi.org/10.1186/s12877-021-02051-x.

Chapters 3.1, 3.7, and 3.11 of the methods section have been published in *Frontiers in Aging Neuroscience* in the scientific article "Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial, pp. 3–6". The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185. All changes from the original publications are indicated in brackets.

3.1 Study design and setting

The present study was a randomized controlled superiority trial with two parallel groups to investigate the effects of a 3 week foreign language course on executive attention and EF in healthy community-dwelling older adults. We randomly assigned participants in a 1:1 ratio to one of the two study arms: a language learning group (LLG) and a waiting list control group (WLCG). Data were collected at the Network Aging Research (NAR), Heidelberg University, Germany, and the SRH University

Heidelberg, Germany. The study protocol followed the Consolidated Standard of Reporting Trials (CONSORT) statement (Moher et al., 2010).

[Ethical approval was obtained from the Ethics Committee of the Faculty of Behavioural and Empirical Cultural Studies of the Ruprecht Karl University of Heidelberg, reference number 2018/1-2, on 10 January 2019. All methods were performed in accordance with the relevant guidelines and regulations of the ethics approval. Participants gave written informed consent prior to participation.]

Due to the COVID-19 pandemic outbreak, the trial had to be temporarily suspended in March 2020. Initially, we aimed to resume recruitment as soon as possible. However, given the ongoing COVID-19 pandemic and the "at-risk" population in terms of age enrolled in our study, we decided to end our trial early in July 2020. We believe that the study participants' overall well-being and best interest should be prioritized.

3.2 Recruitment

Participants [were] recruited from the general public in Heidelberg, Germany, and the surrounding regions through advertisements in local newspapers, advertising in trains, at public lectures, and on the NAR homepage. We (…) also inform[ed] individuals who agreed to be notified about ongoing studies via an e-mailing list of the NAR. Additionally, we (...) distribute[d] flyers and posters in senior and sports centres, language schools, medical practices, and pharmacies. [For logistical reasons, the study team collected data in three waves. In each wave, an equal proportion of subjects were enrolled in both study arms. (Grossmann et al., 2023, p. 4)]

3.3 Eligibility criteria

Eligibility criteria for participation are listed in **Table 2** [emphasis added].

3.4 Interventions

This study comprise[d] two trial arms: a [*sic*] LLG and a WLCG.

Table 2. Eligibility criteria for participation

Inclusion criteria	
-	Aged between 65 and 80 years
-	Native language German
-	Living independently
-	Motivated to participate; Signing written informed consent for study participation confirmed by the local ethics committee
Exclusion criteria	
-	Spanish level ≥ A1.1 according to the Joint European Reference Frame for Languages or having already attended a regular Spanish course for beginners (≥ 18 h of teaching)
-	Bilingualism or multilingualism (def.: fluent command [level ≥ C1 as defined by questions derived from the Joint European Reference Frame for Languages] and frequent usage of any foreign language)
-	Any Romanic language (French, Italian, Latin, Portuguese) level ≥ B1 (as defined by questions derived from the Joint European Reference Frame for Languages)
-	Cognitive functions below cut-off (Cognitive Functions Dementia [CFD; Jahn & Heßler, 2017; Schuhfried: no subtest z ≤ -1.5 [R. C. Petersen et al., 1999]) according to age and where possible additionally to sex and education
-	Impaired/not-corrected vision
-	Colour blindness
-	Wearing a hearing aid or impaired sense of hearing as measured by the whispered voice test (Swan & Browning, 1985)
-	Part-time employment with 20 or more hours of working activity per week
-	Self-reported former or current neurological disease (e.g. stroke, mild cognitive impairment, dementia, Parkinson's disease, epilepsy, multiple sclerosis, encephalitis, cerebral tumour)
-	Self-reported current/diagnosed mental health disorder (e.g. anxiety disorder, major depression, schizophrenia, alcoholism or other addiction)
-	Transient loss of consciousness for more than five minutes
-	Current musical activity for more than five hours per week
-	Psychotropic medication within the last six months before the start of the study
-	Surgery within the last month before the start of the study
-	Other constraints hindering attendance at both the assessments and intervention appointments

3.4.1 Language learning group (LLG)

The LLG (…) stud[ied] Spanish in a three-week course for beginners. Daily lessons of 90 min [were] scheduled in the morning from Monday until Friday. In sum, participants (…) receive[d] 7.5 h of teaching per week, resulting in a total course duration of 22.5 h. The idea of a relatively short but high-intensity training [stemmed] from a study by Bak et al. (2016), who found a positive effect of foreign language learning on attentional switching after only 1 week of 14 h of foreign language instruction.

A qualified teacher [gave] the lessons face-to-face at a language school in the centre of Heidelberg, which is easily accessible by car and public transportation. Beyond attending regular classes, participants [were allowed to] do homework and (...) to practice at home to consolidate the newly learned content. Following the example of previous studies (Kliesch et al., 2018; Ramos et al., 2017), we (…) limit[ed] the group size to a maximum of ten participants per group, to enable the teacher to address all participants and to give them sufficient opportunity to participate in class. A commonly used work book in adult education (Goerrissen, 2016) (...) serve[d] as teaching material. By the end of the course, approximately three chapters [were] completed. Participants (…) learn[ed] accent and pronunciation as well as grammatical rules of the Spanish language and (...) acquire[d] elementary communication skills in various topics (e.g. introducing themselves, asking for directions).

3.4.2 Waiting list control group (WLCG)

The WLCG [did] not receive any treatment during the intervention phase but [took] part in a three-week waiting period during which subjects (…) follow[ed] their usual daily routine. To reduce barriers to participation due to possible group preferences (Mills et al., 2006), the WLCG (…) attend[ed] a control group programme after their study completion. The programme consist[ed] of the same language course intervention as the LLG and an additional voluntary examination after the end of the course. This appointment [was] not included in the main study design as the scope for participation should differ as little as possible between groups.

3.5 Outcomes

Baseline variables, as well as primary, secondary, and language course outcome measures are presented in **Figure 2** [emphasis added]. Unless otherwise stated, primary and secondary outcome measures are valid and reliable tests from the Vienna Test System (VTS; Schuhfried). The VTS is a computerised test system for the assessment of neuropsychological functions and has become a well-established tool in clinical practice. Each cognitive outcome [began] with an instruction phase, followed by a practice phase, in which participants (…) receiv[ed] immediate feedback on their errors. If necessary, the practice phase [could] be repeated to ensure a sufficient understanding of the task.

3.5.1 Baseline variables

Baseline variables include[d] socio-demographic variables such as date of birth, gender, marital and occupational status, as well as an assessment of global cognitive functions. Additionally, cognitive reserve, foreign language knowledge and usage, and handedness [were] examined.

Cognitive reserve

We (…) evaluate[d] cognitive reserve using the Cognitive Reserve Index questionnaire (CRIq; Nucci et al., 2012). The CRIq is a valid and reliable half-structured interview, which quantifies cognitive reserve over an individual's lifetime using three established measures of cognitive reserve: education (CRI-Education), occupation (CRI-WorkingActivity), and engagement in cognitively stimulating activities (CRI-LeisureTime; Opdebeeck et al., 2016). The three sub-scores can be combined into an overall score (CRI-Index). All measures are age-adjusted to allow for comparisons between different age groups.

Figure 2. SPIRIT figure for the schedule of enrolment, interventions, and assessments

		STUDY PERIOD								
		Enrolment		Allocation	Post-allocation					
	TIMEPOINT	$-t_2$	$-t_1$	0	t_1	t_2	t_3	t_4	t_5	t_6
	ENROLMENT									
	Eligibility screening	◆——	——◆							
	Informed consent	X								
	Allocation			X						
	INTERVENTIONS									
	LLG					X				
	WLCG								X	
	ASSESSMENTS									
	Baseline variables									
	Foreign language knowledge (LSBQ)	X	X							
	Socio-demographic variables	X	X							
	Cognitive functions (CFD)		X							
	Cognitive reserve (CRIq)		X							
	Handedness (FLANDERS)		X							
	Outcome variables									
P	Divided attention (WAFG)				X		X	X		X
P	Inhibition (STROOP)				X		X	X		X
S	Alertness (WAFA)				X		X	X		X
S	Attention span (DSF)				X		X	X		X
S	Shifting (SWITCH)				X		X	X		X
S	Shifting (TMT-B – TMT-A)				X		X	X		X
S	Information processing speed (TMT-A)				X		X	X		X
S	Inhibition (INHIB)				X		X	X		X
S	Semantic verbal fluency (WIWO)				X		X	X		X
S	Lexical verbal fluency (WIWO)				X		X	X		X
S	Updating (DSB)				X		X	X		X
S	Updating (NBV)				X		X	X		X
	Adherence					X			X	
	Homework and learning time					X			X	
	Course evaluation					X			X	
	Vocabulary test					X			X	

Note. P = primary outcome measure; S = secondary outcome measure; -t2 telephone screening; -t1 = face-to-face screening; t1 = pre-assessment, t2 = language course/waiting period, t3 = post-assessment, t4 = 3-month follow-up assessment, t5 = language course, t6 = 4-month follow-up assessment (waiting list control group only). Abbreviations: CFD = Cognitive Functions Dementia; CRIq = Cognitive Reserve Index questionnaire; DSB = Digit Span Backwards; DSF = Digit Span Forwards; FLANDERS = Flinders Handedness Survey; INHIB = Response Inhibition; LLG = Language learning group; LSBQ = The Language and Social Background Questionnaire; NBV = N-Back Verbal; STROOP = Stroop Interference Test; SWITCH = Task Switching; TMT-A/-B = Trail-Making Test – Langensteinbach Version part A/part B; WAFA = Perception and Attention Function Battery – Alertness; WAFG = Perception and Attention Function Battery – Divided Attention; WIWO = Vienna Verbal Fluency Test; WLCG = Waiting list control group.

Foreign language knowledge and usage

Foreign language knowledge and usage [were] determined using the Language and Social Background Questionnaire (LSBQ; J. A. E. Anderson et al., 2018). The LSBQ is a validated and reliable questionnaire for self-assessment of foreign language skills and usage in various contexts. A composite factor score quantifies the degree of bilingualism on a continuum from clearly monolingual to highly bilingual.

Global cognitive functions

We (…) assess[ed] global cognitive functioning using the Cognitive Functions Dementia test set (CFD, test form: S1 [touchscreen operation]; Jahn & Heßler, 2017; Schuhfried) to exclude participants with suspected cognitive decline according to Petersen's criteria (no subtest $z \leq -1.5$; R. C. Petersen et al., 1999).

The CFD is a comprehensive computer-based assessment from the VTS and takes about 60 min to perform. Previous studies in this field of research only used brief screening measures such as the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) or the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). Both instruments can distinguish pathological cognitive decline from healthy ageing. However, they are less accurate than comprehensive multidimensional neuropsychological inventories (Roalf et al., 2013). In the CFD, cognitive performance is measured by eleven tasks falling into five cognitive domains: attention, verbal long-term memory, EF, expressive speech, and perceptual motor functions.

Handedness

The Flinders Handedness survey (FLANDERS; Nicholls et al., 2013) is a short standardised questionnaire to measure hand preference. It contains ten items describing ten different activities (e.g. writing). Participants are asked to indicate whether they prefer the left, the right or either hand for each activity.

3.5.2 Primary outcomes

The Stroop Interference Test (STROOP, test form: S7; Schuhfried, 1999) and the Divided Attention, a subtest of the Perception and Attention Function Battery (WAFG, test form: S3; Sturm, 2006b), [were] the two primary outcome measures.

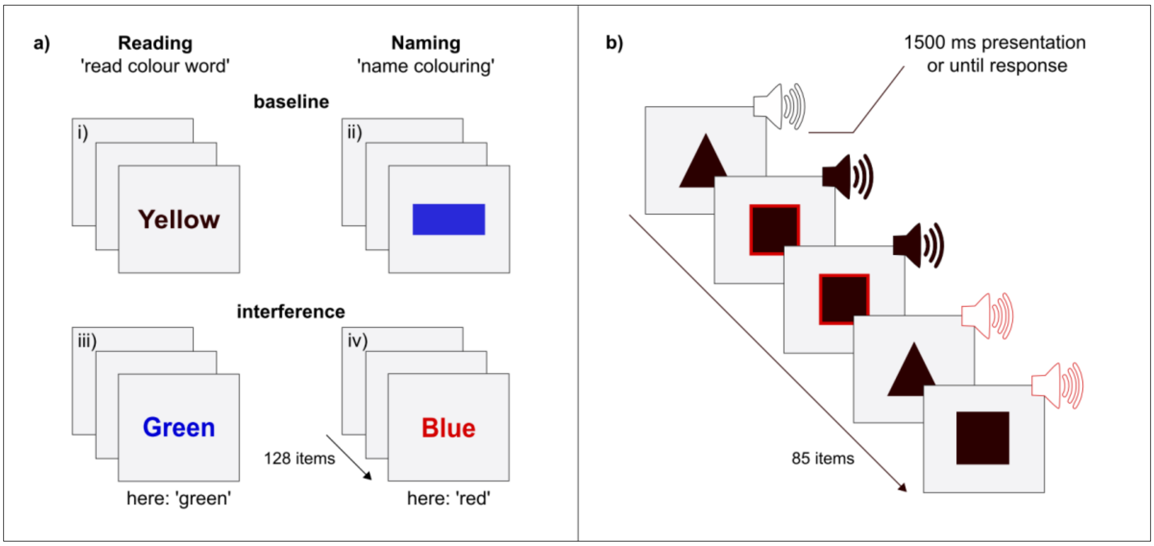
Stroop Interference Test (STROOP)

The STROOP is a sensory-motor speed test [(and) a valid, reliable, and widely used measure of selective attention, representing the monitoring aspect of executive attention (Melrose et al. [2017], as cited in Grossmann et al. [2023])]. In a small pilot study older adults performed significantly better in this task after having participated in a foreign language training (Pfenninger & Polz, 2018). The task consists of two baseline and two interference conditions (see **Figure 3a**) [emphasis added]. In the (i) reading baseline condition, participants read one of four German colour words (blue, green, red, yellow) printed in black one after the other on a computer screen. Similarly, in the (ii) naming baseline condition, subjects see coloured bars (blue, green, red, yellow). The task is to select the corresponding colour on the response panel as quickly as possible. In the two interference conditions, participants see German colour words printed in a different colour (e.g. the written word 'yellow' printed in red). In the (iii) reading interference condition, participants must read the word and ignore the colouring. Conversely, in the (iv) naming interference condition, participants must respond to the colouring and ignore the meaning of the word. Reactions in the two interference conditions are usually slower since the processing of the two simultaneously presented stimuli (the written word and the colour in which it is printed) requires attention. This leads to delayed processing and thus to longer reaction times. Each test part consists of 128 stimuli, which are presented until a response is given. Immediately afterwards, the next stimulus appears. The total test duration is approximately 15 min. The 'naming interference tendency' [was] the primary outcome variable. It [was] calculated from the difference between 'median of reaction times – naming interference condition' and 'median of reaction times – naming baseline'. Using the median reaction time ensures that asymmetries caused by circumstances, such as 'getting stuck' on certain items, do not distort results. The reliability of the naming interference tendency measure is $\alpha = .97$.

Divided Attention (WAFG)

The WAFG is a measure of cross-modal (visual/auditory) divided attention [and represents the switching aspect of executive attention (S. E. Petersen and Posner [2012]; van Zomeren and Brouwer [1994], as cited in Grossmann et al. [2023])]. In the WAFG, subjects are simultaneously confronted with a visual and an auditory

Figure 3. Illustration of primary outcomes, a) the Stroop Interference Test (STROOP) and b) the Divided Attention (WAFG)



Note. White speakers represent the high-pitched sound. Black speakers depict the low-pitched sound. Squares or speakers outlined in red indicate the targets of the visual and the auditory channels. These targets are either two squares or two high-pitched sounds which immediately follow each other.

channel (see **Figure 3b**) [emphasis added]. On the former one, either a triangle or a square appears on a computer screen. On the auditory one, either a low- or a high-pitched tone is emitted. The task is to press a key on the response panel as quickly as possible when either two squares or two high-pitched tones immediately follow each other. The total number of items is 85, of which 21 are relevant. Stimuli are presented until a response is given or after a maximum of 1500 ms followed by an inter-stimulus interval of 1000 ms. If no reaction has occurred within the presentation period, an omission error is counted. Reaction times below 100 ms [were] not (...) considered, as physiological responses below 100 ms are not possible (Di Martino et al., 2008). The total test duration is 9 min. The outcome variable [was] the logarithmic mean reaction time, which accounts for the expected skewness in the distribution of reaction times. The WAFG, test form S3, obtains a reliability score of α = .84.

3.5.3 Secondary outcomes

[As secondary outcomes, performance on a broad range of tasks from the three core domains of EF - inhibition, shifting, and updating – was assessed (Miyake et al. [2000], as cited in Grossmann et al. [2023]).] (…) For comparison reasons, we (...) also assess[ed] verbal fluency and non-executive components of attention. As these predominantly refer to linguistic and attentional functions, they are less likely to load on EF (Shao et al., 2014; Suchy, 2009). Details of the respective tasks for each domain of EF and the corresponding dependent variables are presented in **Table 3** [emphasis added]. Every domain [was] measured by at least two tasks to better capture each construct (Schmiedek et al., 2014).

Inhibition

Inhibition is the ability to suppress unwanted reactions. It [was] assessed using the Response Inhibition (INHIB; Kaiser et al., 2010) and the 'reading interference tendency', which is another outcome measure of the STROOP, as already described in the 'Primary outcomes' section.

Shifting

Shifting refers to the ability to switch flexibly between tasks or mental sets. We (…) measure[d] shifting using the SWITCH (Gmehlin et al., 2012) and the Trail-Making Test – Langensteinbach Version part B minus part A (TMT-B – TMT-A; Rodewald et al., 2012).

Updating

Updating (as a core working memory process) reflects the ability to maintain and continuously update information. It [was] evaluated using the Digit Span Backwards (DSB) from the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Petermann, 2012) and the N-Back Verbal (NBV; Schelling & Schuri, 2009).

Verbal fluency

Verbal fluency is defined as a person's ability to find words of a specific characteristic in the mental lexicon (Shao et al., 2014). To assess verbal fluency, we (…)

conduct[ed] the Vienna Verbal Fluency Test (WIWO; Jahn, 2016). The test distinguishes between two dimensions: semantic and lexical verbal fluency.

Attention

Attention [was] represented by a measure of alertness, a basal process of short-term attention activation. It [was] measured using the Alertness, a subtest of the Perception and Attention Function Battery (WAFA; Sturm, 2006a). Attention span and information processing speed [were] assessed using the Digit Span Forwards (DSF) from the WAIS-IV (Petermann, 2012) [, which can be considered as a measure of attention besides short-term memory span (Cambridge Cognition (2023), as cited in Grossmann et al. (2023))] and the TMT-A (Rodewald et al., 2012), respectively.

3.5.4 Language course outcomes

We (…) collect[ed] the following language course outcomes to evaluate the intervention:

Adherence

Adherence to the language course [was] documented daily by the Spanish language teacher. The teacher (…) inform[ed] the study team if a participant [did] not show up for class without giving prior notice.

Homework and learning time

We (…) collect[ed] information about the time spent on additional learning activities in Spanish at home. In a table, participants (…) indicate[d] in minutes, how much time they invested each day, including at weekends, in doing homework and voluntary learning activities in Spanish. Records (…) start[ed] from the beginning of the course and (…) end[ed] on the last day of the course. The outcome [was] the total time spent on additional learning activities at home.

Course evaluation

A questionnaire [was] completed anonymously on the last day of the course to evaluate the perceived quality and acceptability of the intervention. On a four-point Likert

Table 3. Overview of secondary outcomes

Domain	Task	Test form	Response mode	Description	Dependent variable
Inhibition	INHIB (Kaiser et al., 2010)	S3, (go/no-go)	Response panel	A series of circles and triangles is displayed in succession on the screen. Whenever a triangle appears, the subject must press a green button. Circles do not require a response and occur rarely, whereas triangles appear frequently. This builds a dominant response tendency on the triangles. Hence, reactions to circles require inhibition. A total of 250 stimuli are presented, comprising 202 triangles and 48 circles.	Number of commission errors (number of false reactions to circles)
	STROOP (Schuhfried, 1999)	S7	Response panel	The task is described in detail in the primary outcomes section.	Reading interference tendency
Shifting	SWITCH (Gmehlin et al., 2012)	S1	Response panel	A sequence of triangles and circles appears one by one on the screen. The figures are either dark or light grey. The task is to categorise each stimulus alternately by shape or brightness. After every two stimuli the feature to be attended changes. In case of an incorrect response, the feature to be considered in the next stimulus is displayed (e.g. shape). This hint helps the respondent return to the task after a possible loss of overview. In total, 160 stimuli are presented.	Task switching speed, task switching accuracy
	TMT-B – TMT-A (Rodewald et al., 2012)	S2, S4	Touch-screen	This task consists of two parts: A (S2) and B (S4). In part A, numbers from 1 to 25 are randomly displayed on the screen. The task is to connect the numbers in ascending order as quickly as possible. In part B, numbers (1 – 13) and letters (A – L) are presented. The task is now to alternately link numbers and letters in ascending order (e.g. 1 – A – 2 – B etc.).	Difference score (working time B – A)

Table 3. continued, page 2/2

Domain	Task	Test form	Response mode	Description	Dependent variable
Updating	DSB (Petermann, 2012)	N/A	Verbal	The test leader is reading a series of numbers aloud. The participants must repeat the numbers in reverse order. The task consists of eight items with a constant increase of one number per item. Each item may be attempted twice. If not at least one sequence of numbers of an item can be repeated correctly, the task stops.	Number of correct trials
	NBV (Schelling & Schuri, 2009)	S1, S3	Response panel	Consonants are displayed one after the other on the screen. If a consonant is identical to the one that has been displayed two places (S1) or three places back (S3), the respondent must react by pressing the green button. S1 contains 100 consonants. In S3, 140 consonants are shown.	Number correct
Verbal fluency	WIWO (Jahn, 2016)	S2, S4	Verbal	In these two tasks, the subject must name within two minutes as many words as possible that belong to the category of first names (S2) or begin with the letter K (S4).	Number of correct words
Attention	WAFA (Sturm, 2006b)	S1	Response panel	Twenty-five black circles are presented one by one on the screen. The task requires to react as quickly as possible by pressing the green button when a circle appears.	Logarithmic mean reaction time
	DSF (Petermann, 2012)	N/A	Verbal	The procedure is identical to that of the DSB *(see above)*. The task is to re-peat the numbers in the same order.	Number of correct trials
	TMT-A (Rodewald et al., 2012)	S2	Touch-screen	*See above.*	Working time

Note. DSB = Digit Span Backwards; DSF = Digit Span Forwards; INHIB = Response Inhibition; N/A = Not applicable; NBV = N-Back Verbal; STROOP = Stroop Interference Test; SWITCH = Task Switching; TMT-A/-B = Trail-Making Test – Langensteinbach Version part A/part B; WAFA = Perception and Attention Function Battery – Alertness; WIWO = Vienna Verbal Fluency Test.

scale from ‘fully agree’ to ‘fully disagree’ or ‘cannot judge’, participants [gave] their opinions on the lessons, the teacher, the textbook, their degree of motivation, and their general satisfaction with the course. Additionally, three open format questions allow[ed] participants to provide more specific feedback on aspects they appreciated, disliked, or would recommend for improvement.

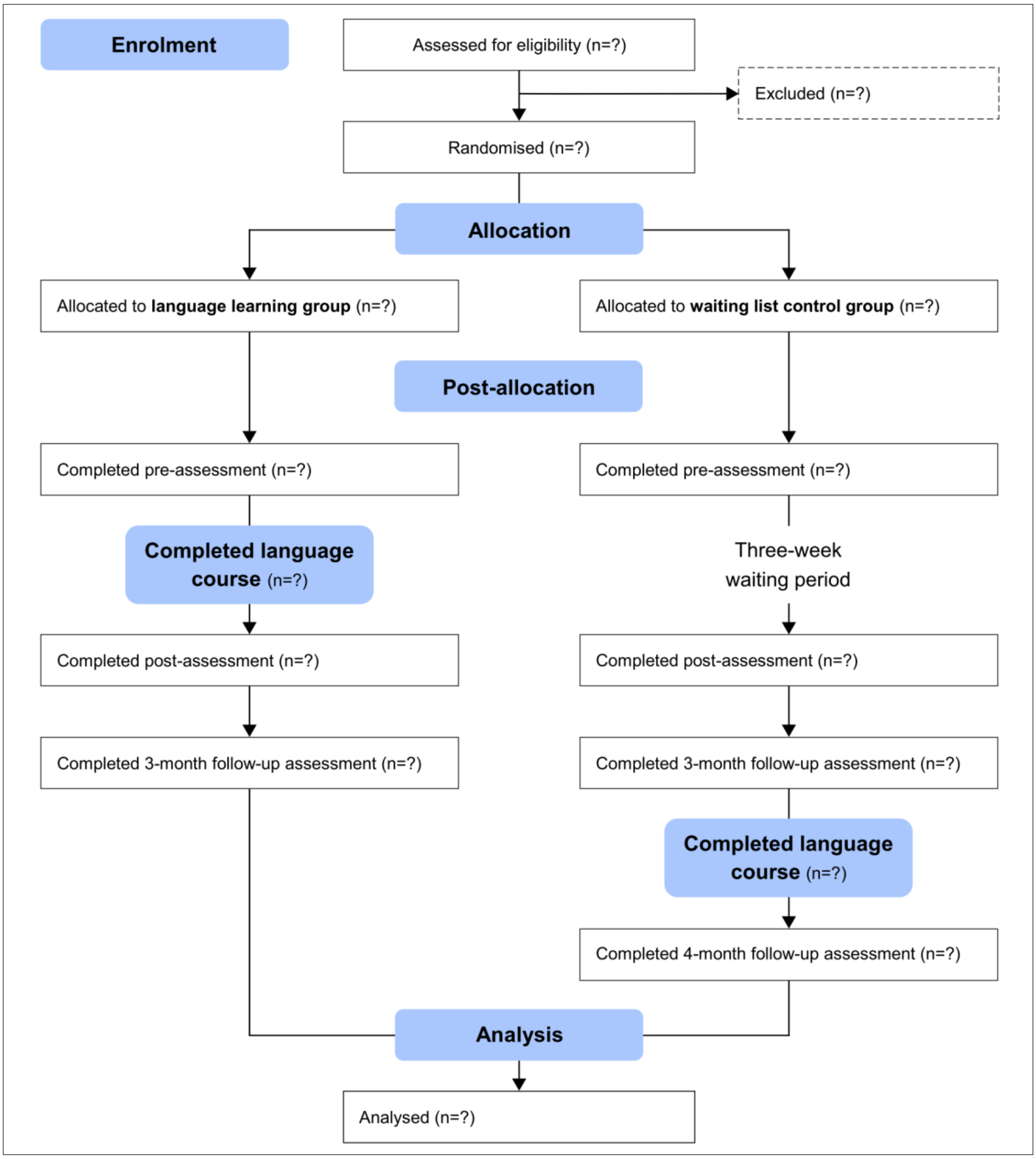
Vocabulary test

To assess the degree of acquired Spanish language skills, participants [took] a vocabulary test on the last day of the course. Following the example of Berggren et al. (2020), the vocabulary test (…) include[d] 108 out of around 577 words learn[t] during class. Two points [were] awarded if the word translated from German into Spanish [was] translated and spelt correctly. One point [was] granted if it ha[d] the correct meaning but [was] misspelt, ignoring punctuation. If the translation [was] wrong or no answer [was] given, zero points [were] assigned. Thus, a maximum of 216 points [could] be achieved.

3.6 Participant timeline

Figure 4 [emphasis added] illustrates the flow of participants through the trial. Eligibility for participation [was] ascertained via a telephone and a subsequent face-to-face screening. In the telephone screening, initial inclusion and exclusion criteria regarding demographic, medical and health-related aspects (see **Table 2**) [emphasis added] as well as availability to participate in the intervention and assessments [were] clarified. Moreover, self-reported foreign language skills and usage [were] surveyed. Participants who claim[ed] to possess few previous skills in Spanish and have not yet attained level A1.1, or who have attended less than one regular course (< 18 hours of classes), [were] asked a list of questions to determine their Spanish skills. The questions [were] derived from units 1-5 of the course book (Goerrissen, 2016). An appointment for the face-to-face screening [was] scheduled if they [met] all eligibility criteria clarified by telephone.

Figure 4. CONSORT flow diagram of participants



The face-to-face screening [took] around two hours and (…) evaluate[d] the remaining exclusion criteria. The CFD [was] applied to exclude participants with suspected cognitive decline and the whispered voice test (Swan & Browning, 1985) [was] used to ascertain sufficient hearing. Participants who indicated in the telephone screening to already have slight experiences in Spanish [took] a placement test (Ernst Klett Verlag, 2004) to ensure that their previous skills [were] low enough

(< level A1.1) for a Spanish course for beginners. Additionally, further baseline data [were] collected (see section ‘Baseline variables’).

Subjects that [met] eligibility criteria after the face-to-face screening [were] randomly assigned to either the LLG or the WLCG. Within 2 weeks before the start of the language course or the three-week waiting period, both groups (…) participate[d] in a pre-assessment. Both post- and 3-month follow-up assessments [were] carried out within 2 weeks immediately (post) and 3 months (3-month follow-up) after the end of the intervention or the waiting period. After this, the LLG (...) completed study participation while the WLCG (…) attend[ed] the control group programme consisting of the language course and a voluntary 4-month follow-up assessment.

Pre-, post-, and follow-up assessments (…) last[ed] around 2 h each and [were] scheduled at the same time of day, either in the morning or in the afternoon, to control for day-dependent variability in cognitive performance. Assessments (…) include[d] both the primary and secondary outcome measures listed in **Figure 2** [emphasis added]. The test order [was] randomised by a computer for each individual. However, the WAFA, as the most straightforward task, [was] always (...) applied first to allow participants to become familiar with the computerised test system.

To enhance participant retention, we (…) arrange[d] appointments for pre-, post-, and follow-up assessments with each participant individually before the start of the study. A confirmation document with the scheduled dates and times [was] sent to participants. Additionally, participants (…) receive[d] reminder phone calls 1 week before each assessment. There [was] no financial compensation apart from the language course, which [was] offered for free. However, participants (…) obtain[ed] information about their results and the conclusion of the study after having completed participation.

3.7 Sample size

We pre-estimated the required sample size based on the first hypothesis and the primary outcomes (improvement of the LLG in either the STROOP naming interference tendency or in the WAFG between pre- and post-assessment compared to the

WLCG) (Moher et al., 2010). Our calculation run on G*Power (version 3.1.9.2) resulted in 42 participants being required. We also considered a 30% drop-out rate (Pfenninger & Polz, 2018), making a total of 60 participants deemed necessary. We set the parameters in G*Power as follows: α = 0.05, corrected for multiple comparisons using the Bonferroni-Holm correction, power to detect significant differences of 0.8. Based on effect sizes found in some previous studies (Bak et al., 2016; Pfenninger & Polz, 2018), we considered a medium effect size of d = 0.25 to be reasonable (Cohen, 1988). At default, the correlation between repeated measures was set very conservatively at an r = .5. However, as mentioned above, we ended our study due to the Corona pandemic after 37 participants (ignoring drop-outs and non-adherence to the course) had completed the post-assessment. Since the correlation between repeated measures in the STROOP naming interference tendency and the WAFG was higher than estimated at r > .6 and higher correlations increase test power (Caldwell et al., 2022), G*Power allowed for a lower sample size of 34 participants, or 17 participants per group. Therefore, we decided to terminate our trial early in order to not to expose any participant to the risk of COVID-19 infection.

3.8 Allocation and blinding

Participants who provide[d] written informed consent and pass[ed] all screening procedures [were] randomly assigned at a 1:1 ratio to one of the two trial arms. A researcher (...), who [was] not involved in the implementation of the current trial, (…) generate[d] the randomisation sequence using a web-based randomisation system (Sealed Envelope). To avoid predictability of group allocation, the system (…) buil[t] permuted blocks of random sizes, namely two, four, six, and eight, with a list length of 20 (Chan et al., 2013). The randomisation sequence and the block sizes [were] not (...) disclosed to trial implementers. To ensure concealment, [the researcher who generated the randomization sequence] prepare[d] and store[d] sequentially numbered, sealed, opaque envelopes containing information on group allocation for each participant. She [did] not open the envelope assigned to a participant until he or she ha[d] been included in the trial (. ...) [I] inform[ed] participants by telephone about their group assignment.

Due to the nature of our study design, participants [could not] be blinded to treatment allocation. There [could] also be no blinding of assessors and teachers. However, assessments [were] predominantly computer-based and standardised to reduce test leader effects.

3.9 Data collection and management

(...) Trained research assistants [and I] conduct[ed] the assessments. To limit the impact of the test leader on results, we aim[ed] to ensure that for each participant at least pre- and post-assessments [were] conducted by the same assessor. For every assessor, the number of assessments [was] balanced between groups.

Data [were] saved and stored according to the DSGVO. At the start of the study, every participant [was] allotted a unique identification number. Only persons involved in study conduct [had] access to the connection code and the data. As assessments [were] mainly conducted via computer, most data [were] generated and saved automatically. Data from paper documents [were] continuously entered into the database throughout the study period and [were] kept in a locked cabin. All electronic data [were] stored on a secure server and password-protected computers of the NAR. Additional password protection for electronic documents with identifying participant information [was] provided. To promote data quality, we (…) check[ed] data randomly throughout study conduct. Missing data [was] addressed as quickly as possible. Before the start of the analysis, we (…) recheck[ed] all data for completeness and correctness. We (...) fully anonymis[ed] data upon completion of data collection, at the latest by 31 March 2021.

3.10 Monitoring

Adverse events [were] assessed as part of routine monitoring. No formal data monitoring committee [was] established, as this study (…) involve[d] mostly standardised cognitive assessments. Also, foreign language training pose[d] only a minimal risk of adverse events and [was] of short duration. [I was] responsible for trial conduct. [I] continuously discuss[ed] with the immediate trial team all aspects concerning participant safety, study design, and data management. Participants [were] informed that they can contact the study team at any time if they ha[d] questions or concerns about the study. Any adverse event related or potentially related to study participation [was] instantaneously reported to the immediate trial team, the sponsor, and the local ethics committee.

3.11 Statistical analysis

The study team conducted all statistical analyses using IBM SPSS Statistics 26 (IBM Corporation: Armonk, NY, USA). We proceeded with the per-protocol dataset, including all participants who completed at least pre- and post-assessment and a minimum of 14 h of formal language instruction (> 62% of the total course duration; Bak et al., 2016). Data from two participants in the LLG were missing at 3 month follow-up due to the COVID-19 pandemic. We replaced these missing data with multiple imputation, assuming that they were at least missing at random. Primary and secondary outcomes were analyzed using 2 (group) x 3 (time) repeated measures analyses of variance (ANOVAs) with *group* as the between-subjects factor and *time-point of assessment* (pre-, post-, 3 month follow-up) as the within-subjects factor. In case of a significant interaction, we conducted *post hoc* tests to evaluate hypotheses 1 and 2. We predicted a significant *group x time interaction* and hypothesized that the LLG would outperform the WLCG at post- (hypothesis 1) or at 3 month follow-up (hypothesis 2) relative to pre-assessment. All hypotheses were tested two-sided with adjustments made for multiple comparisons using the Bonferroni-Holm correction for primary outcomes. No adjustments for multiplicity were undertaken for secondary endpoints because, in smaller studies of an exploratory nature, rejection of the alternative hypothesis is more of a constraint than a type 1 error (Schoenfeld, 1980). Partial eta square (η_p^2) was applied as an effect size indicator.

To assess the robustness of the results, we investigated whether the short-term effects of foreign language learning on cognitive measures also applied to the WLCG after completing the control group program. To this end, the WLCG served as its control. We predicted a significant difference between the change scores from 3- to 4-month follow-up compared to the change scores from pre- to post-assessment of the WLCG using *t*-tests for dependent measures.

For the LLG, we performed exploratory subgroup analyses to evaluate whether effects of foreign language learning on cognitive outcomes would depend on different levels of CR (hypothesis 3), foreign language knowledge and usage (hypothesis 4), or baseline cognition (hypothesis 5). We also incorporated the vocabulary test score as a marker of learning success (Lim et al., 2020). Initially, we planned to conduct multiple regression analyses for hypothesis 3. However, we refrained from doing so since our sample was not large enough to meet the sample size requirement for multiple regression analysis. Instead, we created a correlation matrix between age, CR indices (CRI-Education, CRI-Working Activity, CRI-Leisure Time, and CRI-Index) (hypothesis 3), the LSBQ-score (hypothesis 4), the CFD-Index score (hypothesis 5; defined as independent variables in the further moderation analyses), and the vocabulary test score. We plotted these variables against change scores of the primary and secondary outcomes (defined as the dependent variable in the regression analysis). Change scores for each test were calculated by subtracting the post- from the pre-assessment result. Kendall's tau-b was used as a correlation index, as this marker is generally recommended with small samples and a considerable number of tied ranks (Field, 2009). If the requirements for regression were met, we further analyzed whether group moderated the association between predictor and outcome to evaluate whether the significant association was specific to the LLG. Moderation models were run using model 1 of Hayes's PROCESS macro v. 4.0 (2021).

4. Results

The following chapters on results written by my co-authors and me are published in *Frontiers in Aging Neuroscience* in the scientific article "Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial, pp. 6–11". I have adapted the numbering and labeling of headlines, figures, and tables to the format of this dissertation. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185.

4.1 Participants

A total of 54 participants were randomized into one of the two study arms. However, due to the COVID-19 pandemic and the consequent termination of the study, we lost 14 participants, who had already been assigned to one of the intervention groups but had not yet participated in the pre-assessment. Two participants in the WLCG were excluded from the trial because they did not meet eligibility criteria after the initial inclusion (CFD one subtest $z \leq -1.5$, according to age and, where possible, sex, and education). Three participants from the LLG were excluded from the analysis because they did not finish the language course. One dropped out after two days, and two participants discontinued training after the first week because they experienced the course as too demanding. Another participant in the LLG was excluded from the final analysis due to suspected dyslexia. The final sample included in the analysis consisted of 34 participants who completed the intervention and the pre- and post-assessment. **Figure 5** provides detailed information on participant flow and reasons for drop-out and losses due to the COVID-19 pandemic.

Demographic and language characteristics of participants are provided in **Table 4**. There were no notable differences between groups at baseline. The mean age was 69.47 (*SD* = 3.36) years. Nearly all participants were retired (*M* = 91.18%), the rest worked part-time at most. No participant scored below the cognitive threshold in any of the subtests of the CFD ($z > -1.5$). CR was in general high among both groups (*M* = 135.18, *SD* = 10.51). Regarding foreign language skills, participants can be considered monolingual on average (LSBQ: *M* = -3.57, *SD* = 1.73). None of

Figure 5. CONSORT flow diagram of participants [- Results]

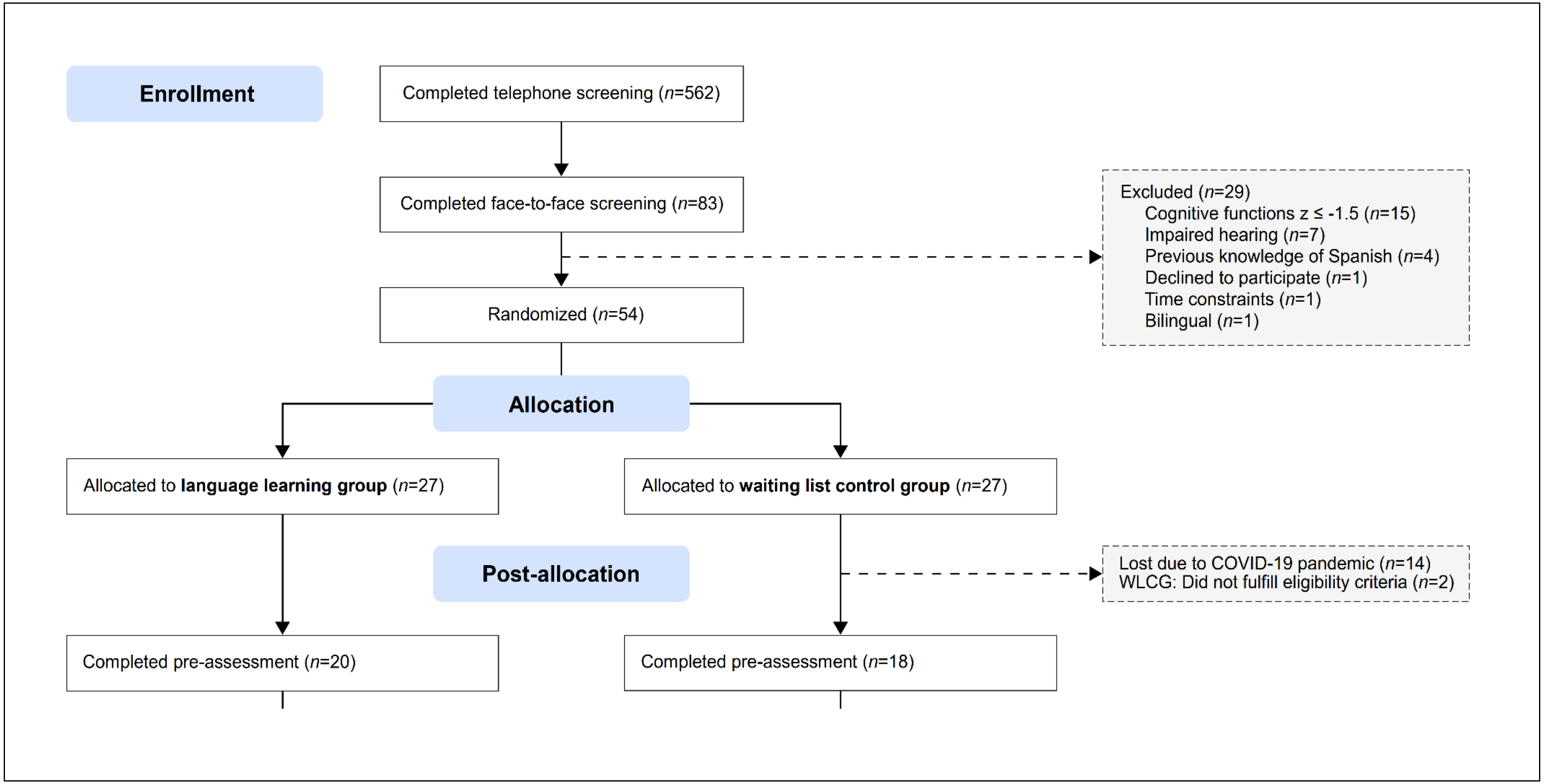


Figure 5. continued, page 2/2

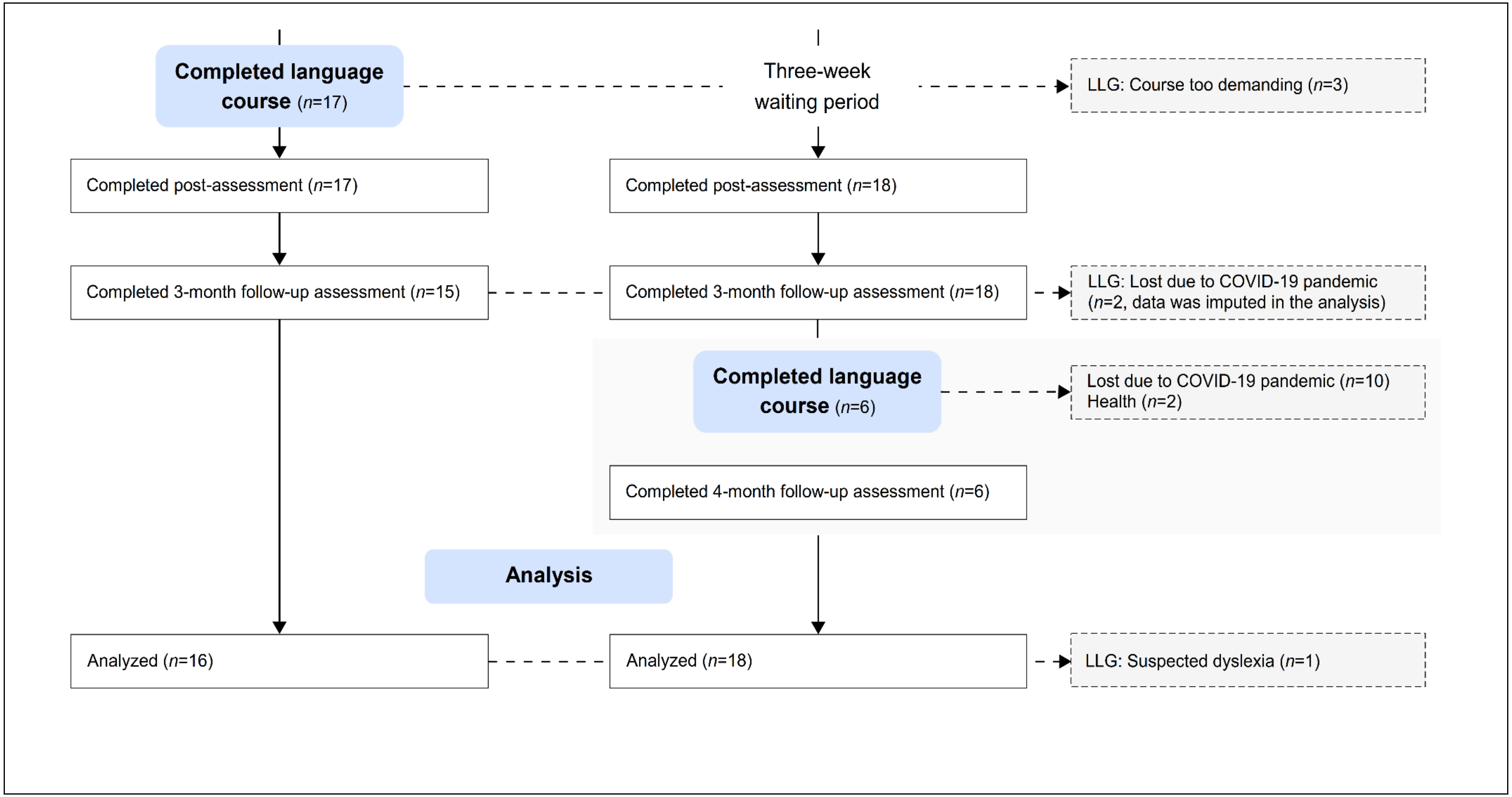


Table 4. Demographic and linguistic characteristics of participants in the two study arms

Baseline characteristic	**LLG (*n* = 16)**		**WLCG (*n* = 18)**	
	Mean (*SD*) / *n* (%)	IQR (Q1, Q3)	Mean (*SD*) / *n* (%)	IQR (Q1, Q3)
Age *(years)*	69.00 (3.10)	(67.00, 70.00)	69.89 (3.61)	(66.00, 72.25)
Sex *(f)*	10 (62.50%)		10 (55.56%)	
Retired	15 (93.75%)		16 (88.89%)	
Handedness *(right)*	15 (93.75%)		16 (88.89%)	
CFD *(z-score)*[a]	0.61 (0.70)	(0.13, 0.99)	0.55 (0.59)	(0.04, 1.08)
CRIq-Index[b]	138.69 (10.83)	(131.75, 146.75)	132.06 (9.43)	(124.75, 137.25)
Education	125.50 (13.00)	(116.50, 138.50)	119.89 (12.56)	(107.75, 127.50)
Working activity	118.31 (11.72)	(107.25, 129.50)	114.28 (17.10)	(103.25, 128.00)
Leisure time	144.19 (15.18)	(130.25, 160.25)	138.39 (15.45)	(131.75, 149.25)
LSBQ[c]	-3.41 (1.60)	(-4.83, -2.58)	-3.72 (1.88)	(-5.00; -2.40)
Romanic languages	13 (81.25%)		13 (72.22%)	
Spanish	7 (43.75%)		4 (22.22%)	

Note: CFD = Cognitive Functions Dementia; CRIq = Cognitive Reserve Index questionnaire; LLG = Language learning group; LSBQ = Language and Social Background Questionnaire, WLCG = Waiting list control group.

[a] z-score: ≤-1.5 = below average; -1.5< average >1.5; ≥1.5 = above average.
[b] CRIq: ≤70 = low; 71-85 = low – medium; 86-114 = medium; 115-129 = medium – high; ≥130 = high.
[c] LSBQ: <-3.13 = monolingual; -3.13-1.23 = not strongly differentiated; >1.23 = bilingual.

them classified themselves as bilingual. Most reported having some basic knowledge of a Romanic language such as French and Italian (*n* = 26, 76.47%), and only a few had previous experience with Spanish (*n* = 11, 32.35%), e.g., having learned some basic vocabulary on holiday.

In the WLCG, *n* = 6 out of 18 participants completed the language course and subsequent assessment. Ten participants had to discontinue their course after the first week due to the COVID-19 pandemic, and two participants had to stop for health reasons. No study-related adverse events were reported.

4.2 Primary outcomes

Data and results for tests of executive attention (primary outcomes), including exact *F*-, η_p^2- and *p* values for the two-way interactions between time and group of the 2 (group) x 3 (time) ANOVAs, are presented in **Table 5**. The analysis of the per-protocol and multiple imputation datasets revealed no significant interaction for the STROOP naming interference or the WAFG ($p > .05$). Additionally, there was no significant difference in the sub-analysis of the WLCG who completed the control group program ($p > .05$).

4.3 Secondary outcomes

Results of the secondary outcomes are presented in **Table 6**. We detected a significant interaction between time and group in the 2 (group) x 3 (time) ANOVA for the WAFA, $p = .03$, $\eta_p^2 = .10$. As discernible in **Figure 6a**, there was a slight group imbalance at pre-assessment, indicating worse performance for the LLG. To account for regression to the mean due to this imbalance, we calculated univariate Analyses of Covariance (ANCOVAs) for hypotheses 1 – short term effects – and 2

Table 5. Means (standard deviations) of primary outcome measures by trial arm and time point

Outcome		**LLG**	**WLCG**	***F*(1, 32)**	η_p^2	***p***	***p*†**
STROOP Naming interference	Pre	211.44 (147.04)	253.61 (226.65)	0.016	<.001	.984	.984
	Post	214.13 (115.53)	256.39 (214.21)				
	FU	220.50 (91.93)	255.89 (162.59)				
WAFG Reaction time	Pre	517.22 (126.55)	524.35 (130.07)	0.974	.030	.383	.766
	Post	522.13 (138.02)	497.95 (82.85)				
	FU	529.86 (156.84)	494.87 (95.42)				

Note: F-values represent interaction effects between *time* and *group* calculated from 2 (group) x 3 (time) ANOVAs with repeated measures; FU = Three-month follow-up assessment, LLG = Language learning group; Pre = Pre-assessment; Post = Post-assessment; STROOP = Stroop Interference Test – Naming interference; WAFG = Perception and Attention Function Battery – Divided Attention; WLCG = Waiting list control group.

† Bonferroni-Holm adjusted *p* value.

Table 6. Means (standard deviations) of secondary outcome measures by trial arm and time point

Outcome		LLG	WLCG	*F*(1, 32)	η_p^2	*p*
INHIB Commission errors	Pre	14.19 (7.99)	13.22 (5.76)	1.599	.048	.214
	Post	11.75 (8.05)	11.06 (2.53)			
	FU	10.19 (6.51)	11.50 (4.48)			
STROOP Reading interference	Pre	104.88 (99.25)	142.72 (99.78)	0.030	.001	.970
	Post	109.00 (70.90)	150.44 (100.61)			
	FU	104.88 (69.42)	149.67 (106.07)			
SWITCH Switching speed	Pre	199.44 (208.26)	212.72 (195.95)	0.710	.022	.495
	Post	150.31 (180.78)	206.06 (182.37)			
	FU	160.56 (143.77)	198.33 (210.89)			
TMT Working time part B-A	Pre	21.09 (8.74)	26.71 (22.21)	0.418	.013	.660
	Post	15.50 (7.86)	18.60 (9.57)			
	FU	20.46 (15.20)	21.19 (17.16)			
DSB Correct trials	Pre	8.69 (1.62)	9.33 (1.94)	0.476	.015	.624
	Post	8.69 (1.70)	9.50 (2.04)			
	FU	9.63 (2.09)	9.94 (2.18)			
NBV S1 Number correct	Pre	12.69 (2.06)	11.39 (2.50)	1.566	.047	.217
	Post	12.63 (1.78)	11.50 (2.68)			
	FU	12.44 (2.19)	12.28 (2.35)			
NBV S3 Number correct	Pre	10.06 (2.17)	8.67 (2.61)	0.567	.017	.570
	Post	10.00 (3.01)	9.17 (2.87)			
	FU	10.50 (1.55)	8.83 (3.20)			
WIWO S2 Number correct	Pre	43.63 (6.80)	38.56 (7.72)	0.460	.014	.583
	Post	40.94 (7.63)	38.44 (7.96)			
	FU	43.94 (9.45)	40.56 (8.89)			
WIWO S4 Number correct	Pre	23.38 (7.59)	19.50 (6.96)	0.365	.011	.696
	Post	25.69 (8.75)	20.06 (7.48)			
	FU	25.00 (8.19)	20.17 (6.08)			
WAFA Reaction time	Pre	224.40 (23.61)	211.01 (23.10)	3.692	.103	**.030***
	Post	212.63 (17.16)	215.88 (26.30)			
	FU	219.36 (25.41)	219.20 (25.22)			
DSF Number correct trials	Pre	10.06 (1.18)	9.39 (1.85)	.016	.001	.984
	Post	9.94 (1.12)	9.28 (1.99)			
	FU	10.25 (1.61)	9.67 (1.75)			

Table 6. continued, page 2/2

Outcome		**LLG**	**WLCG**	***F*(1, 32)**	**η_p^2**	***p***
TMT Working time part A	Pre	14.91 (2.59)	15.79 (3.08)	.351	.011	.705
	Post	14.71 (2.05	16.24 (5.45)			
	FU	14.45 (3.43)	16.18 (3.70)			

Note: F-values represent interaction effects between time and group calculated from 2 (group) x 3 (time) ANOVAs with repeated measures; DSB = Digit Span Backwards; DSF = Digit Span Forwards; FU = Three-month follow-up assessment; INHIB = Response Inhibition task; NBV = N-Back Verbal; LLG = Language learning group; Post = Post-assessment; Pre = Pre-assessment; STROOP = Stroop Interference Test – Reading interference; SWITCH = Task Switching test; TMT = Trail-Making Test – Langensteinbach Version; WAFA = Perception and Attention Function Battery – Alertness; WIWO = Vienna Verbal Fluency Test; WLCG = Waiting list control group.

* $p < .05$. Significant interaction effects are highlighted in bold.

– long term effects of foreign language learning – separately, using the pre-assessment score as covariate and the post- and the 3 month follow-up assessment score, respectively, as dependent variable (Vickers & Altman, 2001). Both for hypothesis 1, $F(1, 31) = 2.55$, $p = .12$, $\eta_p^2 = .07$, and for hypothesis 2, $F(1, 31) = 3.64$, $p = .07$, $\eta_p^2 = .11$, the ANCOVAs missed significance. These results indicate no significant difference between groups in intrinsic alertness after controlling for baseline imbalance, neither immediately nor 3 months after the intervention.

In the sub-analysis of the WLCG, who completed the control group program, there likewise were no significant differences between the predefined change scores ($p > .05$). Yet, there was also a slight, but non-significant, $t(5) = -2.38$, $p = .06$, improvement in the WAFA from 3- to 4-month follow-up compared to the change score from pre- to post-assessment, suggesting that the WLCG also tended to benefit from foreign language learning in intrinsic alertness.

Note that in our protocol, we intended to conduct an intention-to-treat analysis in addition to the per-protocol analysis. However, the intention-to-treat analysis was meant to test the robustness of the results (Tripepi et al., 2020). We refrained from doing so as there were no significant results for either the primary or secondary endpoints.

4.4 Moderation analysis

In a first step, we ran Kendall's tau-b correlations among participants in the LLG to identify significant associations between predictors (age, indices of the CRIq, LSBQ score, CFD index score, vocabulary test) and outcomes (change scores between pre- and post-assessment in primary and secondary outcomes). The correlation matrix is depicted in **Table 7**. Significant associations are highlighted in bold. For significant associations, we further conducted moderation analyses with group as moderator to determine whether the associations in the LLG differed significantly from those in the WLCG.

The moderation model of the prediction of the change score of the WAFG by the CRIq Leisure time sub-score, moderated by group missed significance, $F(3, 30) = 2.40$, $p = .09$, $R^2 = .23$. Also, the model of the change score of the TMT-A predicted by the LSBQ, $F(3, 30) = 3.17$, $p = .18$, $R^2 = .10$, and of the SWITCH predicted by the CRIq Education sub-score, $F(3, 30) = 1.96$, $p = .14$, $R^2 = 25$ [*sic*], were not significant. However, for the change score of the INHIB predicted by the CFD, the overall model was significant: $F(3, 30) = 36.46$, $p < .001$, $R^2 = .43$. **Figure 6b** shows the relationship between predictor and outcome divided by group. Group and the CFD score did not significantly predict the change score of the INHIB ($p > .05$).

However, the interaction term was highly significant $b = 8.75$, $t(30) = 3.86$, $p < .001$. The regression slope was only significant for the LLG, $b = -6.24$, $t(30) = -10.39$, $p < .001$; WLCG: $b = 2.51$, $t(30) = 1.15$, $p = .26$, meaning that for the WLCG, the CFD did not predict change from pre- to post-assessment in the INHIB. In the LLG, lower CFD scores predicted stronger improvement from pre to post in the INIHB.

The vocabulary test score was not entered in the moderation analyses as only participants from the LLG conducted a vocabulary test. For the vocabulary test score, there was a significant negative correlation with the LSBQ, $\tau_b = -.38$, $p = .01$, indicating better performance in the vocabulary test among individuals with higher LSBQ scores. The vocabulary test score also correlated strongly negatively with the change score of the WAFG, $\tau_b = -.44$, $p = .005$, implying that individuals with lower vocabulary scores showed more improvement in the WAFG. And there was a significant negative correlation with the change score of the INHIB, $\tau_b = -.38$, $p = .02$,

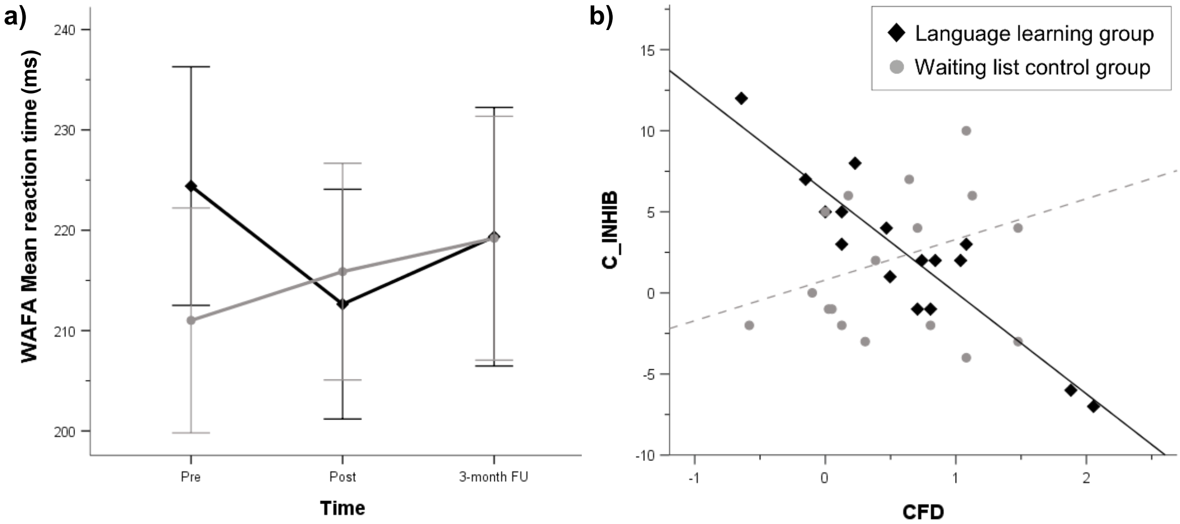
Table 7. Correlation matrix of measures of the language learning group (*n* = 16)

Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1. Age	—							
2. CFD	.30	—						
3. CRIq-Index	.09	.05	—					
4. CRIq-Education	.18	.03	**.57****	—				
5. CRIq-Working act.	-.07	.03	.36	.29	—			
6. CRIq-Leisure time	.17	.11	.30	.03	-.30	—		
7. LSBQ	.12	.29	.31	.23	.25	.12	—	
8. Vocabulary test	-.09	.14	.25	.00	.06	.31	**.38***	—
C_STROOP_N	.08	.24	.00	.03	.10	-.11	.06	.13
C_WAFG	.23	-.33	.03	.14	.18	**-.44***	-.20	**-.44****
C_INHIB	-.20	**-.64*****	-.09	-.04	.13	-.23	-.19	**-.38***
C_STROOP_R	.10	.13	-.29	-.06	-.21	-.07	-.02	.01
C_SWITCH	.08	.18	-.16	**-.39***	-.04	-.05	.10	.07
C_TMT B-A	.05	.01	.03	.19	-.06	.03	.20	.10
C_DSB	-.20	-.08	.01	.16	.29	-.26	.19	-.18
C_NBV S1	.14	-.02	-.04	-.18	.10	.02	-.09	-.22
C_NBV S3	-.11	-.14	-.13	-.18	.00	.08	-.13	.29
C_WIWO S2	-.05	.08	-.07	.01	-.10	.01	-.14	-.01
C_WIWO S4	.23	.28	.02	.10	.26	-.16	.34	-.13
C_WAFA	-.03	-.01	.03	-.01	-.18	.17	.22	.06
C_DSF	.13	.08	-.06	.27	.14	-.21	-.09	-.09
C_TMT-A	-.33	-.16	-.16	-.23	-.11	-.20	**-.50****	-.20

Note. Kendall's tau-b correlations between baseline variables and the vocabulary test score, and change scores of primary and secondary outcomes between pre- and post-assessment; CFD = Cognitive Function Dementia; CRIq = Cognitive Reserve Index questionnaire; LSBQ = Language and Social Background Questionnaire; C_DSB = Change score of the Digit Span Backwards; C_DSF = Change score of the Digit Span Forwards; C_INHIB = Change score of the Response Inhibition task; C_NBV S1 = Change score of the NBV N-Back Verbal S1; C_NBV S3 = Change score of the NBV N-Back Verbal S3; C_STROOP_N = Change score of the Stroop Interference Test – Naming interference; C_STROOP_R = Change score of the STROOP – Reading interference; C_SWITCH = Change score of the Task Switching test; C_TMT B-A = Change score of the Trail-Making Test – Langensteinbach Version part B – part A; C_TMT-A = Change score of the Trail-Making Test – Langensteinbach Version part A, C_WAFA = Change score of the Perception and Attention Function Battery – Alertness; C_WAFG = Change score of the Perception and Attention Function Battery – Divided Attention; C_WIWO S2 = Change score of the Vienna Verbal Fluency Test S2; C_WIWO S4 = Change score of the Vienna Verbal Fluency Test S4.

* *p* < .05, ** *p* < .01. *** *p* < .001. All significant correlations are highlighted in bold.

Figure 6. a) Mean reaction time of the Alertness subtest of the Perception and Attention Function Battery (WAFA) for each time point of assessment, divided by group. b) Relationship between the change score between pre- and post-assessment of the Inhibition task (C_INHIB) commission errors and the index score of the Cognitive Functions Dementia (CFD), divided by group



again indicating a more marked improvement in the INHIB for individuals with lower vocabulary test scores.

4.5 Language course outcomes

4.5.1 Adherence

Overall, adherence to the intervention was high, with a mean of 14.63 (*SD* = 0.50) course days attended. Most participants (62.5%) completed the full course duration.

4.5.2 Homework and learning time

Participants spent *M* = 12.25 (*SD* = 5.79) hours with additional homework and learning activities at home. In this regard, we detected considerable disparities between participants (range = 4.42 – 23.58 h). *Post-hoc* analyses indicated no significant correlation between time spent with additional homework and learning activities at home and performance on cognitive variables at baseline ($p > .05$).

4.5.3 Course evaluation

Participants' opinions on the language course were based on the total proportion of the sample that completed the language course and submitted an evaluation. Since the evaluation was anonymous, two participants who dropped out from the language course but submitted an evaluation are included here. Overall, the responses were very positive, with the teacher and the quality of the lessons being rated highest [lesson: M = 3.88 (SD = 0.23), teacher: M = 3.98 (SD = 0.08)]. Also, opinions about the textbook, participants' motivation, and satisfaction were favorable [textbook: M = 3.65 (SD = 0.59), motivation: M = 3.32 (SD = 0.59), satisfaction M = 3.64, (SD = 0.45)].

Additionally, in open format questions, participants indicated that they were most positive about the quality of teaching. For example, one participant said, "[The teacher] gave clear and understandable explanations and always patiently corrected mistakes. I would very much like to continue the course with her." Aspects participants disliked or would recommend for improvement were mainly related to the textbook, the classroom, the intensity and duration of the course, and the different learning paces within the groups. One participant perceived the textbook as somewhat complicated, e.g., when searching for vocabulary. Some participants would also have preferred to learn more about Spain. The classroom was perceived as too small to create a good learning atmosphere and allow group work. Therefore, the second half of the participants were taught in a larger room. Regarding the intensity of the course, one participant claimed that the course was too intensive. Two participants felt the course was too short, and another would have liked to continue, but only twice a week. Four participants rated the differences in learning pace between participants and prior knowledge of related languages (e.g., French) as too large. Some considered the pace too fast and would have welcomed more time for repetition, whereas another participant said the lessons could have proceeded more quickly.

4.5.4 Vocabulary test

The mean score of the post-intervention vocabulary test was M = 160.88 (SD = 41.17), indicating that, on average, participants had acquired a basic Spanish vocabulary. However, again, we detected major differences between participants, with scores ranging from 84 to 210. *Post-hoc* analysis revealed no significant correlation between the result of the vocabulary test and the homework and learning time (τ_b = -.04, p = .82).

5. Discussion

The discussion written by my co-authors and me has recently been published in *Frontiers in Aging Neuroscience* in the scientific article "Grossmann, J. A., Aschenbrenner, S., Teichmann, B. & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial, pp. 11–14". I have adapted the numbering of headlines and the citation style to the format of this dissertation. The original publication is available at https://doi.org/10.3389/fnagi.2023.1123185.

In older adults aged 65–80 years, a three-week-long intensive Spanish course for beginners did not elicit improvements in executive attention relative to a passive control group neither immediately nor 3 months after the training. We also saw no significant increase in our secondary outcomes encompassing measures of EF, verbal fluency, and attention. Intrinsic alertness improved in the LLG compared to the WLCG after participation in the course. However, when we considered differences between groups in the pre-measurement, this effect failed to reach significance. The sub-analysis of the WLCG who completed the control group program also did not reveal any significant change in cognitive performance after completing the course. However, examining factors that influenced cognitive responses in language learners, moderation analyses showed that differences in global cognition at baseline predicted changes in response inhibition. The lower the language learners' baseline cognition, the greater were the gains in response inhibition. This association was not evident in individuals of the WLCG. Similarly, correlation analyses revealed that lower vocabulary test scores as a marker of learning were associated with more remarkable improvement in response inhibition after the end of the course. This relationship could be due to higher cognitive engagement among those with lower vocabulary test scores.

5.1 Primary and secondary outcomes

For our primary outcome measures – the Stroop naming interference and the WAFG – representing two sub-domains of executive attention, namely selective and divided attention, respectively, we did not find beneficial effects of foreign language learning. Also, we did not observe any general training-related gains in our secondary outcome measures immediately or 3 months after the course. As participation in the course was generally high, and learners' assessments of the teacher and the quality of the lessons were generally very positive, we contend that the overall non-significant results in our primary and secondary outcomes are not the result of a lack of participant engagement. Participants also stated that they were generally motivated to participate in the training. In addition, the outcome of the vocabulary exam demonstrates that participants gained fundamental knowledge while taking the course.

The lack of improvement is consistent with findings from previous studies on the cognitive benefits of foreign language learning (Berggren et al., 2020; Kliesch et al., 2022; Ramos et al., 2017; Valis et al., 2019; Ware et al., 2017). Our results may underline the notion of missing far-transfer effects of foreign language acquisition on supposedly affected cognitive domains. The difficulty of detecting far-transfer effects is a considerable problem accompanying many studies trying to demonstrate the cognitive benefits of a particular intervention. In general, most cognitive gains rarely extend beyond the specific domain being practiced. An example of studies showing near-transfer effects are cognitive training studies in which a particular cognitive domain is intentionally trained. Foreign language learning studies are referred to as far transfer because they do not specifically target any particular cognitive domain. In a second-order meta-analysis, Sala et al. (2019) found that far transfer effects of cognitive training were either small or non-existent. Thus, it might be that far-transfer effects of foreign language training are minor, if existent, and therefore difficult to detect.

However, some studies oppose the assumption of missing far transfer effects of foreign language learning (Bak et al., 2016; Bubbico et al., 2019; Long et al., 2020; Meltzer et al., 2021; Pfenninger & Polz, 2018; Wong et al., 2019). For example, a fairly recent RCT conducted by Meltzer et al. (2021) found significant gains in the naming interference condition of the Stroop task and working memory following 16 weeks of Spanish instruction. The reason we did not discover benefits in the executive attention domain equally in either selective or divided attention could be attributed to the format of our language course. Unlike Meltzer et al. (2021), we did not apply app-based language training. This kind of training allows to better meet participants at their performance level than a group-based face-to-face program and thus might have led to higher cognitive engagement in Meltzer et al.'s study.

Among language learners, we also, for instance, observed an improvement in intrinsic alertness after the course, which did not hold significance after controlling for differences between groups at pre-test. Meltzer et al. (2021) similarly found a medium, albeit non-significant, effect on processing speed in foreign language learners compared to a passive control group, which, like intrinsic alertness, measures intensity-related aspects of attention. Intrinsically maintained tonic alertness is a core function of the cingulo-opercular network, which is as part of the executive attention network also responsible for monitoring behavior, task-set maintenance, and salience detection (Sadaghiani & D'Esposito, 2015). The cingulo-opercular network is among the main cognitive networks that degenerate with aging. In addition, recent behavioral and functional data suggest that the bilingual advantage in older adults may be particularly pronounced in the alerting dimension (Dash et al., 2019). Future research should thus test this hypothesis even though our data did not show a substantial impact of learning a foreign language on intrinsic alertness.

5.2 Moderation analysis

The moderation analyses showed no significant impact of the CR indices or the foreign language proficiency and usage index on change scores of cognitive outcome measures. The prediction of the change score in divided attention by the CRIq Leisure time sub-score in the moderation model fell short of significance. The significant correlations between the change score in task switching and the CRIq education sub-score as well as between the change score in information processing speed and self-rated foreign language proficiency and usage also remained non-significant when entered into the moderation model. These insignificant findings are surprising given that in Mondini et al. (2016), individuals with mild to moderate probable dementia and lower CR improved more in global cognition after cognitive training than those with higher CR. Similarly, Long et al. (2020) demonstrated that healthy younger and older language learners who had lower Gaelic knowledge and consequently were placed in a Gaelic beginner class improved more on attentional switching after intensive language training than those in advanced courses.

However, our moderation analysis using baseline cognitive performance as a predictor was significant. Those Spanish learners who scored lower on baseline cognition were more likely to improve in response inhibition measured by a go/nogo task paradigm than language learners with higher baseline cognitive performance.

The baseline performance-dependent improvement in response inhibition is striking in several ways. First, it is in line with a recent publication by Kliesch et al. (2022). They detected that in healthy older adults, lower baseline cognition was related to more substantial gains in cognitive outcomes exclusively in language learners and most intensely for WM accuracy. In contrast, the authors found no such significant associations in the active and passive control groups. Their results underline the unique benefits of foreign language learning for subjects with lower baseline cognition who are more vulnerable to cognitive decline. According to the CR hypothesis, these individuals might benefit more from engaging in cognitively stimulating activities as their CR is lower and, therefore, their brains have more room for improvement (Valenzuela & Sachdev, 2006). The reason we did not find a significant moderation model in the proxy measures of CR (education, occupation, and leisure activity) but only for global cognition as a predictor may have the reason that global cognition is usually considered a more direct representation of CR (Opdebeeck et al., 2016). As the CR hypothesis predicts, individuals with a lower level of cognitive functioning are more likely to develop dementia than those with better cognitive performance (Whalley et al., 2004). Thus, the influence of CR on cognitive outcomes might have just been more readily apparent with a more direct representation of CR.

Second, while no study in this field on healthy older adults has included a go/nogo task paradigm, response inhibition may be promising as a core cognitive domain affected by foreign language learning. The go/nogo task paradigm is the only one, apart from the Stop-Signal Task, to show a marked age-related deficit in inhibition compared to other tasks frequently applied in bilingual research, such as the Stroop, Flanker, or Simon task (Rey-Mermet & Gade, 2018). According to early work by Persad et al. (2002), inhibition deficits are a precocious hallmark of cognitive decline. Also, inhibition deficits may underlie impairments in other cognitive domains, such as attention and episodic memory. Consequently, older adults are less able to suppress dominant responses while simultaneously maintaining two task sets (go- vs. nogo-stimulus). Applied to the context of foreign language learning, learners' brains may be trained to suppress dominant responses (i.e., their mother tongue) while processing the less dominant stimulus (i.e., the foreign language). This view also aligns with the Adaptive Control Hypothesis [see chapter 1.4.2]. According to this hypothesis, language learners in a dual-language context, which is commonly given in a classroom setting, require selective response inhibition when communicating in the less dominant language (Green & Abutalebi, 2013), since both languages are always active in the brain (Kroll et al., 2015). Learning a foreign language, therefore, may be among the most effective interventions to prevent cognitive decline and dementia, as it is likely to directly affect response inhibition, which is probably fundamental to healthy cognitive aging (Persad et al., 2002).

We interpret our finding in the context of higher cognitive engagement among those with lower global cognitive baseline scores. On the one hand, we derived this assumption from a highly significant negative association between the vocabulary test result and the change score in response inhibition. According to this correlation, individuals with lower vocabulary test scores improved more in response inhibition. The fact that there was no significant correlation between additional hours spent with homework activities and the vocabulary test result rules out the plausible explanation that individuals with lower test results just engaged less. It may instead indicate greater cognitive strain among those same individuals. On the other hand, we derived our conclusion from qualitative feedback from the attendees. Their feedback supports the view that participants were differently challenged: Four participants judged the differences in their learning rates and their past exposure to related languages (for example, French) as excessively great. While one participant felt the sessions could have gone more rapidly, others thought the pace was too fast and would have appreciated additional time for repetition. Aside from large differences in global cognition between participants, disparities in previous foreign language skills, e.g., in Romanic languages, may imply that participants were differentially challenged. Even though we made sure that previous skills in Romanic languages were low (< B1 as defined by questions derived from the Joint European Reference Frame for Languages) and that participants were not bilingual, there was a significant correlation between the vocabulary test score and the LSBQ score. Individuals who had fewer foreign language skills, as assessed by the LSBQ, also performed worse on the vocabulary test. Therefore, these individuals had less prior knowledge to draw on during the course and thus may have been more challenged.

5.3 Strengths and limitations

Several strengths and limitations of our study design and analysis must be respected when interpreting our results.

We conducted the first study to implement a short but intensive foreign language course in a randomized-controlled design. Moreover, we essentially implemented the recommendations that Ware et al. (2021) established regarding the conduct of foreign language studies. First, we chose our primary outcomes theory- and evidence-based. Second, we included a broad test battery of change-sensitive cognitive outcomes from the domains of executive attention, EF, verbal fluency, and attention. Third, we conducted an objective foreign language assessment of the language taught to exclude bilingual individuals and those with more than basic knowledge of Spanish, who are thus not suitable for an A1.1 course for beginners. Furthermore, we excluded participants with suspected cognitive impairment using a comprehensive neuropsychological test battery, and we administered a vocabulary test at the end of the intervention to assess learning gains.

One drawback is our modest sample size. Due to the COVID-19 pandemic, the last language course, for which ten participants were scheduled, could not be conducted. Thus, together with the missing participants from the WLCG, the pandemic outbreak resulted in a total loss of 20 participants. Hundreds of clinical trials were affected by the pandemic (Carlisle, 2020). The fact that these studies could not all reach their initially estimated sample size should be considered when evaluating them. However, we claim that our sample size was still large enough to detect medium effects in one of our primary outcomes, as the correlation between repeated measures was higher than pre-estimated in our study protocol. Therefore, even with a smaller sample size, it was still feasible to detect medium effects in our primary outcomes, an effect size also found in other related studies in the field (Bak, 2016; Pfenninger & Polz, 2018; Wong et al., 2019). Furthermore, no small effect was apparent for most of the primary and secondary outcomes, making it unlikely that 20 more participants would have materially altered the results. It also should be noted that the current study was meant to be exploratory to identify cognitive domains and tasks altered by foreign language learning. The cognitive effects of the early stages of foreign language acquisition in older adults have received little attention thus far. Therefore, preliminary studies must first be carried out before conducting large-scale studies.

One could also argue that our intervention was not long or the proficiency level attained was not high enough to elicit more noticeable cognitive changes. However, we posit that Bak et al. (2016) and Long et al. (2020) found an improvement in attentional switching after even 1 week of high-frequency foreign language training. Moreover, lengthier interventions, e.g., 8 months (Kliesch et al., 2022; Ramos et al., 2017), with thus higher proficiency level attained, were not necessarily associated with cognitive improvements. Additionally, none of the low-frequency courses resulted in significant cognitive gains (Bubbico et al., 2019; Valis et al., 2019; Ware et al., 2017). Consequently, it seems reasonable that the frequency of training held over a specified period is more relevant than its actual length or the proficiency level reached. In addition, from cognitive training studies, it is well known that even significantly less training than delivered in our study can lead to significant improvements in cognition in older adults (Borella et al., 2010: three training sessions; Carretti et al., 2013: six training sessions completed within 2 weeks; Nouchi et al., 2012: game playing for 15 min per day, at least 5 days per week for 4 weeks; see the review of Kelly et al. [2014] for more examples of studies). Furthermore, a very recent publication showed that benefits of cognitive training reach a plateau after 12–14 h of training (Belleville et al., 2021), a number of training sessions we provided in our study. In sum – even though our study differs from other related studies in terms of intervention, design, outcomes, or population, there is sound evidence in the literature that the dosage of our intervention in terms of frequency and proficiency level achieved was high enough to affect cognition generally. Additionally, we actually found an improvement in response inhibition in individuals with lower baseline cognition. Another recent study also supports the importance of baseline cognition for cognitive improvements related to foreign language learning (Kliesch et al., 2022). Given that our sample was generally well educated, as can be inferred from the education sub-score of the CRIq, and higher education is linked to better cognitive performance (Opdebeeck et al., 2016), cognitive improvements could have been obscured among those with higher education. Concluding from our results, we thus consider finding a way to adequately address all participants intellectually in a treatment plan more crucial than its actual dosage.

Another caveat is that we did not include an active comparator. We justify this decision on economic grounds. As previous studies have not conclusively identified cognitive domains addressed by foreign language learning, it is essential to evaluate these before adjusting for alternative explanations (Puhan et al., 2008). Thus, we cannot rule out that the benefit in response inhibition after foreign language learning for subjects with lower baseline cognition is due to general cognitive stimulation through social interaction (Kelly et al., 2017). Excluding alternative explanations was also not the aim of our exploratory study, as we first wanted to uncover cognitive domains and tasks potentially impacted by foreign language learning. However, one indication that the effect found could be due to foreign language learning is that the vocabulary test score, which is highly language-related, strongly correlated with changes in response inhibition. Still, we urge that future studies should include an active control group to account for other possible explanations.

5.4 Conclusion

We did not observe any overarching benefits of foreign language learning in executive attention or EF. However, we found evidence that foreign language learning may improve response inhibition, a domain particularly affected by cognitive aging, in older adults with lower global cognitive baseline levels. Based on our findings, future studies should consider individual differences and target participants with lower baseline cognition, who are thus more vulnerable to cognitive decline and dementia. Systematically selecting participants with lower baseline performance and less foreign language experience might be similar to physical training studies that exclude subjects who are too physically fit for a particular intervention (Jansen et al., 2021). In parallel, interventions must be more flexible to learners' abilities and needs. For example, applying technology-assisted learning of grammatical rules and vocabulary, e.g., using an app (Meltzer et al., 2021), and individually guided and planned by a specialized teacher, would better allow meeting participants at their performance level. At the same time, frustration due to excessive demands, which is more likely in a group setting with different learning paces, would be avoided. If combined with informal group activities covering recreational topics such as culture or traveling, social aspects of language learning would also be addressed to promote individual well-being and motivation through social integration (Pfenninger & Singleton, 2019). In view of these aspects, we might find more conclusive evidence of the impact of foreign language learning on executive attention and EF. This presupposes that we are able to determine the precise training dose, both in terms of length, frequency, and difficulty, and cognitive domains addressed by foreign language learning. Therefore, we will need well-controlled studies with large sample sizes in this field to delineate the specificity of the effects of foreign language learning on cognition as opposed to other cognitively stimulating interventions. Meanwhile, as much as other cognitively enriching leisure time activities that are unlikely to cause harm foreign language learning should be recommended to strengthen CR.

References

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

Abutalebi, J., Guidi, L., Borsa, V., Canini, M., Della Rosa, P. A., Parris, B. A., & Weekes, B. S. (2015). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia*, *69*, 201–210. https://doi.org/10.1016/j.neuropsychologia.2015.01.040

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*(2), 207–245. https://doi.org/10.3102/0034654310368803

Al-Marzouki, S., Roberts, I., Evans, S., & Marshall, T. (2008). Selective reporting in clinical trials: Analysis of trial protocols accepted by The Lancet. *The Lancet*, *372*(9634), 201. https://doi.org/10.1016/S0140-6736(08)61060-0

Anderson, J. A. E., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Anderson, V., Jacobs, R., & Anderson, P. J. (Eds.). (2014). *Studies on neuropsychology, neurology, and cognition*. *Executive functions and the frontal lobes: A lifespan perspective*: Psychology.

Ansaldo, A. I., Ghazi-Saidi, L., & Adrover-Roig, D. (2015). Interference control in elderly bilinguals: Appearances can be misleading. *Journal of Clinical and Experimental Neuropsychology*, *37*(5), 455–470. https://doi.org/10.1080/13803395.2014.990359

Antón, E., Fernández García, Y., Carreiras, M., & Duñabeitia, J. A. (2016). Does bilingualism shape inhibitory control in the elderly? *Journal of Memory and Language*, *90*, 147–160. https://doi.org/10.1016/j.jml.2016.04.007

Antoniou, M., Gunasekera, G. M., & Wong, P. C. M. (2013). Foreign language training as cognitive therapy for age-related cognitive decline: A hypothesis for future research. *Neuroscience and Biobehavioral Reviews*, *37*(10 0 2), 2689–2698. https://doi.org/10.1016/j.neubiorev.2013.09.004

Antoniou, M., & Wright, S. M. (2017). Uncovering the mechanisms responsible for why language learning may promote healthy cognitive aging. *Frontiers in Psychology*, *8*, 2217. https://doi.org/10.3389/fpsyg.2017.02217

Baddeley, A. D. (1986). *Working memory*, Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485–493. https://doi.org/10.1037/0894-4105.8.4.485

Bak, T. H. (2016). Language lessons to help protect against dementia. *BMJ (Clinical Research Ed.)*, *354*, i5039. https://doi.org/10.1136/bmj.i5039

Bak, T. H., Long, M. R., Vega-Mendoza, M., & Sorace, A. (2016). Novelty, challenge, and practice: The impact of intensive language learning on attentional functions. *PloS One*, *11*(4), e0153485. https://doi.org/10.1371/journal.pone.0153485

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*, *75*(6), 959–963. https://doi.org/10.1002/ana.24158

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging*, *12*(1), 12–21. https://doi.org/10.1037/0882-7974.12.1.12

Belleville, S., Cloutier, S., Mellah, S., Willis, S. L., Vellas, B., Coley, N., Andrieu, S., & Ngandu, T. (2021). When is more better? Modeling the effect of dose on the efficacy of the MAPT multidomain intervention as a function of individual characteristics. *Alzheimer's & Dementia*, *17*(S10). https://doi.org/10.1002/alz.054948

Berggren, R., Nilsson, J., Brehmer, Y., Schmiedek, F., & Lövdén, M. (2020). Foreign language learning in older age does not improve memory or intelligence: Evidence from a randomized controlled study. *Psychology and Aging*, *35*(2), 212–219. https://doi.org/10.1037/pag0000439

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bialystok, E., Craik, F. I. M., Klein, R. M., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. https://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(4), 859–873. https://doi.org/10.1037/0278-7393.34.4.859

Borella, E., Carretti, B., Riboldi, F., & Beni, R. de (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging*, *25*(4), 767–778. https://doi.org/10.1037/a0020683

Branzi, F. M., Della Rosa, P. A., Canini, M., Costa, A., & Abutalebi, J. (2016). Language control in bilinguals: Monitoring and response selection. *Cerebral Cortex*, *26*(6), 2367–2380. https://doi.org/10.1093/cercor/bhv052

Brini, S., Sohrabi, H. R., Hebert, J. J., Forrest, M. R. L., Laine, M., Hämäläinen, H., Karrasch, M., Peiffer, J. J., Martins, R. N., & Fairchild, T. J. (2020). Bilingualism is associated with a delayed onset of dementia but not with a lower risk of developing it: A systematic review with meta-analyses. *Neuropsychology Review*, *30*(1), 1–24. https://doi.org/10.1007/s11065-020-09426-8

Bubbico, G., Chiacchiaretta, P., Parenti, M., Di Marco, M., Panara, V., Sepede, G., Ferretti, A., & Perrucci, M. G. (2019). Effects of second language learning on the plastic aging brain: Functional connectivity, cognitive decline, and reorganization. *Frontiers in Neuroscience*, *13*, 423. https://doi.org/10.3389/fnins.2019.00423

Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., & Bedenbaugh, P. H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging & Mental Health*, *11*(4), 464–471. https://doi.org/10.1080/13607860601086504

Caldwell, A. R., Lakens, D., Parlett-Pelleriti, C. M., Prochilo, G., & Aust, F. (2022). *Power Analysis with Superpower*. Retrieved March 6, 2023, from https://aaroncaldwell.us/SuperpowerBook/

Cambridge Cognition. (2023). *Digit Span (DGS)*. Retrieved March 6, 2023, from https://www.cambridgecognition.com/cantab/cognitive-tests/memory/digit-span-dgs

Carlisle, B. G. (2020). *Over 1,200 clinical trials worldwide have ground to a halt due to COVID-19*. Retrieved April 2, 2023, from https://www.transparimed.org/single-post/2020/05/08/Discontinued-clinical-trials-COVID

Carretti, B., Borella, E., Zavagnin, M., & Beni, R. de (2013). Gains in language comprehension relating to working memory training in healthy older adults. *International Journal of Geriatric Psychiatry*, *28*(5), 539–546. https://doi.org/10.1002/gps.3859

Celik, S., Kokje, E., Meyer, P., Frölich, L., & Teichmann, B. (2020). Does bilingualism influence neuropsychological test performance in older adults? A systematic review. *Applied Neuropsychology*, *29*(4), 855–873. https://doi.org/10.1080/23279095.2020.1788032

Chan, A.-W., Tetzlaff, J. M., Gøtzsche, P. C., Altman, D. G., Mann, H., Berlin, J. A., Dickersin, K., Hróbjartsson, A., Schulz, K. F., Parulekar, W. R., Krleza-Jeric, K., Laupacis, A., & Moher, D. (2013). Spirit 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ (Clinical Research Ed.)*, *346*, e7586. https://doi.org/10.1136/bmj.e7586

Clare, L., Whitaker, C. J., Martyr, A., Martin-Forbes, P. A., Bastable, A. J., Pye, K. L., Quinn, C., Thomas, E. M., Mueller Gathercole, V. C., & Hindle, J. V. (2016). Executive control in older Welsh monolinguals and bilinguals. *Journal of Cognitive Psychology*, *28*(4), 412–426. https://doi.org/10.1080/20445911.2016.1148041

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition), Hillsdale, NY: Erlbaum.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149. https://doi.org/10.1016/j.cognition.2009.08.001

Darowski, E. S., Helder, E., Zacks, R. T., Hasher, L., & Hambrick, D. Z. (2008). Age-related differences in cognition: The role of distraction control. *Neuropsychology*, *22*(5), 638–644. https://doi.org/10.1037/0894-4105.22.5.638

Dash, T., Berroir, P., Joanette, Y., & Ansaldo, A. I. (2019). Alerting, orienting, and executive control: The effect of bilingualism and age on the subcomponents of attention. *Frontiers in Neurology*, *10*, 1122. https://doi.org/10.3389/fneur.2019.01122

de Bruin, A., Della Sala, S., & Bak, T. H. (2016). The effects of language use on lexical processing in bilinguals. *Language, Cognition and Neuroscience*, *31*(8), 967–974. https://doi.org/10.1080/23273798.2016.1190024

Degirmenci, M. G., Grossmann, J. A., Meyer, P., & Teichmann, B. (2022). The role of bilingualism in executive functions in healthy older adults: A systematic review. *International Journal of Bilingualism*, *26*(4), 426–449. https://doi.org/10.1177/13670069211051291

Di Martino, A., Ghaffari, M., Curchack, J., Reiss, P., Hyde, C., Vannucci, M., Petkova, E., Klein, D. F., & Castellanos, F. X. (2008). Decomposing intra-subject variability in children with attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *64*(7), 607–614. https://doi.org/10.1016/j.biopsych.2008.03.008

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Donnelly, S. (2016). *Re-examining the bilingual advantage on interference-control and task-switching tasks: A meta-analysis*. Dissertation.

Donnelly, S., Brooks, P. J., & Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review*, *26*(4), 1122–1147. https://doi.org/10.3758/s13423-019-01567-z

Duncan, H. D., Nikelski, J., Pilon, R., Steffener, J., Chertkow, H., & Phillips, N. A. (2018). Structural brain differences between monolingual and multilingual patients with mild cognitive impairment and Alzheimer disease: Evidence for cognitive reserve. *Neuropsychologia*, *109*, 270–282. https://doi.org/10.1016/j.neuropsychologia.2017.12.036

Ernst Klett Verlag. (2004). *Placement test Spanish A1*. Retrieved March 14, 2023, from https://tests.online-platform.net/test_es.php

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive*

Neuroscience, *14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports*, *11*(1), 1382. https://doi.org/10.1038/s41598-020-80866-1

Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)* (3. ed.), Los Angeles: Sage Publ.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Fournet, N., Roulin, J.-L., Vallet, F., Beaudoin, M., Agrigoroaei, S., Paignon, A., Dantzer, C., & Desrichard, O. (2012). Evaluating short-term and working memory in older adults: French normative data. *Aging & Mental Health*, *16*(7), 922–930. https://doi.org/10.1080/13607863.2012.674487

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

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. https://doi.org/10.1037/0096-3445.137.2.201

Gallo, F., DeLuca, V., Prystauka, Y., Voits, T., Rothman, J., & Abutalebi, J. (2022). Bilingualism and aging: Implications for (delaying) neurocognitive decline. *Frontiers in Human Neuroscience*, *16*, 819105. https://doi.org/10.3389/fnhum.2022.819105

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., Hernández, M., Costa, A., & Avila, C. (2010). Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *NeuroImage*, *53*(4), 1272–1278. https://doi.org/10.1016/j.neuroimage.2010.05.078

Gauthier, S., Rosa-Neto, P., Morais, J.A., & Webster, C. (2021). *World Alzheimer Report 2021: Journey through the diagnosis of dementia*, London, England: Alzheimer's Disease International.

Gmehlin, D., Stelzel, C., Weisbrod, M., Kaiser, S., & Aschenbrenner, S. (2012). *Manual Task Switching*, Moedling, Austria: SCHUHFRIED GmbH.

Goerrissen, M. (2016). *Caminos hoy A1: Kurs- und Übungsbuch mit MP3-CD und DVD* (1. Auflage), Stuttgart: Ernst Klett Sprachen.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grossmann, J. A., Aschenbrenner, S., Teichmann, B., & Meyer, P. (2023). Foreign language learning can improve response inhibition in individuals with lower baseline cognition: Results from a randomized controlled superiority trial. *Frontiers in Aging Neuroscience*, *15*, 1–16. https://doi.org/10.3389/fnagi.2023.1123185

Grossmann, J. A., Koelsch, V. M., Degirmenci, M. G., Aschenbrenner, S., Teichmann, B., & Meyer, P. (2021). Effects of foreign language learning on executive functions in healthy older adults: Study protocol for a randomised controlled trial. *BMC Geriatrics*, *21*(1), Article 122, 1–14. https://doi.org/10.1186/s12877-021-02051-x

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*(3), 325–340. https://doi.org/10.1177/0267658316678286

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *17*(1), 163–169. https://doi.org/10.1037//0278-7393.17.1.163

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, *22*, 193–225. https://doi.org/10.1016/S0079-7421(08)60041-9

Head, D., Buckner, R. L., Shimony, J. S., Williams, L. E., Akbudak, E., Conturo, T. E., McAvoy, M., Morris, J. C., & Snyder, A. Z. (2004). Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: Evidence from diffusion tensor imaging. *Cerebral Cortex*, *14*(4), 410–423. https://doi.org/10.1093/cercor/bhh003

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, *18*(4), 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hilchey, M. D., Saint-Aubin, J., & Klein, R. M. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. H. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 586–613): Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107447257.026

Iizuka, A., Suzuki, H., Ogawa, S., Kobayashi-Cuya, K. E., Kobayashi, M., Takebayashi, T., & Fujiwara, Y. (2019). Can cognitive leisure activity prevent cognitive decline in older adults? A systematic review of intervention studies. *Geriatrics & Gerontology International*, *19*(6), 469–482. https://doi.org/10.1111/ggi.13671

Jahn, T. (2016). *Manual Vienna Verbal Fluency Test*, Moedling, Austria: SCHUHFRIED GmbH.

Jahn, T., & Heßler, J. B. (2017). *Cognitive Functions Dementia*, Moedling, Austria: SCHUHFRIED GmbH.

Jansen, C.-P., Nerz, C., Labudek, S., Gottschalk, S., Kramer-Gmeiner, F., Klenk, J., Dams, J., König, H.-H., Clemson, L., Becker, C., & Schwenk, M. (2021). Lifestyle-integrated functional exercise to prevent falls and promote physical activity: Results from the LiFE-is-LiFE randomized non-inferiority trial. *The International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 115. https://doi.org/10.1186/s12966-021-01190-z

Kaiser, S., Aschenbrenner, S., Pfueller, U., Roesch-Ely, D., & Weisbrod, M. (2010). *Manual Response Inhibition*, Moedling, Austria: SCHUHFRIED GmbH.

Katzman, R., Terry, R., DeTeresa, R., Brown, T., Davies, P., Fuld, P., Renbing, X., & Peck, A. (1988). Clinical, pathological, and neurochemical changes in dementia: A subgroup with preserved mental status and numerous neocortical plaques. *Annals of Neurology*, *23*(2), 138–144. https://doi.org/10.1002/ana.410230206

Kelly, M. E., Duff, H., Kelly, S., McHugh Power, J. E., Brennan, S., Lawlor, B. A., & Loughrey, D. G. (2017). The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: A systematic review. *Systematic Reviews*, *6*(1), 259. https://doi.org/10.1186/s13643-017-0632-2

Kelly, M. E., Loughrey, D. G., Lawlor, B. A., Robertson, I. H., Walsh, C., & Brennan, S. (2014). The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: A systematic review and meta-analysis. *Ageing Research Reviews*, *15*, 28–43. https://doi.org/10.1016/j.arr.2014.02.004

Kliesch, M., Giroud, N., Pfenninger, S. E., & Meyer, M. (2018). Research on second language acquisition in old adulthood: What we have and what we need. In D. Gabryś-Barker (Ed.), *Second language acquisition: Vol. 120. Third age learners of foreign languages :* Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781783099412-006

Kliesch, M., Pfenninger, S. E., Wieling, M., Stark, E., & Meyer, M. (2022). Cognitive benefits of learning additional languages in old adulthood? Insights from an intensive longitudinal intervention study. *Applied Linguistics*, *43*(4), 653–676. https://doi.org/10.1093/applin/amab077

Kousaie, S., & Phillips, N. A. (2012). Ageing and bilingualism: Absence of a "bilingual advantage" in stroop interference in a nonimmigrant sample. *Quarterly Journal of Experimental Psychology*, *65*(2), 356–369. https://doi.org/10.1080/17470218.2011.604788

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147. https://doi.org/10.1037//0882-7974.15.1.126

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, *1*, 377–394. https://doi.org/10.1146/annurev-linguist-030514-124937

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. https://doi.org/10.1037/bul0000142

Li, T., Boutron, I., Al-Shahi Salman, R., Cobo, E., Flemyng, E., Grimshaw, J. M., & Altman, D. G. (2016). Review and publication of protocol submissions to Trials - what have we learned in 10 years? *Trials*, *18*(1), 34. https://doi.org/10.1186/s13063-016-1743-0

Lim, Y. Y., Baker, J. E., Bruns, L., Mills, A., Fowler, C., Fripp, J., Rainey-Smith, S. R., Ames, D., Masters, C. L., & Maruff, P. (2020). Association of deficits in short-term learning and Aβ and hippocampal volume in cognitively normal adults. *Neurology*, *95*(18), e2577-e2585. https://doi.org/10.1212/WNL.0000000000010728

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., . . . Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Long, M. R., Vega-Mendoza, M., Rohde, H., Sorace, A., & Bak, T. H. (2020). Understudied factors contributing to variability in cognitive performance related to language learning. *Bilingualism: Language and Cognition*, *23*(4), 801–811. https://doi.org/10.1017/S1366728919000749

Luk, G., Bialystok, E., Craik, F. I. M., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of Neuroscience*, *31*(46), 16808–16813. https://doi.org/10.1523/JNEUROSCI.4563-11.2011

Luk, G., de Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, *14*(4), 588–595. https://doi.org/10.1017/S1366728911000010

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2011). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, *27*(10), 1479–1488. https://doi.org/10.1080/01690965.2011.613209

Maurer-Karattup, P., Neumann, O., Danneil, W., & Thöne-Otto, A. I. T. (2022). Leitlinienbasierte Standards zur Struktur- und Prozessqualität neuropsychologischer Diagnostik und Therapie. *Zeitschrift Für Neuropsychologie*, *33*(3), 115–128. https://doi.org/10.1024/1016-264X/a000362

Melrose, R. J., Young, S., Weissberger, G. H., Natta, L., Harwood, D., Mandelkern, M., & Sultzer, D. L. (2017). Cerebral metabolic correlates of attention networks in Alzheimer's Disease: A study of the Stroop. *Neuropsychologia*, *106*, 383–389. https://doi.org/10.1016/j.neuropsychologia.2017.10.020

Meltzer, J. A., Kates Rose, M., Le, A. Y., Spencer, K. A., Goldstein, L., Gubanova, A., Lai, A. C., Yossofzai, M., Armstrong, S. E. M., & Bialystok, E. (2021). Improvement in executive function for older adults through smartphone apps: A randomized clinical trial comparing language learning and brain training. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 1–22. https://doi.org/10.1080/13825585.2021.1991262

Mills, E. J., Seely, D., Rachlis, B., Griffith, L., Wu, P., Wilson, K., Ellis, P., & Wright, J. R. (2006). Barriers to participation in clinical trials of cancer: A meta-analysis and systematic review of patient-reported factors. *The Lancet Oncology*, *7*(2), 141–148. https://doi.org/10.1016/S1470-2045(06)70576-9

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. https://doi.org/10.1177/0963721411429458

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

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., & Altman, D. G. (2010). Consort 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed.)*, *340*, c869. https://doi.org/10.1136/bmj.c869

Mondini, S., Madella, I., Zangrossi, A., Bigolin, A., Tomasi, C., Michieletto, M., Villani, D., Di Giovanni, G., & Mapelli, D. (2016). Cognitive reserve in dementia: Implications for cognitive training. *Frontiers in Aging Neuroscience*, *8*, 84. https://doi.org/10.3389/fnagi.2016.00084

Monnier, C., Boiché, J., Armandon, P., Baudoin, S., & Bellocchi, S. (2021). Is bilingualism associated with better working memory capacity? A meta-analysis. *International Journal of Bilingual Education and Bilingualism*, 1–27. https://doi.org/10.1080/13670050.2021.1908220

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414. https://doi.org/10.1212/WNL.43.11.2412-a

Morris, J. C. (2005). Early-stage and preclinical Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *19*(3), 163–165. https://doi.org/10.1097/01.wad.0000184005.22611.cc

Mukadam, N., Sommerlad, A., & Livingston, G. (2017). The relationship of bilingualism compared to monolingualism to the risk of cognitive decline or dementia: A systematic review and meta-analysis. *Journal of Alzheimer's Disease*, *58*(1), 45–54. https://doi.org/10.3233/JAD-170131

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, *15*(10), 453–459. https://doi.org/10.1016/j.tics.2011.07.011

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nguyen, L., Murphy, K., & Andrews, G. (2019). Immediate and long-term efficacy of executive functions cognitive training in olde adults: A systematic review and meta-analysis. *Psychological Bulletin*, *145*(7), 698–733. https://doi.org/10.1037/bul0000196

Nicholls, M. E. R., Thomas, N. A., Loetscher, T., & Grimshaw, G. M. (2013). The Flinders Handedness survey (FLANDERS): A brief measure of skilled hand preference. *Cortex*, *49*(10), 2914–2926. https://doi.org/10.1016/j.cortex.2013.02.002

Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., Abdoli, A., Abualhasan, A., Abu-Gharbieh, E., Akram, T. T., Al Hamad, H., Alahdab, F., Alanezi, F. M., Alipour, V., Almustanyir, S., Amu, H., Ansari, I., Arabloo, J., Ashraf, T., . . . Vos, T. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*, *7*(2), e105-e125. https://doi.org/10.1016/S2468-2667(21)00249-8

Nouchi, R., Taki, Y., Takeuchi, H., Hashizume, H., Akitsuki, Y., Shigemune, Y., Sekiguchi, A., Kotozaki, Y., Tsukiura, T., Yomogida, Y., & Kawashima, R. (2012). Brain training game improves executive functions and processing speed in the elderly: A randomized controlled trial. *PloS One*, *7*(1), e29676. https://doi.org/10.1371/journal.pone.0029676

Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive Reserve Index questionnaire (CRIq): A new instrument for measuring cognitive reserve. *Aging Clinical and Experimental Research*, *24*(3), 218–226. https://doi.org/10.3275/7800

Opdebeeck, C., Martyr, A., & Clare, L. (2016). Cognitive reserve and cognitive function in healthy older people: A meta-analysis. *Aging, Neuropsychology, and Cognition*, *23*(1), 40–60. https://doi.org/10.1080/13825585.2015.1041450

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037/0882-7974.17.2.299

Park, D. C., & Reuter-Lorenz, P. A. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*, 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Perani, D., Farsad, M., Ballarini, T., Lubian, F., Malpetti, M., Fracchetti, A., Magnani, G., March, A., & Abutalebi, J. (2017). The impact of bilingualism on brain reserve and metabolic connectivity in Alzheimer's dementia. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(7), 1690–1695. https://doi.org/10.1073/pnas.1610909114

Persad, C. C., Abeles, N., Zacks, R. T., & Denburg, N. L. (2002). Inhibitory changes after age 60 and their relationship to measures of attention and memory. *The Journals of Gerontology: Psychological Sciences and Social Sciences*, *57*(3), P223-32. https://doi.org/10.1093/geronb/57.3.p223

Petermann, F. (2012). *WAIS-IV: Wechsler Adult Intelligence Scale - Fourth Edition - German adaptation after David Wechsler*, Frankfurt: Pearson.

Petersen, R. C. (2011). Clinical practice. Mild cognitive impairment. *The New England Journal of Medicine*, *364*(23), 2227–2234. https://doi.org/10.1056/NEJMcp0910237

Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S. D., Ganguli, M., Gloss, D., Gronseth, G. S., Marson, D., Pringsheim, T., Day, G. S., Sager, M., Stevens, J., & Rae-Grant, A. (2018). Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology*, *90*(3), 126–135. https://doi.org/10.1212/WNL.0000000000004826

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(6), 760. https://doi.org/10.1001/archneur.56.6.760

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Pfenninger, S. E., & Polz, S. (2018). Foreign language learning in the third age: A pilot feasibility study on cognitive, socio-affective and linguistic drivers and benefits in relation to previous bilingualism of the learner. *Journal of the European Second Language Association*, *2*(1), 1–13. https://doi.org/10.22599/jesla.36

Pfenninger, S. E., & Singleton, D. (2019). A critical review of research relating to the learning, use and effects of additional and multiple languages in later life. *Language Teaching*, *52*(4), 419–449. https://doi.org/10.1017/S0261444819000235

Pillai, J. A., Hall, C. B., Dickson, D. W., Buschke, H., Lipton, R. B., & Verghese, J. (2011). Association of crossword puzzle participation with memory decline in persons who develop dementia. *Journal of the International Neuropsychological Society*, *17*(6), 1006–1013. https://doi.org/10.1017/S1355617711001111

Pliatsikas, C. (2020). Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition*, *23*(2), 459–471. https://doi.org/10.1017/S1366728919000130

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Puhan, M. A., Vollenweider, D., Steurer, J., Bossuyt, P. M., & Ter Riet, G. (2008). Where is the supporting evidence for treating mild to moderate chronic obstructive pulmonary disease exacerbations with antibiotics? A systematic review. *BMC Medicine*, *6*, 28. https://doi.org/10.1186/1741-7015-6-28

Ramos, S., Fernández García, Y., Antón, E., Casaponsa, A., & Duñabeitia, J. A. (2017). Does learning a language in the elderly enhance switching ability? *Journal of Neurolinguistics*, *43*, 39–48. https://doi.org/10.1016/j.jneuroling.2016.09.001

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., & Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex*, *15*(11), 1676–1689. https://doi.org/10.1093/cercor/bhi044

Raz, N., & Rodrigue, K. M. (2006). Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience & Biobehavioral Reviews*, *30*(6), 730–748. https://doi.org/10.1016/j.neubiorev.2006.07.001

Reijnders, J., van Heugten, C., & van Boxtel, M. (2013). Cognitive interventions in healthy older adults and people with mild cognitive impairment: A systematic review. *Ageing Research Reviews*, *12*(1), 263–275. https://doi.org/10.1016/j.arr.2012.07.003

Reuter-Lorenz, P. A., & Park, D. C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review*, *24*(3), 355–370. https://doi.org/10.1007/s11065-014-9270-9

Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review*, *25*(5), 1695–1716. https://doi.org/10.3758/s13423-017-1384-7

Roalf, D. R., Moberg, P. J., Xie, S. X., Wolk, D. A., Moelter, S. T., & Arnold, S. E. (2013). Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimer's & Dementia*, *9*(5), 529–537. https://doi.org/10.1016/j.jalz.2012.10.001

Rodewald, K., Weisbrod, M., & Aschenbrenner, S. (2012). *Manual Trail-Making Test - Langensteinbach Version*, Moedling, Austria: SCHUHFRIED GmbH.

Sadaghiani, S., & D'Esposito, M. (2015). Functional characterization of the cingulo-opercular network in the maintenance of tonic alertness. *Cerebral Cortex*, *25*(9), 2763–2773. https://doi.org/10.1093/cercor/bhu072

Sajeev, G., Weuve, J., Jackson, J. W., VanderWeele, T. J., Bennett, D. A., Grodstein, F., & Blacker, D. (2016). Late-life cognitive activity and dementia: A systematic review and bias analysis. *Epidemiology*, *27*(5), 732–742. https://doi.org/10.1097/EDE.0000000000000513

Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology*, *5*(1), Article 18. https://doi.org/10.1525/collabra.203

Salloway, S., Sperling, R. A., Fox, N. C., Blennow, K., Klunk, W., Raskind, M., Sabbagh, M., Honig, L. S., Porsteinsson, A. P., Ferris, S., Reichert, M., Ketter, N., Nejadnik, B., Guenzler, V., Miloslavsky, M., Wang, D., Lu, Y., Lull, J., Tudor, I. C., . . . Brashear, H. R. (2014). Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *The New England Journal of Medicine*, *370*(4), 322–333. https://doi.org/10.1056/NEJMoa1304839

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295x.103.3.403

Schelling, D., & Schuri, U. (2009). *Manual N-Back Verbal*, Moedling, Austria: SCHUHFRIED GmbH.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*, 1475. https://doi.org/10.3389/fpsyg.2014.01475

Schoenfeld, D. (1980). Statistical considerations for pilot studies. *International Journal of Radiation Oncology*Biology*Physics*, *6*(3), 371–374. https://doi.org/10.1016/0360-3016(80)90153-4

Schuhfried, G. *Vienna Test System* (Version 7.3.00) [Computer software]: SCHUHFRIED GmbH. Moedling, Austria.

Schuhfried, G. (1999). *Manual Stroop Interference Test*, Moedling, Austria: SCHUHFRIED GmbH.

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I. M., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex*, *48*(8), 991–996. https://doi.org/10.1016/j.cortex.2011.04.009

Sealed Envelope. *Simple Randomiser*. Retrieved December 9, 2020, from https://www.sealedenvelope.com/simple-randomiser/v1/lists

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*, 772. https://doi.org/10.3389/fpsyg.2014.00772

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., . . . Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 280–292. https://doi.org/10.1016/j.jalz.2011.03.003

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*(10), 2015–2028. https://doi.org/10.1016/j.neuropsychologia.2009.03.004

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology*, *11*(11), 1006–1012. https://doi.org/10.1016/S1474-4422(12)70191-6

Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M. J., Vemuri, P., & Vuoksimaa, E. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*, *16*(9), 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219

Sturm, W. (2006a). *Manual WAF Perception and Attention Function Battery - Alertness*, Moedling, Austria: SCHUHFRIED GmbH.

Sturm, W. (2006b). *Manual WAF Perception and Attention Function Battery - Divided Attention*, Moedling, Austria: SCHUHFRIED GmbH.

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin*, *95*(1), 3–28. https://doi.org/10.1037/0033-2909.95.1.3

Suchy, Y. (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *Annals of Behavioral Medicine*, *37*(2), 106–116. https://doi.org/10.1007/s12160-009-9097-4

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, *22*(2), 401–415. https://doi.org/10.1017/S1366728917000682

Swan, I. R., & Browning, G. G. (1985). The whispered voice as a screening test for hearing impairment. *The Journal of the Royal College of General Practitioners*, *35*(273), 197.

Tripepi, G., Chesnaye, N. C., Dekker, F. W., Zoccali, C., & Jager, K. J. (2020). Intention to treat and per protocol analysis in clinical trials. *Nephrology*, *25*(7), 513–517. https://doi.org/10.1111/nep.13709

United Nations. (2020). *World population ageing 2020 highlights: Living arrangements of older persons*, New York: United Nations.

Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and dementia: A systematic review. *Psychological Medicine*, *36*(4), 441–454. https://doi.org/10.1017/S0033291705006264

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(1), 3–24. https://doi.org/10.1017/S1366728914000522

Valis, M., Slaninova, G., Prazak, P., Poulova, P., Kacetl, J., & Klimova, B. (2019). Impact of learning a foreign language on the enhancement of cognitive functions among healthy older population. *Journal of Psycholinguistic Research*, *48*(6), 1311–1318. https://doi.org/10.1007/s10936-019-09659-6

van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S.-H., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, *9*(3). https://doi.org/10.3390/bs9030027

van Zomeren, A. H., & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*, New York: Oxford Univ. Press.

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition*, *137*, 40–46. https://doi.org/10.1016/j.cognition.2014.12.008

Vickers, A. J., & Altman, D. G. (2001). Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ*, *323*(7321), 1123–1124. https://doi.org/10.1136/bmj.323.7321.1123

Ware, C., Damnee, S., Djabelkhir, L., Cristancho, V., Wu, Y.-H., Benovici, J., Pino, M., & Rigaud, A.-S. (2017). Maintaining cognitive functioning in healthy seniors with a technology-based foreign language program: A pilot feasibility study. *Frontiers in Aging Neuroscience*, *9*, 42. https://doi.org/10.3389/fnagi.2017.00042

Ware, C., Dautricourt, S., Gonneaud, J., & Chételat, G. (2021). Does second language learning promote neuroplasticity in aging? A systematic review of cognitive and neuroimaging studies. *Frontiers in Aging Neuroscience*, *13*, Article 706672. https://doi.org/10.3389/fnagi.2021.706672

Whalley, L. J., Deary, I. J., Appleton, C. L., & Starr, J. M. (2004). Cognitive reserve and the neurobiology of cognitive aging. *Ageing Research Reviews*, *3*(4), 369–382. https://doi.org/10.1016/j.arr.2004.05.001

Wong, P. C. M., Ou, J., Pang, C. W. Y., Zhang, L., Tse, C. S., Lam, L. C. W., & Antoniou, M. (2019). Language training leads to global cognitive improvement in older adults: A preliminary study. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2411–2424. https://doi.org/10.1044/2019_JSLHR-L-18-0321

World Health Organisation. (2021). *Dementia*. Retrieved June 1, 2022, from https://www.who.int/news-room/fact-sheets/detail/dementia

Zhang, Y., Cai, J., An, L., Hui, F., Ren, T., Ma, H., & Zhao, Q. (2017). Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. *Ageing Research Reviews*, *35*, 1–11. https://doi.org/10.1016/j.arr.2016.12.003

Zhou, B., & Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychonomic Bulletin & Review*, *23*(4), 1221–1230. https://doi.org/10.3758/s13423-015-0981-6

List of abbreviations

ACC	Anterior cingulate cortex
AD	Alzheimer's Dementia
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
CDR	Clinical Dementia Rating Scale
CFD	Cognitive Functions Dementia
CONSORT	Consolidated Standard of Reporting Trials
CR	Cognitive reserve
CRIq	Cognitive Reserve Index questionnaire
DSB	Digit Span Backwards
DSF	Digit Span Forwards
EF	Executive Functions
FLANDERS	Flinders Handedness Survey
IFG	Inferior frontal gyrus
INHIB	Response Inhibition
LLG	Language learning group
LSBQ	The Language and Social Background Questionnaire
MCI	Mild cognitive impairment
MMSE	Mini-Mental Status Examination
MoCA	Montreal Cognitive Assessment
NAR	Network Aging Research
NBV	N-Back Verbal
RCT	Randomized-controlled trial
STAC-r	Scaffolding theory of aging and cognition - revised
STROOP	Stroop Interference Test
SWITCH	Task Switching
TMT	Trail-Making Test Langensteinbach Version
VTS	Vienna Test System
WAFA	Perception and Attention Function Battery - Alertness
WAFG	Perception and Attention Function Battery - Divided Attention
WAIS-IV	Wechsler Adult Intelligence Scale - Fourth Edition

WIWO	Vienna Verbal Fluency Test
WLCG	Waiting list control group
WM	Working memory

List of figures

List of tables

Declaration in accordance to § 8 (1) c) and d) of the doctoral degree regulation of the Faculty

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

FAKULTÄT FÜR VERHALTENS- UND EMPIRISCHE KULTURWISSENSCHAFTEN

Promotionsausschuss der Fakultät für Verhaltens- und Empirische Kulturwissenschaften der Ruprecht-Karls-Universität Heidelberg / Doctoral Committee of the Faculty of Behavioural and Cultural Studies of Heidelberg University

Erklärung gemäß § 8 (1) c) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften / Declaration in accordance to § 8 (1) c) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe. / I declare that I have made the submitted dissertation independently, using only the specified tools and have correctly marked all quotations.

Erklärung gemäß § 8 (1) d) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften / Declaration in accordance to § 8 (1) d) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe. / I declare that I did not use the submitted dissertation in this or any other form as an examination paper until now and that I did not submit it in another faculty.

Vorname Nachname / First name Family name	Judith Alina Großmann
Datum / Date	17.04.2023
Unterschrift / Signature	Dem Dekanat der Fakultät für Verhaltens- und Empirische Kulturwissenschaften liegt eine unterschriebene Version dieser Erklärung vom 17.04.2023 vor.