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 publication-based thesis

The Linguistic and Cultural Aspects of Neuropsychological Assessment in People with Dementia

presented by
Simge Celik

year of submission
2022

Dean: Prof. Dr. Guido Sprenger
Advisor: Prof. Dr. Dr. h.c. Andreas Kruse

Table of Contents

Acknowledgments

I would like to thank each of my supervisors, Prof. Dr. Patric Meyer, and Dr. Birgit Teichmann, for their advice, discussions, and support in improving the research methodologies included in this thesis and their help in finding solutions to the obstacles experienced along the way of my doctoral studies. Additionally, I would like to thank Prof. Dr. Hans-Werner Wahl for his assistance and support during the challenging recruitment process. I am also grateful to Prof. Dr. Dr. h.c. Andreas Kruse for agreeing to review my thesis at short notice.

I would like to extend my utmost appreciation to Prof. Dr. Özgür Onur, Prof. Dr. Josef Kessler, and Prof. Dr. Görsev Yener, without whom I would still be collecting data to this day. Your kind attitudes, support, and interest in my research projects were a source of great motivation in challenging times. I am very grateful for your immediate help and support in the recruitment of study participants and in procedures for the ethical approval of the research project.

I would also like to offer my appreciation to the participants, their families, or close relatives who were willing to participate in my testing procedure, without whom this thesis would not have been possible. I am incredibly thankful for their willingness to undergo a long-duration neuropsychological evaluation while facing Alzheimer's disease and to share stories of experiences with me. Thanks to you, I have learned more than I could have ever imagined over the past few years.

I would also like to thank my parents, brother, and partner for their endless support and encouragement throughout my academic endeavors. This journey would not have been possible without your unwavering support and guidance. Additionally, I loved being a part of an interdisciplinary team from diverse cultural backgrounds in Network Aging Research (NAR).

Firstly, I would like to express my gratitude to my lovely colleagues and officemates, Julia Kirch and Julia Schneider, for creating a motivating and supportive working atmosphere in our office. Secondly, I would like to extend my sincere and special thanks to my colleagues and friends, Eesha Kokje and Merve Gül Degirmenci, who have shaped so much of my Ph.D. experience for the better. Work breaks with you were always a delightful change from the long hours in the office. It was a great pleasure to work with you all, and thanks so much for making the office a place to dwell. Finally, I am grateful for the Robert Bosch Foundation Stuttgart's financial support for a doctoral fellowship program titled “People with Dementia in Acute Care Hospitals.” This scholarship provided me with the time and freedom to complete my thesis, particularly during delays or barriers associated with the COVID-19 pandemic.

List of Publications for the Publication-Based Dissertation

1st Publication (Study 1)

- Kokje, E., **Celik, S.,** Wahl, H. W., & von Stutterheim, C. (2021). Can discourse processing performance serve as an early marker of Alzheimer's disease and mild cognitive impairment? A systematic review of text comprehension. European journal of ageing, 19(1), 3–18. https://doi.org/10.1007/s10433-021-00619-5.

2nd Publication (Study 2)

- **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: Adult*, 1–19. https://doi.org/10.1080/23279095.2020.1788032.

3rd Publication (Study 3)

- **Celik, S.,** Onur, O., Yener, G., Kessler, J., Özbek, Y., Meyer, P., Frölich, L., & Teichmann, B. (2022). Cross-cultural comparison of MMSE and RUDAS in German and Turkish patients with Alzheimer's disease. Neuropsychology, 36(3), 195–205. https://doi.org/10.1037/neu0000764.

Abbreviations

General Abbreviations

AD	Alzheimer's Disease
AI	Assimilation Index
AoA	Age of Language Acquisition
BNT	Boston Naming Test
CDR	Clinical Dementia Rating Scale
CR	Cognitive Reserve
CO	Orientation towards Culture of Origin
CSF	Cerebrospinal Fluid
DSM-5	Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition
EDAT	Early-Stage Dementia of the Alzheimer's type
FRAKK	Frankfurt Acculturation Scale
GDS	Geriatric Depression Scale
GP	General Practitioner
GR-AD	German participants with Alzheimer's Disease
HC	Orientation towards Host Culture
L1	First (Native) Language
L2	Second Language
LSBQ	Language Background Questionnaire
MCI	Mild Cognitive Impairment
a-MCI	amnestic MCI

na-MCI	non-amnestic MCI
MDAT	Moderate Stage Dementia of the Alzheimer's type
MMSE	Mini-Mental State Examination
NP	Neuropsychological Test Performance
NIA-AA	National Institute on Aging-Alzheimer's Association
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QualSyst	The Standard Quality Assessment Criteria for Evaluating Primary Research Papers: Quality Scoring for Quantitative Studies
RUDAS	Rowland Universal Dementia Assessment Scale
TEA	Test of Everyday Attention
TMT	Trail-Making Test
TR-AD	Turkish participants with Alzheimer's Disease
TR-IM-AD	Turkish immigrants with Alzheimer's Disease
WEIRD	Western, Educated, Industrialized, Rich, and Democratic

Extended Abstract

Objectives: Given the public health crisis that Alzheimer's disease (AD) has become (Naylor et al., 2012), neuropsychological assessment tools that provide timely and accurate identification of cognitive decline in older adults have gained increasing focus in the scientific literature. Accurate evaluation of cognitive function and early identification of cognitive changes are paramount to understanding the disease course of AD and improving effective treatments and patients' quality of life. To this end, language offers a cognitive neuropsychological approach to identifying cognitive decline in the early stages of AD. Moreover, it represents a multi-dimensional variable that may influence the neuropsychological test performance of older adults due to its potential contribution to cognitive reserve. Therefore, the present thesis aims at combining two aspects of language to explore its potential in the early detection of AD and its association with neuropsychological test performance in older adults and cross-cultural neuropsychology. Study 1 assessed the currently available studies to explore whether discourse processing, particularly macro-structural discourse comprehension, offers a novel approach to neuropsychological testing in distinguishing normal cognitive aging from AD pathology-related decline. Study 2 evaluated the results of the studies that examined the impact of bilingualism on neuropsychological test performance in monolingual and bilingual older adults to inform the neuropsychological evaluation of these groups in clinical practice. Study 3 investigated the influence of bilingualism and its associated factors, namely, cultural background and acculturation, on cognitive screening tests in three clinically diagnosed AD patient groups to identify a cross-culturally/linguistically appropriate measure of cognition.

Method: Data of Study 1 and Study 2 were based on the original research studies published in English investigating discourse comprehension and bilingualism in healthy older adults,

individuals with mild cognitive impairment (MCI), and AD. A literature search focusing on these topics with participant groups aged 60 years and over was conducted in PubMed, Web of Science, and PsycINFO databases. Study 1 included eight articles consisting of studies only with cross-sectional designs. Study 2 was comprised of twenty-seven articles, of which sixteen articles had cross-sectional designs. On the other hand, Study 3 was original research based on a cross-sectional design targeting culturally/linguistically diverse patients diagnosed with AD. Specifically, the study sample consisted of Turkish immigrant (n=21, $M \pm SD_{age}$ =71.62 ±7.41) and monolingual, non-immigrant German (n=20, $M \pm SD_{age}$ =74.70±7.50) and Turkish (n=24, $M \pm SD_{age}$=70.33±7.73) patients with AD. All participants were administered the Mini-Mental State Examination (MMSE), Rowland Universal Dementia Assessment Scale (RUDAS), a dementia severity rating scale, and a self-report measure of depression. Additionally, self-report measures of bilingualism and acculturation were conducted with Turkish-immigrant participants with AD.

Results: Study 1 revealed that people with AD and MCI have significant deficits in discourse comprehension, which are not observed in cognitively normal older adults of any age. On five of six discourse comprehension measures, groups with AD were significantly worse than healthy older adults, with one measure yielding mixed findings. Furthermore, compared to the cognitively healthy groups, individuals with MCI showed significant performance deficits in discourse comprehension measures similar to those with AD. Study 2 indicated better performance for bilingual older adults on executive function tests when compared to their monolingual counterparts. On the other hand, bilinguals were found to perform poorer than monolinguals on tests assessing the language domain. However, these findings did not remain robust when the impact of bilingualism on test performance was investigated longitudinally. Lastly, Study 3 provided further evidence on the linguistic and educational bias of the MMSE when employed in

culturally and linguistically diverse individuals with AD. Bilingualism was linked to better performance on the MMSE in the Turkish immigrant group. German patients with AD obtained higher scores on this test than the other two groups. Furthermore, RUDAS was shown to be a better alternative for assessing global cognition in German and Turkish individuals with AD.

Conclusion: The macro-structural discourse comprehension assessment paradigm has shown promising results in identifying the preclinical stages of AD. Further research on this paradigm may help develop a diagnostic tool with a clinical value that can be utilized for differential diagnosis, predicting conversion from MCI to dementia in research and clinical settings. On the other hand, another aspect of linguistic skills, namely, the evaluation of research on the link between bilingualism and neuropsychological test performance, did not provide definitive answers to the question of bilingual advantages and disadvantages addressed in the second study due to methodological challenges in the field. However, it identified a comprehensive and critical list of clinically and empirically relevant bilingualism-associated variables which may guide future research and neuropsychological practice. In light of the Study 2 findings, Study 3 filled an important gap in the literature by exploring cultural, demographic, and immigration-related factors that may influence neuropsychological testing experiences in Germany. The study findings may help the field of cross-cultural neuropsychology serve culturally and linguistically diverse populations more efficiently. Overall, the present thesis contributed to the literature by highlighting the importance and potential of linguistic abilities in the clinical diagnosis and neuropsychological evaluation of individuals with dementia.

Keywords: Discourse, Comprehension, Language, Bilingualism, Neuropsychological Assessment, Alzheimer's disease, Mild cognitive impairment, Aging, Immigration, Racial/ethnic minorities.

1. Chapter 1 – General Introduction

The rise in life expectancy increases the average age of the older adult population, thus, growing the incidence and prevalence rates of Alzheimer's disease and other dementias. Dementia is one of the most common diseases in the older population. It has become a major cause of concern internationally due to its substantial effect on informal, social, and health care costs, particularly in view of changing demographic patterns (Wimo et al., 2017). In 2010, 35.6 million people were predicted to live with dementia, and this number is projected to double in 20 years and to reach 115.4 million in 2050 (Prince et al., 2013).

Among dementia subtypes, Alzheimer's disease (AD) is the most prevalent form of dementia, making up 60% to 80% of cases ("2020 Alzheimer's disease facts and figures," 2020), with fewer than half being pure AD and the majority estimated to be mixed dementias (DeTure & Dickson, 2019). AD has been predicted to affect 4.4% of individuals aged 65 years and over in the European population, based on the pooled data of population-based studies (Lobo et al., 2000). The prevalence and economic burden of AD lead to a shift in focus of research towards neuropsychological detection and characterization of cognitive deficits in the earliest stages of AD since neuropsychological assessment helps provide reliable cognitive markers of AD that are essential for early diagnosis (Weintraub, Wicklund, & Salmon, 2012). Early diagnosis is critical to benefit from treatments that can help manage symptoms in people with dementia and provide preventive strategies that may reduce the risk or postpone the clinical onset of AD.

Memory decline and impairments in other cognitive domains, including executive functions and language, are core features of AD and its preceding stage of mild cognitive impairment (MCI; Budson & Solomon, 2015). Memory and executive functions in MCI and AD have been widely studied using neuropsychological and brain-imaging measures in relation to the differentiation

between MCI and AD (Bondi, 2002; Kirova, Bays, & Lagalwar, 2015; Traykov et al., 2007). However, language functions can be a more sensitive indicator of early cognitive deficits in AD and provide valuable information to assist in clinical diagnosis and monitoring of disease progression, as language deficits are frequently observed in the very early stages of AD (Taler & Phillips, 2008). On the other hand, another perspective on language in AD, that is, constant use of two or more languages is considered as a cognitively stimulating activity that may help to counteract the adverse effects of cognitive decline or dementia and delay the onset of AD (Bialystok, Craik, & Freedman, 2007). This potential impact of bilingualism on cognitive functioning has important clinical implications for neuropsychological detection of AD, as it points out that performance on neuropsychological tests that provide diagnostic information for, and differential diagnosis of dementia is susceptible to environmental and experiential factors. As one of those environmental factors, the effect of culture on cognition poses an additional significant challenge to the studies of bilingualism in relation to neuropsychological testing when the primary language of the participants is different than English, and they come from a culturally diverse background than the Western culture where the tests were developed and validated (Chan, Shum, & Cheung, 2003). To this end, culturally and linguistically appropriate assessment of cognitive functioning becomes of utmost importance to deliver competent neuropsychological services and to provide reliable research results in ethnically or linguistically diverse populations.

In light of these considerations, this thesis brings together two areas of dementia research relevant to language; language-related aspects in the assessment of cognitive deterioration in patients with MCI and AD and culturally/linguistically appropriate neuropsychological assessment tools for the accurate characterization of cognitive profile in AD. Language is a primary source of cross-cultural heterogeneity, which can affect cognition in both minor and substantial ways (Calvo, Ibáñez,

Muñoz, & García, 2018). However, the importance of language and its associated variables in the diagnostic process of dementia-related cognitive impairments have been largely ignored in two growing lines of research, namely, bilingualism and language deficits as AD markers. Therefore, the overarching aim of this thesis is to foster progress in these fields by assessing the available findings and providing evidence as to the potential impact and implications of these topics on research and clinical practice. More specifically, this thesis seeks to elucidate the role of language in the prodromal and dementia stages of AD and underscore the importance of its features for a more reliable cognitive characterization of AD in clinical and research settings.

Study 1 summarizes available research on a specific language deficit in AD, precisely, discourse comprehension, to evaluate its potential as a novel approach in neuropsychological testing. Study 2 gives an overview of studies investigating performance on neuropsychological tests used to assess dementia in bilingual older adults in order to provide directions for neuropsychological assessment of these individuals in clinical practice. This study also sheds light on factors that come along with bilingual experiences, such as cultural background, immigration status, and language of test administration, and discusses their impact on the outcome of studies. The study population of this review was in particular, people with cognitive deficits affected by dementia to comprehensively assess and present findings with regard to the effects of bilingualism on test performance as participant groups in this research topic involved both people with AD and other types of dementia. Study 3 shifts the focus onto an alternative neuropsychological test for the assessment of dementia that is less affected by the factors outlined in the second study and can be applied to people from diverse language and cultural backgrounds. These three individual papers examine whether language profiles in older adults help determine a more sensitive assessment tool for the clinical detection of dementia The primary focus of this thesis is on Alzheimer's type

dementia as the study populations of Studies 1, and 3 are primarily patients with AD, and the participant groups of Study 2 commonly involve individuals diagnosed with AD.

The introduction part of the thesis begins with an overview of the definition of AD and MCI due to AD and their biomarkers, diagnostic criteria, and risk factors. Thereafter, it provides a summary of neuropsychological aspects of language impairments in AD and its preceding stage of MCI using standardized and non-standardized language measurement tools. It highlights the utility of discourse comprehension as a predictive linguistic marker of MCI and AD. Then, it addresses another aspect of language abilities in AD: the benefits and challenges of bilingualism in cognitive aging and its implications for neuropsychological testing. The aims, research questions, methodological approaches, and the study design of this thesis are summarized in Chapters 2, 3, and 4. Finally, the thesis concludes with a general discussion of the study findings, their implications for theory and practice, and future directions.

1.1. Pathological Cognitive Aging: Alzheimer's Disease as a Continuum from a Preclinical State of Mild Cognitive Impairment to Dementia

AD is an age-related multi-factorial neurodegenerative disease, characterized by several neuropathological markers and gradually deteriorating cognitive abilities that influence individuals' activities of daily living (McKhann et al., 2011; Rentz et al., 2013; Sperling et al., 2011; Sperling, Karlawish, & Johnson, 2013). The neuropathological features underlying AD are marked as the accumulation of beta-amyloid protein in plaques and tau deposition in neurofibrillary tangle (Jack et al., 2011; Jack et al., 2013). Additional neuropathological features in AD include loss of neurons and synapses, reactive microgliosis, white matter decline, and granulovacuolar degeneration (Hyman et al., 2012). The earliest changes in amyloid plaques and neurofibrillary tangles that are associated with the neuropathology of AD generally appear in

medial temporal structures, including, the hippocampus and entorhinal cortex (Braak & Braak, 1991; Drachman, 2006). These structures affect the neural network for episodic memory which is essential for the recall and acquisition of new information (Braak & Braak, 1991; Drachman, 2006). Thus, memory loss is considered the most common and prominent clinical hallmark of AD, along with impairments in at least one other cognitive domain, namely, language deficits, visuospatial difficulties, and executive dysfunction (McKhann et al., 2011).

Although beta-amyloid plaques and neurofibrillary tangles are the two main features of AD pathology, most AD patients have numerous pathologies and various forms of brain proteinopathies (Molinuevo et al., 2018). For instance, since cerebrovascular pathology becomes more prevalent with age, approximately 30% of AD patients show concomitant cerebrovascular disease (Toledo et al., 2013). On the other hand, nearly 40% of patients who have dementia with Lewy bodies have AD pathology, as evident from cerebrospinal fluid (CSF) biomarkers (Lemstra et al., 2017). Given the heterogeneity observed in this disease, a definitive diagnosis of AD is only achieved through a post-mortem verification of neuropathological hallmarks (Humpel, 2011).

Biomarkers have been a key component of understanding the biology of AD. A descriptive biomarker classification system has been proposed to be used as an effective and unbiased tool in cognitive aging and AD research (Jack et al., 2016). For this classification system, seven key AD biomarkers have been grouped into three binary categories depending on the pathophysiology that each assesses using the “A/T/N” system. In this “A/T/N” system (Jack et al., 2016), “A” indicates the value of an amyloid-beta biomarker (amyloid PET or CSF amyloid-beta 42), “T” stands for the value of a tau biomarker (CSF phospho tau, or tau PET) and “N” for the biomarkers of neurodegeneration (fluorodeoxyglucose–PET, structural MRI, or CSF total tau). Positive or negative ratings are assigned to each biomarker category in this system. With this framework,

individuals are categorized as positive (+) or negative (–) for A, T, and N, resulting in eight potential A/T/N profiles. This system has been developed to employ validated biomarkers to distinguish AD from non-AD causes of cognitive impairment at the individual level. Thus, it classifies the AD spectrum based on its biological manifestations rather than clinical assessments of cognitive status.

Recent advances in biomarkers have resulted in a growing recognition that AD is a multidimensional process on a continuous continuum rather than a distinct and defined clinical stage (Aisen et al., 2017). Since pathopsychological changes in AD occur several years before the clinical manifestations of the disease, AD is considered a biological and clinical continuum encompassing both the preclinical (asymptomatic individuals with signs of AD pathology) and clinical (symptomatic) phases of the disease. Thus, efforts have been undertaken to capture the disease continuum by taking into account multiple components, namely, pathophysiological processes, biomarker findings, and clinical symptoms. The clinical classification of the AD continuum includes cognition and function trajectories, with cognitive impairment occurring before and predicting functional impairment. More specifically, this process has been suggested as the accumulation of pathopsychological changes at first, which then proceeds with progressive loss of cognitive and functional skills, resulting in no clear boundaries between different clinical stages.

These developments and the resulting progress in understanding the AD continuum have caused several changes to the diagnostic criteria. The main criteria for the diagnosis of AD have been the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS–ADRDA) criteria, since 1984 (McKhann et al., 1984). Significant scientific progress during the last few decades gave rise to a revision of these

diagnostic criteria for AD-type dementia (Jack et al., 2011). The new criteria; published by the National Institute on Aging-Alzheimer's Association (NIA-AA) in 2011, defined AD as a clinical-biomarker construct where biomarkers are utilized to help with diagnosis, but the clinical diagnosis was still required (Ahmed, Ahmed, & Imtiaz, 2021). Furthermore, it brought changes in conceptualization with regard to the clinical spectrum of the disease and characterized not only the dementia stage of AD with diagnoses of "possible", "probable," and "definite" Alzheimer's disease (McKhann et al. 2011), but acknowledge a broader spectrum of cognitive aging (Albert et al., 2011). A transitional state between normal cognitive aging and dementia, named "mild cognitive impairment" (MCI), has been introduced to this spectrum (Albert et al., 2011). MCI has been shown as a valuable label in clinical settings for identifying people who are at risk of progression to AD (Reitz & Mayeux, 2014). However, three different stages (Preclinical, MCI, and Dementia) specified by NIA-AA in the 2011 criteria were amended in 2018 and the new framework identified AD as a biologic construct detected by biomarkers in the individual (Ahmed et al., 2021; Jack et al., 2018). This has resulted in blurred boundaries between the distinct phases of AD (Jack et al., 2018). Thus, these phases have been transformed into a continuum that may be detected in preclinical stages utilizing biomarkers A, T, and N categories (Figure 1). However, since this new framework suggested a biological rather than clinical definition of AD for use in clinical research, it has been criticized for not taking into account both clinical (e.g., impaired cognition) and biomarker evidence (amyloid and tau pathology) to support a diagnosis (Dubois et al., 2021).

The operational definition of MCI has received many revisions and is still an evolving diagnosis (Vega & Newhouse, 2014). As MCI research advanced, it became evident that there are numerous clinical subtypes of the condition, each with various underlying etiologies (Vega & Newhouse, 2014). According to the revised Mayo Clinic Criteria in 2003 (Petersen et al., 2014; Winblad et

al., 2004), performance impairments on neuropsychological tests of episodic memory were classified as amnestic MCI (a-MCI), whereas, lower performance on neuropsychological tests of non-memory domains of cognition was categorized as non-amnestic MCI (na-MCI). These impairments were suggested to be restricted to a single cognitive domain (MCI single domain) or to multiple domains (MCI multiple domains). In 2011, the NIA-AA suggested criteria, particularly for MCI due to AD, to identify people who are symptomatic but not demented and have AD as the underlying cause (Vega & Newhouse, 2014). Therefore, a-MCI closely mimics the neurobiological profile of clinically diagnosed AD. Furthermore, individuals with a-MCI and those diagnosed with AD have numerous characteristics in common, including cognitive, behavioral, and neuropsychological test performance, as well as genetic, neuroimaging, and CSF biomarker characteristics (Tarawneh & Holtzman, 2012). In line with these findings, in a meta-analysis of 33 MCI studies, multiple domains and single domain a-MCI were more likely to evolve to AD, whereas, multiple and single-domain na-MCI had a low likelihood of progressing to AD (Oltra-Cucarella et al., 2018). In sum, categorization of MCI into subtypes has been considered helpful in establishing more homogeneous groups to address predictive accuracy for dementia. However, it has also been argued that findings from progression to AD show variety when data is categorized by community vs. specialist clinic patients (Tahami Monfared, Byrnes, White, & Zhang, 2022). For this reason, the clinical utility of MCI subtypes in predicting who would develop dementia remains unclear (Glynn et al., 2021), and the incorporation of biomarkers estimates into the diagnostic process is recommended to overcome the limitations (Tahami Monfared et al., 2022).

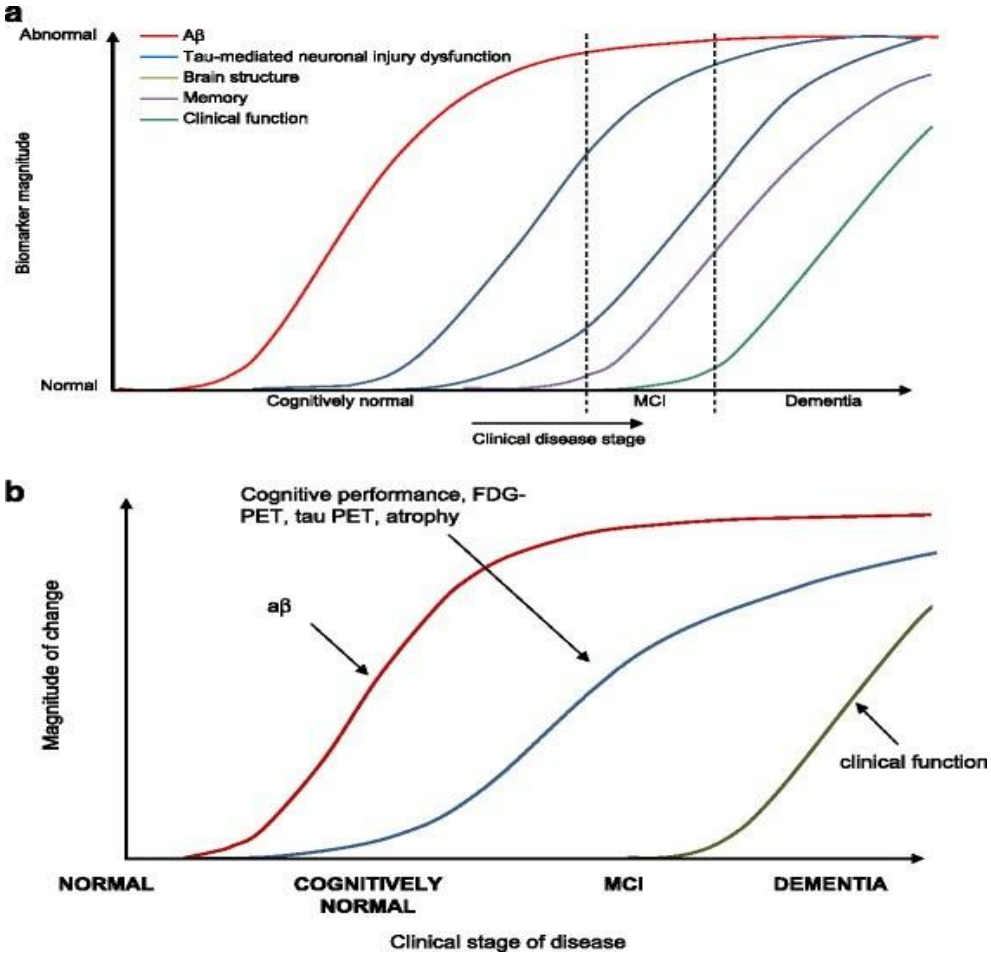


Figure 1- Reprinted from Aisen et al. (2017): Change in biomarkers over time. a. A change in measures of AD over time. b. Modified graph indicating that, although changes in cognitive performance, FDG-PET, tau PET, and MRI atrophy follow a consistent, progressively steepening curve, amyloid accumulation begins early, and functional deterioration emerges later in the continuum of AD (as before). ***Aβ*** *amyloid beta,* ***FDG*** *fluorodeoxyglucose,* ***PET*** *positron emission tomography,* ***MCI*** *mild cognitive impairment due to Alzheimer's disease.*

1.1.2 Diagnostic Criteria and Risk Factors for MCI due to Alzheimer's Disease and Alzheimer's Disease

Several different diagnostic criteria exist to be used for AD. The most commonly used two diagnostic criteria for AD are the Diagnostic and Statistical Manual of Mental Disorders (DSM; Edition 5) and the National Institute on Aging- Alzheimer's Association Criteria, NIA-AA (Albert et al., 2011; McKhann et al., 2011). The following are the elements of AD that each has in common

(Budson & Solomon, 2016): the presence of dementia, impairments in two or more cognitive domains, insidious onset, a steadily deteriorating state, and no indication of another etiology that could significantly contribute to the dementia. On the other hand, the shared elements for MCI due to AD in two criteria are as follows (Budson & Solomon, 2016): the presence of MCI, impairments in one or more cognitive domains, including memory for DSM-5, and generally including memory for NIA-AA and gradual onset and progression, no indication as to another etiology that could significantly contribute to the MCI.

The revised criteria, namely, NIA-AA (McKhann et al., 2011), also suggest including biomarker evidence in diagnosing AD, as opposed to the DSM-5. Two categories of biomarkers for AD diagnosis that are currently used are markers of amyloid-beta protein deposition in the brain and downstream neurodegeneration (e.g., elevated CSF tau, reduced metabolism in the temporal and parietal cortex on PET, atrophy on MRI in the temporal and medial parietal cortex). However, these biomarkers are not included in the core clinical diagnostic criteria for AD and are considered as additional factors that help increase the certainty of an AD diagnosis (Mantzavinos & Alexiou, 2017; Weissberger et al., 2017). The recommended use of biomarker evidence to diagnose AD is commonly applicable to research settings (Weintraub et al., 2012). This renders neuropsychological testing a more impactful role in the early diagnostic process of AD and the clinical diagnostic work-up as a more accessible, less invasive cognitive biomarker of AD.

The underlying cause of these irreversible pathological changes observed in AD has remained to be understood (Reitz & Mayeux, 2014). However, it is considered that the association between increasing age and the presence of AD may bring along the accumulative impact of multiple protective and risk factors, such as complex interactions between genetic predisposition, biological factors, psychosocial variables, and environmental factors (Qiu, Kivipelto, & von Strauss, 2009).

The major risk factors for AD are the inheritance of a genetic component: APOE ε4 allele and family history of dementia in first-degree relatives, low educational attainment (Budson & Solomon, 2016; Qiu et al., 2009). The other putative risk factors are considered as midlife high blood pressure and high body/mass index, diabetes mellitus, cerebrovascular disease, traumatic brain injury, female gender, smoking, and Down's syndrome (Budson & Solomon, 2016; Qiu et al., 2009). Recently, researchers have focused on the role of psychosocial factors in the prevention or amelioration of clinical symptoms of AD and how these factors may mediate changes in cognitive processes in older adults due to their modifiable potential through lifestyle and behavior change. These are discussed in the following sections in relation to bilingualism and cognitive reserve.

1.2 Language Impairment in Mild Cognitive Impairment and Alzheimer's Disease: Neuropsychological Aspects

Several studies have demonstrated the importance of changes in language abilities for the early detection of AD (Szatloczki, Hoffmann, Vincze, Kalman, & Pakaski, 2015; Taler & Phillips, 2008) as language deficits have also been indicated to appear early in the disease course in addition to impairments in episodic memory (Cuetos, Arango-Lasprilla, Uribe, Valencia, & Lopera, 2007; Taler & Phillips, 2008). Furthermore, the language profile of people with AD has been shown to provide specific deterioration patterns depending on the stages of disease severity (Szatloczki et al., 2015).

Language deficits in the early stages of AD are reported to be in lexical-semantic abilities, with difficulties in word-finding, naming objects (e.g., anomie), and semantic paraphasias (Catricalà et al., 2015; Forbes-McKay & Venneri, 2005; Klimova & Kuca, 2016; Taler & Phillips, 2008).

Impairments at the pragmatic level of language processing, namely, changes in discourse (e.g., fluent but empty speech), have also been indicated (Drummond et al., 2015; Fleming & Harris, 2008). Specifically, the changes in discourse in the early stages of AD compared to healthy older adults have been identified by incomplete narratives and by the lack of relevant details for the meaning of the message to be conveyed (Drummond et al., 2015). On the other hand, the syntactic and phonological abilities of people with AD have been often shown to be intact in this stage of the disease (Kertesz, 1994; Light & Burke, 1993). In the moderate and severe stages of AD, impairments in verbal fluency, reduced comprehension, and difficulty in the semantic processing of sentences have been indicated to become pervasive (Ferris & Farlow, 2013). With disease progression into the late stages of AD, a severe language impairment has been reported to be more prominent, with expression being limited to the repetitions of meaningless words, phrases, and echolalia (Ferris & Farlow, 2013). Patients with aMCI who are generally regarded to be a prodromal stage of AD have been reported to exhibit similar patterns of language impairments to those identified in the early disease course of AD, including lexical, semantic, and pragmatic domains of language (Boschi et al., 2017; Taler & Phillips, 2008; Tsantali, Economidis, & Tsolaki, 2013). The similar patterns of linguistic deficits in aMCI and AD underscore the significance of examining changes in language abilities of aMCI as a potential avenue to assist in the early diagnosis of AD (Jokel, Seixas Lima, Fernandez, & Murphy, 2019). Measurement tools that examine these abilities may be particularly well suited to characterizing early language impairments in AD (Taler & Phillips, 2008). Thus, it may help provide critical information for the early detection of AD.

1.2.1. Standardized Neuropsychological Evaluation of Language in Alzheimer's Disease

Standardized language evaluation of dementia most commonly includes the assessments of verbal fluency, naming, repetition of words/sentences, and discourse ability (the ability to comprehend and communicate in a given language).

The verbal fluency test, which assesses the ability to produce words verbally, based on a set of criteria in a given time, is a widely used measure of verbal functioning in the diagnosis of AD (Taler & Phillips, 2008). This test consists of two tasks, assessing semantic and phonemic abilities (Bertola et al., 2014), and they are referred to as category (e.g., naming as many animals as possible) and letter fluency (generating words beginning with certain letters of the alphabet), respectively. Verbal fluency tasks have been shown to be reliable markers of linguistic deficits in the detection of AD (Laws, Duncan, & Gale, 2010; Lonie et al., 2009; Taler & Phillips, 2008). In a meta-analysis of 153 studies comparing the extent of deficits on letter and category fluency tasks in AD patients compared to healthy controls, AD patients have shown to exhibit greater impairments on category fluency tasks than on letter fluency tasks (Henry, Crawford, & Phillips, 2004). These results suggest that semantic memory impairments are more pronounced in people with AD, since the category fluency task is more reliant on semantic memory (Henry et al., 2004).

Similarly, in another study examining verbal fluency task performance in AD, aMCI and healthy controls, aMCI and AD patients have shown to be more impaired on the category fluency task than healthy older adults (Murphy, Rich, & Troyer, 2006). On the other hand, letter fluency performance was relatively intact in the a-MCI group. Numerous studies conducted with MCI patients have shown poorer performance on the category fluency task in individuals who converted to dementia than those who did not develop dementia (Taler & Phillips, 2008). For this reason,

verbal fluency performance has been considered as a potential prognostic predictor for AD and a diagnostic indicator for MCI (Maseda et al., 2014; Taler & Phillips, 2008). However, various methods of administration and analysis have been introduced for verbal fluency tasks, including the count of repetitions, errors, and calculation of scores in every 10 seconds of the tasks. There remains a lack of consistency regarding which methods of administration and analysis of this test are the most valid for obtaining the most accurate diagnostic information (Bertola et al., 2014; Hall, Harvey, Vo, & O'Bryant, 2011).

Word-finding difficulty is one of the most prominent symptoms of AD (Fisher, Rourke, & Bieliauskas, 1999). Impairments in this domain are commonly investigated through confrontational naming or word retrieval tests. The Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983) is a widely used test for evaluating confrontational naming in AD. BNT comprises 60, 30, or 15 (depending on the version of the test) line drawings of objects with varying levels of difficulty, ranging from everyday objects (e.g. tree) to less familiar objects (e.g. abacus). Patients are asked to recall the shown drawings verbally and spontaneously within 20 seconds, and in case of no response, they are given two types of prompting cues, namely phonemic and semantic. In studies examining BNT performance in AD, aMCI and healthy controls, AD patients have been found to perform poorer than aMCI patients (Balthazar, Cendes, & Damasceno, 2008; Choi et al., 2016; Willers, Feldman, & Allegri, 2008), whereas aMCI and healthy controls had similar total BNT scores (Balthazar et al., 2008; Willers et al., 2008). However, the diagnostic utility of the BNT in the early detection of AD has been shown to be controversial after taking into consideration of performance on the delayed recall impairment measures (Testa et al., 2004). Furthermore, since correlations between the BNT and category fluency have been found to be stronger than letter fluency (Henry et al., 2004), the verbal fluency task has been considered to be

a more sensitive measure of semantic deterioration in the disease course of AD due to its additional involvement in executive function (Sala, Lorenzi, Spinnler, & Zuffi, 1993).

Discourse ability is routinely assessed by picture description tasks in patients with AD in order to gain insight into patients' ability of naming, sentence production, or comprehension in a connected speech context (Cummings, 2019). A frequently used standardized test for this domain in AD is the "Cookie Theft Picture" from the Boston Diagnostic Aphasia Examination (Goodglass, Kaplan & Weintraub, 2001). In this task, participants are shown a picture of a scene taking place in the kitchen with two children and a mother. Then, participants are asked to report a series of events depicted in this picture. A number of studies have demonstrated that individuals with mild AD often produce shorter texts/narratives than healthy controls with repetitions, words containing empty content, and fewer relevant details (Sajjadi, Patterson, Tomek, & Nestor, 2012; Taler & Phillips, 2008). However, it has been suggested that the Cookie Theft Picture has little benefit over other language assessment tools, such as verbal fluency or the BNT in discriminating between different stages of AD, mainly in distinguishing MCI patients from healthy controls (Bschor, Kühl, & Reischies, 2005).

1.2.2. Non-Standardized Measures of Language in Alzheimer's Disease

Research on language performance of individuals with MCI and AD using non-standardized measures is relatively scarce as extant literature has been mainly focused on various aspects of language abilities in these groups by employing standardized tests of language functions (Taler & Phillips, 2008). The focus of studies investigating the language profile of MCI and AD patients with non-standardized measures of language functioning ranges from individual words and sentences to discourse processing. Among these, discourse processing has received substantial attention in language and aging research. It has been indicated as a valuable source for detecting

language deterioration in dementia since it is a natural form of communication (Drummond et al., 2019). Furthermore, it requires a high-order linguistic ability and a broad range of cognitive functions, including the administration of lexical-semantic operations, memory, organization, planning of information, and interpretation of meanings (Chapman et al., 2004; Copland, Chenery, & Murdoch, 2002; Duong, Giroux, Tardif, & Ska, 2005). These abilities have been shown to be impaired in individuals with AD to some extent (Duong et al., 2005); thus, a few studies have suggested the inclusion of discourse processing in the diagnostic process (Ehrlich, 1994; Orange & Kertesz, 2000).

Age-related difficulties in discourse processing have been linked to declines in cognitive processes. Over the last two decades, several theoretical approaches have been proposed to elucidate age-related changes in this domain. More specifically, the information-universal theories suggest that declines in general cognitive abilities such as processing speed, working memory, and inhibitory function contribute to language impairments in old age (Burke, MacKay, & James, 2000). In contrast, the information-specific theories focus on the particular changes in linguistic representations and their connections in the brain, which may cause the weakened interplay between comprehension and production. Among four leading theoretical approaches, three types of theories suggest that deterioration in linguistic abilities of older individuals is due to their diminished cognitive ability (information-universal factor). These three theories have addressed the poorer performance in discourse processing in old age in relation to generalized cognitive slowing (Salthouse, 1996), working memory deficits that impact storage and manipulation of information (Miyake, Just, & Carpenter, 1994), or to weakened inhibitory processes for the irrelevant information (Hasher & Zacks, 1988; Zacks & Hasher, 1997). On the other hand, a fourth approach, the transmission deficit hypothesis, has postulated that the impairments in older

individuals' language ability may be information-specific and hence influenced by linguistic characteristics (Burke, MacKay, Worthley, & Wade, 1991). According to this theory, the decline in language capabilities among older individuals is caused by a weakening connection between the semantic and phonological levels (or the orthographic levels) in the language system. The strength of connections between these systems degrades with time as individuals age (Abrams & Farrell, 2011; Burke & Mackay, 1997; MacKay & Burke, 1990), which affects the speed and quantity of activation that is conveyed across the systems.

These three theories, mainly, generalized deterioration, working memory and inhibition deficit theories have received criticism due to their information-universal nature. For instance, numerous studies have revealed that language production ability in older individuals deteriorates more significantly than comprehension ability, pointing to an asymmetric pattern, instead of a generalized deterioration in this domain (Burke et al., 2000). Several studies have also provided evidence that the inhibitory deficit plays a role in impairments related to language comprehension and production (Burke & College, 1997). Similarly, working memory impairment has been indicated as a factor that compromises the ability to comprehend discourse (De Beni, Borella, & Carretti, 2007). In sum, both information-specific and information-universal factors have been shown to contribute to the deterioration of linguistic ability(Wu, Yu, Wang, & Zhang, 2020). However, studies in this field have shifted to prioritize empirical findings over the development of theories for understanding the patterns of change in this domain in old age (Burke et al., 2000). Overall, the theories on the nature of age-related changes in language have been discussed within broad frameworks instead of the systematic development of a theoretical approach (Thornton & Light, 2006).

Discourse processing is generally analyzed along two dimensions, namely, micro-linguistic and macro-linguistic levels. The micro-linguistic dimension of analysis is a study of language on a small scale, focusing on phonetic, phonological, morphological, and grammatical skills that are required to form words and sentences (Messer, 2017). On the other hand, the macro-linguistic dimension analysis addresses the speaker's ability to comprehend relevant information and/or generate the central theme of a narrative discourse or the gist of a conversation by connecting sentences in a cohesive, coherent way (Messer, 2017). The patterns of micro-and macro-structural processing impairments have been mainly investigated using experimental tasks in one component of discourse processing, namely, discourse production (de Lira, Ortiz, Campanha, Bertolucci, & Minett, 2011; Fleming, 2014; Fleming & Harris, 2008, 2009; Harris, Kiran, Marquardt, & Fleming, 2008). In these tasks, participants were either asked to describe how they would spend their time in New York in case they went there for vacation or shown a series of pictures that depicted a story in order for them to tell a narration described in the pictures. The micro and macro-linguistic levels of analyses in discourse production have been found to be helpful in discriminating clinical populations, namely in studies comparing MCI and healthy control groups or AD and healthy older adult groups (de Lira et al., 2011; Fleming & Harris, 2008; Harris et al., 2008). However, another aspect of discourse processing, namely, discourse comprehension, is also a crucial part of communication as extracting meaning from a speech or a text is necessary for daily life activities and quality of life in older adults. The extant literature has demonstrated that older adults compensate for age-related declines in discourse comprehension ability by allocating cognitive resources differently than younger adults and by employing their linguistic expertise and world knowledge (Messer, 2017; Wingfield & Grossman, 2006). Since this ability may be more resilient

to healthy aging, it may offer substantial evidence for distinguishing normal cognitive aging from AD pathology-related cognitive decline.

1.3. The Importance of New Neuropsychological Techniques for the Detection of Language Impairments in Alzheimer's Disease

Currently available standardized language assessments are considered relatively insensitive for distinguishing normal brain aging from the development of AD. One of the underlying causes of this is that the majority of language assessment tools employed in individuals with AD originate from vascular aphasia batteries, which are not particular to neurodegenerative diseases (Pistono et al., 2019). Another explanation for the difficulty in distinguishing healthy aging from pathological brain aging indicative of AD is the lack of ecological validity of traditional neuropsychological tests. Specifically, standardized language assessments are unable to account for communication and macro-linguistic characteristics of expressive and receptive language (Pistono et al., 2019). For instance, one study has investigated the word-finding abilities of young adults (ages 18-22 years), young-old adults (ages 58-74 years), and old-old adults (Schmitter-Edgecombe, Vesneski, & Jones, 2000) by employing the BNT and a spontaneous discourse production test. The results have revealed that the group of older adults performed worse than younger adults on the discourse production task. In contrast, the performance on the BNT was more accurate in older adults than in younger adults. These findings suggest that discourse processing tasks may yield a better overview of language abilities that are more reflective of the actual skills required for daily life activities or everyday conversation in older adults (Messer, 2017).

Furthermore, language impairments have been extensively studied among AD patients, but they have not been well-understood in MCI patients (Mueller, Hermann, Mecollari, & Turkstra, 2018).

A few existing studies have pointed out impairments in sentence or discourse-level processing in individuals with MCI (Taler & Phillips, 2008). However, previous studies on discourse in MCI patients provide limited understanding concerning how participants with MCI vary from cognitively healthy participants in terms of different discourse processing outcomes (B. S. Kim, Kim, & Kim, 2019). Considering the urgent need for sensitive measures that can detect cognitive changes in the early stage of AD (Snyder et al., 2014), assessment tools focusing on discourse processing in patients groups with MCI and AD when compared to cognitively healthy older adults may be beneficial for informing the clinical diagnosis of AD at very early stages of the disease. A systematic overview of performance on different measures of discourse processing in these groups can reveal important information, which may help inform diagnosis and intervention. Specifically, discourse comprehension at a macro-linguistic level in MCI patients may allow an additional layer of information for neuropsychological testing as a critical component of discourse processing abilities. The majority of studies that focused on aging and discourse processing have examined the micro-linguistic level analysis of discourse comprehension (Williams, McAuliffe, & Theys, 2021). This has resulted in the characterization of particular linguistic skills in patients with AD and MCI without providing a holistic account of the patterns of language changes along the AD continuum. Hence, discourse comprehension at a macro linguistic level may present a novel and ecologically valid paradigm in understanding the linguistic deterioration of the prodromal and dementia stages of AD. A summary of available findings may help understand the importance of this ability in the detection of AD as a potential field for future investigations.

1.4. The Other Side of the Coin: Language as a Cognitive Reserve in Aging and Alzheimer's Disease

Language abilities represent two sides of the same coin to AD, offering convergent evidence to serve as an early cognitive marker of AD and as a protective individual or environmental factor against cognitive decline associated with AD. Accumulating evidence suggests that the neuropathology of AD occurs decades before the appearance of AD symptoms (K. Kim et al., 2020). Several studies have revealed that most individuals who had AD pathology and fulfilled the clinical criteria for AD maintained their cognitive function within the normal range (Mortimer, 1997; Shaw et al., 2009; Valenzuela & Sachdev, 2006). This indicates that some individuals with normal functioning have progressively increasing pathology (Antoniou, Gunasekera, & Wong, 2013). Variation in time between the neuropathological process and clinical manifestation of AD is considered to be affected by individual factors or environmental exposures, such as cognitive reserve (Jack et al., 2013; Stern, 2012).

The cognitive reserve (CR) theory was developed to explain the disparity between the burden of brain pathology and cognitive performance. The concept of CR is characterized as resilience to the effects of disease-related pathology and maintenance of higher function by the involvement of compensation mechanisms in the brain despite brain deterioration (Stern, Barnes, Grady, Jones, & Raz, 2019). CR theory postulates that specific variables boost the brain's ability to cope with changes or damage, alleviating its influence on cognitive performance or day-to-day function (Stern, 2009). The primary variables or socio-behavioral proxies related to and contributing to the improvement of CR have been suggested as education, IQ, occupational complexity, leisure, and physical activity (Cabeza et al., 2018; Stern et al., 2019). Of these factors, a higher educational level has been shown as the most consistent factor associated with a higher cognitive reserve. In a

meta-analysis of 133 studies examining the evidence of education and its influence on dementia, higher educational levels were found to be associated with a reduced prevalence and incidence of AD and other dementia types (Meng & D'Arcy, 2012). Additionally, the analyses of studies focusing on the neural basis of cognitive reserve have indicated that individuals with higher educational levels had greater brain degeneration, but this damage did not appear initially as a poorer cognitive function (Meng & D'Arcy, 2012).

Language experience, specifically second or multiple language knowledge referred to as 'bilingualism,' has also recently been put forward as a contributing factor to CR as the educational level. Several studies that compared monolingual and bilingual individuals have shown that bilingualism was associated with a later onset of MCI (Bialystok, Craik, Binns, Ossher, & Freedman, 2014; Ossher, Bialystok, Craik, Murphy, & Troyer, 2013) and AD symptoms ((Bialystok et al., 2007; Bialystok et al., 2014; Chertkow et al., 2010; Craik, Bialystok, & Freedman, 2010; Woumans et al., 2014), irrespective of gender, years of education, and immigration status (Alladi et al., 2013). Interestingly, bilingual patients with AD were shown to experience greater brain atrophy than monolinguals, despite being matched to monolinguals with AD for disease severity, education, and cognitive performance (Schweizer, Ware, Fischer, Craik, & Bialystok, 2012). The authors suggest that bilingualism may provide a buffer against the increased neuropathology or neurodegeneration before the clinical expression of the disease. One other study carried out with cognitively healthy early and late bilinguals and monolinguals has reported that early bilinguals had reduced CSF total-tau, which are AD biomarkers, and a lower preclinical AD prevalence than the other groups, particularly at older ages (Estanga et al., 2017). These results indicate that bilingualism is one factor that acts as a CR, and it may moderate the relationship between age and CSF AD-biomarkers (Estanga et al., 2017).

A cognitive model has been suggested to explain the protective effects of bilingualism against cognitive aging. It has been proposed that the ability to monitor or control two (or more) languages may lay the foundation for various language-independent cognitive benefits. Precisely, it has been argued that increased cognitive demands associated with controlling two languages and selecting the appropriate language in a particular context may extend to domain-general cognition, specifically to some domains of executive functions (Abutalebi & Green, 2007). This has been termed as 'bilingual language control' (Abutalebi & Green, 2007). This theory assumes that inhibition of non-target language and monitoring for any intrusions of the other language (s) are necessary when the target language is in use since the presence of two languages builds a mental competition. Thereby, the constant practice in language selection and control contributes to the executive functions of conflict monitoring, updating, interference suppression, and working memory. This theory was supported by the structural neuroimaging studies of bilingualism in a group of aging bilinguals. Two studies have reported that bilinguals had increased gray matter density in regions associated with cognitive control and monitoring conflicting information, such as in the anterior cingulate cortex (Abutalebi, Canini, Della Rosa, Green, & Weekes, 2015) and the inferior parietal lobules bilaterally (Abutalebi, Guidi, et al., 2015), compared to monolinguals.

Further evidence came from the studies that reported the recruitment of the same neural networks for bilingual language control and executive functions in linguistic and non-linguistic tasks (Calabria, Costa, Green, & Abutalebi, 2018). With these findings, bilingual language control and executive functions have been shown to be connected not only by behavior but also by neural substrates (Voits, Pliatsikas, Robson, & Rothman, 2020).

Overall, considerable experience in controlling two languages has often been reported to enhance the executive control system for bilinguals, particularly in older adults (Valian, 2014). It has been

claimed that this improvement in executive functions helps to enhance CR, which may eventually delay the onset of dementia. Although the exact mechanisms are not well-understood, and there is evidence to the contrary that limits the generalizability of these claims (Antoniou & Wright, 2017), the findings regarding the impact of bilingualism on cognitive functions indicate that bilingualism may be a potential confounding factor in neuropsychological evaluation and may obscure the objective characterization of cognitive profiles. Therefore, there remains a question as to whether and how these cognitive differences between monolinguals and bilinguals impact neuropsychological test performance (Paplikar et al., 2021).

1.5. The Implications of Bilingualism for Neuropsychological Evaluation of Older Adults

In recent years, there has been significant growth in experimental investigations aiming at determining the effects of bilingualism on language functioning and other cognitive skills. The effects of bilingualism on cognitive processing reveal two mechanisms that are crucial to understanding the performance differences between bilinguals and monolinguals (Rivera Mindt et al., 2008). Besides the model associated with competition or interference between languages, the reduced frequency of use has been proposed as the other model specific to bilingualism.

The reduced frequency of use hypothesis, referred to as ‘weaker links’ (Gollan, Montoya, Cera, & Sandoval, 2008) postulates that bilinguals utilize each language less frequently than monolinguals since bilinguals divide the use between two languages. Thereby, the usage of words specific to each language in bilinguals is less frequent than in monolinguals. Due to the strong association between frequency of use and lexical accessibility, high-frequency words are accessible more rapidly and correctly than low-frequency ones. The support for this account came from a study

with younger and older Spanish-English bilinguals and English monolinguals that examined picture-naming performance in English and Spanish (if bilingual) (Gollan et al., 2008). The findings have revealed that bilingualism and aging were associated with a slower naming performance, with more pronounced language group differences for low-frequency items than high ones. Furthermore, the difference between naming times in the dominant and non-dominant languages was shown to be more evident in the production of low-frequency words in the bilingual group.

The weaker links and competition accounts predict poorer performance in lexical retrieval abilities in bilingual older individuals, with different results depending upon the language dominance in individuals (Sullivan, Poarch, & Bialystok, 2018). The weaker links account assumes that increased experience of older adults may result in strengthened connections between lexical items and concepts; thereby, it may help diminish language group differences. On the other hand, the competition account suggests that language group differences persist across ages because the need for selection and conflict resolution does not diminish with years (Sullivan et al., 2018).

Consistent with these accounts, some studies on bilingualism have shown distinct patterns of neuropsychological test performance in comparisons of language groups or participants with varying degrees of second language proficiency. Specifically, in a study conducted with neurologically intact Marathi-Hindi bilingual speakers, higher levels of bilingualism were linked to better performance on inhibitory control and executive function skills as assessed by the Color-Trails and Stroop-Color Word Test (Kamat et al., 2012). On the other hand, in a study that examined naming performance in both languages of Spanish-English older bilinguals, balanced bilinguals performed worse on the BNT than unbalanced bilinguals. However, they named more

pictures accurately in both languages when the pictures had cognate names (Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007).

In conclusion, the effects of bilingualism on executive control, verbal abilities, and brain structure have important implications for the neuropsychological evaluation of older bilingual individuals. Neuropsychological tests are integral to the detection and characterization of cognitive deficits and differentiation between various brain diseases and mental health disorders. However, they are typically developed and normed to be used in monolingual individuals (Gasquoine & Gonzalez, 2012). This leads to the use of normative data, which are applicable to monolinguals, in bilingual groups to determine whether they perform within the monolingual range. Therefore, the use of neuropsychological tests on bilinguals that were normed in monolinguals raises critical issues in the competent evaluation of bilinguals since it may not accurately reflect the cognitive functioning of bilinguals (Gasquoine & Gonzalez, 2012). And thereby, it limits neuropsychological tests' utility in clinical practice. Considering that bilingual people make up approximately half of the world's population (Bhatia & Ritchie, 2006), the synthesis of findings from the effects of bilingualism on commonly used standardized neuropsychological tests in clinical settings has become imperative to the practice of clinical neuropsychology. Although a number of studies have been conducted with bilingual older adults, efforts to address these performance differences in language groups more in-depth and to incorporate bilingualism into neuropsychological evaluation have been limited.

1.6. Bilingualism-Related Neuropsychological Testing Considerations in Older Adults: Immigration, Cultural and Linguistic Diversity, and Culturally Appropriate Assessment Tools

The robustness of the cognitive benefits of bilingualism has been debated since a growing body of studies failed to find an association between bilingualism and the advantages observed in cognitive functions or the delayed onset of dementia. The role of bilingualism in cognitive aging and enhanced CR are not widely acknowledged due to many confounding factors involved in this research area (Bak & Alladi, 2016; van den Noort, Struys, & Bosch, 2019; Woumans et al., 2014). Researchers exploring the role of bilingualism on CR, and cognitive functioning point towards the issue of the variations in bilingual experience (Chertkow et al., 2010), the instruments employed to measure cognitive functions (Calvo, García, Manoiloff, & Ibáñez, 2015), immigration (Quinteros Baumgart & Billick, 2018) and cultural backgrounds of participants involved in the samples (Samuel, Roehr-Brackin, Pak, & Kim, 2018).

In studies of bilingualism and cognitive aging, the bilingual group varies from the monolingual group not just in terms of language but also in terms of immigrant status and/or ethnic, religious, and cultural backgrounds, which are frequently linked to significant lifestyle disparities (Bak, 2016). For instance, in several studies carried out in Canada and the USA, bilinguals were more likely to be immigrants. In contrast, monolinguals consisted of a population that lived in the same region for many years (Bak, 2016). This brings a potential selection bias that may result from the "healthy immigrant effect" (Watson, Manly, & Zahodne, 2016), namely, "a self-selection in which healthy people are more inclined to migrate" (Fuller-Thomson & Kuh, 2014). Previous research suggests that immigrants tend to have better general health and lower mortality rates despite lower educational levels, occupational status, and income than their native-born counterparts (Thomson,

Nuru-Jeter, Richardson, Raza, & Minkler, 2013). Hence, the "healthy immigrant effect" may have contributed to the cognitive advantages associated with bilingualism in some studies (Hill, Angel, Balistreri, & Herrera, 2012). Indeed, a few studies have shown an association between immigrant status and better cognitive functioning (Hill et al., 2012; Kopec, Williams, To, & Austin, 2001).

On the other hand, it has been argued that the health advantages associated with immigrant status may diminish with time spent in the host country (Antecol & Bedard, 2006; Lopez-Gonzalez, Aravena, & Hummer, 2005). More specifically, it has been suggested that the health conditions of immigrants may deteriorate over time due to the stress associated with language acquisition and new cultural contact or working in substandard conditions. Consequently, acculturative stress and working conditions may indulge in risky health behaviors such as smoking and poor dietary habits in the immigrant population (Kaestner, Pearson, Keene, & Geronimus, 2009; Markides & Gerst, 2011). This argument has also been supported, with findings showing a significant association between smoking status and cognitive decline in Mexican Americans (Collins, Sachs-Ericsson, Preacher, Sheffield, & Markides, 2009). Further studies carried out with immigrant and non-immigrant groups point to either no evidence of immigration advantage on cognitive function (Kavé, Eyal, Shorek, & Cohen-Mansfield, 2008; Sheffield & Peek, 2009) or worse cognitive functioning in immigrant groups (Black et al., 1999). Overall, there is still considerable ambiguity with regard to whether and how immigrant status and cognitive functions are linked (Xu, Zhang, & Wu, 2017). However, acculturation has been proposed as a crucial factor and a way to comprehensively measure immigration to gain a better insight into the cognitive changes associated with immigration (Xu et al., 2017).

Acculturation is broadly defined as intercultural adaptation processes with psychological and cultural changes due to prolonged contact with a new host culture (Berry, 1997). Alternatively, it

has been described as a phenomenon having two distinct dimensions, with adapting to the host culture and maintaining one's own culture (Ryder, Alden, & Paulhus, 2000). The acculturation processes are categorized as integration, marginalization, separation, and assimilation based on the individual's cultural identity maintenance and participation with the host country (Berry, 1997). These acculturation outcomes have been helpful in explaining the within-group variability reported in neuropsychological test performance between ethnically diverse individuals (Ferraro, 2016; Moyerman & Forman, 1992). For instance, a recent systematic review that included 21 studies investigating the impact of acculturation on neuropsychological test performance has reported that in some studies, higher adoption toward the host culture (e.g. integration) was associated with better test performance beyond the effect of education and age (Tan & Burgess, 2020). However, the construct of acculturation has several domains (e.g., language, ethnic identity) and dimensionality that render it challenging to understand what constitutes cognitive tests with respect to acculturation. Although some studies demonstrated that the language components of acculturation had a considerably larger influence on test performance (Arentoft et al., 2012; Hasson, Wu, & Fine, 2019), using a multi-domain approach to measuring acculturation is necessary for a complete understanding of which elements of acculturation are connected to test performance.

Immigration also creates an increasing demand for neuropsychological testing in relation to the provision of culturally competent neuropsychological services to ethnically diverse individuals. Due to the increase in international migration, the potential impact of bilingualism on cognitive functions is further complicated by the addition of the cultural background of the participants involved in the samples. Culture has been recognized as one of the most critical potential confounding variables in studies evaluating the bilingual advantage hypothesis, since several

studies have examined groups that vary in cultural and linguistic backgrounds, with bilinguals consisting of one cultural background and monolinguals of another (Samuel et al., 2018). For instance, in a study that compared monolingual and bilingual groups of Korean and British young adults, bilingualism was not associated with enhanced executive control performance. However, when participants were divided into two different cultural groups, Korean adults outperformed their British counterparts on this task (Samuel et al., 2018). The authors have suggested that cultural practices in Korea, rather than bilingualism, promote self-regulation and inhibition, resulting in better performance on tasks requiring inhibitory control. Thus, the inclusion of participants with diverse cultural backgrounds, particularly in bilingual groups, complicates the association with neuropsychological performance, making it difficult to determine the exclusive impact of bilingualism on test performance.

Cultural diversity also brings issues related to linguistic heterogeneity in studied samples. Linguistic variety in groups requires the consideration of the cross-linguistic applicability of neuropsychological tests and differences in language practices and exposure among bilinguals, including language proficiency and age of language acquisition. Especially, the participants' language proficiency is crucial to determining the validity of the bilingual cognitive advantage hypothesis and the testing language. To illustrate, bilinguals who spoke English as a second language were considered proficient in English to perform neuropsychological tests in some studies. However, these individuals may have different proficiency levels in both languages, and consequently, they may underperform if they are not evaluated in their dominant language (Manuel-Dupont, Ardila, Rosselli, & Puente, 1992). The effect of language of test administration on neuropsychological test performance has been well documented (Boone, Victor, Wen, Razani, & Pontón, 2007; Gasquoine, Croyle, Cavazos-Gonzalez, & Sandoval, 2007; Kisser, Wendell,

Spencer, & Waldstein, 2012). For instance, a study compared neuropsychological test performance in cognitively healthy Hispanic American adults who were divided into Spanish-dominant, balanced, and English-dominant bilingual groups (Gasquoine et al., 2007). It has been found that the language of test administration was associated with performance differences in tests involving significant language demands in Spanish and/or English dominant bilinguals. However, there were no significant variations in test scores among balanced bilinguals between the Spanish and English language administrations.

Another issue related to linguistic heterogeneity in samples is the applicability of neuropsychological tests across participants of different language backgrounds. The neuropsychological test procedure must be modified and standardized for use across various languages to provide reliable and valid evidence regarding cognitive functioning (Fernandez & Abe, 2018), especially in the context of bilingualism, by considering the language proficiency and dominance of participants. Some test items may not be familiar to certain cultures, and words used in one language may not have the same meaning in another language (Hofmann, 2017). For instance, the semantic scope of the English word "vegetables" is slightly different than the Spanish term "vegetales." "Vegetales" is a word in Spanish that refers to all plants, but the noun in English refers to herbaceous plants utilized for food (Rosselli et al., 2002). Thus, the neuropsychological tools used in studies of bilingualism need to be culturally and linguistically appropriate to the population being tested.

In conclusion, bilingualism is not a unitary or categorical construct (Luk & Bialystok, 2013). Instead, it requires a comprehensive understanding of the relevant factors involved in its entity, such as context and experience. Furthermore, bilingual individuals differ on several dimensions, including linguistic, cognitive, experiential, educational, and other variables, all of which must be

considered when interpreting test performance (Kroll & Bialystok, 2013). Therefore, bilingualism needs a method that requires further than a categorical assignment to the language groups, such as monolinguals and bilinguals (Kroll & Bialystok, 2013). Effective analysis and clarification of these mentioned components of bilingualism can help better understand the cognitive consequences of bilingualism from a holistic perspective. However, due to the interplay between culture and bilingualism, without a thorough knowledge of the influence of culture on neuropsychological test performance, the accuracy and validity of findings regarding the cognitive advantages of bilingualism remain controversial. Therefore, a cross-cultural perspective, namely, neuropsychological assessment tools suitable for use with different cultural and linguistic groups, is necessary to elucidate the increasing discrepancy in bilingual literature concerning cognitive advantages.

1.7. Neuropsychological Assessment of Culturally Dissimilar Older Adults: Cross-Cultural Neuropsychology

The demographic characteristics of the population in Europe have shifted drastically, resulting in fast growth in the cultural and linguistic diversity across the continent. Especially, the number of persons over 64 years old with an immigration background in the EU increased dramatically from 4.73 million in 2000 to 7.37 million in 2017, and this population is projected to grow (Schmachtenberg, Monsees, & Thyrian, 2021). Among people with an immigration background, dementia is under-diagnosed to a larger extent in the 60+ age range and over-diagnosed in the 60+ age group (Nielsen, Vogel, Phung, Gade, & Waldemar, 2011). Early and accurate identification of dementia in ethnic minority patients in Europe presents a challenge for general practitioners (GPs) and specialists, mainly due to a lack of linguistically and culturally sensitive assessment and diagnostic tools (Georges et al., 2019; Nielsen, Vogel, Riepe, et al., 2011).

The majority of standardized neuropsychological tests are developed for and normed in Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Henrich, Heine, & Norenzayan, 2010). Therefore, their applicability in ethnic minority groups or culturally dissimilar groups has been controversial. Several studies have reported that ethnic minority groups underperform or exhibit below cut-off performance compared to native-born groups (Gasquoine, 2009; Puente, Perez-Garcia, Lopez, Hidalgo-Ruzzante, & Fasfous, 2013). Normative data, which are used to estimate the relative performance of a particular cultural group, are lacking for some populations (Gasquoine, 2009). The absence of applicable normative data and variability in test scores between cultural groups have a detrimental impact on clinical diagnoses and access to appropriate care for culturally diverse populations (Rivera Mindt, Byrd, Saez, & Manly, 2010). The study examining whether the use of North American neuropsychological tests influences diagnostic accuracy in cognitive impairments of participants from Colombia, Morocco, and Spain provides further evidence for the effect of culture on clinical outcomes (Daugherty, Puente, Fasfous, Hidalgo-Ruzzante, & Pérez-Garcia, 2017). The results revealed that tests developed for a specific population to assess individuals from various cultures caused significant false positives. Furthermore, the prevalence of misdiagnosis was shown to occur up to 20% of the time, depending on the ethnicity.

Given that Germany has been considered Europe's most important destination for immigration (Green, 2013), growing cultural diversity in Germany poses a significant challenge to the timely recognition of dementia, dementia diagnosis, and neuropsychological assessment of culturally diverse individuals. Currently, individuals with an immigration background comprise 18.6 million people in Germany, accounting for 23% of the population (Monsees et al., 2020). Among this immigrant population, 1.86 million people were 65 years and older, and approximately 5.2% of

them were reported to have dementia (Monsees, Hoffmann, & Thyrian, 2019). Within this cultural diversity, individuals from Turkey make up one of the largest groups of people with an immigrant background in Germany, as well as the group of people with dementia (Monsees et al., 2020). For this reason, the improvement in the reliability and validity of neuropsychological measures and the diagnostic accuracy of dementia in this population requires urgent attention.

Among the neuropsychological measures, cognitive screening represents a primary step in determining the need for additional clinical assessment in individuals with suspected cognitive impairment (Lischka, Mendelsohn, Overend, & Forbes, 2012). A recent study examining the impact of cognitive screening for dementia on the formal diagnosis rate in 108 GP practices in Germany has underscored the importance of cognitive screening tools in identifying dementia (Eichler et al., 2015). The results showed that 49% of previously undiagnosed patients were formally diagnosed with dementia after a positive screening outcome at their GP practice. However, since some diagnoses were received on the day of screening, the authors have pointed out that some GPs were likely to formally diagnose dementia based on only a positive screening result and no other differential diagnostic assessment. Given the substantial role of cognitive screening tools in the diagnostic process of dementia, there is a pressing need for novel cognitive screening measures that are suitable for use among specific populations or across cultural groups.

1.7.1. A Culturally Sensitive Cognitive Screening Tool for Dementia: RUDAS

The Mini-Mental State Examination, MMSE (Folstein, Folstein, & McHugh, 1975), has been the most widely used cognitive screening tool in Europe. As a result, it has become rather a gold standard for detecting early indications of dementia (Georges et al., 2019). However, this tool has been shown to be biased by age, the limited language proficiency of immigrants, educational level, the language of test administration, and culture (Escobar et al., 1986; Parker & Philp, 2004;

Ramirez, Teresi, Holmes, Gurland, & Lantigua, 2006). Particularly, the impact of cultural background on MMSE performance has been reported in numerous studies. For instance, in studies conducted with African-American and White patients diagnosed with AD (Welsh et al., 1995) and with African-American, Hispanic, and White groups of dementia and non-dementia patients (Bohnstedt, Fox, & Kohatsu, 1994), the cultural background was significantly associated with the MMSE performance even after adjustment for education.

Instead of depending on a total score, various strategies, including age and education corrected normative data for particular ethnic groups, (Blesa et al., 2001; Mungas, Marshall, Weldon, Haan, & Reed, 1996) and/or interpretation of the performance based on the individual items of the MMSE (Matallana et al., 2010) have been suggested and demonstrated to be a useful method to compensate for performance differences across ethnic groups (Milman, Faroqi-Shah, Corcoran, & Damele, 2018). However, there is accumulating evidence showing that experiential factors, such as length of residence in a new country, acculturation, and bilingualism, may also impact neuropsychological test performance (Boone et al., 2007; Gasquoine, 1999; Gasquoine et al., 2007; Manly, Byrd, Touradji, & Stern, 2004). These factors are commonly not taken into account when developing norms for a population (Manly & Espino, 2004; Sayegh, 2015). In societies with diverse cultures and languages, the development of normative data may be inadequate unless specific test items or tests representing a range of cultural experiences are generated concurrently with normative data (Boone et al., 2007).

The Rowland Universal Dementia Assessment Scale, RUDAS (Storey, Rowland, Basic, Conforti, & Dickson, 2004) was designed using a simultaneous development process by culture and health advisory groups for the selection of culturally appropriate test items (Fernandez & Abe, 2018). This measure has been suggested as the most promising culturally sensitive cognitive screening

tool (Fernandez & Abe, 2018). It has been shown to be relatively uninfluenced by language use and cultural background (Naqvi, Haider, Tomlinson, & Alibhai, 2015). Moreover, a study (Goudsmit et al., 2018) that compared the diagnostic accuracy of the RUDAS with the MMSE among different ethnic minority groups in a geriatric outpatient clinic has supported the use of the RUDAS in a highly illiterate, culturally varied population.

Considering the cross-cultural potential of the RUDAS across a variety of cultural groups, an essential avenue for research in this domain is to explore the applicability of the RUDAS by comparing test performance in older Turkish and German adults. In addition, the Turkish older adults living in Germany provide a unique opportunity to examine variances in test scores in relation to the immigration-related variables, namely, acculturation and bilingualism. By examining all stated factors above, comparing the performance on RUDAS with the MMSE, a test developed for the WEIRD societies, may also shed light on possible differences between the tests. Thus, it may help improve the diagnostic accuracy and utility of neuropsychological measures across all ethnic and linguistic groups.

1.8. Research Gaps in the Literature and Aims of the Current Thesis

As summarized above, there is sparse evidence of language deficits in the disease process of AD, specifically on the usefulness of the macro-structural discourse comprehension domain as a sensitive cognitive marker of AD. In addition, as another aspect of language, it is still an open question how widely used neuropsychological assessment tools in older adults may be impacted by bilingualism and its associated factors, as does the extent to which these effects may have diagnostic implications. This thesis strives to cover these critical research gaps in previous research on language and neuropsychological assessment of dementia.

Gaining a better understanding of language impairments and bilingualism in the disease course of AD is not only needed for neuropsychology literature but holds substantial implications for timely and accurate diagnosis of AD, appropriate treatment plans, and effective interventions for culturally diverse patients with dementia. Therefore, this thesis links language in dementia to neuropsychological assessment of culturally diverse populations. Thus, it aims to suggest new and more rigorous methods for the accurate characterization of cognitive functioning in individuals at risk of developing AD or in patients with AD, thereby informing research and clinical practice.

Firstly, as stated above, diagnostic procedures of linguistic functioning play a vital role in identifying cognitive impairments at various levels. Language impairment is a prominent feature even early in the disease process of AD. It has also been included in recently published criteria for the clinical diagnosis of AD at the onset in one of the subtypes (McKhann et al., 2011). However, the question of whether the characterization of the language profiles of MCI/AD patients can aid in the early detection of AD remains unanswered. Studies have shown that discourse comprehension at a macro-structural level remains unaffected in normal cognitive aging, as older adults utilize this type of processing more to compensate for declines in detail-level processing of information (Davis, Alea, & Bluck, 2015; Martin et al., 2018). Furthermore, the deficits reported in micro and macro-structural processing have been found to help distinguish clinical populations such as, individuals with MCI due to AD from those with MCI due to non-AD pathologies (Mazzon et al., 2019). Therefore, Study 1 targets studies including individuals with AD and its preceding stage of MCI compared to cognitively healthy older adults to evaluate whether the discourse comprehension paradigm may be used as a unique approach to neuropsychological testing. It evaluates the present state of research focusing on changes in discourse comprehension in prodromal and manifest AD. Additionally, it assesses the relationship between standardized

neuropsychological tests used in clinical settings and measures of discourse comprehension to address whether neuropsychological assessment tools are sensitive to discourse comprehension. Consequently, this systematic review may shed light on the potential of this domain in detecting AD-related cognitive impairments.

Secondly, language and culture represent significant components of neuropsychological evaluation. Yet, although these factors have sparked attention in neuropsychology literature, there remain gaps in research. It is currently unclear how the bilingual advantages/disadvantages experienced by bilinguals, specifically in older adults, translate into performance on neuropsychological tests used in clinical settings. Based on the previous evidence showing the potential impact of bilingualism on cognitive reserve, executive functions, and verbal abilities in older adults, the primary goal of this research is to provide a summary of previous studies on bilingualism and neuropsychological test performance and to evaluate whether or not there is a bilingual advantage/disadvantage in neuropsychological test performance in older adults. Therefore, this study targets the studies that include healthy bilingual older adults or bilinguals with cognitive decline due to dementia and a monolingual comparison group to provide a more comprehensive review of findings in this field.

Study 2 also focuses on the available evidence on bilingualism and neuropsychological test performance in older adults by highlighting the inconsistencies in findings and methodological variables. This is done by characterizing various factors underlying bilingualism and several possible confounding variables, including, but not limited to, the immigrant status of participants and the language of test administration that may have affected the results. More specifically, this systematic review focuses on the criteria to define bilingualism, variances in sample characteristics, and language experiences of bilinguals to critically evaluate the impact of

bilingualism on test performance. The interactions between these factors so far have seldom been addressed in the literature, and the cognitive consequences of bilingualism may not be fully understood without considering the multi-faceted nature of bilingualism (Yow & Li, 2015).

Lastly, as an important extension to bilingualism and neuropsychological test performance literature, the relevance of bilingualism research is further underlined by rapidly increasing cultural diversity due to immigration patterns globally. The growth in the number of culturally and linguistically diverse people poses substantial obstacles to the quality and accuracy of findings in clinical and research settings. Therefore, there needs to be a responsive approach to the changing and diverse profile of the population by considering the cultural, linguistic, and educational background of the people who are assessed (Manly, 2008). This approach should not be limited to the neuropsychological measures or normative data used, but immigration, acculturation, and bilingualism should also be taken into account in the neuropsychological evaluation process (Llorente, 2007; Manly, 2008). Thus, Study 3 examines the influence of cultural, educational background, and immigration-related factors, namely, acculturation and bilingualism, on the performance of two cognitive screening tools. This study compares the performance on the culturally sensitive measure- RUDAS to the MMSE in three groups with clinically diagnosed AD to shed light on the applicability of the RUDAS in culturally diverse populations. For this aim, Turkish immigrants and non-immigrant, monolingual Turkish and German patients with AD were included in the study. Previous studies have almost exclusively focused on comparing test performance in immigrant and host country groups of cognitively healthy participants (Krist et al., 2019; Nielsen, Vogel, Gade, & Waldemar, 2012). This paper addresses this need by including a clinical sample and native-born comparison group, namely, Turkish participants without an immigration background, so far lacking in the scientific literature. Thus, the effects of

immigration-related variables on the performance of cognitive screening tools are further analyzed without complicating the findings with differences in the cultural and linguistic background of the participants involved in the study samples.

In conclusion, this thesis broadens the boundaries of single disciplines and brings in linguistic, cultural, neuropsychological, and even sociological perspectives due to immigration trends in Europe. Table 1 gives an overview of the specific research questions and hypotheses which were addressed in the three studies of the current thesis. The three individual studies are presented in Chapters 2-4, and the findings from the three studies are combined to provide an integrated discussion in Chapter 5.

Table 1. Summary of the Research Questions and Hypotheses Addressed in the Current Thesis

Research Questions/Hypotheses	**Study/Chapter**
• What is the potential of a discourse comprehension paradigm as a unique method for neuropsychological testing, and what can it bring to present testing procedures? • How are the measures of macro-structural discourse comprehension characterized in relevant studies? • What is the link between measures of discourse comprehension and neuropsychological test measures that are frequently used in clinical settings?	Study 1/Chapter 2
• Do bilingual older adults show advantages/disadvantages in neuropsychological test performance? • Are the bilingual advantages/disadvantages observed in specific cognitive domains of neuropsychological tests? • Are the findings influenced by the language of test administration and participants' language proficiency, age of language acquisition, and immigrant status?	Study 2/Chapter 3
• What is the influence of demographic and immigration-related variables on the performance of the MMSE and RUDAS in the manifestation of AD? **H1:** The performance of the native-born Germans with AD would be better on the MMSE than Turkish immigrant and native-born Turkish patients with AD due to this test's lack of cross-cultural potential. **H2:** Turkish immigrants with AD would underperform compared to the native-born Turkish group with AD due to immigration-related experiences. **H3:** Performance differences in the RUDAS would be less evident between groups, and test results would be less affected by education, gender, age, acculturation, and bilingualism. • Are there performance differences in the individual items of the MMSE and RUDAS across groups?	Study 3/Chapter 4

2. Chapter 2: Can Discourse Processing Performance Serve as an Early Marker of Alzheimer's Disease and Mild Cognitive Impairment? A Systematic Review of Text Comprehension (1st Publication)

Eesha Kokje[a], Simge Celik[a], Hans-Werner Wahl[a] and Christiane von Stutterheim[b]

[a] *Network Aging Research, Heidelberg University, Germany*

[b] *Institut für Deutsch als Fremdsprachenphilologie, Heidelberg University, Germany*

Kokje, E., Celik, S., Wahl, H. W., & von Stutterheim, C. (2021). Can discourse processing performance serve as an early marker of Alzheimer's disease and mild cognitive impairment? A systematic review of text comprehension. *European journal of ageing*, 19(1), 3–18. https://doi.org/10.1007/s10433-021-00619-5

European Journal of Ageing (2022) 19:3–18
https://doi.org/10.1007/s10433-021-00619-5

REVIEW

Can discourse processing performance serve as an early marker of Alzheimer's disease and mild cognitive impairment? A systematic review of text comprehension

Eesha Kokje[1] · Simge Celik[1] · Hans-Werner Wahl[1] · Christiane von Stutterheim[2]

Accepted: 1 April 2021 / Published online: 20 April 2021

Abstract
A number of linguistic and cognitive deficits have been reported during the course of Alzheimer's disease (AD) and its preceding stage of mild cognitive impairment (MCI), with some deficits appearing years before onset of clinical symptoms. It continues to be a critical task to identify tools that may serve as an early marker of pathology that are also reliably able to distinguish AD from normal ageing. Given the limited success of classic psychometric cognitive testing, a novel approach in assessment is warranted. A potentially sensitive assessment paradigm is discourse processing. The aim of this review was to synthesize original research studies investigating comprehension of discourse in AD and MCI, and to evaluate the potential of this paradigm as a promising avenue for further research. A literature search targeting studies with AD or MCI groups over 60 years of age was conducted in PubMed, Web of Science, and PsycINFO databases. Eight articles with good quality were included in the review. Six measures of discourse comprehension—naming latency, summary, lesson, main idea, proportion of inferential clauses, true/false questions—were identified. All eight studies reported significant deficits in discourse comprehension in AD and MCI groups on five of the six measures, when compared to cognitively healthy older adults. Mixed results were observed for associations with commonly used cognitive measures. Given the consistent findings for discourse comprehension measures across all studies, we strongly recommend further research on its early predictive potential, and discuss different avenues for research.

Keywords Discourse · Alzheimer's disease · Mild cognitive impairment · Language · Comprehension

Introduction

As life expectancy continues to increase, the ageing population continues to grow in number, and so does the prevalence and incidence of age-related disorders. Dementia is one of the most common age-related disorders, and is a major cause of concern worldwide due to its untreatable nature. As of 2018, an estimated 50 million people worldwide live with dementia, with the number expected to be over 152 million by the year 2050 (Patterson 2018). Alzheimer's disease (AD) is the most common type of dementia, accounting for an estimated 60% to 80% of the cases. It results in progressive cognitive and functional decline, which is irreversible, and begins before clinical onset of AD. The clinical manifestation of AD is preceded by a transitional stage of mild cognitive impairment (MCI), which has received considerable attention as a target stage for early detection and interventions.

The long preclinical stage of AD is marked by irreversible neuropathological changes, such as deposition of amyloid plaques and neurofibrillary tangles, which result in neuronal and synaptic loss, and cortical atrophy, as well as subtle cognitive deficits (Bäckman et al. 2005; DeTure and Dickson 2019). Due to the irreversible nature of AD, current possibilities are limited to delaying onset of the disease or slowing its progression. Interventions based on modifiable risk and protective factors (Imtiaz et al. 2014; Livingston et al. 2017; Xu et al. 2015) can only be successful when targeted before significant pathological changes and cognitive decline have

Responsible Editor: Matthias Kliegel.

✉ Eesha Kokje
kokje@nar.uni-heidelberg.de

[1] Network Aging Research, Heidelberg University, Bergheimer Strasse 20, 69115 Heidelberg, Germany

[2] Institut für Deutsch als Fremdsprachenphilologie, Heidelberg University, Heidelberg, Germany

occurred (DeKosky 2003). Cognitive decline resulting from AD pathology occurs in several domains, over a long period of time, up to over a decade before individuals meet clinical criteria for AD (Amieva et al. 2008; Chen et al. 2001). AD is a clinically heterogeneous disease, often difficult to distinguish from normal cognitive ageing in the early and preclinical stages of the disease. Episodic memory impairment is commonly reported in early AD stages. However, an important diagnostic step forward has been that it is no longer seen as the defining symptom (Lim et al. 2020), as impairment may be evident in several other domains, including executive functions, visuospatial ability, or language, in the form of reduced complexity of sentences or anomia (Galton et al. 2000). Considering the heterogeneity in presentation of the disease, the irreversible nature, as well as the increasing emphasis on characterization of clinical and preclinical stages of AD as a continuum (Jack et al. 2018), it is crucial to develop assessment tools that can identify the subtle cognitive changes early on that indicate underlying pathology before AD is clinically evident.

MCI was introduced as a transitional phase between cognitive ageing and dementia, which is characterized by some decline in one or two cognitive domains without marked functional impairment, making it a target stage for interventions. Reported rate of conversion from MCI to dementia varies widely, depending on a number of factors, including, but not limited to, subtype of MCI, level of cognitive impairment, length of follow-up, loss to follow-up, and study setting (Ward et al. 2013). Generally, an annualized conversion rate of 10% to 15% has been widely cited, with this rate being as high as 28% for the amnestic subtype (Schmidtke and Hermeneit 2008). It has, however, been challenging to detect subtle changes occurring due to pathology during this stage, to distinguish MCI from age-related cognitive decline, and to predict conversion to dementia; however, it has been suggested that combining several markers greatly increases predictive power (Devanand et al. 2008). Therefore, continued efforts are required in the detection of MCI and in predicting conversion to dementia.

The role of discourse processing as a potentially important early marker of AD and MCI

Assessment tools that are able to detect pathology-related cognitive decline early in the course of the disease remains a challenging field looking for innovative approaches. Established neuropsychological testing includes the Mini-Mental Status Examination (MMSE) as a screening tool, verbal fluency and the Boston Naming Test (BNT) for measuring language abilities, the logical memory subscale from Weschler's Memory Scale for measuring episodic memory, constructional praxis for measuring visuoconstructive abilities, and the Trail Making Test (TMT) to measure executive functions. Language functions are preserved for longer, and reveal rather low vulnerability during healthy ageing (Park and Reuter-Lorenz 2009). Classic cognitive testing, so far, taps into language-related functions only marginally (Cummings et al. 1988; Taler and Phillips 2008; Verma and Howard 2012; Vuorinen et al. 2000), using tasks involving word retrieval, verbal fluency, and word list memory.

Most studies have suggested impairment primarily in the lexical and semantic components of language (Emery 2000; Henry et al. 2004; Reilly et al. 2011), which is central for relating the concept to the linguistic form. In contrast, syntactic and phonological components appear to be relatively preserved, until the advanced stages of the disease, although syntactic complexity is reduced (Emery 2000; Rochon et al. 1994). These methods for studying language-related functions, however, are rather artificial as they lack any context, and have little ecological validity. There is also considerable heterogeneity in the patterns of cognitive and linguistic decline observed, and different language functions may be variably affected in different individuals (Cummings 2000), which may not always be captured by studying language functions in isolation, such as lexical access, verbal memory, or syntactic complexity.

A more holistic approach is to study language deficits in their interactions with cognitive processes. Linguistic and cognitive processes are highly interdependent, with language shaping cognitive processes—including non-verbal processes, such as visual perception or memory—and cognition, in turn, aiding higher-order linguistic processes (Gerwien and von Stutterheim 2018). Here, we focus on discourse as a highly demanding task involving interdependency of cognitive and linguistic processes. Discourse refers to written or spoken language in a social context, and according to most definitions, encompasses information distributed over more than one sentence. Despite syntactical preservation, production of discourse is impaired very early on in the course of the disease, even before the onset of other clinical symptoms, as evidenced in studies using spontaneous speech and picture description tasks (Mueller et al. 2018; Slegers et al. 2018).

Importantly, discourse processing is qualitatively different from isolated linguistic tasks or even sentence processing. It occurs simultaneously on multiple representational levels, namely, surface code, textbase, and situation model (Fletcher and Chrysler 1990; Graesser et al. 1997). The most basic and superficial level of representation is the *surface code*, which simply preserves the exact syntax and wording of the text, generally for a few seconds only. The *textbase* is a representation of the text at a semantic level, extracting and retaining meaning from the text by inferencing, but not retaining the exact details of the text. Finally, the *situation model* refers to the level of representation wherein overall meaning of the text is interpreted in the wider context of

structured world knowledge. These final two levels of processing require an interaction between cognitive and linguistic processes, as it involves abstraction, organization of information, contextual embedding, accessing appropriate schemata, incorporating relevant knowledge structures, perspective taking, and inferencing (Sparks 2012; Thorndyke 1976). Macrostructural organization is an essential property at the textbase level as well as at the level of the situational model, relevant for establishing global coherence (Kintsch 1988; Kintsch and Rawson 2005). *Macrostructural processing* is a form of higher-level language processing, which involves the representation of the global meaning of discourse in the form of the topic, theme, or gist, as opposed to *microstructural processing*, which is a very local form of processing, involving linguistic structure at the phrasal or sentence level, and meaning of words (Van Dijk 2019).

Considering the complexity of the processing involved at the macrostructural level, it may be particularly susceptible to decline early in the course of AD development. This has in fact been observed in studies using a discourse production paradigm, wherein, macrolinguistic features of discourse production were the most susceptible to decline in the early and prodromal stages of AD (Brandão et al. 2013; Pistono et al. 2019). The patterns of deficits observed in micro- and macrostructural processing have been shown to have utility in distinguishing clinical populations (Ulatowska et al. 1999). They were able to successfully distinguish individuals with MCI due to AD from those with MCI due to non-AD pathologies (Mazzon et al. 2019). Further, studies indicate that macrostructural level comprehension remains intact in normal cognitive ageing; in fact, older adults rely increasingly on this form of processing, in order to compensate for decline in detail-level memory (Radvansky and Dijkstra 2007; Ulatowska et al. 1998). Hence, emerging research targeting discourse comprehension at a macrostructural level may have the potential to add to the ongoing discussion on early markers of pathology, and in distinguishing normal cognitive ageing from AD pathology-related decline. Therefore, a systematic account of the available evidence in this area is needed.

Goals of review

The overarching goal of this review is to evaluate currently available research measuring macrostructural discourse comprehension in the course of AD, and to assess the potential of a discourse comprehension paradigm as a novel approach in neuropsychological testing, in seeing what it may add to current testing practices. The review focuses on studies with individuals with late-onset early stage AD (mild or early moderate) and individuals with MCI, in comparison with cognitively healthy older adults. Subgoals of our review are, first, to systematize and characterize the measures of macrostructural discourse comprehension, applied in relevant studies. Second, we evaluated the associations between measures of discourse comprehension and cognitive and neuropsychological test measures that are commonly in use in clinical settings.

Method

Search strategy

A literature review was performed using the methods specified in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; see S1). We searched PubMed, Web of Science, and PsycINFO for original, peer-reviewed research articles published in English, using combinations of the following search terms: Alzheimer's disease, mild cognitive impairment, discourse, global coherence, macrolinguistic, connected speech, connected language, narrative speech, narrative comprehension.[1] We placed no restrictions based on date of publication of a study (for detailed search strings, see S2).

The searches were completed on January 20, 2020. Two researchers (EK and SC) screened the title and abstract of articles. When abstracts did not contain enough information to determine inclusion or exclusion, the full text of the article was obtained and read. Additionally, the references of included studies were screened to identify any other studies that may meet the inclusion criteria. Any conflicts between the two reviewers were discussed and resolved.

Study selection

For a study to be included in the review, the following criteria had to be met: (1) the study included a group of participants who had a formal diagnosis of Alzheimer's disease or Mild Cognitive Impairment, using well-established criteria; (2) the study included a healthy control group for comparison; (3) mean age of the healthy group was ≥ 60 years, or population was age-matched to the patient group; (4) study consisted of a text followed by outcomes measuring overall comprehension of text; (v) study was published in English in a peer-reviewed journal. The criteria for exclusion were as following: (1) Studies with other types of dementia population; (2) studies measuring verbatim recall of discourse

[1] As per recommendation from one reviewer, we conducted an additional search with the search terms 'gist', 'inference', and 'text comprehension' in combination with 'Alzheimer's disease' and 'Mild Cognitive Impairment' to potentially identify articles we may have missed in our original search. However, this search did not yield any new articles that met our criteria. These search results have not been added to the original search results.

texts or only memory for details within the text; (3) studies measuring spontaneous or picture-elicited discourse production; (4) case studies. No restrictions were placed on the type of study design.

Data extraction

The reviewers (EK and SC) extracted the following data from the articles that were finally included in synthesis: first author's last name, year of publication, participant groups, number of participants, age, country in which study was conducted, language of study, stage of Alzheimer's/MCI, diagnostic criteria used, variables controlled for, task, outcome measures.

Quality assessment

The Standard Quality Assessment Criteria for Evaluating Primary Research Papers: Quality Scoring for Quantitative Studies or 'QualSyst' (Kmet et al. 2004) was used to assess and rate the quality of the studies that were finally included in the analysis. The assessment originally contained a total of fourteen questions, of which, two questions concerning 'intervention' were eliminated, as the review did not include intervention studies. There were three possible scores for each question. A score of '2' indicated the study fulfilled the criteria fully, a score of '1' indicated a partial fulfilment of the criteria, and when criteria was not fulfilled, a score of '0' was given. The score obtained for each study was then divided by the total possible score (24 points), giving a score between 0 and 1. Two raters (EK and SC) scored the studies independently, and a good inter-rater agreement was observed (ICC = 0.87). Any discrepancies in scoring between the two raters were discussed until consensus was reached. The quality score for the individual studies is presented in Table 1. All studies were deemed to be of a fairly good quality (≥ 0.75).

Results

Search results and study characteristics

The search yielded a total of 4716 articles combined from PubMed (1954–2020), Web of Science (1934–2020), and PsycINFO/EBSCO (1934–2020). After removing duplicates 2941 articles remained, for which title and abstract were screened. Additionally, references of included articles were screened, and three additional articles, which met the inclusion criteria, were identified (Chapman et al. 2006; Graville and Rau 1991; MacDonald et al. 2001), making it a total of 2944 articles that were screened for eligibility. Of these, 2895 articles were excluded as they did not pertain to the topic or did not meet inclusion criteria. Full-text screening was conducted, and inclusion and exclusion criteria were applied for the remaining 49 articles. Of these, 41 articles were excluded, with a good inter-rater agreement ($\kappa = 0.81$). The reasons for exclusion are highlighted in Fig. 1. The most common reason for exclusion was 'Outcome not relevant' with most studies being excluded as they investigated spontaneous or picture-elicited discourse production or verbatim recall of text. Finally, a total of eight articles were included in the review, which aimed to measure discourse comprehension at a macrolevel, in adults with Alzheimer's disease or MCI.

An overview of the study characteristics is presented in Table 1. All the studies were cross-sectional, in which AD and/or MCI groups were compared to cognitively healthy older adults. Seven of the eight studies were conducted with native English-speakers, with six of them being conducted in USA, and one in Canada. One study was conducted in Brazil, with a native Brazilian Portuguese-speaking population. The studies were published between the years 1998 and 2019. One study included two groups of healthy older adults, classified as 'young-older adults' (65–80 years) and 'old-older adults' (> 80 years) (Chapman et al. 2006), and one study (Welland et al. 2002) included two AD groups—early stage (EDAT) and moderate stage (MDAT). The total sample sizes ranged from 20 to 84 participants, with their mean ages ranging from 65 to 86. All studies controlled for age, and all but one (Chapman et al. 2002) controlled for education, wherein the different groups were either matched on these variables or the variables were entered as covariates during analysis. Apart from this, six studies also controlled for sex (Chapman et al. 1998, 2002, 2006; Creamer and Schmitter-Edgecombe 2010; Drummond et al. 2019; Schmitter-Edgecombe and Creamer 2010), one study controlled for depression(Chapman et al. 2006), and one study controlled for IQ (Welland et al. 2002). All studies determined cognitive status of the healthy control group using at least one or a combination of several of the following measures—MMSE, self-report, Clinical Dementia Rating (CDR), Global Deterioration Scale (GDS).

Only one study (Drummond et al. 2019) used a test from a standardized battery (MAC battery) (Fonseca et al. 2008), and one (Welland et al. 2002) used a modified form of the Discourse Comprehension Test (DCT) battery (Brookshire and Nicholas 1993) to measure discourse comprehension. In other studies, an experimental task was used to measure discourse comprehension, wherein participants were presented with a series of short texts, usually narrative stories. This was generally followed by a variety of tasks designed to test participants' comprehension of the texts. This involved giving a short summary of the story, stating the lesson or intended main idea of the story, answering true/false questions about the story, a think-aloud paradigm while reading,

Table 1 Characteristics of included studies

References	Population (*N*)	Mean age	Language; country	Stage	Linguistic task	Discourse comprehension measures	Variables controlled for	Diagnostic criteria used (staging)	Quality assessment rating
Almor et al. (2001)	AD (10), NC (10)	AD = 82, NC = 78	English; USA	Mild to moderate	Reading aloud visual target words continuing from auditory stimuli	Naming latencies	(Age, education)*	NINCDS-ADRDA[a] (MMSE[b])	0.77
Chapman et al. (1998)	AD (10), Fluent Aphasia (10), NC (10)	AD = 65, FA = 65, NC = 65	English; USA	Mild to early moderate	Summarizing fables	Gist, lesson of stories, main idea	Age, education, sex	NINCDS-ADRDA (MMSE)	0.75
Chapman et al. (2002)	AD (24), MCI (20), NC (25)	AD = 72.4, MCI = 72.7, NC = 76.1	English; USA	Mild	Summarizing biographical narratives	Summary, main idea, lesson of stories	Age, sex	NINCDS-ADRDA (MMSE, CDR[c]); Petersen et al. 1999	0.83
Chapman et al. (2006)	AD (12), Young OA (12), Old OA (12)	AD = 71.6, YOA = 72.2, OOA = 85.8	English; USA	Mild	Summarizing a narrative; Logical Memory Subtest of WMS-III	Transformed gist, main idea	Education, sex, depression	NINCDS-ADRDA	0.83
Creamer and Schmitter-Edgecombe (2010)	AD (20), NC (20)	AD = 77.2, NC = 76.7	English; USA	Mild	Think-aloud while reading stories	Proportion of inferential clauses	Age, education, sex	NINCDS-ADRDA (CDR)	0.96
Drummond et al. (2019)	AD (14), aMCI (31), NC (39)	AD = 75.3, aMCI = 72.2, NC = 71.8	Portuguese; Brazil	Mild	Summarizing narrative story	Main ideas, comprehension questions, inferential lesson	Age, education, sex	DSM-5[d]; Winblad et al., 2004	0.88
Schmitter-Edgecombe and Creamer (2010)	aMCI (23), NC (23)	MCI = 70.8, NC = 70.6	English; USA	MCI	Think-aloud while reading stories	Proportion of inferential clauses	Age, education, sex	Petersen et al., 2001; CDR	0.96
Welland et al. (2002)	EDAT (8), MDAT (8), NC (8)	EDAT = 78, MDAT = 76.7, NC = 72.2	English; Canada	Mild and moderate	Answering yes/no comprehension questions about narratives	Implied main ideas and implied details questions	Age, education, IQ	NINCDS-ADRDA (MMSE)	0.83

[a]National Institute of Neurological and Communicative Diseases-Alzheimer's Disease and Related Disorders Association (McKhann et al. 1984)

[b]Mini-Mental State Examination (Folstein et al. 1975)

[c]Clinical Dementia Rating (Hughes et al. 1982)

[d]The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

*Entered as covariates

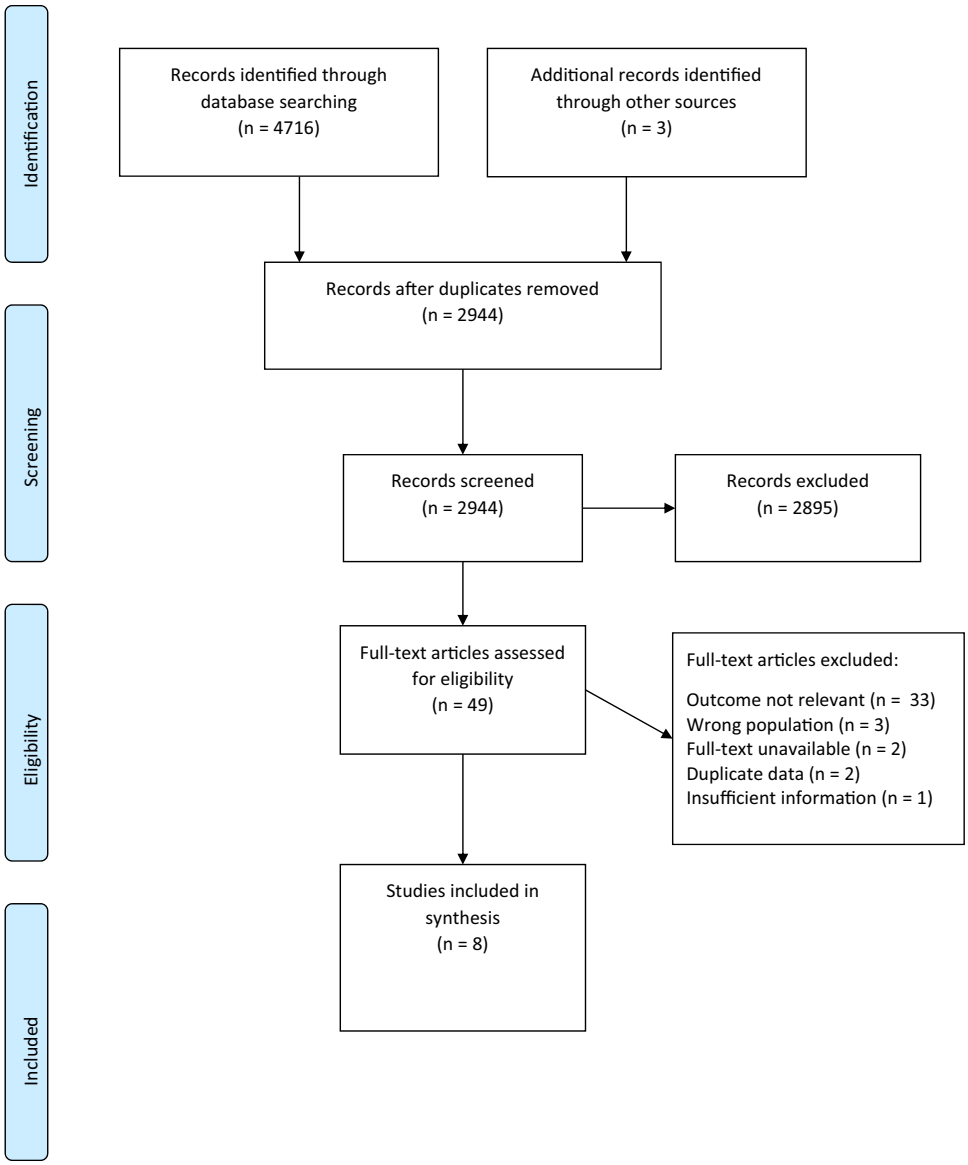


Fig. 1 Flowchart of literature search and study selection process

or reading out loud the last word in the story, which was either congruent or incongruent with previous text. With one exception (Welland et al. 2002), the studies did not report independently on hearing and visual/reading abilities of participants. However, they generally included practice trials before the start of the study to ensure participants understood the task, and were able to perform it successfully. Almost all of the included studies looked at performance of participants on one or more neuropsychological tests (for example, subtests of Boston Diagnostic Aphasia Examination) to ensure that participants were able to follow instructions, in order to be able to perform the task. The outcome measures varied across studies, with some studies measuring the proportion of inferential and non-inferential clauses produced (Creamer and Schmitter-Edgecombe 2010; Schmitter-Edgecombe and Creamer 2010), one study measuring naming latencies for congruent and incongruent pronouns (Almor et al. 2001), and others measuring gist-level retelling in the form of summary, lesson, main ideas (Chapman et al. 1998, 2002, 2006; Welland et al. 2002). Due to this heterogeneity in tasks and reported outcome measures, a meta-analysis was not performed.

Diagnostic criteria

One study (Drummond et al. 2019) used the Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition (DSM-5) criteria for Major Neurocognitive Disorder due to Alzheimer's Disease (Sachdev et al. 2014), for diagnosis of AD. All other studies used the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association

(NINCDS-ADRDA) criteria (McKhann et al. 1984). In all studies, a diagnosis of 'probable AD' was applied, wherein individuals are diagnosed based on clinical and neuropsychological evidence without histopathologic confirmation. As these were cross-sectional studies, they could not follow-up to confirm AD via autopsy. Additionally, all, but one, studies were conducted prior to 2011, when the NINCDS-ADRDA criteria were first revised to the National Institute on Aging-Alzheimer's Association (NIA-AA), to include biomarker evidence in the diagnosis of AD (McKhann et al. 2011). The DSM-5 criteria, which was used in the study by Drummond et al. (2019), does not yet include biomarker evidence in diagnosis of Major Neurocognitive Disorder due to AD. The major difference between the NINCDS-ADRDA and the DSM-5 criteria is that presence of memory impairment is not required for diagnosis in DSM-5; rather, impairment in any two cognitive domains is acceptable. This shows a general trend towards moving away from memory impairment, as is seen in the NIA-AA 2011 criteria too, which was a revision of the NINCDS-ADRDA criteria. For determining the stage of the AD (mild, moderate, severe), studies used either MMSE or CDR scale (Folstein et al. 1975; Hughes et al. 1982). These two scales have been shown to have good agreement for the stages of AD that have been investigated in included studies (Perneczky et al. 2006). Overall, although two different criteria were used for the diagnosis of AD, the criteria were comparable enough that a qualitative synthesis of studies was possible.

For a diagnosis of MCI, one study (Chapman et al. 2002) used the criteria by Petersen et al. (1999); another study (Schmitter-Edgecombe and Creamer 2010) applied the criteria by Petersen et al. (2001). The studies also ruled out other possible causes of cognitive impairment (such as stroke or other neurological or psychological causes) via a series of tests. As with the diagnostic criteria for AD, the criteria for MCI too evolved to shift focus away from memory complaints, towards a more wholesome approach to include all cognitive domains. While the Petersen et al. (1999) criteria required a subjective memory complaint, the subsequent revised criteria from 2001 onwards allowed for complaints in any cognitive domain. Instead, the Petersen et al. (2001) criteria focused on classifying MCI into several subtypes (e.g. amnestic MCI, multi-domain MCI), depending on the cognitive domain(s) in which deficits were observed. Accordingly, studies included in the review that were conducted after the Petersen et al. (2001) criteria were established, have included population specifically with a diagnosis of amnestic MCI (aMCI). Finally, one study (Drummond et al. 2019) applied the Winblad et al. (2004) criteria, which was a revision of the Petersen et al. (2001) criteria. This revision acknowledges that there may be multiple aetiologies for each subtype of MCI, and modifies the stipulation concerning normal daily functioning in previous criteria, to allow for subtle impairment in complex functions. Although different evolving diagnostic criteria have been used in the included studies, the different criteria are not sufficiently different enough so as to affect a qualitative synthesis of these studies.

Measures of discourse comprehension

Due to a lack of standardized tests for measuring discourse comprehension, there was considerable variability in the method used to evaluate comprehension, and consequently in the type of outcome measures used. Most measures used some form of language production to measure comprehension. This implies a general problem which poses a dilemma for comprehension studies in other contexts as well (e.g. language acquisition, pedagogy). We know from studies on language production that patients with AD have deficits in accessing lexical units, though deficits at the morphological and syntactical level are less pronounced. These deficits could affect the validity of the measures for language comprehension.

Relevant outcome measures used in each study were identified. Several of the identified outcome measures were used in multiple studies, and these were grouped together. The names of the outcome measures were derived from the outcomes used in the included studies. However, the terms for certain measures were used interchangeably in the different studies. Therefore, to summarize the results from different studies, the measures were categorized according to the definitions or descriptions of the measure presented in the studies, rather than the terms used. Accordingly, the measures were grouped into the six variables described below. The results for each measure are summarized in Table 2.

Naming latencies

Naming latency was used as an outcome in only one of the studies (Almor et al. 2001). In this study, participants were presented with a short text in an auditory format, in which two entities (antecedents) were introduced in the first sentence. The final sentence referred back to these entities, wherein it mentioned one of the entities and was left incomplete before the other entity is mentioned. Finally the target pronoun was presented visually, which was either congruent with the incomplete sentence or incongruent, based on the singularity or plurality of the antecedent and the pronoun. Participants were to read aloud the pronoun, and their response time was measured. Ideally, when the pronoun is incongruent to the antecedent, response time should be longer compared to when it is congruent, as it would be more difficult to integrate an incongruent word into the passage, indicating adequate processing of cohesive devices. This effect would, however, only be seen if individuals are able to integrate different information units

Table 2 Comparison of group performance on discourse comprehension measures

References	Naming latencies	Summary	Lesson/message	Main idea	Inferential clauses	Comprehension questions
Almor et al. (2001)	AD < NC***	–	–	–	–	–
Chapman et al. (1998)	–	AD < NC†	AD < NC†	AD < NC†	–	AD < NC†
Chapman et al. (2002)	–	AD = MCI < NC**	AD < MCI < NC***	AD < MCI < NC***	–	–
Chapman et al. (2006)	–	AD < OOA*** AD < YOA*** OOA < YOA*	AD < OOA*** AD < YOA*** OOA = YOA	AD < OOA** AD < YOA** OOA = YOA	–	–
Creamer and Schmitter-Edgecombe (2010)	–	–	–	–	AD < NC*	AD < NC***
Drummond et al. (2019)	–	AD < MCI < NC*	AD = MCI < NC*	–	–	AD < MCI < NC*
Schmitter-Edgecombe and Creamer (2010)	–	–	–	–	MCI < NC**	MCI < NC*
Welland et al. (2002)	–	–	–	–	–	MDAT = EDAT < NC**

*$p \leq .05$, **$p \leq .01$, ***$p \leq .001$

†p value not reported

within a macrostructure, indicating the ability to establish coherence relations. Slower reaction times for incongruent trials were seen in healthy older adults, as well as the group with AD. However, the size of the effect was much smaller in the AD group compared to the healthy older adults, meaning that the difference in the reaction times to congruent vs incongruent trials was much higher in the controls than in the AD population, as was expected. This shows that AD patients were less sensitive to incongruent pronouns, indicating a problem in integrating and connecting the presented information.

Summary

In four studies (Chapman et al. 2006, 1998, 2002; Drummond et al. 2019), participants were presented with a short story. Following this, participants were asked to retell the story or give a summary in their own words which involved focusing on important units of information that are required for an overall understanding of the story, and omitting unnecessary details. Participants' performance was scored according to the number of main informational and/or thematic units produced. This measure can be taken to illustrate in how far language production was taken as a measure for comprehension. The linguistic output was not analysed with respect to relevant features of language production (time course, lexical choice, or number of words per sentence), but only at the level of meaning in relation to the stimulus text. AD groups produced fewer synthesized meaningful units of information compared to cognitively healthy adults in all four studies, including the old-older adults. In both studies with MCI population (Chapman et al. 2002; Drummond et al. 2019), the MCI group performed significantly worse than the healthy older adults. Between the AD and MCI groups, AD group scored significantly lower than the MCI group in one study (Drummond et al. 2019); however, the performance of the two groups was comparable in another study (Chapman et al. 2002). Additionally, there was a small but significant difference in the performance of old-older adults compared to young-older adults. This was the only measure for which such a difference was observed.

Lesson/message

Another probe following the presentation of a short story, employed in four studies (Chapman et al. 2006, 1998, 2002; Drummond et al. 2019), was the lesson or message probe, wherein participants were to formulate a lesson or a title that could be inferred from the story. AD and MCI patients scored significantly lower than healthy adults, focusing on unimportant details from the story rather than an overall lesson. Additionally, the AD group performed significantly worse than old-older adults. When performances of MCI and AD groups were compared, the results were mixed,

wherein one study (Drummond et al. 2019) reported no significant difference in their performance, whereas another study (Chapman et al. 2002) reported that the AD group scored significantly lower than the MCI group. This measure required maximum inferential processing, as participants need to be able to synthesize a large amount of information, condense it, and make interpretations about what message it carries.

Main idea

This probe, also administered following a short story in three of the studies (Chapman et al. 2006, 1998, 2002), measured the ability of participants to summarize the story in one sentence i.e. the primary concept of the story, which required substantial condensation of information and abstraction into one generalized idea. Both AD and MCI groups performed significantly worse than the control group. Furthermore, a significant difference was observed between the performance of AD and MCI groups, with the AD group scoring lower than the MCI group. AD and MCI patients were generally prone to giving more unimportant information or details rather than summarizing statements, although individuals' responses varied to some extent. Additionally, as was also observed for previous measures, the AD group's performance was significantly worse compared to the old-older adults.

Inferential clauses

Two studies used a think-aloud procedure (Creamer and Schmitter-Edgecombe 2010; Schmitter-Edgecombe and Creamer 2010), wherein participants were given a short narrative text to read, and were asked to vocalize their thoughts about the story simultaneously while reading the narrative text. Every utterance of participants was classified either as an 'inferential clause' or a 'non-inferential clause', by two assessors, one of whom was blinded to the diagnostic status. The classification system used by Trabasso and Magliano (1996) was employed, wherein, statements that were either explanations, predictions, or formed associations, were categorized as 'inferential', and other statements (e.g. repetitions or paraphrases) were classified as 'non-inferential'. Although, overall, all groups uttered more inferential clauses compared to non-inferential, both AD and MCI groups uttered significantly fewer inferential clauses compared to cognitively healthy adults.

Comprehension questions

One included study (Welland et al. 2002) used Yes/No questions as the only outcome to measure comprehension following story narration. The format used in this study was adapted from the standardized discourse comprehension test developed by Brookshire and Nicholas (1993). The questions were categorized based on the level of detail—main idea and details, and the type of information—implied or stated. Both patient groups—EDAT and MDAT—performed significantly worse on all types of questions, compared to the healthy group, but the performance of the two patient groups did not differ from one another on any measure. All groups generally performed better on 'main idea' questions compared to 'details', and on 'stated' information compared to 'implied'. Three other studies (Creamer and Schmitter-Edgecombe 2010; Drummond et al. 2019; Schmitter-Edgecombe and Creamer 2010) included comprehension questions following the other retelling and 'think-aloud' tasks, to test for comprehension of the narrative passage. In two studies, half of the True/False questions were based on information that needed to be inferred from the text and half of the questions were based on facts that were explicitly stated in the text. AD and MCI groups answered fewer questions correctly, overall, compared to controls, in all studies. However, when performance on inferential questions was examined specifically, in the two studies that made this distinction, AD and MCI groups did not differ significantly from controls. Therefore, in these studies, this measure was relatively less informative, as the nature of the questions (True/False) poses two problems. First, there is a 50% chance of answering the question correctly, irrespective of how well one may or may not have understood the narrative. This can be observed in the AD group's performance, which was in fact at chance level. Second, there may be possible ceiling effects in the healthy adults group's performance, as can be observed in the high means across all the studies. It is also possible that performance on this task was made easier by reliance on recognition memory, rather than recall. Therefore, this method may not be optimal in terms of appropriateness and complexity in investigating the current question.

Overall, a deficit in discourse comprehension in individuals with AD and MCI was consistently observed across all studies, pointing to a robust effect. These result show that, with the exception of one measure, discourse comprehension measures are able to reliably distinguish early stage AD and MCI patients from cognitively healthy older adults.

Association between discourse comprehension measures and cognitive measures

In addition to examining the discourse comprehension differences between AD, MCI, and cognitively healthy older

adults, the review also aimed to examine whether performance on the discourse comprehension task correlated with performance on commonly used neuropsychological tests. The purpose of this was twofold: the first was to examine which cognitive processes, if any, are able to predict performance on a discourse comprehension task, giving an indication of the underlying mechanisms involved; the second was to determine whether discourse comprehension tasks are able to tap into processes beyond what traditionally used neuropsychological tests measure. Studies used tests such as RAVLT, WAIS-III, listening span, D-KEFS, MMSE to measure verbal memory, working memory, executive functions. However, all these measures were not consistently used across all included studies. Therefore, it was somewhat challenging to draw robust conclusions about their association with discourse comprehension. For measures that were employed in multiple studies, the results were mostly mixed. When the association between MMSE scores and performance on the experimental task were examined, one study (Chapman et al. 2002) found a significant correlation ($r = 0.65$), whereas another study (Almor et al. 2001) found only a marginally significant correlation between the two measures, which disappeared when working memory was accounted for. In another study (Welland et al. 2002), MMSE scores did not significantly predict discourse comprehension when episodic memory or working memory were added to the regression model. Similarly, working memory measures were associated significantly ($r = 0.64$, $r = -0.83$) with discourse comprehension in two studies (Almor et al. 2001; Welland et al. 2002), but two other studies (Creamer and Schmitter-Edgecombe 2010; Schmitter-Edgecombe and Creamer 2010) found no association. It is important to note that different studies used different tests to measure working memory (e.g. listening span, WAIS-III, digit span). These varying results may be due to heterogeneity in the different experimental tasks and tests used in different studies. However, both studies that included a verbal memory measure (RAVLT) found a significant, albeit moderate ($r = 0.50$ to $r = 0.64$) correlation with discourse comprehension measures. Only one study (Welland et al. 2002) reported a positive association with episodic memory ($r = 0.91$). Additionally, one study (Creamer and Schmitter-Edgecombe 2010) found significant correlations with TMT-A ($r = 0.58$) and D-KEFS ($r = 0.62$), measuring attention and executive functions, respectively. The study also looked at several other tests of attention and executive functions, as well as tests of language, but none of these showed association with macrostructural measures of discourse comprehension. The moderate correlation with verbal memory, and the moderate or non-significant correlations with other measures indicate that discourse comprehension tasks tap into additional processes that are not assessed by neuropsychological tests used routinely in the clinical diagnosis of AD. This warrants investigation of discourse comprehension tasks as a possibly more comprehensive assessment tool.

Discussion

The purpose of this review was to synthesize results of studies investigating whether individuals with mild AD or MCI experience significant deficits in macrostructural discourse comprehension, in comparison with cognitively healthy older adults. In the included studies, participants were presented with short narratives, which were accompanied either by a think-aloud procedure, or were followed by a retelling of the story in short, along with questions which measured comprehension of the story. Six measures were identified from these studies—naming latencies, global synopsis, lesson, main idea, inferential clauses, and comprehension questions. Despite some variations in the methods and outcome measures across the eight studies included in the review, significant deficits in macrostructural discourse comprehension were observed in AD and MCI groups across all, but one, measures in all studies, in comparison with cognitively healthy older adults. These findings also receive additional support from results of neuroimaging and biomarkers employed in the study by Drummond et al. (2019), where they observed that performance on the discourse task was associated with the degree of neurodegeneration observed, in terms of reduced white matter integrity and neuronal loss. Although the number of studies in this review was limited, we observed a very consistent pattern of findings across the studies, indicating a rather robust effect.

The groups with AD performed significantly worse than healthy older adults on five of six measures, with one measure (comprehension questions) showing mixed results. Moreover, individuals with MCI similarly displayed significant deficits in performance when compared to the healthy groups. In studies that included both, AD and MCI groups, a direct comparison of their performance showed mixed results. On the measure of 'main idea', MCI group outperformed the AD group. However, for the 'lesson' measure, performance of the two groups was comparable in one study, whereas AD group performed worse than the MCI group in another study. Similarly, for the 'summary' measure, AD group performed worse than MCI group in one study, whereas their performance was comparable to the MCI group in another study. Most notably, however, one study compared performance of the AD group with the 'old-older adults' group (> 80 years), and found that the AD group's performance was significantly worse on all three outcome measures included in the study. This is noteworthy, as the mean age of the 'old-older adults' group was significantly higher than that of the AD group. Although compared to younger adults,

macrolevel comprehension shows some decline in older adults (Cohen 1979), over time it stabilizes, and is seen to be fairly preserved in the old-old, even though memory for details is generally seen to deteriorate (Radvansky and Dijkstra 2007; Ulatowska et al. 1998).

In addition to the discourse comprehension task, the studies also included some commonly used standardized cognitive and neuropsychological tests. The only measure for which an association was observed across the limited number of studies that employed it, was verbal memory, which was measured using RAVLT. A deficit in verbal memory measures has also been observed in the preclinical stage of the disease, in earlier studies (Bondi et al. 1994; Howieson et al. 1997). Even so, the strength of the correlation was moderate. It is noteworthy that all studies (Creamer and Schmitter-Edgecombe 2010; Drummond et al. 2019; Schmitter-Edgecombe and Creamer 2010) that employed commonly used verbal tasks—verbal fluency and BNT—did not find a significant association with macrostructural discourse comprehension measures, even though the measures use some form of language production, and previous discourse production studies have reported word-finding difficulties (Slegers et al. 2018). Only one study reported on correlations with episodic memory. Although the correlation was strong, the measure for which the correlation was reported was 'Yes/No' comprehension questions. It would be of interest to see whether there is a correlation between episodic memory performance and more complex measures such as summarizing or giving the main idea of the text. For working memory measures and MMSE, the associations produced mixed results; and when a significant association was observed, it was a moderate association. While the inconsistencies in associations may be in part due to the varying methodologies and tests used in different studies, the strength of the associations do indicate that a discourse comprehension task measures constructs beyond what classic neuropsychological tests are able to measure.

These findings highlight the need to go beyond classic cognitive and linguistic tasks (e.g. verbal fluency, confrontation naming), for a more comprehensive approach, in the neuropsychological assessment of MCI and AD. A discourse comprehension task is more representative of everyday communication and thus gives a more well-rounded picture of cognitive and linguistic deficits, over tasks measuring isolated linguistic functions. The complexity of such an assessment paradigm also means that it is perhaps a more sensitive indicator of AD pathology in the preclinical stage, although that remains to be seen, and should be an avenue for future research. Additionally, breakdown of communication is a major issue in the latter stages of AD, and is a moderating variable in determining functional independence of individuals. A discourse comprehension-based assessment tool may help track the level of functional impairment as disease progresses, and serve as a tool for targeting interventions to maintain communication ability.

The findings of this review are also notable considering that syntax and phonology are preserved in production of language during the early or even early moderate stage of AD (Kavé and Levy 2003). Evidence from studies examining spontaneous or picture-elicited discourse production shows a similar pattern of breakdown, wherein participants produce syntactically and phonologically sound sentences. However, the discourse produced was severely lacking in information content, coherence, and cohesion (Chenery and Murdoch 1994; Laine et al. 1998; Toledo et al. 2018), critical macrolinguistic features of discourse. The preservation of syntactic structure in production indicates that language processing abilities are preserved at a local, sentence-based level. Tracking information and establishing links across sentences are tasks in which the deficits show. This suggests that the comprehension deficits seen in AD patients are also more reflective of impairment in cognitive functioning, and consequently in areas where language and cognition interact. Therefore, there is a need to go beyond testing paradigms that study linguistic and cognitive functions independently of the other.

While there has been considerable research looking at patterns of language impairment in AD, this research has been conducted primarily using laboratory tasks such as word lists, confrontational naming, and word definitions, which measure individual language functions in isolation from others. These same testing paradigms are then used for assessment of linguistic functions in clinical practice too. Such paradigms do not transfer to situations that people encounter in everyday life, lacking ecological validity. They give us limited insight into individual language functions, such as lexical access or semantic fluency, but no insight into the multi-level processing of language use. Therefore, the impairments seen in AD patients during communication are often attributed to lexico-semantic deficits (Price et al. 1993; Reilly et al. 2011). Considering that deficits were observed on macrostructural measures of comprehension, as shown in this review, we cannot attribute communication deficits in AD to simply one linguistic component.

Everyday communication occurs in the form of situated discourse, which involves more than simple retention and retrieval of word lists in a contextual vacuum. Production and comprehension of discourse necessitates higher-order information processing, which requires interaction of linguistic and cognitive processes. This includes integration of context, accessing the appropriate schema, understanding goals and intentions of the communicative counterpart, merging of information in the text and semantic knowledge, generating inferences, or simply deletion of superfluous or redundant details (Kintsch and Van Dijk 1978). Such an assessment paradigm that is rooted in the practicalities of

everyday interactions and experiences, provides a holistic approach in understanding cognitive and linguistic deficits in AD, offering a new dimension to neuropsychological testing practices and interventions. Previous studies with individuals with Traumatic Brain Injury (TBI) have also reported macrolevel abstraction and comprehension deficits in this population (Vas et al. 2015). They showed tasks employing macrolevel processing to have high sensitivity and specificity in TBI due to the complexity of processing required (Vas et al. 2016). With processing occurring simultaneously on multiple levels, any number of variables could be manipulated in order to pinpoint the areas where interventions should be targeted. Emerging evidence indicates that cognitive training in MCI patients that targets macrolevel processing not only benefits abstraction ability, but also extends to other general cognitive functions like attention and executive functions (Chapman and Mudar 2014; Das et al. 2019), and is also linked to brain changes (Mudar et al. 2019).

Finally, as identified from previous studies, executive functions, episodic memory, semantic memory, and working memory play important roles in discourse comprehension (Calvo 2001; Cohen 1979; Daneman and Merikle 1996; Just and Carpenter 1992). It is possible that deficits seen in macrostructural comprehension may be in part due to impairment in any one of these, or possibly even multiple processes. There is evidence already that these processes are impaired in AD (Belleville et al. 2007; Huntley and Howard 2010). And, although, possibly all of these processes may be implicated in the deficits observed, which of these play a greater role remains to be seen.

Limitations

There are several limitations to this review. First, the review was limited to studies published in English, which may also somewhat limit the countries where the included studies were primarily conducted. Another major limitation is the low number and types of studies, due to the limited literature existing in this area of research, reflecting the low emphasis on studying interaction of linguistic and cognitive processes in AD.

A further limitation is the heterogeneity of the tasks used in the studies. Due to a lack of standardized tests measuring discourse comprehension, the studies varied in the procedure and measures implemented. As a result, a meta-analysis was not conducted, which somewhat limits synthesis of the results. Further, there is a lack of consistency in the neuropsychological tests applied in the different studies. Therefore, it was difficult to draw robust conclusions about the association between cognitive abilities and discourse comprehension, and which abilities contribute to the deficits observed. Future studies should closely examine these associations.

A major limitation of the literature is the lack of longitudinal studies. Although the review placed no restriction on the type of study design, none of the studies followed-up with participants to track their trajectory. This would be especially crucial with MCI patients, as it is presently difficult to predict conversion to dementia. Another possible limitation in studying macrostructural comprehension lies in the tediousness of the procedure for analysing discourse. The linguistic expertise required to meet the standards in this field is often not available. However, there have been efforts in the past few years to simplify the procedure and to develop standardized measures for discourse analysis (Dalton et al. 2019). Additionally, recent advances in computational linguistics are promising, with major components of the analyses being automatized, making the process less time consuming and less error-prone (Aluisio et al. 2016; Clarke et al. 2020).

Finally, as addressed previously, most of the included studies employ tasks which use some form of language production to measure comprehension. This is disadvantageous to individuals whose comprehension ability may be unaffected, but who may be experiencing deficits in production of language. This issue can be resolved using tasks which do not involve production, or even entirely non-verbal tasks that measure macrostructural processing by using other cognitive domains such as in visual world paradigms.

Future directions

This review highlights the potential of discourse comprehension measures as such a novel, comprehensive approach towards neuropsychological assessment that is able to capture cognitive and linguistic variables at multiple levels—microstructural, macrostructural, pragmatic, grammatical. Given the consistent findings despite some methodological variations across studies, its sensitivity during the early and preclinical stage of AD (MCI), and its advantage over classic cognitive tests, it warrants further research with more linguistically and culturally diverse populations, and an attempt to establish a standardized format for the test, with the aim of early detection of pathology.

In one study, it was observed that individuals with AD that scored in the normal range on MMSE showed difficulties in discourse comprehension. Additionally, two studies reported that MMSE scores were not associated with performance on discourse comprehension measures. This indicates that task paradigms such as those used in the studies included in this review may be more sensitive in the early stage of the disease. This is also evident in the performance of the MCI group, which was significantly worse than the healthy group, in all the studies that included these patients. Such paradigms for assessment may also be advantageous when considering individuals with a high cognitive reserve

(CR), who take longer to show clinical indication of AD, when tested using classic neuropsychological assessment tools. It has, however, been suggested that using more complex and challenging tasks may be better able to detect the presence of pathology in this challenging group (Stern 2012, 2013).

In recent years, a number of reliable biomarkers of AD have been identified (Khoury and Ghossoub 2019). Consequently, this has opened up the possibility of detecting AD in its preclinical stage, when individuals show no cognitive deficits on standard neuropsychological assessments (Haldenwanger et al. 2010; Villemagne et al. 2011). The preclinical stage of AD is, however, characterized by subtle cognitive deficits. Although standard neuropsychological assessments, using simple, isolated tests of language and cognition may not able to detect AD pathology during the preclinical stage, this is not necessarily the case for more complex cognitive tasks. In some recent studies that used cognitive tasks requiring more complex processing (e.g. face name association task, memory binding task), significant deficits in performance were observed in preclinical AD population (Rentz et al. 2013; Tort-Merino et al. 2017). In the study by Drummond et al. (2019), which was included in this review, it was observed that severity of deficits on discourse task correlated with the degree of neurodegeneration, as measured through neuroimaging and CSF biomarkers, in the AD group. A combination of biomarkers and comprehensive cognitive testing has shown more promise in predicting clinical outcomes, over biomarkers alone (Bondi and Smith 2014). Future studies should aim for a translational approach to investigate discourse comprehension ability in preclinical AD population and its association with AD biomarkers, for the potential development of a robust assessment tool for the early detection of AD pathology in clinical settings, where biomarker use is uncommon.

Additionally, in studies in this review that included both MCI and AD groups, performance of the two groups was comparable on some measure, but significantly different on other measures. Upon closer examination, it was observed that just over half of the individuals with MCI displayed deficits in discourse comprehension, whereas the performance of the rest of the group was comparable to the healthy older adults. Previous research has shown that MCI patients who go on to convert to dementia show more severe impairment in some linguistic and cognitive domains, compared to those who do not convert (Celsis 2000). Another study also showed disparate profiles of MCI patients in a text comprehension task (Chesneau et al. 2016). It is of interest to find predictors of conversion, and this approach shows preliminary promise.

Finally, it has been suggested that neuropsychological testing should move into a new direction, focusing on novel approaches, especially in populations in prodromal stages of the disease, when classic neuropsychological tests are unable to detect underlying pathology (Rentz et al. 2013). Macrostructural processing, which taps into top-down processes, seems to be a promising area for such research. A multi-dimensional approach, combining several biological and cognitive-linguistic predictors, also helps to track cognitive changes over time and our ability to predict clinical outcomes (Bondi et al. 2008). While discourse processing is one paradigm that taps into these processes, other approaches for testing comprehension at a macrostructural level, extending to non-verbal paradigms as well, are warranted to measure and understand the decline from the prodromal stage of AD to the clinical stage.

Conclusion

Individuals with AD and MCI experience significant deficits in discourse comprehension, which are not otherwise seen in cognitively normally ageing adults, irrespective of their age. These deficits are present in the early stage of AD, and only show moderate correlation with verbal memory and working memory capacity measures, indicating that they tap into additional constructs. With the increasing emphasis on identifying and characterizing the preclinical stages of AD in order to target interventions, more studies are focusing on such novel approaches, which have shown promising results. Studying impairment in AD using tasks which require multi-level cognitive processing, integrating knowledge from different sources and modalities, could reveal deficits which do not show in less complex processes, at this stage. We conclude on the basis of the results obtained that studies which use measures that tap into top-down processes rather than studying individual linguistic and cognitive components might serve this purpose, finally leading to a diagnostic tool with clinical utility in early detection. Such an approach has utility in research and clinical settings for differential diagnosis, for predicting conversion from MCI to dementia, and also as a tool for training intervention in older adults who experience a subjective decline in cognitive functions. Longitudinal studies, beginning before clinical onset of AD, are required to determine the potential of this assessment paradigm to identify indicators of AD pathology during the preclinical stage. Additionally, further studies to increase reliability and validity of this measure, and translational studies which include neuroimaging and biomarkers, are warranted to investigate the potential of discourse comprehension assessment paradigm for these purposes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10433-021-00619-5.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was supported by the Klaus Tschira Foundation. The funding sources had no involvement in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

References

*Almor A, MacDonald MC, Kempler D, Andersen ES, Tyler LK (2001) Comprehension of long distance number agreement in probable Alzheimer's disease. Lang Cognit Process 16:35–63

Aluisio S, Cunha A, Toledo C, Scarton C (2016) Computational tool for automated language production analysis aimed at dementia diagnosis. In: International conference on computational processing of the Portuguese language, demonstration session

Amieva H et al (2008) Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. Ann Neurol 64:492–498

Bäckman L, Jones S, Berger A-K, Laukka EJ, Small BJ (2005) Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. Neuropsychology 19:520

Belleville S, Chertkow H, Gauthier S (2007) Working memory and control of attention in persons with Alzheimer's disease and mild cognitive impairment. Neuropsychology 21:458

Bondi MW, Smith GE (2014) Mild cognitive impairment: a concept and diagnostic entity in need of input from neuropsychology. J Int Neuropsychol Soc 20:129–134. https://doi.org/10.1017/S1355617714000010

Bondi MW, Monsch AU, Galasko D, Butters N, Salmon DP, Delis DC (1994) Preclinical cognitive markers of dementia of the Alzheimer type. Neuropsychology 8:374

Bondi MW, Jak AJ, Delano-Wood L, Jacobson MW, Delis DC, Salmon DP (2008) Neuropsychological contributions to the early identification of Alzheimer's disease. Neuropsychol Rev 18:73–90. https://doi.org/10.1007/s11065-008-9054-1

Brandão L, Lima TM, Parente MAdMP, Peña-Casanova J (2013) Discourse coherence and its relation with cognition in Alzheimer's disease. Rev Psicol Pesq 7:99–107

Brookshire RH, Nicholas LE (1993) Discourse comprehension test. Communication Skill Builders

Calvo MG (2001) Working memory and inferences: evidence from eye fixations during reading. Memory 9:365–381. https://doi.org/10.1080/09658210143000083

Celsis P (2000) Age-related cognitive decline, mild cognitive impairment or preclinical Alzheimer's disease? Ann Med 32:6–14. https://doi.org/10.3109/07853890008995904

Chapman SB, Mudar RA (2014) Enhancement of cognitive and neural functions through complex reasoning training: evidence from normal and clinical populations. Front Syst Neurosci 8:69

*Chapman SB, Highley AP, Thompson JL (1998) Discourse in fluent aphasia and Alzheimer's disease: linguistic and pragmatic considerations. J Neurolinguistics 11:55–78

*Chapman SB, Zientz J, Weiner M, Rosenberg R, Frawley W, Burns MH (2002) Discourse changes in early Alzheimer disease, mild cognitive impairment, and normal aging. Alzheimer Dis Assoc Disord 16:177–186

*Chapman SB, Anand R, Sparks G, Cullum CM (2006) Gist distinctions in healthy cognitive aging versus mild Alzheimer's disease. Brain Impair 7:223–233

Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST, Ganguli M (2001) Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study. Arch Gen Psychiatry 58:853–858

Chenery HJ, Murdoch BE (1994) The production of narrative discourse in response to animations in persons with dementia of the Alzheimer's type: preliminary findings. Aphasiology 8:159–171

Chesneau S, Lepage É, Giroux F, Belleville S (2016) Trouble léger de la cognition: profils variés en compréhension de texte = Mild cognitive impairment: varied texts comprehension profiles. Revue canadienne d'orthophonie et d'audiologie = Can J Speech-Lang Pathol Audiol 40:51–65

Clarke N, Foltz P, Garrard P (2020) How to do things with (thousands of) words: computational approaches to discourse analysis in Alzheimer's disease. Cortex 129:446–463

Cohen G (1979) Language comprehension in old age. Cogn Psychol 11:412–429

*Creamer S, Schmitter-Edgecombe M (2010) Narrative comprehension in Alzheimer's disease: assessing inferences and memory operations with a think-aloud procedure. Neuropsychology 24:279

Cummings JL (2000) Cognitive and behavioral heterogeneity in Alzheimer's disease: seeking the neurobiological basis. Neurobiol Aging 21:845–861. https://doi.org/10.1016/S0197-4580(00)00183-4

Cummings JL, Darkins A, Mendez M, Hill MA, Benson D (1988) Alzheimer's disease and Parkinson's disease: comparison of speech and language alterations. Neurology 38:680–680

Dalton SG, Hubbard HI, Richardson JD (2019) Moving toward non-transcription based discourse analysis in stable and progressive aphasia. In: Seminars in speech and language. Thieme Medical Publishers

Daneman M, Merikle PM (1996) Working memory and language comprehension: a meta-analysis. Psychon Bull Rev 3:422–433

Das N et al (2019) Cognitive training and transcranial direct current stimulation in mild cognitive impairment: a randomized pilot trial. Front Neurosci 13:307

DeKosky S (2003) Early intervention is key to successful management of Alzheimer disease. Alzheimer Dis Assoc Disord 17:S99–S104

DeTure MA, Dickson DW (2019) The neuropathological diagnosis of Alzheimer's disease. Mol Neurodegener 14:32. https://doi.org/10.1186/s13024-019-0333-5

Devanand DP et al (2008) Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. Biol Psychiat 64:871–879

*Drummond C et al (2019) Narrative impairment, white matter damage and CSF biomarkers in the Alzheimer's disease spectrum. Aging 11:9188–9208. https://doi.org/10.18632/aging.102391

Emery VOB (2000) Language impairment in dementia of the Alzheimer type: a hierarchical decline? Int J Psychiatry Med 30:145–164

Fletcher CR, Chrysler ST (1990) Surface forms, textbases, and situation models: recognition memory for three types of textual information. Discourse Process 13:175–190
Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198
Fonseca R, Parente M, Côté H, Ska B, Joanette Y (2008) Introducing a communication assessment tool to Brazilian speech therapists: the MAC battery. Pro-fono 20:285–291
Galton CJ, Patterson K, Xuereb JH, Hodges JR (2000) Atypical and typical presentations of Alzheimer's disease: a clinical, neuropsychological, neuroimaging and pathological study of 13 cases. Brain 123:484–498. https://doi.org/10.1093/brain/123.3.484
Gerwien J, von Stutterheim C (2018) Event segmentation: cross-linguistic differences in verbal and non-verbal tasks. Cognition 180:225–237
Graesser AC, Millis KK, Zwaan RA (1997) Discourse comprehension. Annu Rev Psychol 48:163–189
Graville DJ, Rau MT (1991) Reading comprehension of directly stated and inferred information in paragraph-length material by non-demented and demented elderly subjects clinical. Aphasiology 19:77–86
Haldenwanger A, Eling P, Kastrup A, Hildebrandt H (2010) Correlation between cognitive impairment and CSF biomarkers in amnesic MCI, non-amnesic MCI, and Alzheimer's disease. J Alzheimer's Dis 22:971–980
Henry JD, Crawford JR, Phillips LH (2004) Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. Neuropsychologia 42:1212–1222
Howieson DB, Dame A, Camicioli R, Sexton G, Payami H, Kaye JA (1997) Cognitive markers preceding Alzheimer's dementia in the healthy oldest old. J Am Geriatr Soc 45:584–589
Hughes CP, Berg L, Danziger W, Coben LA, Martin RL (1982) A new clinical scale for the staging of dementia. Br J Psychiatry 140:566–572
Huntley J, Howard R (2010) Working memory in early Alzheimer's disease: a neuropsychological review. Int J Geriatr Psychiatry 25:121–132
Imtiaz B, Tolppanen A-M, Kivipelto M, Soininen H (2014) Future directions in Alzheimer's disease from risk factors to prevention. Biochem Pharmacol 88:661–670
Jack CR et al (2018) NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement 14:535–562
Just MA, Carpenter PA (1992) A capacity theory of comprehension: individual differences in working memory. Psychol Rev 99:122
Kavé G, Levy Y (2003) Morphology in picture descriptions provided by persons with Alzheimer's disease. J Speech Lang Hear Res 46:341–352
Khoury R, Ghossoub E (2019) Diagnostic biomarkers of Alzheimer's disease: a state-of-the-art review. Biomark Neuropsychiatry 1:100005. https://doi.org/10.1016/j.bionps.2019.100005
Kintsch W (1988) The role of knowledge in discourse comprehension: a construction-integration model. Psychol Rev 95:163
Kintsch W, Rawson KA (2005) Comprehension. In: Snowling MJ, Hulme C (eds) The science of reading: a handbook. Blackwell, Oxford, pp 211–226. https://doi.org/10.1002/9780470757642.ch12
Kintsch W, Van Dijk TA (1978) Toward a model of text comprehension and production. Psychol Rev 85:363
Kmet LM, Cook LS, Lee RC (2004) Standard quality assessment criteria for evaluating primary research papers from a variety of fields
Laine M, Laakso M, Vuorinen E, Rinne J (1998) Coherence and informativeness of discourse in two dementia types. J Neurolinguistics 11:79–87
Lim YY et al (2020) Association of deficits in short-term learning and Aβ and hippoampal volume in cognitively normal adults. Neurology 95:e2577–e2585
Livingston G et al (2017) Dementia prevention, intervention, and care. Lancet 390:2673–2734
MacDonald MC, Almor A, Henderson VW, Kempler D, Andersen ES (2001) Assessing working memory and language comprehension in Alzheimer's disease. Brain Lang 78:17–42
Mazzon G et al (2019) Connected speech deficit as an early hallmark of CSF-defined Alzheimer's disease and correlation with cerebral hypoperfusion pattern. Curr Alzheimer Res 16:483–494. https://doi.org/10.2174/1567205016666190506141733
McKhann GM, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34:939–939
McKhann GM et al (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's Dement 7:263–269
Mudar RA, Nguyen LT, Eroh J, Chiang H-S, Rackley A, Chapman SB (2019) Event-related neural oscillation changes following reasoning training in individuals with mild cognitive impairment. Brain Res 1704:229–240
Mueller KD, Hermann B, Mecollari J, Turkstra LS (2018) Connected speech and language in mild cognitive impairment and Alzheimer's disease: a review of picture description tasks. J Clin Exp Neuropsychol 40:917–939. https://doi.org/10.1080/13803395.2018.1446513
Park DC, Reuter-Lorenz P (2009) The adaptive brain: aging and neurocognitive scaffolding. Annu Rev Psychol 60:173–196
Patterson C (2018) World Alzheimer report 2018: the state of the art of dementia research: new frontiers. Alzheimers Disease International (ADI), London
Perneczky R, Wagenpfeil S, Komossa K, Grimmer T, Diehl J, Kurz A (2006) Mapping scores onto stages: mini-mental state examination and clinical dementia rating. Am J Geriatr Psychiatry 14:139–144. https://doi.org/10.1097/01.JGP.0000192478.82189.a8
Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 56:303–308
Petersen RC et al (2001) Current concepts in mild cognitive impairment. Arch Neurol 58:1985–1992
Pistono A, Jucla M, Bezy C, Lemesle B, Le Men J, Pariente J (2019) Discourse macrolinguistic impairment as a marker of linguistic and extralinguistic functions decline in early Alzheimer's disease. Int J Lang Commun Disord 54:390–400. https://doi.org/10.1111/1460-6984.12444
Price BH, Gurvit H, Weintraub S, Geula C, Leimkuhler E, Mesulam M (1993) Neuropsychological patterns and language deficits in 20 consecutive cases of autopsy-confirmed Alzheimer's disease. Arch Neurol 50:931–937
Radvansky GA, Dijkstra K (2007) Aging and situation model processing. Psychon Bull Rev 14:1027–1042
Reilly J, Troche J, Grossman M (2011) Language processing in dementia. In: Budson AE, Kowall NW (eds) The handbook of Alzheimer's disease and other dementias. Blackwell, Oxford, pp 336–368
Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S (2013) Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: a selective review. Alzheimer's Res Ther 5:58. https://doi.org/10.1186/alzrt222
Rochon E, Waters GS, Caplan D (1994) Sentence comprehension in patients with Alzheimer′ s disease. Brain Lang 46:329–349

Sachdev PS, Blacker D, Blazer DG, Ganguli M, Jeste DV, Paulsen JS, Petersen RC (2014) Classifying neurocognitive disorders: the DSM-5 approach. Nat Rev Neurol 10:634

Schmidtke K, Hermeneit S (2008) High rate of conversion to Alzheimer's disease in a cohort of amnestic MCI patients. Int Psychogeriatr 20:96–108. https://doi.org/10.1017/S1041610207005509

*Schmitter-Edgecombe M, Creamer S (2010) Assessment of strategic processing during narrative comprehension in individuals with mild cognitive impairment. J Int Neuropsychol Soc 16:661–671

Slegers A, Filiou R-P, Montembeault M, Brambati SM (2018) Connected speech features from picture description in Alzheimer's disease: a systematic review. J Alzheimers Dis 65:519–542. https://doi.org/10.3233/JAD-170881

Sparks JR (2012) Language/discourse comprehension and understanding. In: Seel NM (ed) Encyclopedia of the sciences of learning. Springer, Boston, pp 1713–1717

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

Stern Y (2013) Cognitive reserve: implications for assessment and intervention. Folia Phoniatr Logop 65:49–54. https://doi.org/10.1159/000353443

Taler V, Phillips NA (2008) Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. J Clin Exp Neuropsychol 30:501–556

Thorndyke PW (1976) The role of inferences in discourse comprehension. J Verbal Learn Verbal Behav 15:437–446

Toledo CM, Aluísio SM, dos Santos LB, Brucki SMD, Trés ES, de Oliveira MO, Mansur LL (2018) Analysis of macrolinguistic aspects of narratives from individuals with Alzheimer's disease, mild cognitive impairment, and no cognitive impairment. Alzheimer's Dement 10:31–40

Tort-Merino A et al (2017) Early detection of learning difficulties when confronted with novel information in preclinical Alzheimer's disease stage 1. J Alzheimer's Dis 58:855–870

Trabasso T, Magliano JP (1996) Conscious understanding during comprehension. Discourse Process 21:255–287

Ulatowska HK, Chapman SB, Highley AP, Prince J (1998) Discourse in healthy old-elderly adults: a longitudinal study. Aphasiology 12:619–633

Ulatowska HK, Chapman S, Johnson J, Branch C (1999) Macrostructure and inferential processing in discourse of aphasic patients. Psychol Lang 3:3–20

Van Dijk TA (2019) Macrostructures: an interdisciplinary study of global structures in discourse, interaction, and cognition. Routledge

Vas AK, Spence J, Chapman SB (2015) Abstracting meaning from complex information (gist reasoning) in adult traumatic brain injury. J Clin Exp Neuropsychol 37:152–161

Vas AK, Spence JS, Eschler B, Chapman SB (2016) Sensitivity and specificity of abstraction using gist reasoning measure in adults with traumatic brain injury. J Appl Biobehav Res 21:216–224

Verma M, Howard RJ (2012) Semantic memory and language dysfunction in early Alzheimer's disease: a review. Int J Geriatr Psychiatry 27:1209–1217. https://doi.org/10.1002/gps.3766

Villemagne VL et al (2011) Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. Ann Neurol 69:181–192. https://doi.org/10.1002/ana.22248

Vuorinen E, Laine M, Rinne J (2000) Common pattern of language impairment in vascular dementia and in Alzheimer disease. Alzheimer Dis Assoc Disord 14:81–86

Ward A, Tardiff S, Dye C, Arrighi HM (2013) Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a systematic review of the literature. Dement Geriatr Cogn Disord Extra 3:320–332. https://doi.org/10.1159/000354370

*Welland RJ, Lubinski R, Higginbotham DJ (2002) Discourse comprehension test performance of elders with dementia of the Alzheimer type. J Speech Lang Hear Res 45:1175–1187

Winblad B et al (2004) Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 256:240–246. https://doi.org/10.1111/j.1365-2796.2004.01380.x

Xu W et al (2015) Meta-analysis of modifiable risk factors for Alzheimer's disease. J Neurol Neurosurg Psychiatry 86:1299–1306

3. Chapter 3: Does Bilingualism Influence Neuropsychological Test Performance in Older Adults? A Systematic Review (2nd Publication)

Simge Celik[a], Eesha Kokje[a], Patric Meyer[a,b], Lutz Frölich [a,c] & Birgit Teichmann[a]

[a] *Network Aging Research, Heidelberg University, Heidelberg, Germany*

[b] *Faculty of Applied Psychology, SRH University Heidelberg, Heidelberg, Germany*

[c] *Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany*

Celik, S., Kokje, E., Meyer, P., Frölich, L., & Teichmann, B. (2022). Does bilingualism influence neuropsychological test performance in older adults? A systematic review. *Applied neuropsychology*. Adult, 29(4), 855–873. https://doi.org/10.1080/23279095.2020.1788032

APPLIED NEUROPSYCHOLOGY: ADULT
2022, VOL. 29, NO. 4, 855–873
https://doi.org/10.1080/23279095.2020.1788032

Routledge
Taylor & Francis Group

Does bilingualism influence neuropsychological test performance in older adults? A systematic review

Simge Celik[a], Eesha Kokje[a], Patric Meyer[a,b], Lutz Frölich[a,c], and Birgit Teichmann[a]

[a]Network Aging Research, Heidelberg University, Heidelberg, Germany; [b]Faculty of Applied Psychology, SRH University Heidelberg, Heidelberg, Germany; [c]Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany

ABSTRACT

Objective: Using standardized tests which have been normed on monolinguals for the assessment of bilinguals presents challenges to the accurate characterization of cognitive profile as the literature provides compelling evidence for the influence of bilingualism on cognitive abilities. However, little is known about the generalizability of these findings to clinical neuropsychology. The aim of this review was to address this gap by summarizing current evidence on the performance of bilingual older adults on standardized tests routinely used in clinical practice.

Method: A systematic search of Web of Science, PsycINFO and PubMed was conducted. 27 cross-sectional and longitudinal studies which use at least one standardized neuropsychological test for cognitive impairment were included in the review. Potential demographic (cultural/linguistic background of the participants, immigrant status), clinical (diagnostic status), and methodological confounders (language of test administration, components of bilingualism) were also examined. The review protocol was registered at the PROSPERO International Prospective Register of Systematic Review with registration number CRD42018114658.

Results: The results of this review revealed some bilingual advantage on measures of inhibitory control and bilingual disadvantage on measures of verbal fluency in cross-sectional studies. Bilingualism status was not associated with test performance in longitudinal studies. However, findings lack consistency due to demographic variables and methodological differences across studies.

Conclusion: Neuropsychological tests assessing language domains and, to some extent executive function act as clinically relevant features of bilingualism for neuropsychological evaluation. However, immigration status, acculturation level and language of test administration needs to be taken into account when assessing bilingual older adults.

KEYWORDS
Aging; cognition; dementia; language; neuropsychological assessment

Introduction

The accuracy of clinical diagnosis of dementia and determination of its underlying cause play a crucial role in prognosis and early, suitable and effective application of disease-specific treatments (Salmon & Bondi, 2009). Neuropsychological assessment, a fundamental tool used in the diagnosis of mild cognitive impairment (MCI) and dementia, is considered to be primary means of guiding differential diagnosis, monitoring of the disease state and measuring post-treatment outcome (Casaletto & Heaton, 2017; Gracey & Morris, 2007). Normative comparison as a critical concept in neuropsychological assessment allows for more accurate detection of disease-related changes in cognitive function since it enables a comparison between the performance of an individual and reference groups of the same age, gender, ethnicity and educational attainment (Harvey, 2012). Due to the significant independent effects of these demographic factors on test performance, the extant literature provides demographically adjusted normative data in a variety of neuropsychological tests (Casaletto et al., 2015; Heaton et al., 2009; Norman et al., 2000). However, demographic characteristics of the population in Europe and the United States have changed over the years because of increased migration. This has resulted in increased diversity in the population, in terms of ethnicity and spoken languages (Llorente et al., 1999).

Employing standardized tests, which have been normed on monolingual populations for the assessment of bilinguals, presents challenges to clinical diagnostic processes (Gasquoine & Gonzalez, 2012). Bilinguals' possible advantages over monolinguals in executive tasks, and disadvantages in language-based tasks (Mindt et al., 2008) may produce biased results. Specifically, the cognitive advantages associated with bilingualism has been shown as smaller costs in task switching (Prior & MacWhinney, 2010) and smaller conflict effects (Bialystok et al., 2004; Coderre et al., 2013)

CONTACT Simge Celik celik@nar.uni-heidelberg.de Network Aging Research, Heidelberg University, Bergheimer Strasse 20, Heidelberg 69115, Germany.
Supplemental data for this article can be accessed at publisher's website.
This article has been republished with minor changes. These changes do not impact the academic content of the article.

across a variety of executive functioning experimental tasks and the Stroop task in populations ranging from children (Bialystok & Feng, 2009) to young adults, with the most robust advantage reported in older adults (Bialystok et al., 2012). On the other hand, cognitive disadvantages in relation to bilingualism has been indicated as reduced category fluency (Portocarrero et al., 2007), poorer naming performance (Roberts et al., 2002) and retrieval failures (Golan & Brown, 2006) with majority of these studies focusing on young adults.

Bilingualism may influence neuropsychological test performance (NP) beyond what is estimated by normative corrections for age, education, sex and ethnicity (Suarez et al., 2014). Thus, it is imperative to understand the impact of bilingualism on NP for the effective interpretation of test results, in order to determine whether obtained scores truly reflect individual's cognitive abilities, or the effects of life experiences such as bilingualism, particularly given the lack of normative data for this population (Mindt et al., 2008).

The role of confounding variables

The findings in studies of cognition and bilingualism are controversial and inconclusive due to large number of potential confounding variables stated in the extant literature (Bak, 2016)

Language proficiency constitutes one of the critical confounding variables in this field as it provides implications for operationalization of language dominance (Hulstijn, 2012) and covers multiple aspects of language experience for its definition, such as vocabulary size, language use in a specific pattern or age of acquisition (AoA). The impact of language proficiency on cognition was reported in the Stroop task, with its modulating effect on inhibition (Singh & Mishra, 2013), conflict resolution and goal maintenance (Tse & Altarriba, 2012)

AoA is considered as a factor modulating language proficiency in studies assessing the relation between bilingualism and cognition. However, its role in the neural organization of the bilingual brain has been more commonly investigated, with findings showing a differential pattern of brain activation associated with lexical retrieval (Perani et al., 2003) and syntax (Mahendra et al., 2003) in early language acquisition.

Variations in language proficiency and amount of exposure to the language (e.g., AoA) play a crucial role in determining the language of testing in neuropsychological evaluation of bilingual individuals. The impact of testing language on NP has been consistently shown on neuropsychological tests requiring significant language demands, namely, verbal fluency (Boone et al., 2007; Kisser et al., 2012), Boston Naming Test (BNT) and Digit Span (Boone et al., 2007) in a sample of culturally diverse young and middle-aged adults comparing native versus non-native English speakers. In the study by Kisser et al. (2012), native language status was additionally associated with a better performance on the Trail-Making Test (TMT-A). However, the associations between native language status and letter fluency and TMT-A did not remain significant with the inclusion of ethnicity as a covariate in the analyses. In another study, which included a more culturally homogenous group of Spanish/English bilingual adults (Gasquoine et al., 2007), no significant differences were found in test scores between Spanish and English language administrations in balanced bilinguals. However, the impact of language of test administration was observed on tests that exert a higher demand on language domain, such as, letter fluency, Woodcock-Munoz Language Survey-Revised (WMLS-R), Stroop Test and Story Memory in Spanish and English dominant bilinguals Therefore, a critical understanding regarding language of test administration is needed to interpret the differences in NP for both clinical and research settings.

Immigration status plays a distinct role in bilingualism research as prolonged contact between culturally dissimilar groups or members not only involves acquisition of another language but also leads to cultural, psychological and cognitive changes in terms of adapting to the cognitive styles of the host culture, a process referred to as acculturation (Berry, 2007; Park & Huang, 2010).The findings regarding immigration status and cognition are mixed. Regardless of bilingualism, some studies showed improved cognitive function and slower rate of cognitive decline in older adult immigrants, (Hill et al., 2012). In contrast, others studies reported either no significant differences in cognitive function and decline (Sheffield & Peek, 2009) or poorer performance on measures of abstract reasoning, verbal fluency, and naming in immigrant groups (Touradji et al., 2001). Hence, immigration status is a significant background variable which may elucidate some differences on NP attributed to bilingualism.

In light of these findings, language proficiency, AoA, language of test administration and immigration status have all been explored as possible modifiers of NP in bilingual older adults. The influence of education and socioeconomic status on NP was not addressed in relation to findings as they were beyond the scope of this review.

Objectives of the current review

The aim of the present work is to summarize current evidence on the performance of healthy bilingual older adults or bilinguals with cognitive decline due to dementia on standardized neuropsychological assessments routinely used in clinical practice. Specifically, this systematic review addresses whether (1) bilingual older adults display advantages or disadvantages in NP, (2) the advantage/disadvantage is found in specific cognitive domains and (3) the findings are influenced by language proficiency, AoA, immigration status and language of test administration.

Method

Search strategy and study selection

A systematic review of the literature was conducted using the following databases: Web of Science, PsycINFO and

PubMed. The search terms included were neuropsychological assessment, neuropsychology, cognitive tests, cognitive assessment, cognitive performance combined with second language, language proficiency, bilingualism, trilingualism and multilingualism. Using these search terms, published peer reviewed articles which were relevant to this review were identified and collected for further review. No restrictions on study design or date of publication were applied. The final search was carried out on March 24th, 2020. Additionally, manual searches were conducted in which reference lists of reviews and other articles were examined in order to identify relevant studies not detected by database searches. This review was conducted in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) and it was registered with the PROSPERO International Prospective Register of Systematic Review (Registration No: CRD42018114658).

Inclusion and exclusion criteria

Cross-sectional and longitudinal studies which use at least one standardized neuropsychological assessment routinely used in the diagnosis of dementia or cognitive impairment were included. Neuropsychological tools were identified in this review as brief screening measures, psychometric instruments, intelligence tests or computer-based cognitive assessments. The criteria used for selection of the studies were as follows: (1) the study investigated the impact of bilingualism on NP in older adults, (2) the study administered standardized neuropsychological tests or cognitive screening instruments which are commonly used in clinical practice, (3) the study included a group of cognitively healthy older adults, or individuals with MCI or any type of dementia, with a mean age of 60 years and older, (4) the study included a monolingual group for comparison. Studies were excluded if: (1) samples studied included comorbid neurological or psychiatric conditions which did not allow clear differentiation of findings by suspected etiology, (2) a modified version of the neuropsychological tests or a novel test was used, (3) conference abstracts, reviews, case studies, or PhD dissertations, (4) studies were published in a language other than English, or (5) studies aiming to provide normative data in bilingual older adults.

The abstracts were screened for eligibility and potentially relevant articles were reviewed in full via Covidence by two independent reviewers, namely, doctoral students in psychology (SC, EK). Disagreements between the two independent reviewers were discussed and resolved or a third reviewer (BT; PhD in biology) was consulted.

Data extraction

Data extraction process was guided by an extraction table designed by the first author tailored to the aims of this review. The extracted data included information on geographical region, study design, participant characteristics (sample sizes for all groups, age, immigration status of participants, diagnostic criteria used), the neuropsychological tool/s used, the language of test administration, results and conclusions. With regard to the outcome measures, only measures of cognition based on standardized tests were extracted from eligible studies and measures of cognition based on experimental tasks were excluded from this review. Additionally, in order to examine the extant literature comprehensively, in studies that included both younger and older adults, only data of older adults was included in the analysis.

Quality assessment

The methodological quality of the studies was evaluated by two reviewers (SC, EK) using a modified version of the Standard Quality Assessment Criteria for Evaluating Primary Research Papers: Quality Scoring for Quantitative Studies or "QualSyst" (Kmet et al., 2004). The Qualsyst tool consists of 14 items, of which, three items regarding interventional studies were removed as they were irrelevant to included studies. Additionally, one item related to adjustment of potential confounders was eliminated. A large number of potential confounding variables reported in this research area did not allow their effects to be accounted for in each study (Bak, 2016). The remaining 10 items assess studies based on objective, design, subject selection, subject characteristics, outcome measures, sample size, analytic methods, estimate of variance, results and conclusion supported by the results. Each item was scored as 2 (fully met the quality criterion), 1 (partially met the quality criterion) or as 0 (did not meet the quality criterion). A summary score was calculated for each study by dividing the obtained total score by the maximum possible score. As defined by Lee et al. (2008), quality of the studies was categorized as: limited (<0.50), adequate (0.50–0.70), good (0.71–0.79) or high (>0.80). The studies were not excluded on the basis of their score on the quality checklist, rather, this tool was used to identify strengths and weaknesses of the current literature and provide recommendations for future research.

Results

The search strategy identified 2,553 citations from the three electronic databases and 7 citations from hand searching. After adjusting for duplicates, 2,168 citations remained in the pool of papers. After screening titles and abstracts, 85 citations were potentially eligible and full reports were retrieved and analyzed. 58 studies were excluded based on a detailed examination of the full text. A total of 27 studies were identified and included in the review. The study selection process is summarized in Figure 1.

Study characteristics

Of the 27 included studies, 18 were cross-sectional and 9 were longitudinal studies. Tables 1 and 2 provide an overview of the main findings on bilingualism and its impact on the NP based on the study design.

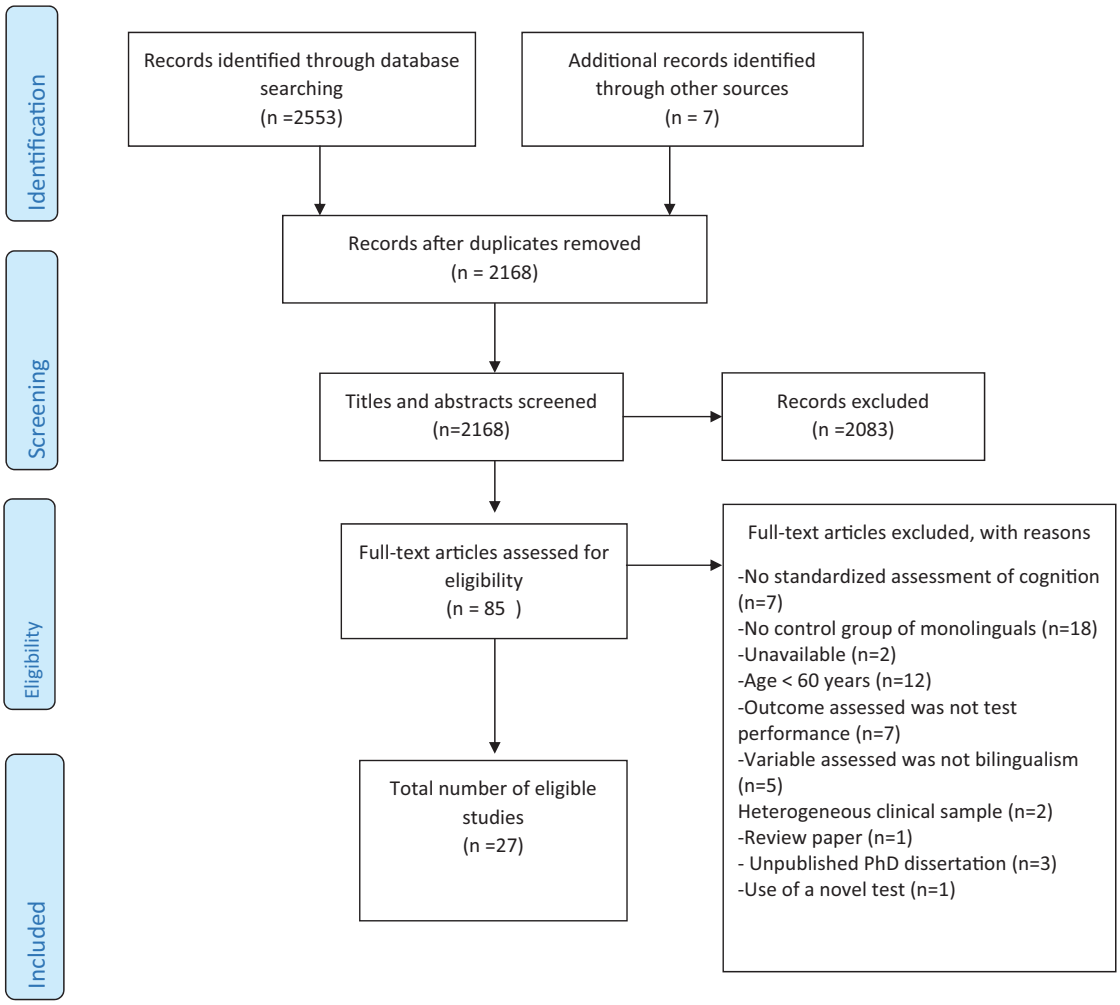


Figure 1. Flowchart of literature search and study selection process.

Participants in cross-sectional studies were recruited mainly from community settings (Anderson et al., 2017; Clare, Whitaker, Martyr, et al., 2016; Kousaie et al., 2014; Luo et al., 2013; Nielsen et al., 2019; Ossher et al., 2013; Papageorgiou et al., 2019; Rosselli et al., 2000; Sheppard et al., 2016; Soltani et al., 2019), participant pools (Bialystok et al., 2008; Bialystok, Poarch, et al., 2014; Friesen et al., 2015; Ihle et al., 2016) and referrals to specialist memory clinics (Clare, Whitaker, Craik, et al., 2016; Kowoll et al., 2015; Rosselli et al., 2019). Sample sizes ranged from 24 to 2812. Additionally, of the eighteen cross-sectional studies, eleven comprised only of cognitively healthy older adults, one study included only patients with AD (Clare, Whitaker, Craik, et al., 2016) and two other studies included participants with amnestic MCI (aMCI; Ossher et al., 2013; Padilla et al., 2016; Rosselli et al., 2019). In longitudinal studies, participants were included from referrals to specialist memory clinics (Bialystok, Craik, et al., 2014; Chertkow et al., 2010; Costumero et al., 2020), community samples (Yeung et al., 2014; Zahodne et al., 2014) or individuals who were part of on-going studies (Bak et al., 2014; Cox et al., 2016; Padilla et al., 2016; Mungas et al., 2018). Sample sizes in longitudinal studies ranged from 90 to 1459. Additionally, the samples predominantly included MCI and dementia patients, with only one study focusing solely on healthy older adults (Cox et al., 2016). The follow-up duration in the longitudinal studies ranged from 6 months to 63 years.

Two studies did not report on the cognitive status of participants (Bialystok et al., 2008; Ihle et al., 2016). Among twelve cross-sectional and longitudinal studies that included cognitively healthy older adults, three studies provide no information on the instruments used for cognitive screening, but reported that the participants were cognitive healthy (Bialystok, Poarch, et al., 2014; Kowoll et al., 2015; Luo et al., 2013). The absence of neurological and psychiatric condition was the criteria used in some of the studies (Anderson et al., 2017; Papageorgiou et al., 2019; Sheppard et al., 2016). Other studies reported on the instruments they employed to test for cognitive status of the participants, such as, the Mini Mental State Examination (MMSE; Folstein et al., 1975), Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), Modified MMSE (3MS; Teng & Chui, 1987), Beck Depression Inventory-II (BD-II; Beck et al., 1996) and Hospital Anxiety and Depression

Table 1. Summary of findings of cross-sectional studies.

First author, year	Neuropsychological tests	Controlled for	Main findings
Anderson et al. (2017)	• D-KEFS	• Age • Education (years) • Cognitive status (MMSE)	The results indicate that BL[a] healthy group had poorer performance than their ML[b] counterparts on the Stroop and category fluency subtests. BLs with AD[c] had lower scores in letter fluency task and reduction in inhibition/switching errors on the Stroop task than their ML peers. Despite a lower baseline performance, BLs showed less decline in cognitive scores from HA[d] to MCI[e] than their monolingual counterparts.
Bialystok et al. (2008)	• Corsi blocks (Forward and Backward) • PPVT-III, Form B • CVF and LVF (Milner, 1964)	• Age of English acquisition • Vocabulary knowledge	BLs and MLs performed equivalently on the Corsi block task. The results from lexical access tasks indicated consistent monolingual advantages.
Bialystok, Poarch, et al. (2014)	Study 1 • Shipley Vocabulary Test • Cattell Culture Fair Test • Stroop Task Study 2 • Shipley Vocabulary Test • Cattell Culture Fair Test	———	Study 1 BLs performed poorer than MLs on the Shipley Vocabulary Test and BLs had significantly faster completion times and less interference than MLs in the Stroop Task. Study 2 BLs obtained lower scores than MLs on the Cattell Culture Fair Test and Shipley Test.
Clare, Whitaker, Craik, et al. (2016)	• MMSE • NART • Spot-the-Word Test • BPVS • BNT • D-KEFS • WMS-III (Spatial and Digit Span) • TEA (Elevator Counting with Distraction) • Stroop	• Educational level • Use of acetylcholinesterase-inhibiting medication (AChEI) • MMSE scores at time of diagnosis	The findings show that BLs performed worse than MLs on the MMSE at the time of diagnosis and BLs did not demonstrate a clear advantage in EF[f], but they may have some relative advantage compared to MLs in the domain of inhibition and management of response conflict (TEA).
Clare, Whitaker, Martyr, et al. (2016)	• NART • Spot-the-Word Test • BPVS • BNT • D-KEFS • WMS-III (Spatial and digit span) • TEA (Elevator counting with distraction, Visual Elevator) • Stroop	• Age • Gender • Educational level • SES • Health status • Functional ability • Mood • Cognitive status (MMSE) • Word reading ability • Cognitive reserve	The findings show that there were no overall differences in EF tasks performance between MLs and BLs. For each measure, there was a trend for MLs to have a better performance, which was less pronounced in the domain of inhibition and management of response conflict.
Friesen et al. (2015)	• Shipley Vocabulary Test • D-KEFS (VF)	• WCST performance • Vocabulary knowledge	Category fluency performance showed decline in older adults and was not influenced by bilingualism status, whereas, better letter fluency performance was associated with bilingualism status.
Ihle et al. (2016)	• MHVS • TMT (A and B)	• Language of test administration • Educational level • Number of leisure activities • Occupational complexity • Gainful activity	Higher number of languages spoken was associated with better performance in verbal abilities and processing speed, but not after considering educational level and cognitive level of job.
Kousaie et al. (2014)	• Stroop • WMS-III (Forward and Backward Digit Span • WCST • BNT • COWAT (CVF and LVF)	• Age • Education (years) • Cognitive status (MoCA)	The results do not support a bilingual advantage on the EF tasks and no bilingual disadvantage was observed in language tasks. A small bilingual advantage was observed only in the Stroop interference, with BLs showing smaller interference than both ML groups.
Kowoll et al. (2015)	• CERAD	• Gender • Depression	The results do not indicate significant differences in cognitive performance between MLs and BLs across diagnostic groups.

(*continued*)

Table 1. Continued.

First author, year	Neuropsychological tests	Controlled for	Main findings
		• German as a non-dominant language	
Luo et al. (2013)	• Shipley Vocabulary Test • Cattell Culture Fair Test • Corsi Block Test	• Education (years) • English vocabulary level • Nonverbal intelligence	The results reveal that bilingualism was associated with better performance in forward and backward conditions of Corsi Block Test.
Massa et al. (2020)	• Stroop • TMT (A and B) • VF	• Age • Education (years) • Cognitive status (MMSE)	The findings show higher number of correct responses in the Stroop test and higher number of words in the letter fluency for BLs as compared to MLs.
Nielsen et al. (2019)	• CTT • FDT • MMSE (language items) • VF (Semantic) • Copying of a Greek cross and star • CRT	• Education • Gender • Ethnicity • Years of residence in a new country	After controlling for the covariates, bilingualism was associated with a better performance in executive functions and task switching ability.
Ossher et al. (2013)	• MMSE • D-KEFS (Color-Word Naming) • VF	• Education (years) • Gender	The findings indicate no significant difference between language groups in each aMCI[g] subgroup on the MMSE and D-KEFS except for the slower completion time on the color-word naming condition in bilingual multiple domain aMCI group.
Papageorgiou et al. (2019)	Background measures • BPVS-III • RPMT • WAIS-IV (Digit span backward and forward) Main measure • TOL	• Gender • Age • Educational level • SES	The results do not indicate an effect of bilingualism on the TOL and there was a trend toward a bilingual disadvantage in response times in this test.
Rosselli et al. (2000)	• LVF • CVF (fruits and animals) • BDAE (The Cookie Theft) • MAE (Sentence Repetition)	• Age • Education (years) • Cognitive status (MMSE)	The results found no protective or detrimental effect of bilingualism on language test performance except for the lower scores in the semantic fluency task in BLs.
Rosselli et al. (2019)	Background measures • WAIS-R (Block Design) • MINT • TMT-A • Stroop Main measure • ROCF (Benson Figure)	• Age • Gender • Education (years) • Cognitive status (MMSE and MOCA)	The results indicate no differences in performance between MLs and BLs on the Benson Figure Test.
Sheppard et al. (2016)	• BNT	• Age • Education (years)	BLs performed better in the language of their choice administration than in the English-only administration, with lowest performance in the French-only administration.
Soltani et al. (2019)	• VF	———	The results showed no differences in verbal fluency performance between MLs and BLs, when BLs were tested in their second and most frequently used language.

D-KEFS: Delis-Kaplan Executive Function System; PPVT: Peabody Picture Vocabulary Test; CVF: category verbal fluency; LVF: letter verbal fluency; MMSE: Mini Mental State Exam; NART: National Adult Reading Test; BPVS: British Picture Vocabulary Scale; BNT: Boston Naming Test; TEA: Test of Everyday Attention; WMS: Wechsler Memory Scale; TMT: Trail Making Test; Color-CTT: Trails Test; FDT: Five Digit Test; CRT: Clock-Reading Test; MHVS: Mill Hill Vocabulary Scale; WCST: Wisconsin Card Sorting Test; COWAT: Controlled Oral Word Association Test; MoCA: Montreal Cognitive Assessment; CERAD: Consortium to Establish a Registry for Alzheimer's Disease; CTT: Color Trails Test; FDT: Five Digit Test; CRT: Clock Reading Test; VF: Verbal Fluency; WAIS: Wechsler Adult Intelligence Scale; RPMT: Raven's Progressive Matrices Task, TOL: Tower of London; BDAE: Boston Diagnostic Aphasia Examination; MAE: Multilingual Aphasia Examination; MINT: Multilingual Naming Test; ROCF: Rey–Osterrieth Complex Figure Test.
[a]Bilinguals; [b]monolinguals; [c]Alzheimer's disease; [4]healthy older adults; [e]mild cognitive impairment; [f]executive function; [g]amnestic mild cognitive impairment.

Table 2. Summary of findings of longitudinal studies.

First author, year	Neuropsychological tests	Controlled for	Baseline findings	Longitudinal findings
Bak et al. (2014)	• WAIS-III (Letter–Number Sequencing, Matrix Reasoning, Block Design, Digit Symbol and Symbol Search) • WMS-III (Logical Memory, Spatial Span, Verbal Paired Associates, Digit Span Backward) • MHT • NART • LVF	• Age at time of testing • Gender • Socio-economic class (participants' and their fathers)	———	The results of the cohort who had been first tested in 1947 and subsequently between 2008 and 2010 suggest a protective effect of bilingualism on the NART in BLs[a]. As compared to bilingualism, multilingualism showed a greater advantage in the NART, general intelligence and verbal fluency, independent of childhood intelligence, even if the second language is acquired in adulthood.
Bialystok, Craik, et al. (2014)	• MMSE • BNA • D-KEFS (TMT, CWIT, VF)	• Lifestyle factors • Immigration • Cognitive status (MMSE)	BLs had poorer performance than MLs[b] on the MMSE and BNA, however, BLs were significantly older than MLs. First session BLs in the MCI[d] group performed slower on the TMT-B, with no language difference in the AD[c] group. There was no difference between MLs and BLs on the CIWT, but BLs showed a smaller Stroop effect than MLs in both diagnostic groups.	After 6.2 months, the results revealed that BLs and MLs had comparable declining scores through disease progression on the subtests of D-KEFS.
Chertkow et al. (2010)	• MMSE	• Gender • Immigrant/native status • Education (years)	The results showed that rate of cognitive decline in MMSE scores were similar between language groups at the diagnostic visit.	The performance in one-year follow-up visit was comparable between a subset of multilingual and ML patients with AD.
Costumero et al. (2020)	• MMSE • BNT • VF • CDT	• Age • Gender • Education (years) • Immigration • Cultural background • Cognitive status (MMSE)	MLs and BLs with aMCI were matched on neuropsychological test results.	A follow-up of 7 months revealed a greater a decline on the MMSE and letter fluency task for MLs with aMCI when compared to BLs with aMCI.
Cox et al. (2016)	• D-KEFS (Tower Test) • MHT	• Age • MHT (IQ at age 11) • Social class (participants' and their fathers') • L2 acquisition before age 11	———	The results did not indicate an effect of bilingualism on the Tower Test over the 63-year period.
Mungas et al. (2018)	• 3MS • SEVLT	• Age • Gender • Education (years)	When language use and proficiency included in the analysis, Spanish MLs had poorer performance on the 3MS, but there was no significant differences in	The results revealed that bilingualism status did not have an impact on the rate of cognitive decline over six years.

(*continued*)

			the 3MS and SEVLT between English MLs and BLs.	
Padilla et al. (2016)	• 3MS • SEVLT	• Age • Gender • Education (years) • Household income • Depression	The results revealed a bilingual advantage on the 3MS, particularly on the items of language, praxis, EF. Language/executive and language/praxis items distinguished between MLs and BLs.	With or without adjustment for demographic variables and depressive symptoms, there were no significant differences in rates of cognitive decline on the 3MS and SEVLT learning trials between MLs and BLs over approximately 6 years.
Yeung et al. (2014)	• 3MS	• Age • Education (years) • Gender • Subjective memory loss • Literacy	TIME 1 and TIME 2 The results indicated that the group with English as a L2 had a lower performance on the 3MS at baseline and follow-up of 5 years than those whose spoke English as a L1.	FROM TIME 1 to TIME 2 Bilingualism was not associated with a change in cognitive test scores over the 5-year period.
Zahodne et al. (2014)	• SRT • BNT • BDAE • WAIS-R (Similarities) • MDRS (Identities and Oddities, LVF, CTT)	• Country of origin • Gender • Education (years) • Amount of time spent in the USA • Recruitment wave	The findings suggest that independent of covariates, higher bilingualism scores were associated with better baseline scores on the measures of EF[e] (WAIS-R, MDRS), the task switching (CTT) and episodic memory (SRT).	Bilingualism status did not appear to have an independent protective effect against age-related cognitive decline in a follow-up of 23 years.

WAIS: Wechsler Adult Intelligence Scale; WMS-III: Wechsler Memory Scale; MHT: Moray House Test; NART: National Adult Reading Test; LVF: Letter verbal fluency; MMSE: Mini Mental State Examination; BNA: Behavioral Neurology Assessment; D-KEFS: Delis-Kaplan Executive Function System; TMT: Trail Making Test; CWIT: Color-Word Interference Test; VF: Verbal Fluency Test; CDT: Clock Drawing Test; 3MS: Modified Mini-Mental State Examination; SEVLT: Spanish English Verbal Learning Test; SRT: Selective Reminding Test; BNT: Boston Naming Test; BDAE: Boston Diagnostic Aphasia Examination; WAIS-R: Wechsler Adult Intelligence Scale-Revised; MDRS: Mattis Dementia Rating Scale; CTT: Color Trails Test.

———— indicates no data available; [a]bilinguals; [b]monolinguals; [c]Alzheimer's disease; [d]mild cognitive impairment; [e]executive function.

Scale (Zigmond & Snaith, 1983) scores. Nine of the eleven studies that included clinical samples reported detailed information about diagnostic criteria. The diagnostic criteria used in the studies, study characteristics and quality rating scores are summarized in Supplementary Tables 1 and 2.

Different definitions of bilingualism used in the studies

The definition for bilingualism was found to vary considerably in the studies. Specifically, five studies (Anderson et al., 2017; Bialystok, Craik, et al., 2014; Chertkow et al., 2010; Kowoll et al., 2015; Ossher et al., 2013) defined bilingualism based on the criteria from Bialystok et al. (2007) namely, bilinguals were defined as "individuals who spent the majority of their lives, at least from early adulthood, speaking two or more languages fluently". Eighteen studies did not define the criteria for bilingualism in sufficient detail (Bak et al., 2014; Bialystok et al., 2008; Bialystok, Poarch, et al., 2014; Costumero et al., 2020; Cox et al., 2016; Friesen et al., 2015; Massa et al., 2020; Mungas et al., 2018; Nielsen et al., 2019; Ihle et al., 2016; Padilla et al., 2016; Papageorgiou et al., 2019; Rosselli et al., 2000, 2019; Sheppard et al., 2016; Soltani et al., 2019; Yeung et al., 2014; Zahodne et al., 2014). Of those, six studies (Bialystok et al., 2008; Bialystok, Poarch, et al., 2014; Friesen et al., 2015; Massa et al., 2020; Sheppard et al., 2016; Soltani et al., 2019) did not specify their criteria for bilingualism. In general, the definitions used for bilingualism varied based on the goal of the study and the bilingual population that were being tested. Bak et al. (2014) and Cox et al. (2016) referred to bilingualism as being able speak in a second language, Yeung et al. (2014) and Costumero et al. (2020) described bilingual groups based on the participants' first language (e.g., English bilinguals, for a group whose first language was English or bilinguals, for a group whose native language is Catalan), whereas Ihle et al. (2016) identified bilingual or multilingual groups based on the number of languages participants spoke. Three study identified bilingualism based on the frequency of second language use (Mungas et al., 2018; Padilla et al., 2016; Papageorgiou et al., 2019), two based on the level of proficiency (Rosselli et al., 2000, 2019), and two other based exclusively on the second language proficiency, instead of proficiency in each language (Nielsen et al., 2019; Zahodne et al., 2014). The remaining four studies (Clare, Whitaker, Craik, et al., 2016; Clare, Whitaker, Martyr, et al., 2016; Kousaie et al., 2014; Luo et al., 2013) used more specific definitions of the bilingual groups showing some similarities with the definition used by Bialystok et al. (2007).

Factors influencing neuropsychological test performance in bilinguals

Language proficiency. The level of proficiency reported, and tools used to measure language proficiency differ considerably among studies. In sixteen studies, proficiency was only measured through self-report questionnaires wherein participants rated their proficiency in the domains of reading, writing, comprehension and speaking in either one or both languages, based on objective or subjective measures. Eight studies did not report proficiency level. Four studies conducted interviews to assess participants' language history (Chertkow et al., 2010; Costumero et al., 2020; Nielsen et al., 2019; Yeung et al., 2014) and one study (Chertkow et al., 2010) categorized participants based on the number of languages spoken by participants instead of language proficiency. In seven studies, proficiency was determined by formal assessment in addition to the self-report questionnaire. Of these, five studies administered the measures to both bilingual and monolingual groups. Two studies (Kousaie et al., 2014; Sheppard et al., 2016) applied Animacy Judgment Task (Segalowitz & Frenkiel-Fishman, 2005), one study (Papageorgiou et al., 2019) employed British Picture Vocabulary Scale-III (Dunn & Dunn, 1997) and one other study (Massa et al., 2020) used verbal fluency task to both monolingual and bilingual groups. Additionally, one study included BNT as a measure of bilingualism (Rosselli et al., 2000) for both bilingual and monolingual groups. The other two studies only assessed proficiency for the bilingual group, which included a combination of BNT and semantic verbal fluency task in one study (Kowoll et al., 2015), and the Wide Range Achievement Test-Version 3 (WRAT-3) in the other (Zahodne et al., 2014).

The variation in measuring language proficiency makes comparisons between studies difficult. Seven studies which included highly proficient bilinguals reported an association between bilingualism and NP, with four studies showing a bilingual advantage in inhibition domain, namely, on the Stroop task (Anderson et al., 2017; Bialystok, Poarch, et al., 2014; Kousaie et al., 2014), and the elevator counting with distraction subtest of the Test of Everyday Attention (TEA; Clare, Whitaker, Craik, et al., 2016). The remaining three studies showed better NP on the Corsi Block Test (Luo et al., 2013), letter fluency (Friesen et al., 2015), and Color Trails and Five Digit Test (Nielsen et al., 2019). Additionally, a study (Zahodne et al., 2014) demonstrated a better performance on executive function, memory and language composite of neuropsychological tests for an increase in self-reported bilingualism. However, other studies which reported high level of language proficiency in bilingual participants did not show any significant differences in neuropsychological test scores between monolinguals and bilinguals (Kowoll et al., 2015; Papageorgiou et al., 2019). One study which did not apply strict proficiency criteria found no significant differences in NP of bilinguals and monolinguals (Padilla et al., 2016).

Age of second language acquisition. AoA was defined differently in the studies. It was identified as the age of learning a language (Bak et al., 2014; Costumero et al., 2020; Cox et al., 2016; Rosselli et al., 2019; Sheppard et al., 2016; Yeung et al., 2014), the age of active use of a language (Anderson et al., 2017; Bialystok, Craik, et al., 2014; Clare, Whitaker, Craik, et al., 2016; Whitaker, Martyr, Clare et al., 2016; Luo et al., 2013), as the age of exposure to the second language (Papageorgiou et al., 2019; Rosselli et al., 2000) or the age of arrival in a new country (Bialystok et al., 2008;

Zahodne et al., 2014) in studies which included immigrants. Information on AoA was not obtained from language history questionnaires or was not reported in twelve studies (Anderson et al., 2017; Bialystok, Craik, et al., 2014; Chertkow et al., 2010; Costumero et al., 2020; Ihle et al., 2016; Kousaie et al., 2014; Mungas et al., 2018; Nielsen et al., 2019; Ossher et al., 2013; Padilla et al., 2016; Soltani et al., 2019; Zahodne et al., 2014).

Additionally, cutoff ages used to distinguish early and late language acquisition varied between studies, depending on the characteristics of included participants, with some studies using 11 (Clare, Whitaker, Craik, et al., 2016), 12 (Rosselli et al., 2000), 18 (Bak et al., 2014), and 8 versus 12 or older (Bialystok et al., 2008) as cutoff points. This confounds the interpretation of outcome comparisons across studies.

Four studies included AoA in their analysis. Of these, the study by Bak et al. (2014) showed an effect of early AoA on the tests for general intelligence and reading (NART) and of late AoA on the tests of general intelligence, processing speed (Symbol Search and Digit Symbol subtests of WAIS-III) and reading. On the other hand, the studies conducted by Bialystok et al. (2008), Papageorgiou et al. (2019) and Rosselli et al. (2000) did not find any significant differences in performance of early and late bilinguals in any of the cognitive domains. However, in the study by Rosselli et al. (2000), there was an interaction between AoA and the language of test administration on tests of repetition, naming and verbal fluency. Bilinguals who acquired English before the age of 12 performed better in their second language, namely, in the English version of the tests.

Language of test administration and the use of culturally appropriate tools. Only five studies used validated neuropsychological assessment tools or reported conducting culturally or linguistically appropriate neuropsychological tests for bilinguals (Massa et al., 2020; Mungas et al., 2018; Nielsen et al., 2019; Padilla et al., 2016; Rosselli et al., 2000). The potential impact of cultural, ethnic, and/or language factors in test performance was rarely considered. Of eighteen cross-sectional studies, seven included culturally and/or linguistically diverse participants in bilingual groups who were either tested in their non-dominant or second language. Monolingual groups were more likely to share the similar cultural and/or linguistic background (Anderson et al., 2017; Bialystok, Poarch, et al., 2014; Bialystok et al., 2008; Kowoll et al., 2015; Luo et al., 2013; Ossher et al., 2013; Papageorgiou et al., 2019). Additionally, the lack of use of neuropsychological tests standardized for the cultural background of the bilingual participants was commonly observed in these studies.

Three studies assessed the language groups in their native language or language of preference (Ihle et al., 2016; Nielsen et al., 2019; Rosselli et al., 2019). Eight studies administered the tests in both languages of participants (Clare, Whitaker, Craik, et al., 2016; Clare, Whitaker, Martyr, et al., 2016; Kousaie et al., 2014; Kowoll et al., 2015; Massa et al., 2020; Rosselli et al., 2000; Sheppard et al., 2016; Soltani et al., 2019). Among the studies analyzing the impact of language of test administration on NP, Clare, Whitaker, Martyr, et al. (2016) indicated no significant differences in performance between English and Welsh administration of Stroop task in healthy bilinguals. Similarly, Rosselli et al. (2000) reported similar test performance in healthy bilinguals on English and Spanish language tests. The only exception was the Cookie Theft task from the Boston Diagnostic Aphasia Examination, on which bilinguals produced a greater number of words when using their second language, English. On the other hand, in the study by Soltani et al. (2019) monolinguals performed better than bilinguals in the verbal fluency test when bilinguals were tested in their native language. However, no significant difference was found in this test between monolinguals and bilinguals when bilinguals were assessed in their second and mostly frequently used language. Furthermore, Kowoll et al. (2015) compared the performance of bilinguals across different diagnostic groups and tested them in their dominant and non-dominant language on the verbal fluency task and BNT. Bilingual MCI patients had a poorer performance when tested in their dominant language, whereas bilingual AD patients had a poorer performance when tested in their non-dominant language.

Two other studies employed a different testing procedure and tested French-English healthy bilinguals in three conditions, namely, English, French and bilingual (either-language) test administration wherein, participants could perform the test using both languages simultaneously (Kousaie et al., 2014; Sheppard et al., 2016). Both studies reported an advantage for the BNT in monolingual English speakers compared to monolingual French speakers and bilinguals. The findings from these studies suggest that the items of the BNT may not be equivalent in terms of difficulty or familiarity in French and English. Furthermore, Sheppard et al. (2016) showed that majority of bilingual participants performed better on the BNT when bilinguals used both languages during testing.

Of the nine longitudinal studies, three studies comprised of participants with mixed linguistic and cultural background and participants were tested in their second-language (Bialystok, Craik, et al., 2014; Chertkow et al., 2010; Yeung et al., 2014). In the study by Yeung et al. (2014), the participants whose second language was English performed worse on a test of global cognitive function (3MS) than those whose first language was English. Three studies included more culturally and linguistically homogenous samples, with Spanish-English speaking Hispanic and Latino American individuals and assessed the participants in their preferred language (Padilla et al., 2016; Mungas et al., 2018; Zahodne et al., 2014). In the study by Zahodne et al. (2014), bilinguals demonstrated better baseline performance on executive function and verbal episodic memory (Selective Reminding Test) than monolinguals when tested in their language of choice. The studies by Padilla et al. (2016) and Mungas et al. (2018) used a validated neuropsychological measure for the cultural and linguistic background of

participants and reported no effect of language of test administration on the 3MS in bilinguals.

Immigration status. The included studies were limited geographically, and thereby, culturally. With the exception of two studies, conducted in Israel and Iran, all other studies were conducted in Europe and North America. In half of the studies from Canada and four from the USA, knowledge of second language was linked to immigration status. Additionally, in majority of studies that included immigrant populations, the age at immigration (Bialystok, Craik, et al., 2014; Chertkow et al., 2010; Padilla et al., 2016; Papageorgiou et al., 2019; Sheppard et al., 2016) and the length of residence in the new country (Anderson et al., 2017; Bialystok, Craik, et al., 2014; Bialystok et al., 2008; Chertkow et al., 2010; Kowoll et al., 2015; Mungas et al., 2018; Padilla et al., 2016; Papageorgiou et al., 2019; Sheppard et al., 2016) were not taken into account. Among cross-sectional studies, seven studies did not report on immigration status of participants (Bialystok, Poarch, et al., 2014; Friesen et al., 2015; Massa et al., 2020; Ihle et al., 2016; Luo et al., 2013; Ossher et al., 2013; Sheppard et al., 2016; Soltani et al., 2019). Two studies that included nonimmigrants showed a bilingual advantage only in the domain of response inhibition and management of response conflict on the TEA and Stroop task (Clare, Whitaker, Craik, et al., 2016; Kousaie et al., 2014). Of the nine longitudinal studies, one study (Yeung et al., 2014) did not specify the immigration status of participants. In three studies including nonimmigrant language groups, one study (Cox et al., 2016) did not report any effect of bilingualism on a reasoning/planning test, whereas one study showed a multilingual advantage on tests of verbal fluency, general intelligence, vocabulary/reading (Bak et al., 2014) and one other indicated a bilingual advantage on tests of global cognitive function, letter fluency (Costumero et al., 2020). Three studies included immigrant participants in both monolingual and bilingual groups in which two of them consisted of higher number of immigrants in the bi/multi-lingual group (Bialystok, Craik, et al., 2014; Chertkow et al., 2010; Mungas et al., 2018). Lastly, the other two studies included samples comprising of entirely immigrants in both language groups (Padilla et al., 2016; Zahodne et al., 2014). Bilingualism was associated with a better initial performance on language, executive function and praxis items of a global cognitive functioning test, before and after controlling for immigration status (Padilla et al., 2016), as well as, on tests of verbal episodic memory (Selective Reminding Test) and executive function without controlling for immigration status (Zahodne et al., 2014).

Discussion

The primary rationale for this review was to assess the findings from the extant literature to determine whether bilingualism confers an advantage and/or disadvantage in NP in older adults and, thereby, to give an insight into the assessment of bilingual older adults in clinical practice as little is known about the generalizability of research on NP of bilingual older adults to clinical neuropsychology. Language proficiency, AoA, immigrant status, and language of test administration were factors addressed in this review to examine whether participant and language related variables influence the degree of advantage or disadvantage that bilingual older adults may experience during test performance.

Does bilingualism in older adults offer advantages or disadvantages in the neuropsychological test performance?

In cross-sectional studies that included bilingual healthy older adults, a bilingual disadvantage was observed in lexical access tests involving vocabulary knowledge and verbal fluency (Anderson et al., 2017; Bialystok et al., 2008; Bialystok, Poarch, et al., 2014; Clare, Whitaker, Martyr, et al., 2016; Luo et al., 2013; Rosselli et al., 2000), particularly in category fluency subtests (Anderson et al., 2017; Bialystok et al., 2008; Rosselli et al., 2000). This is consistent with previous literature showing poorer performance among bilinguals in comparison to monolinguals on verbal tasks which require vocabulary knowledge and faster lexical access across different age groups (Kroll & Gollan, 2013). Explanation for these findings might lie in the frequency lag (Gollan et al., 2011; Gollan et al., 2008) and competition account (Abutalebi & Green, 2007; Green, 1998; Kroll & Gollan, 2013). According to the frequency lag account, bilinguals are exposed to each word less frequently, as they speak each language less often than monolinguals. Therefore, there is a weaker connection between the words and the concepts represented by the words, leading to reduced accessibility of words in each language (Gollan et al., 2008).

On the other hand, the competition account assumes that the intention to speak prompts simultaneous activation of both languages in bilinguals, resulting in greater competition for selection of the target language (Green, 1998; Kroll et al., 2008). The need to resolve this competition between languages in bilinguals could slow down retrieval of target language words. This may lead to a decrease in the number of correct responses in bilinguals than in monolinguals (Sandoval et al., 2010), particularly in a semantic fluency task which relies primarily on lexical retrieval speed (Gordon et al., 2018). Semantic fluency task is more constrained as it is highly dependent on the connections between conceptual and semantic representations as opposed to the letter fluency task, where such connections are not required (Kavé & Knafo-Noam, 2015; Meinzer et al., 2009). Additionally, semantic fluency is more likely influenced by non-target language interference since translation equivalents is in the same semantic category (Giezen & Emmorey, 2017). The study by Sheppard et al. (2016) provides additional support for this account with the findings showing an association between either-language test administration and improved test performance in a non-speeded measure of semantic memory, namely the BNT. This is perhaps not surprising given that bilinguals may need to control interference from the non-target language and resolve competition

in lexical selection and BNT and semantic fluency put demands on semantic memory (Henry et al., 2004). With the simultaneous use of both languages, there may be no need to inhibit any language. Therefore, no additional resources are used in inhibition, making it easier for bilinguals to access the lexicon more quickly than they would in a single-language test administration. Better performance on the BNT when using either-language test administration was found in bilingual healthy adults (Gollan et al., 2007).

Further, in cross-sectional and baseline results of longitudinal studies, there was some evidence showing a bilingual advantage with respect to inhibition and management of response conflict domain assessed by the Stroop task in healthy older adults, (Bialystok, Poarch, et al., 2014; Kousaie et al., 2014; Massa et al., 2020) as well as in MCI and AD patients (Anderson et al., 2017; Bialystok, Craik, et al., 2014). These findings are consistent with the results of a previous study (Kousaie & Phillips, 2017) using both behavioral and electrophysiological measures in older adults showing a superior performance in the Stroop task in bilinguals compared to monolinguals .The competition account may provide an explanation for both the observed disadvantage in lexical access and the advantage in executive control. The mechanism of inhibitory control used by bilinguals regularly to reduce interference from the non-target language (Green, 1998; Linck et al., 2008) may result in enhanced cognitive control. The results from two studies with healthy older adults (Friesen et al., 2015; Massa et al., 2020) may also lend some support to this notion with findings showing a better performance in bilinguals in letter fluency which demands greater executive control functioning (Gordon et al., 2018; Shao et al., 2014). The recruitment of executive control in letter fluency may be a possible explanation of the observed bilingual advantages in these studies as it is a verbal task which requires retrieval of words starting with a specific letter and inhibition of words beginning with different letters, suggesting involvement of conflict-resolution skills (Blumenfeld et al., 2016; Luo et al., 2010). This result is in line with a previous study wherein the Stroop performance as a measure of inhibitory control and conflict-resolution was correlated with verbal fluency measures and bilingual healthy adults outperformed monolinguals in letter fluency task, especially, where there was a higher demand for executive control (Patra et al., 2020). However, on the other hand, the studies by Bialystok et al. (2008) and Soltani et al. (2019) with healthy older adults, provide contradictory findings, with monolinguals showing a better performance than bilinguals in letter fluency. A possible explanation for these contradictory findings may be due to language of test administration and executive control abilities in bilinguals. In the studies by Friesen et al. (2015) and Massa et al. (2020), majority of bilingual participants were assessed in their dominant language and they were either matched on executive function performance or a better performance was reported in Stroop test for bilinguals. Furthermore, in the study by Soltani et al. (2019), the bilingual disadvantage disappeared when bilinguals were tested in their most frequently used language. These findings suggests that bilingual advantages may occur in verbal tasks which have a higher role for executive control once the language of proficiency has been accounted for (Luo et al., 2010), particularly in language of testing. Overall, these results indicate that the advantages and disadvantages associated with cross-linguistic interference and inhibition may differ for tasks which impose demands on lexical access and executive control.

The longitudinal studies examining healthy older adults (Cox et al., 2016), older adults who are either with intact cognition or cognitive impairment at baseline (Mungas et al., 2018; Padilla et al., 2016), MCI and AD patients (Bialystok, Craik, et al., 2014) and conversion to dementia (Yeung et al., 2014; Zahodne et al., 2014) yield consistent pattern of findings showing no bilingualism advantage on test performance over time except for one study. The longitudinal data from Costumero et al. (2020) which included multiple-domain aMCI patients reported a greater cognitive decline in letter fluency and global cognition in monolinguals compared to bilinguals. This finding is in contrast to the findings from a longitudinal study with MCI patients (Bialystok, Craik, et al., 2014), which did not find any differences between monolinguals and bilinguals on the letter fluency task. The languages of bilinguals in the study by Costumero et al. (2020) consisted of two typologically similar languages, namely, Spanish and Catalan and monolinguals were identified as passive bilinguals who could speak Spanish and understand Catalan with poor fluency. Furthermore, bilinguals and monolinguals did not differ in their baseline performance on several tests. On the other hand, in the study by Bialystok, Craik, et al. (2014), bilinguals were participants with various language combinations and cultural backgrounds who were tested in their second language. For this reason, interpretation of these studies and their apparently conflicting findings is limited as a consequence of a number of factors, namely, differences in typology, structure and culture between languages (Eng et al., 2019), language of test administration and definitions used for bilingualism. These contradictory and inconclusive results highlight the need for more longitudinal studies investigating differences in performance on verbal fluency tasks in monolingual and bilingual MCI patients.

In general, the evidence in favor of the existence of a bilingual advantage is weak and observed more often in the cross-sectional studies using measures of inhibitory control or on baseline performance of bilinguals assessed in the longitudinal studies, but the findings lack consistency. A possible reason for these contradictory findings may be methodological differences across studies, namely, the sample (size, inclusion/exclusion criteria, inclusion of different clinical samples), different study designs, lack of consistency in standardized tests used to assess cognitive domains and language and participant related variables (see Figure 2).

What factors affect the findings of neuropsychological test performance in older bilingual adults?

A number of methodological differences might explain the apparent discrepancies in results across studies. First, the

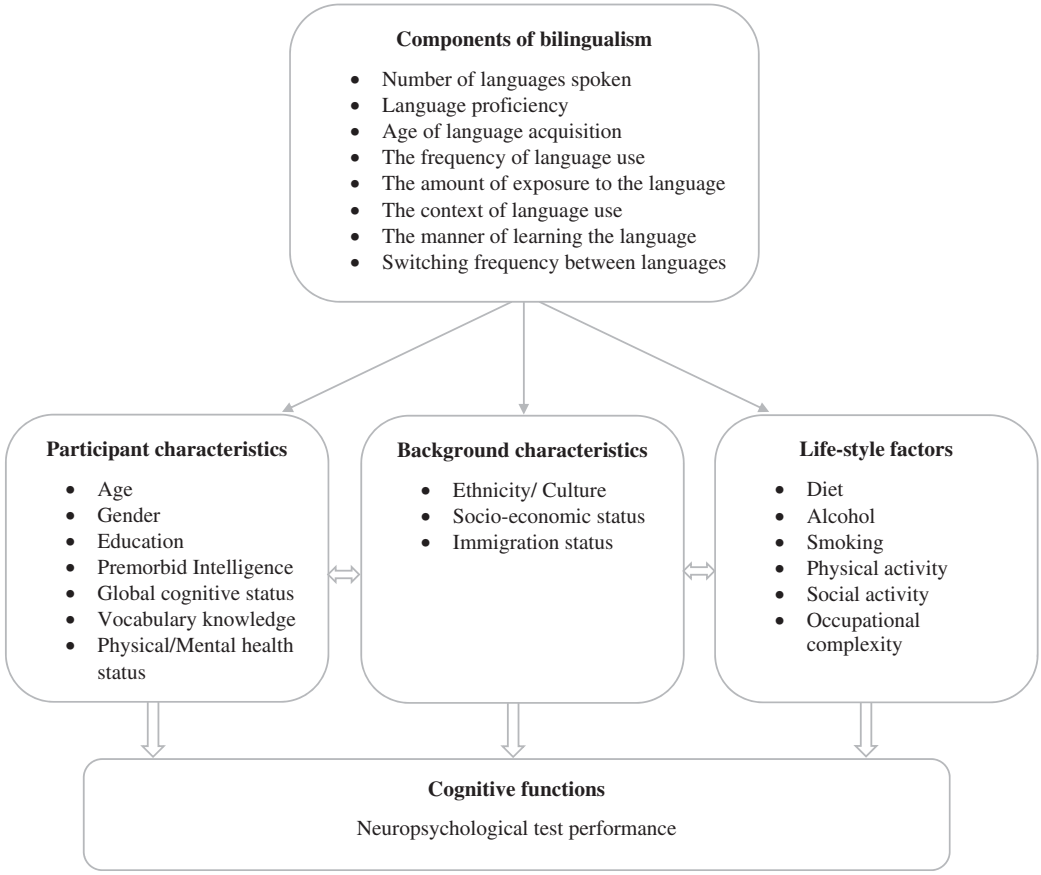


Figure 2. Confounding variables stated in the articles.

characteristics of bilingual participants varied across studies due to lack of a standardized definition of bilingualism (Anderson et al., 2018), leading to inconsistencies in defining bilingualism across studies. In order to gain a clear understanding of the nature of the relationship between bilingualism and cognitive outcomes, it is important to explore specific factors associated with bilingualism that may be responsible for the observed effect (Kroll, 2009).

Language proficiency comprises a core aspect of language experience in bilingualism research. However, only 19 studies reported the level of language proficiency of bilinguals. Studies showed that higher level of language proficiency was associated with better performance, particularly on tests measuring inhibition. This corresponds well with the findings of other studies using Stroop tasks (Tzelgov et al., 1990; Zied et al., 2004) which indicate that the size of Stroop effect is influenced by the proficiency in the second language such that a minimum level of language proficiency is necessary to elicit interference effects and higher levels of proficiency results in better controlled processing. Additionally, the finding regarding the higher level of language proficiency and letter fluency task is in line with the results reported in young adults (Luo et al., 2010). A possible explanation for this may be that the competition evoked by higher language proficiency may result in more engagement in executive control due to increasing demand imposed by managing cross-linguistic influences. However, language proficiency alone does not fully account for the advantage reported in the studies as a few studies including highly proficient bilinguals indicated no significant differences in test scores between language groups (Kowoll et al., 2015; Papageorgiou et al., 2019).

Another variable related to bilingualism, namely, AoA may provide further insight and help explain the inconclusive findings. AoA is a complex variable as it is not only associated with the level of input received according to years of exposure, but also varied patterns of language use between speakers with early and late acquisition (Antoniou & Wright, 2017). Indeed, some studies included in this review revealed a different pattern of test performance among early and late bilinguals in various cognitive domains. However, a major confounder is the different ages used as cutoffs for the classification of early versus late acquisition. The studies investigating the effect of AoA on the neural organization of the bilingual brain commonly use 6 years of age as cutoffs (Wattendorf & Festman, 2008). It is important for studies on the NP of bilinguals to determine AoA for the categorization of early and late bilinguals, based on available literature and examine its effect on NP for the integration of findings. However, simultaneous bilinguals who acquired two languages from birth can be considered as a potentially distinct group since their experience may vary from bilinguals with early AoA (Sabourin & Vīnerte, 2015).

Inconsistent results in NP of language groups may further be explained by a methodological difference in the

comparison of monolingual versus bilingual groups, namely, language of test administration. In a few studies (Clare, Whitaker, Martyr, et al., 2016; Rosselli et al., 2000), scores on the Stroop task and language tests did not differ depending upon whether the tests were administered in different languages. Upon closer examination of the characteristics of samples, it can thus be hypothesized that inclusion of bilinguals with comparable proficiency in each language may help eliminate the effects related to language of test administration (Gasquoine et al., 2007). Furthermore, contradictory findings in letter fluency performance highlight the need to consider the testing language in bilinguals.

Language of test administration is also closely linked to the issue of linguistically and culturally equivalent version of the neuropsychological tests used in the studies. In some studies, monolingual English speakers performed better than monolingual French speakers on the BNT, pointing to a limitation for cross-linguistic and cultural equivalence of this test. This is further evident through studies which found no impact of language of administration on test performance when using a validated assessment tool targeting the cultural and linguistic background of the test-takers (Padilla et al., 2016; Mungas et al., 2018). Taken together, the results point to an interplay between language proficiency, language of test administration and the use of culturally/linguistically appropriate tools in the assessment of bilinguals. Furthermore, these results highlight the need to evaluate cross-cultural and cross-linguistic potential of neuropsychological tests and emphasize the importance of using validated assessment tools appropriate for the cultural/linguistic background of the participants in order to obtain more accurate results of NP in bilinguals.

Besides language-related confounding variables, another critical factor that has been neglected in the studies is the demographically diverse backgrounds of monolingual and bilingual groups, such as different cultures/ethnicities and language families. Specifically, studies were more likely to compare a homogenous group of English monolinguals with group of bilinguals comprised of varying cultural, ethnic, and linguistic background (Bialystok et al., 2008; Luo et al., 2013; Ossher et al., 2013; Papageorgiou et al., 2019). The effect of cultural variables on neuropsychological tests is well-established in the extant literature (Agranovich & Puente, 2007; Ardila, 2005, 2007a; Loewenstein et al., 1994; Walker et al., 2010). Cultural variables may exert a greater impact on test performance in individuals with a non-Western or non-English background as majority of the tests used in clinical settings are developed by Western, English speaking countries (Wong, 2000). Performance of culturally/ethnically diverse bilinguals on neuropsychological tests may be mitigated by an interaction of various linguistic and cultural variables and not necessarily reflect a direct relationship between cognition and test performance.

The diverse cultural and linguistic backgrounds in bilingual groups also bring out another confounding factor in the studies; immigration status of participants. In studies examining nonimmigrant language groups (Clare, Whitaker, Craik, et al., 2016; Clare, Whitaker, Martyr, et al., 2016; Cox et al., 2016; Kousaie et al., 2014), there was no difference in NP between monolinguals and bilinguals on tests of executive function, except for a small advantage in bilinguals on the inhibition domain reported in cross-sectional studies (Clare, Whitaker, Martyr, et al., 2016; Kousaie et al., 2014). The findings in longitudinal studies were mixed; showing either no long-term effect of bilingualism on NP or better performance at baseline on tests of global cognitive functioning., verbal episodic memory and executive function. The association between immigration status and cognitive function is still poorly understood (for a review, see Xu et al., 2017). The inconsistent findings may be partially explained by level of acculturation. There is a growing body of evidence indicating that acculturation might be a salient factor in NP differences among various ethnic/cultural groups (Arentoft et al., 2012). Higher acculturation level is significantly linked to better NP in a variety of cognitive domains, including executive function, verbal fluency, naming, processing speed and global cognitive functioning (Arentoft et al., 2012; Boone et al., 2007; Coffey et al., 2005; Manly et al., 1998). However, acculturation was measured in a few studies in cursory terms by focusing on country of birth and length of residence in the new country without referring to them as proxies of acculturation. Acculturation is a multidimensional construct and these proxies yield only an indirect measure of acculturation (Abe-Kim et al., 2001; Lopez-Class et al., 2011). It could be that different levels of acculturation as a consequence of immigration status may have distinct effects on NP in language groups. Hence, the implementation of acculturation measures has strong potential to benefit research in NP of bilingual older adults.

Taken together, all the factors listed reveal the difficulty in disentangling the interplay between linguistic and, participant-related variables on NP.

Clinical implications

Despite the mixed findings, the results of this review can be utilized as a tool that might aid in making clinical interpretations with bi/multi-lingual populations in both clinical and research settings. In order for accurate assessment of individuals' cognitive abilities, researchers and neuropsychologists should take into account factors such as, acculturation level of the individuals, culturally corrected norms or cross-cultural neuropsychological assessment tools for participants from diverse cultural backgrounds, as well as language proficiency. Additionally, given the impact of language of test administration, consideration of alternate assessment approaches and identification of strategies for increasing the utility of neuropsychological testing is of critical importance when testing bilingual populations with differing proficiency in each language. Lastly, due to contradictory findings on letter fluency between cross-sectional and longitudinal studies, it remains an open question whether a differential performance emerges between monolinguals and bilinguals.

Limitations and future recommendations

The findings of this review indicate that AoA and proficiency may play a distinct role in bilingualism research. Although they are highly correlated, it has been shown that AoA and proficiency influence the brain in different ways, with AoA influencing regions used in grammatical processing and proficiency affecting those involved in semantic judgements (Wartenburger et al., 2003). A more rigorous investigation of these two factors may help determine the weight of their impact on test performance, specifically on tests measuring language domain.

This review revealed that another critical language-related factor, which has been largely overlooked in this area of research is context of language use and language switching pattern. Studies included in the review predominantly addressed language proficiency or the number of languages spoken by the participants. It is possible that the context in which language is used and switching frequency between languages might be associated with a different pattern of findings. According to the Adaptive Control Hypothesis proposed by Green and Abutalebi (2013), context is a key factor in determining which language can be used and the extent to which bilinguals can switch between languages. For this reason, it has been proposed that the demand on cognitive control processes can vary as a result of bilingual language context. However, there is a paucity of research in this area and future studies should closely examine the relationship between context of language use, language switching pattern and NP.

Furthermore, variables accounted for and neuropsychological assessments employed vary considerably across studies. For this reason, a meta-analysis was not possible. Besides this, there was a lack of consensus on categorization of cognitive domains assessed by neuropsychological tests in the studies.

Finally, due to the scarcity of longitudinal studies, it is currently difficult to fully understand the changes in NP that occur with the progression of the disease in bilingual MCI and AD patients. The use of longitudinal studies, rather than cross-sectional studies seems to shed more light on the relationship between language knowledge and cognition (Calvo et al., 2016).

Taken together, efforts should be taken to address or examine moderating variables identified in this review and to use more standardized neuropsychological tests that can be appropriately used with culturally and linguistically diverse bilingual individuals in order to elucidate the effects of bilingualism on cognitive domains more comprehensively.

Conclusions

The research on NP in bilingual older adults is riddled with complexity and discrepancies regarding the conceptualization and measurement of bilingualism and its association with immigration status, language of test administration and the standardized assessment tools appropriate for the cultural background of the tested population. This review demonstrated that neuropsychological tests in the language domain, specifically in verbal fluency and to some extent executive function (particularly, inhibitory control) represents a core and clinically relevant feature of bilingualism which needs to be taken into account during the evaluation of participants in clinical and research settings.

There is an increasing need for measures that accurately and efficiently contribute to the neuropsychological assessment of bilingual populations. The cross-culturally and cross-linguistically applicable tests, namely, the Cross-Linguistic Naming Test (Ardila, 2007b), European Cross-Cultural Neuropsychological Test Battery (Nielsen et al., 2018) or tests (Five Digit Test, Stick Design Test, Rowland Universal Dementia Assessment Scale) which are less influenced by differences in cultural and language background (Franzen et al., 2020) may utilize research in neuropsychological performance of bi/multilingual older adult groups consisting of various cultural/language backgrounds. Furthermore, the Language Background Questionnaire (LSBQ; Anderson et al., 2018) which was developed to quantify bilingualism based on a composite score in culturally diverse populations and to provide with evidence-based classifications of language groups may help compare the results of various studies effectively. Alternatively, language proficiency is a significant indicator of individuals' level of acculturation as it characterizes the degree of integration in a new country and the amount of exposure to the dominant language of the host culture (Schumann, 1986). Thus, evaluation of acculturation level may aid in determining the language in which participants can be tested (Pontón) and it can allow for addressing the contextual variables which may maintain or improve language proficiency in bilinguals.

A comprehensive theoretical model is necessary to portray which cognitive domains are influenced by which components of bilingualism and to suggest testable predictions (Antoniou, 2019). Furthermore, components of bilingualism and participant-related variables stated in this review (see Figure 2) highlight the lack of a thorough and systematic understanding of the underlying factors contributing to the inconsistent findings in the field and may provide an initial step toward studies with more rigorous methodology.

Disclosure statement

The authors declare no conflict of interest.

Funding

The study was supported by the Robert Bosch Foundation (grant no. 11.5.G411.0005.0) Stuttgart, Germany (Robert Bosch Stiftung) and the interdisciplinary Graduate Program "People with Dementia in Acute-care Hospitals", located at the Network Aging Research (NAR), Heidelberg University, Germany.

ORCID

Simge Celik http://orcid.org/0000-0002-7964-9376

References

Abe-Kim, J., Okazaki, S., & Goto, S. G. (2001). Unidimensional versus multidimensional approaches to the assessment of acculturation for Asian American populations. *Cultural Diversity & Ethnic Minority Psychology*, *7*(3), 232–246. https://doi.org/10.1037/1099-9809.7.3.232

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242–275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Agranovich, A. V., & Puente, A. E. (2007). Do Russian and American normal adults perform similarly on neuropsychological tests? Preliminary findings on the relationship between culture and test performance. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *22*(3), 273–282. https://doi.org/10.1016/j.acn.2007.01.003

Anderson, J. A., 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, J. A., Saleemi, S., & Bialystok, E. (2017). Neuropsychological assessments of cognitive aging in monolingual and bilingual older adults. *Journal of Neurolinguistics*, *43*(A), 17–27. https://doi.org/10.1016/j.jneuroling.2016.08.001

Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics*, *5*(1), 395–415. https://doi.org/10.1146/annurev-linguistics-011718-011820

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

Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology Review*, *15*(4), 185–195. https://doi.org/10.1007/s11065-005-9180-y

Ardila, A. (2007a). The impact of culture on neuropsychological test performance. In N. W. Nelson & M. O. Pontón (Eds.), *International handbook of cross-cultural neuropsychology* (pp. 23–44). Routledge.

Ardila, A. (2007b). Toward the development of a cross-linguistic naming test. *Archives of Clinical Neuropsychology*, *22*(3), 297–307. https://doi.org/10.1016/j.acn.2007.01.016

Arentoft, A., Byrd, D., Robbins, R. N., Monzones, J., Miranda, C., Rosario, A., Coulehan, K., Fuentes, A., Kubo Germano, K., D'Aquila, E., Sheynin, J., Fraser, F., Morgello, S., & Rivera Mindt, M. (2012). Multidimensional effects of acculturation on English-language neuropsychological test performance among HIV + Caribbean Latinas/os. *Journal of Clinical and Experimental Neuropsychology*, *34*(8), 814–825. https://doi.org/10.1080/13803395.2012.683856

Bak, T. H. (2016). The impact of bilingualism on cognitive ageing and dementia: Finding a path through a forest of confounding variables. *Linguistic Approaches to Bilingualism*, *6*(1–2), 205–226. https://doi.org/10.1075/lab.15002.bak

Bak, T. H., 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

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio*, *78*(2), 490–498.

Berry, J. W. (2007). Acculturation strategies and adaptation. In J. E. Lansford, K. Deater-Deckard, & M. H. Bornstein (Eds.), *Duke series in child development and public policy. Immigrant families in contemporary society* (p. 69–82). Guilford Press.

Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language*, *109*(2–3), 93–100. https://doi.org/10.1016/j.bandl.2008.09.001

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

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

Bialystok, E., Craik, F. I., Binns, M. A., Ossher, L., & Freedman, M. (2014). Effects of bilingualism on the age of onset and progression of MCI and AD: Evidence from executive function tests. *Neuropsychology*, *28*(2), 290–304. https://doi.org/10.1037/neu0000023

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

Bialystok, E., Craik, F., & 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

Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging*, *29*(3), 696–705. https://doi.org/10.1037/a0037254

Blumenfeld, H. K., Bobb, S. C., & Marian, V. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals' verbal fluency. *International Journal of Speech-Language Pathology*, *18*(2), 190–201. https://doi.org/10.3109/17549507.2015.1081288

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, *22*(3), 355–365. https://doi.org/10.1016/j.acn.2007.08.005

Calvo, N., García, A. M., Manoiloff, L., & Ibáñez, A. (2016). Bilingualism and cognitive reserve: a critical overview and a plea for methodological innovations. *Frontiers in Aging Neuroscience*, *7*, 249. https://doi.org/10.3389/fnagi.2015.00249

Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological assessment: Past and future. *Journal of the International Neuropsychological Society*, *23*(9–10), 778–790. https://doi.org/10.1017/S1355617717001060

Casaletto, K. B., Umlauf, A., Beaumont, J., Gershon, R., Slotkin, J., Akshoomoff, N., & Heaton, R. K. (2015). Demographically corrected normative standards for the English version of the NIH Toolbox Cognition Battery. *Journal of the International Neuropsychological Society*, *21*(5), 378–391. https://doi.org/10.1017/S1355617715000351

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: Evidence from a bilingual community. *Alzheimer Disease and Associated Disorders*, *24*(2), 118–125. https://doi.org/10.1097/WAD.0b013e3181ca1221

Clare, L., Whitaker, C. J., Craik, F. I. M., Bialystok, E., Martyr, A., Martin-Forbes, P. A., Bastable, A. J. M., Pye, K. L., Quinn, C., Thomas, E. M., Gathercole, V. C. M., & Hindle, J. V. (2016). Bilingualism, executive control, and age at diagnosis among people with early-stage Alzheimer's disease in Wales. *Journal of Neuropsychology*, *10*(2), 163–185. https://doi.org/10.1111/jnp.12061

Clare, L., Whitaker, C. J., Martyr, A., Martin-Forbes, P. A., Bastable, A. J. M., 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

Coderre, E. L., Van Heuven, W. J., & Conklin, K. (2013). The timing and magnitude of Stroop interference and facilitation in monolinguals and bilinguals. *Bilingualism: Language and Cognition*, *16*(2), 420–441. https://doi.org/10.1017/S1366728912000405

Coffey, D. M., Marmol, L., Schock, L., & Adams, W. (2005). The influence of acculturation on the Wisconsin card sorting test by Mexican Americans. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *20*(6), 795–803. https://doi.org/10.1016/j.acn.2005.04.009

Costumero, V., Marin-Marin, L., Calabria, M., Belloch, V., Escudero, J., Baquero, M., Hernandez, M., Ruiz de Miras, J., Costa, A., Parcet, M.-A., & Ávila, C. (2020). A cross-sectional and longitudinal study

on the protective effect of bilingualism against dementia using brain atrophy and cognitive measures. *Alzheimer's Research & Therapy*, *12*(1), 11–10. https://doi.org/10.1186/s13195-020-0581-1

Cox, S. R., Bak, T. H., Allerhand, M., Redmond, P., Starr, J. M., Deary, I. J., & MacPherson, S. E. (2016). Bilingualism, social cognition and executive functions: A tale of chickens and eggs. *Neuropsychologia*, *91*, 299–306. https://doi.org/10.1016/j.neuropsychologia.2016.08.029

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. American Guidance Service.

Eng, N., Vonk, J. M., Salzberger, M., & Yoo, N. (2019). A cross-linguistic comparison of category and letter fluency: Mandarin and English. *Quarterly Journal of Experimental Psychology*, *72*(3), 651–660. https://doi.org/10.1177/1747021818765997

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

Franzen, S., van den Berg, E., Goudsmit, M., Jurgens, C. K., van de Wiel, L., Kalkisim, Y., Uysal-Bozkir, Ö., Ayhan, Y., Nielsen, T. R., & Papma, J. M. (2020). A systematic review of neuropsychological tests for the assessment of dementia in non-western, low-educated or illiterate populations. *Journal of the International Neuropsychological Society*, *26*(3), 331–351. https://doi.org/10.1017/S1355617719000894

Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience*, *30*(3), 238–250. https://doi.org/10.1080/23273798.2014.918630

Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007). Language of administration and neuropsychological test performance in neurologically intact Hispanic American bilingual adults. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *22*(8), 991–1001. https://doi.org/10.1016/j.acn.2007.08.003

Gasquoine, P. G., & Gonzalez, C. D. (2012). Using monolingual neuropsychological test norms with bilingual Hispanic Americans: Application of an individual comparison standard. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *27*(3), 268–276. https://doi.org/10.1093/arclin/acs004

Giezen, M. R., & Emmorey, K. (2017). Evidence for a bimodal bilingual disadvantage in letter fluency. *Bilingualism: Language and Cognition*, *20*(1), 42–48. https://doi.org/10.1017/S1366728916000596

Gollan, T. H., & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General*, *135*(3), 462–483. https://doi.org/10.1037/0096-3445.135.3.462

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, *13*(2), 197–208. https://doi.org/10.1017/S1355617707070038

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Gordon, J. K., Young, M., & Garcia, C. (2018). Why do older adults have difficulty with semantic fluency? *Aging, Neuropsychology, and Cognition*, *25*(6), 803–828. https://doi.org/10.1080/13825585.2017.1374328

Gracey, D. J., & Morris, R. G. (2007). Neuropsychological assessment in dementia. *Psychiatry*, *6*(12), 498–502. https://doi.org/10.1016/j.mppsy.2007.09.005

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

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

Harvey, P. D. (2012). Clinical applications of neuropsychological assessment. *Dialogues in Clinical Neuroscience*, *14*(1), 91–99.

Heaton, R. K., Ryan, L., & Grant, I. (2009). Demographic influences and use of demographically corrected norms in neuropsychological assessment. *Neuropsychological Assessment of Neuropsychiatric and Neuromedical Disorders*, *3*, 127–155.

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, *42*(9), 1212–1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Hill, T. D., Angel, J. L., Balistreri, K. S., & Herrera, A. P. (2012). Immigrant status and cognitive functioning in late-life: An examination of gender variations in the healthy immigrant effect. *Social Science & Medicine (1982)*, *75*(12), 2076–2084. https://doi.org/10.1016/j.socscimed.2012.04.005

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, *15*(2), 422–433. https://doi.org/10.1017/S1366728911000678

Ihle, A., Oris, M., Fagot, D., & Kliegel, M. (2016). The relation of the number of languages spoken to performance in different cognitive abilities in old age. *Journal of Clinical and Experimental Neuropsychology*, *38*(10), 1103–1114. https://doi.org/10.1080/13803395.2016.1197184

Kavé, G., & Knafo-Noam, A. (2015). Lifespan development of phonemic and semantic fluency: Universal increase, differential decrease. *Journal of Clinical and Experimental Neuropsychology*, *37*(7), 751–763. https://doi.org/10.1080/13803395.2015.1065958

Kisser, J. E., Wendell, C. R., Spencer, R. J., & Waldstein, S. R. (2012). Neuropsychological performance of native versus non-native English speakers. *Archives of Clinical Neuropsychology*, *27*(7), 749–755. https://doi.org/10.1093/arclin/acs082

Kmet, L. M., Cook, L. S., & Lee, R. C. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Retrieved May 11, 2020, from http://www.ihe.ca/download/standard_quality_assessment_criteria_for_evaluating_primary_research_papers_from_a_variety_of_fields.pdf

Kousaie, S., & Phillips, N. A. (2017). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia*, *94*, 23–35. https://doi.org/10.1016/j.neuropsychologia.2016.11.013

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience*, *8*, 250. https://doi.org/10.3389/fnbeh.2014.00250

Kowoll, M. E., Degen, C., Gladis, S., & Schröder, J. (2015). Neuropsychological profiles and verbal abilities in lifelong bilinguals with mild cognitive impairment and Alzheimer's disease. *Journal of Alzheimer's Disease*, *45*(4), 1257–1268. https://doi.org/10.3233/JAD-142880

Kroll, J. F. (2009). The consequences of bilingualism for the mind and the brain. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *10*(3), i–ii. https://doi.org/10.1177/1529100610389314

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416–430. https://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., & Gollan, T. H. (2013). Speech planning in two languages. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production*. Oxford University Press.

Lee, L., Packer, T. L., Tang, S. H., & Girdler, S. (2008). Self-management education programs for age-related macular degeneration: A systematic review. *Australasian Journal on Ageing*, *27*(4), 170–176. https://doi.org/10.1111/j.1741-6612.2008.00298.x

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon*, *3*(3), 349–374. https://doi.org/10.1075/ml.3.3.06lin

Llorente, A. M., Pontón, M. O., Taussig, I. M., & Satz, P. (1999). Patterns of American immigration and their influence on the acquisition of neuropsychological norms for Hispanics. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *14*(7), 603–614. https://doi.org/10.1093/arclin/14.7.603

Loewenstein, D. A., Argüelles, T., Argüelles, S., & Linn-Fuentes, P. (1994). Potential cultural bias in the neuropsychological assessment of the older adult. *Journal of Clinical and Experimental Neuropsychology*, *16*(4), 623–629. https://doi.org/10.1080/01688639408402673

Lopez-Class, M., Castro, F. G., & Ramirez, A. G. (2011). Conceptions of acculturation: A review and statement of critical issues. *Social Science & Medicine (1982)*, *72*(9), 1555–1562. https://doi.org/10.1016/j.socscimed.2011.03.011

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, *28*(1), 28–34. https://doi.org/10.1037/a0030875

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29–41. https://doi.org/10.1016/j.cognition.2009.08.014

Mahendra, N., Plante, E., Magloire, J., Milman, L., & Trouard, T. P. (2003). fMRI variability and the localization of languages in the bilingual brain. *Neuroreport*, *14*(9), 1225–1228. https://doi.org/10.1097/00001756-200307010-00007

Manly, J. J., Miller, S. W., Heaton, R. K., Byrd, D., Reilly, J., Velasquez, R. J., Saccuzzo, D. P., & Grant, I. (1998). The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *Journal of the International Neuropsychological Society*, *4*(3), 291–302. https://doi.org/10.1017/S1355617798002914

Massa, E., Köpke, B., & El Yagoubi, R. (2020). Age-related effect on language control and executive control in bilingual and monolingual speakers: Behavioral and electrophysiological evidence. *Neuropsychologia*, *138*, 107336. https://doi.org/10.1016/j.neuropsychologia.2020.107336

Meinzer, M., Flaisch, T., Wilser, L., Eulitz, C., Rockstroh, B., Conway, T., Gonzalez-Rothi, L., & Crosson, B. (2009). Neural signatures of semantic and phonemic fluency in young and old adults. *Journal of Cognitive Neuroscience*, *21*(10), 2007–2018. https://doi.org/10.1162/jocn.2009.21219

Mindt, M. R., Arentoft, A., Kubo Germano, K., D'Aquila, E., Scheiner, D., Pizzirusso, M., Sandoval, T. C., & Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review*, *18*(3), 255–268. https://doi.org/10.1007/s11065-008-9069-7

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Mungas, D., Early, D. R., Glymour, M. M., Zeki Al Hazzouri, A., & Haan, M. N. (2018). Education, bilingualism, and cognitive trajectories: Sacramento Area Latino Aging Study (SALSA). *Neuropsychology*, *32*(1), 77–88. https://doi.org/10.1037/neu0000356

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

Nielsen, T. R., Antelius, E., & Waldemar, G. (2019). Cognitive advantages in adult Turkish bilingual immigrants–A question of the chicken or the egg. *Journal of Cross-Cultural Gerontology*, *34*(2), 115–129. https://doi.org/10.1007/s10823-019-09375-7

Nielsen, T. R., Segers, K., Vanderaspoilden, V., Bekkhus-Wetterberg, P., Minthon, L., Pissiota, A., Bjørkløf, G. H., Beinhoff, U., Tsolaki, M., Gkioka, M., & Waldemar, G. (2018). Performance of middle-aged and elderly European minority and majority populations on a cross-cultural neuropsychological test battery (CNTB). *The Clinical Neuropsychologist*, *32*(8), 1411–1430. https://doi.org/10.1080/13854046.2018.1430256

Norman, M. A., Evans, J. D., Miller, W. S., & Heaton, R. K. (2000). Demographically corrected norms for the California verbal learning test. *Journal of Clinical and Experimental Neuropsychology*, *22*(1), 80–94. https://doi.org/10.1076/1380-3395(200002)22:1;1-8;FT080

Ossher, L., Bialystok, E., Craik, F. I., Murphy, K. J., & Troyer, A. K. (2013). The effect of bilingualism on amnestic mild cognitive impairment. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(1), 8–12. https://doi.org/10.1093/geronb/gbs038

Padilla, C., Mendez, M. F., Jimenez, E. E., & Teng, E. (2016). Bilingualism in older Mexican-American immigrants is associated with higher scores on cognitive screening. *BMC Geriatrics*, *16*(1), 189. https://doi.org/10.1186/s12877-016-0368-1

Papageorgiou, A., Bright, P., Periche Tomas, E., & Filippi, R. (2019). Evidence against a cognitive advantage in the older bilingual population. *Quarterly Journal of Experimental Psychology*, *72*(6), 1354–1363. https://doi.org/10.1177/1747021818796475

Park, D. C., & Huang, C. M. (2010). Culture wires the brain: A cognitive neuroscience perspective. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *5*(4), 391–400. https://doi.org/10.1177/1745691610374591

Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition*, *23*(1), 204–218. https://doi.org/10.1017/S1366728918001098

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, *19*(3), 170–182. https://doi.org/10.1002/hbm.10110

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415–422. https://doi.org/10.1016/j.acn.2007.01.015

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(2), 253–262. https://doi.org/10.1017/S1366728909990526

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, *16*(4–6), 635–645. https://doi.org/10.1080/02687030244000220

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, *7*(1), 17–24. https://doi.org/10.1207/S15324826AN0701_3

Rosselli, M., Loewenstein, D. A., Curiel, R. E., Penate, A., Torres, V. L., Lang, M., Greig, M. T., Barker, W. W., & Duara, R. (2019). Effects of bilingualism on verbal and nonverbal memory measures in mild cognitive impairment. *Journal of the International Neuropsychological Society*, *25*(1), 15–28. https://doi.org/10.1017/S135561771800070X

Sabourin, L., & Vinerte, S. (2015). The bilingual advantage in the Stroop task: Simultaneous vs. early bilinguals. *Bilingualism: Language and Cognition*, *18*(2), 350–355. https://doi.org/10.1017/S1366728914000704

Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology*, *60*, 257–282. https://doi.org/10.1146/annurev.psych.57.102904.190024

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*(2), 231–252. https://doi.org/10.1017/S1366728909990514

Schumann, J. H. (1986). Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development*, *7*(5), 379–392. https://doi.org/10.1080/01434632.1986.9994254

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: Attention shifting and second-language proficiency. *Memory & Cognition*, *33*(4), 644–653. https://doi.org/10.3758/BF03195331

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

Sheffield, K. M., & Peek, M. K. (2009). Neighborhood context and cognitive decline in older Mexican Americans: Results from the Hispanic Established Populations for Epidemiologic Studies of the Elderly. *American Journal of Epidemiology*, *169*(9), 1092–1101. https://doi.org/10.1093/aje/kwp005

Sheppard, C., Kousaie, S., Monetta, L., & Taler, V. (2016). Performance on the Boston Naming Test in bilinguals. *Journal of the International Neuropsychological Society*, *22*(3), 350–363. https://doi.org/10.1017/S135561771500123X

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: Evidence from Hindi-English bilinguals. *Frontiers in Psychology*, *4*, 322. https://doi.org/10.3389/fpsyg.2013.00322

Soltani, M., Moradi, N., Rezaei, H., Hosseini, M., & Jasemi, E. (2019). Comparison of verbal fluency in monolingual and bilingual elderly in Iran. *Applied Neuropsychology: Adult*, 1–8. Advance online publication. https://doi.org/10.1080/23279095.2019.1594234

Suarez, P. A., Gollan, T. H., Heaton, R., Grant, I., & Cherner, M. (2014). Second-language fluency predicts native language Stroop effects: Evidence from Spanish–English bilinguals. *Journal of the International Neuropsychological Society*, *20*(3), 342–348. https://doi.org/10.1017/S1355617714000058

Teng, E. L., & Chui, H. C. (1987). The Modified Mini-Mental State (3MS) examination. *The Journal of Clinical Psychiatry*, *41*(2), 114–121.

Touradji, P., Manly, J. J., Jacobs, D. M., & Stern, Y. (2001). Neuropsychological test performance: A study of non-Hispanic White elderly. *Journal of Clinical and Experimental Neuropsychology*, *23*(5), 643–649. https://doi.org/10.1076/jcen.23.5.643.1246

Tse, C. S., & Altarriba, J. (2012). The effects of first-and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism: Language and Cognition*, *15*(3), 663–676. https://doi.org/10.1017/S1366728912000077

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(5), 760–771. https://doi.org/10.1037/0278-7393.16.5.760

Walker, A. J., Batchelor, J., Shores, E. A., & Jones, M. (2010). Effects of cultural background on WAIS-III and WMS-III performances after moderate-severe traumatic brain injury. *Australian Psychologist*, *45*(2), 112–122. https://doi.org/10.1080/00050060903428210

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, *37*(1), 159–170. https://doi.org/10.1016/S0896-6273(02)01150-9

Wattendorf, E., & Festman, J. (2008). Images of the multilingual brain: The effect of age of second language acquisition. *Annual Review of Applied Linguistics*, *28*, 3–24. https://doi.org/10.1017/S0267190508080033

Wong, T. M. (2000). Neuropsychological assessment and intervention with Asian Americans. In E. Fletcher-Janzen, T. L. Strickland, & C. R. Reynolds (Eds.), *Handbook of cross-cultural neuropsychology* (pp. 43–53). Springer.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization.

Xu, H., Zhang, Y., & Wu, B. (2017). Association between migration and cognitive status among middle-aged and older adults: A systematic review. *BMC Geriatrics*, *17*(1), 184. https://doi.org/10.1186/s12877-017-0585-2

Yeung, C. M., John, P. D. S., Menec, V., & Tyas, S. L. (2014). Is bilingualism associated with a lower risk of dementia in community-living older adults? Cross-sectional and prospective analyses. *Alzheimer Disease and Assoc Disorders*, *28*(4), 326–332. https://doi.org/10.1097/WAD.0000000000000019

Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., & Manly, J. J. (2014). Bilingualism does not alter cognitive decline or dementia risk among Spanish-speaking immigrants. *Neuropsychology*, *28*(2), 238–246. https://doi.org/10.1037/neu0000014

Zied, M. K., Phillipe, A., Pinon, K., Havet-Thomassin, V., Aubin, G., Roy, A., & Le Gall, D. (2004). Bilingualism and adult differences in inhibitory mechanisms: Evidence from a bilingual Stroop task. *Brain and Cognition*, *54*(3), 254–256. https://doi.org/10.1016/j.bandc.2004.02.036

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–363. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

4. Chapter 4: Cross-Cultural Comparison of MMSE and RUDAS in German and Turkish Patients with Alzheimer's Disease (3rd Publication)

Simge Celik[a], Oezguer Onur[b], Görsev Yener[c,d], Josef Kessler[b], Yagmur Özbek[c], Patric Meyer[e,f], Lutz Frölich[a,f], Birgit Teichmann[a]

[a] *Network Aging Research, Heidelberg University*

[b] *Department of Neurology, Medical Faculty and University Hospital, University of Cologne*

[c] *Department of Neurology, Faculty of Medicine, Dokuz Eylul University*

[d] *Department of Neurology, Medical Faculty, Izmir University of Economics*

[e] *Faculty of Applied Psychology, SRH University Heidelberg*

[f] *Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University*

Celik, S., Onur, O., Yener, G., Kessler, J., Özbek, Y., Meyer, P., Frölich, L., & Teichmann, B. (2022). Cross-cultural comparison of MMSE and RUDAS in German and Turkish patients with Alzheimer's disease.*Neuropsychology*, 36(3), 195–205. https://doi.org/10.1037/neu0000764

AMERICAN PSYCHOLOGICAL ASSOCIATION

Neuropsychology

ISSN: 0894-4105

2022, Vol. 36, No. 3, 195–205
https://doi.org/10.1037/neu0000764

Cross-Cultural Comparison of MMSE and RUDAS in German and Turkish Patients With Alzheimer's Disease

Simge Celik[1], Oezguer Onur[2], Görsev Yener[3, 4], Josef Kessler[2], Yagmur Özbek[5], Patric Meyer[6, 7], Lutz Frölich[1, 7], and Birgit Teichmann[1]

[1] Network Aging Research, Heidelberg University
[2] Department of Neurology, Medical Faculty and University Hospital, University of Cologne
[3] Department of Neurology, Faculty of Medicine, Dokuz Eylul University
[4] Department of Neurology, Medical Faculty, Izmir University of Economics
[5] Department of Neurosciences, Health Sciences Institute, Dokuz Eylul University
[6] Faculty of Applied Psychology, SRH University Heidelberg
[7] Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University

Objective: Given the increasing cultural, linguistic diversity in Europe, there is a growing need for cognitive screening tools that minimize the influence of linguistic, cultural, and demographic differences as they are the first means to determine the need for further clinical evaluation of individuals with suspected cognitive impairment. This cross-sectional study compared performance on the Mini-Mental State Examination (MMSE) and the Rowland Universal Dementia Assessment Scale (RUDAS) in Alzheimer's Disease (AD) patients in relation to cultural, demographic, and immigration-related factors (acculturation, bilingualism). ***Method:*** The study comprised Turkish immigrant ($n = 21$) and monolingual, nonimmigrant German ($n = 20$) and Turkish ($n = 24$) patients with AD. All participants were administered cognitive screening tools, measures of depression, and dementia severity. ***Results:*** The mean MMSE total score was significantly higher in German patients with AD compared to both patient groups, but did not differ between native-born Turkish and Turkish immigrant groups. After adjustment for years of education, differences in MMSE performance were no longer significant between groups. Furthermore, bilingualism was associated with better performance on the MMSE in Turkish-immigrant patients. The mean RUDAS total scores were similar between groups with and without adjustment for educational level. Performance on the RUDAS was not associated with demographic and immigration-related variables. ***Conclusions:*** The findings highlight the need to consider the educational background, linguistic integration of older non-Western immigrants for the objective characterization of cognitive profiles. The results provide support for the use of the RUDAS, particularly, among older Turkish immigrants with lower educational levels and varying degrees of acculturation, bilingualism.

This article was published Online First September 2, 2021.
Simge Celik https://orcid.org/0000-0002-7964-9376
Oezguer Onur https://orcid.org/0000-0001-8336-7075
Görsev Yener https://orcid.org/0000-0002-7756-4387
Lutz Frölich https://orcid.org/0000-0003-1494-0813
Birgit Teichmann https://orcid.org/0000-0001-7301-3409
Lutz Frölich and Birgit Teichmann should be considered joint last authors.

Simge Celik is a member of the Graduate Program "People with Dementia in General Hospitals," located at the Network Aging Research (NAR), Heidelberg University, Germany, and received a Doctoral Fellowship funded by Robert Bosch Foundation, Stuttgart, Germany.

We would like to express our sincere thanks to Lucrezia Hausner, Dora Toth, Annika Knoll, Simone Loesch, Gloria S. Benson, Sandra Schmitt from the Department of Geriatric Psychiatry, Central Institute of Mental Health for their valuable support and help in the recruitment of participants and collection of data. We would also like to thank assistant psychologists of the Neuropsychology Department of the UniKlinik Köln for their assistance in the recruitment of study participants and collection of data.

Simge Celik played lead role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, and writing of original draft. Oezguer Onur played lead role in resources and equal role in project administration and writing of review and editing. Görsev Yener played equal role in project administration, resources, and writing of review and editing. Josef Kessler played supporting role in project administration and resources and equal role in writing of review and editing. Yagmur Özbek played equal role in resources and writing of review and editing. Patric Meyer played lead role in supervision and equal role in conceptualization, methodology and writing of review and editing. Lutz Frölich played lead role in resources and supervision, supporting role in investigation and equal role in conceptualization, methodology, project administration, and writing of review and editing. Birgit Teichmann played lead role in funding acquisition and supervision, supporting role in project administration and equal role in conceptualization, methodology and writing of review and editing.

Correspondence concerning this article should be addressed to Simge Celik, Network Aging Research, Heidelberg University, Bergheimer Strasse 20, 69115, Heidelberg, Germany. Email: celik@nar.uni-heidelberg.de

Key Points

Question: This study examined the influence of cultural background and immigration-related experiences on the Mini-Mental State Examination (MMSE) and Rowland Universal Dementia Assessment Scale (RUDAS) by comparisons of Turkish immigrants and nonimmigrant, monolingual Turkish and German patients diagnosed with Alzheimer's disease (AD). ***Findings:*** Although the educational level and second language knowledge of Turkish immigrants diagnosed with AD influenced performance on the MMSE, no effects due to educational background and immigration-related experiences were observed on the performance of the culturally appropriate screening tool, RUDAS. ***Importance:*** These findings highlight the need to consider the cross-cultural potential of neuropsychological measures for the objective characterization of cognitive profiles in clinical and research settings, given the rapidly changing demographic characteristics of the population in Europe. ***Next Steps:*** Further research regarding neuropsychological measures appropriate for different cultural contexts is needed and future research should further elaborate the role of the educational and cultural background on cognitive test performance.

Keywords: immigration, aging, racial/ethnic minorities, neuropsychology, cognitive assessment

Supplemental materials: https://doi.org/10.1037/neu0000764.supp

Cognitive screening tools are a crucial component of the diagnostic process in dementia as they are used for early detection of cognitive impairment as a first and brief step toward accurate diagnosis. Currently, the most widely used cognitive screening tool for dementia in clinical and research settings worldwide is the Mini-Mental State Examination (MMSE; Folstein et al., 1975; Sheehan, 2012) due to its ease of use and rapid administration. However, the use of MMSE in culturally and linguistically diverse older adults poses challenges for the diagnostic accuracy of dementia detection as numerous studies have demonstrated that performance on the MMSE is influenced by age, educational level, ethnicity, and the language of test administration as well as immigrants' limited proficiency in the host country's language (Escobar et al., 1986; Espino et al., 2001; Jones & Gallo, 2002; Murden et al., 1991; Parker & Philp, 2004; Ramirez et al., 2006). Of these influencing factors, ethnicity has shown increasing importance in studies assessing performance on a variety of cognitive measures due to its effect on the acquisition of certain cognitive abilities and differences in the sociodemographic characteristics of ethnic groups.

The Rowland Universal Dementia Assessment Scale (RUDAS; Storey et al., 2004) was specifically designed to be used for dementia screening in multicultural populations. It has been shown to be less influenced by language, age, gender, and education than the MMSE when employed in culturally and linguistically diverse populations with limited or no formal education (Nielsen et al., 2012, 2013; Nielsen & Jørgensen, 2020). These findings render the RUDAS as a potential alternative to the MMSE for use in culturally and linguistically diverse individuals in clinical and research settings.

A growing body of evidence suggests that experiential factors, such as length of residence in a new country, acculturation, and bilingualism, may also impact neuropsychological test performance (Boone et al., 2007; Gasquoine, 1999; Gasquoine et al., 2007; Manly et al., 2004). The Turkish community living in Germany, which comprises the largest ethnic minority group in Germany (Bongard et al., 2002), is a particularly important population to examine, not only in terms of the effect of cultural and demographic factors but also of experiential variables on cognitive screening instruments. The educational level and literacy skills vary considerably, among this population, particularly, in the first-generation Turkish immigrants. The majority of them have lower levels of formal education than the native German population and their counterparts living in Turkey. This is a result of German Labor Immigration Act from the 1960s and 1970s, which encouraged unskilled or semiskilled workers to migrate to Germany (Söhn & Özcan, 2006). Furthermore, due to long-term residence in Germany and reduced exposure to their first language, first-generation Turkish immigrants tend to mix languages by inserting German words in the Turkish discourse (Yagmur, 2009) and they may experience first language attrition. Therefore, many of them may have limited language proficiency and fluency in both Turkish and German. In addition to the education and language profile, Turkish immigrants in Germany were shown to have greater levels of identification with people of their own ethnicity and lower levels of identification with the mainstream culture than their counterparts living in Australia or the Netherlands due to differences in the multicultural ideology of these countries (Yagmur & van de Vijver, 2012). Several studies have shown that lower levels of acculturation are a significant predictor for diminished cognitive test performance in different cultural groups (Al-Jawahiri & Nielsen, 2020; Coffey et al., 2005; Kennepohl et al., 2004), specifically on tests relying more on verbal items (Acevedo et al., 2007; Razani et al., 2007). Therefore, assessment of acculturation, particularly, in this population, may provide an insight into underlying reasons for differences in cognitive test performance.

Research exploring the influence of cultural background and immigration-related factors (acculturation, bilingualism) on the performance of cognitive screening tools is scarce in Europe. Studies available for immigrant groups in Europe have been based primarily on non-clinical samples without comparison groups (Krist et al., 2019; Nielsen et al., 2012) or samples with a mixed cultural and linguistic background (Goudsmit et al., 2018). There are currently no studies conducted in Europe that examine the influence of these factors on the performance of cognitive screening tools, especially in a clinical sample (Krist et al., 2019). To our knowledge, no study has been conducted on the comparison of test performance in the native-born and immigrant individuals of the same cultural background and native-born individuals of the host

culture. Therefore, given the sociocultural characteristics of older Turkish older immigrants in Germany, the primary aim of this study was to build on the scant evidence regarding the impact of culture, demographic, and immigration-related factors on the performance of two cognitive screening tools: the MMSE and RUDAS. Performance on these tests was examined in three groups of clinically diagnosed Alzheimer's dementia patients (AD); native born-monolingual Turkish older adults, Turkish immigrants living in Germany and native-born, monolingual German older adults. The rationale for the inclusion of patients with AD was to assess whether any influence of education, acculturation, and bilingualism on cognitive screening tools was preserved in the presence of a clinical condition and to determine performance differences in the early manifestation of disease in individuals from different cultural backgrounds. The secondary aim of this study was to examine performance differences in the individual items of the MMSE and RUDAS to gain insight into the item performance across groups.

Based on the available findings, it was hypothesized that (a) native-born Germans with AD (GER-AD) would perform better on the MMSE than Turkish immigrant (TR-IM-AD) and native-born Turkish patients with AD (TR-AD), the difference in performance being due to the development process of this test in a Western/European context. It was also hypothesized that TR-IM-AD would perform worse than a native-born Turkish group with AD due to immigration-related experiences, and that (b) performance differences on the RUDAS would be less pronounced across groups and be less influenced by education, gender, age, acculturation, and bilingualism.

Method

Participants

A total of 65 community-dwelling individuals, with a diagnosis of AD-type dementia (50 years and older), participated in the study and were categorized into three groups, namely, monolingual TR-AD ($n = 24$) and GER-AD ($n = 20$) without immigration background and TR-IM-AD participants ($n = 21$). The participants were a cohort of older adults who were initially seen for clinical evaluation due to memory complaints and other cognitive deficits at the clinics of the Central Institute of Mental Health Mannheim, University of Cologne, Department of Neurology, and Dokuz Eylul University Hospital. They were either recruited through referrals from the neurologists or neurology clinic/old age psychiatry service databases of hospitals. TR-IM-AD and GER-AD participants were recruited from Mannheim and Cologne, where a large population of Turkish immigrants live and TR-AD patients were recruited from Izmir, Turkey. All patients were diagnosed according to the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS–ADRDA) criteria for probable AD (McKhann et al., 1984) applied by an experienced psychiatrist or neurologist. Additionally, the guidelines of the German Society of Neurology (DGN) and the German Society of Psychiatry, Psychotherapy, and Neurology (DGGPN) were used in Germany as a diagnostic workup for dementia. The diagnosis of AD was based on a comprehensive medical, neurological, psychiatric and neuropsychological examination, routine laboratory tests, and MRI (T1, T2 and fluid-attenuated inversion recovery sequences).

Inclusion criteria for GER-AD and TR-AD participants in this study were (a) being of German or Turkish descent with a diagnosis of AD, (b) a global Clinical Dementia Rating Scale (CDR) score of 0.5 or 1, (c) German and Turkish as primary languages with no or basic second language knowledge, and (d) no residence in a foreign country for more than 6 months. Turkish immigrants were eligible to participate in the study if the range of their global CDR score was 0.5–1.0 and they were immigrants from Turkey and have lived in Germany for a minimum of 10 years. All participants were carefully screened for the following factors recorded in the databases and were excluded based on a current or history of neurological diseases (other than AD), delirium, current alcohol and/or substance dependence, severe current psychiatric disorder (e.g., major depression or psychosis at the time of assessment), current known physical impairments that could interfere with test performance (e.g., movement disorders, severe hearing impairment, or vision problems). Inability to read and write and the score on the 15-item Geriatric Depression Scale (GDS-15; Sheikh & Yesavage, 1986) also served as exclusionary criteria for each group.

Instruments

MMSE

The MMSE has a maximum score of 30 points and consists of items assessing various cognitive domains including orientation for time and place (e.g., date, season, location), registration (repetition of three items until they are learned), attention and calculation (e.g., serial subtraction of 7's, spelling "WORLD" backward), memory (recall of the three repeated items), language (e.g., naming common objects, reading, writing), and praxis (copying intersecting pentagons). Standardized German (Kessler et al., 1990) and Turkish (Güngen et al., 2002) versions of the MMSE were used in this study.

RUDAS

The RUDAS assesses six cognitive domains, namely, body orientation, praxis, visuoconstructional drawing, judgment, memory (delayed recall of items presented earlier), and category fluency (animals), and gives a single score with a maximum of 30. The validated Turkish (Ayan et al., 2019) and German (Nielsen et al., 2019) versions of the RUDAS were employed in this study.

GDS-15

Turkish (Durmaz et al., 2018) and German (Gauggel & Birkner, 1999) version of the GDS-15 were reported to be a valid and reliable tool for assessing depression when a cut-off score of 5 and 6 points were used respectively for older Turkish and German populations. The GDS-15 was administered to eliminate clinically significant symptoms of depression in our study population. A score >9 as a cut-off for moderate to severe depression was used in this study due to the prevalence of depressive symptoms reported in patients with AD (Gatz et al., 2005; Holmes et al., 1998; Holtzer et al., 2005), as well as in immigrant populations (Aichberger et al., 2012), particularly among first-generation Turkish immigrants living in Germany (Beutel et al., 2016).

CDR

The CDR (Morris, 1993) is an internationally and widely used scale to stage cognitive decline and dementia in six domains of functioning through a semistructured interview of the patient and informant. Each domain (except for the personal care domain) is scored on a 5-point scale of functioning with 0 = no; 0.5 = questionable; 1 = mild; 2 = moderate; and 3 = severe impairment. The global CDR score which is computed by entering the domain scores into the Washington University online algorithm (http://www.biostat.wustl.edu/adrc/cdrpgm/index.html) was used to characterize the severity of dementia in this study. The CDR Sum of the Boxes (CDR-SoB, range 0–18) was also calculated to provide more information on the severity score of impairment (Hughes et al., 1982).

LSBQ

The LSBQ (Anderson et al., 2018) is a comprehensive self-report bilingualism measure that allows for the computation of a composite score for the classification of bilingual and monolingual participants in culturally diverse populations. Bilingualism was characterized based on the composite score obtained from the participant's evaluations of the community language use behavior, frequency of use for each language and proficiency in their native and second language. The items of the LSBQ were translated and back-translated from English to Turkish or German by bilingual researchers with the removal of some items irrelevant to the older population (e.g., language used with grandparents, parents' education). Therefore, data regarding "language used with grandparents" were not entered into the spreadsheet for the composite score calculation. As suggested by Anderson et al., (2018), a composite score of less than −3.13 is categorized as monolinguals while the score above 1.23 is classified as bilinguals and scores that range between −3.12 and 1.22 may reflect unspecific language backgrounds that cannot be categorized as monolingual or bilingual, such as is the case in receptive bilinguals.

The Frankfurt Acculturation Scale

The Frankfurt Acculturation Scale (FRAKK; Bongard et al., 2007) is a 20-item self-report questionnaire for measuring the acculturation level of adult immigrants living in Germany. The scale evaluates two dimensions of acculturation, categorized as *Orientation towards Culture of Origin* (CO) and *Orientation towards Host Culture* (HC). Each dimension has values between 0 and 60 and consists of 10 items regarding media preference, language use, interethnic social networks, and ethnic identity which are rated on a seven-point Likert scale ranging from 0 = *absolutely* to 6 = *absolutely not*. The scores obtained from CO and HC yield an Assimilation Index (AI), which represents the degree of an individual's socio-cultural adjustment to a host culture. The availability of the FRAKK in the Turkish language and its use in the Turkish immigrant population living in Germany (Bongard et al., 2002; Reiss et al., 2015) made this instrument feasible to be employed for this study.

Design and Procedure

All patients who met the preliminary inclusion criteria (being of German or Turkish descent with a diagnosis of AD, being aged 50 years or older, no exclusion criteria) after the screening of databases were contacted via a telephone call by the research team. Participants were given information about the study aim and procedure. If the participant was interested, a structured phone interview was conducted to obtain more information on language, immigration background, and functional status of patients. Patients who were eligible for study participation were given an appointment at the hospital.

All participants underwent an approximately 120-min assessment procedure including examiner-administered demographic background, GDS, and CDR. TR-IM-AD patients were additionally administered acculturation and bilingualism questionnaires (FRAKK and LSBQ). Since none of the TR-AD participants reported any other additional language knowledge during the testing session, the language background questionnaire was not administered after conducting the MMSE and RUDAS to minimize the cognitive fatigue associated with the testing procedure. All participants were tested individually in a quiet room provided by the participating hospitals and all tests and questionnaires were administered by experienced Turkish and German neuropsychologists. Responses to the questionnaires and CDR were obtained with the help of a close family member of the patient, if necessary. TR-IM-AD was assessed in their language of preference and all of them chose to be tested in Turkish. The order of administration of the MMSE and RUDAS was counterbalanced.

Prior to data collection, this study was preregistered on the DRKS—German Clinical Trials Register (DRKS-ID: DRKS00017380). The study was completed in accordance with the Helsinki Declaration and approved by the ethics committee at the Faculty of Behavioral and Cultural Studies of Heidelberg University (AZ Cel 2018 1/1-A1), the Medical Faculty of the University of Cologne (19-1242_1), and Medical Faculty of the Dokuz Eylul University (2019/17–12). All participants provided signed informed consent and in case of reduced capacity to consent, informed consent by a proxy was obtained.

Statistical Analyses

Group differences in the distribution of categorical variables including gender and global CDR scores were assessed using the Pearson's chi-square tests. For continuous variables (i.e., total MMSE and RUDAS scores), where the assumption of normality (Shapiro–Wilk test) and homogeneity of variances were met, separate one-way analysis of variance (ANOVA) and analysis of covariance (ANCOVA) tests were conducted to compare mean differences in the three groups while controlling for demographic factors when appropriate. Significant ANOVA test results were then further examined using pair-wise comparisons (Fisher's Least Significant Difference; LSD). In case of an absence of normality for some variables (age, years of education, and GDS-score), nonparametric Kruskal–Wallis tests, with pairwise comparisons using Dunn (1964) procedure with a Bonferroni correction were run to investigate whether the three groups differed across those variables. Kruskal–Wallis tests were also performed to explore the pattern of performance differences in the individual and continuous items of the MMSE and RUDAS between the three groups. Effect size values were calculated and reported by using η_p^2, Cohen's d for parametric and r for nonparametric analyses. Chi-square tests were performed for the items that were scored categorically on the MMSE (correct vs. incorrect), when the expected cell count was ≥5, and Fisher's Exact test, when the expected cell count was <5.

Lastly, Spearman's rank-order and point-biserial correlations (gender) were conducted separately in each group to investigate the associations between demographic, immigration-related variables and total scores of the MMSE and RUDAS. Additionally, a linear regression was run to examine how much of the variation in test performance was explained by years of education when a linear relationship between the variables was indicated. All tests were two-tailed, and a p-value $< .05$ was considered to be statistically significant. Statistical analyses were performed using the SPSS for Windows Version 25.0 (IBM SPSS, Chicago, IL, USA).

Results

Participant Characteristics

The sociodemographic and clinical characteristics of each group are summarized in Tables 1 and 2. There were significant differences in years of education among groups, $H(2) = 11.47$, $p = .003$. Pairwise comparisons with adjusted p-values revealed that TR-IM-AD had fewer years of education than GER-AD patients ($r = .52$), but there were no significant differences in years of education between TR-AD and GER-AD ($r = .34$) groups and TR-IM-AD and TR-AD ($r = .17$) participants. Overall, 41.7% ($n = 10$) and 66.8% of participants ($n = 14$) had ≤5 years of education in TR-AD and TR-IM-AD groups, respectively, whereas the GER-AD group received at least 8 years of education. With regard to the level of depressive symptoms, there were significant differences in the distribution of GDS-scores among groups, $H(2) = 6.84$, $p = .033$. However, pairwise comparisons with Bonferroni adjustment indicated that groups did not differ significantly in GDS-scores (GER-AD vs. TR-AD: $r = .33$; GER-AD vs. TR-IM-AD: $r = .36$; TR-AD vs. TR-IM-AD: $r = .03$). Age, distribution of sex, $\chi^2(2) = .345$, $p = .841$, and dementia severity as measured by global-CDR scores, $\chi^2(2) = .811$, $p = .667$, were comparable across groups.

Group Differences in Total Scores and Individual Items of the MMSE and RUDAS

The detailed results for performance differences in terms of total MMSE and RUDAS scores are presented in Table 3. One-way ANOVA results revealed that the main effect of group was significant for total MMSE score, $F(2, 62) = 4.213$, $p = .019$, $\eta_p^2 = .120$. Post-hoc analysis indicated that both TR-AD, $p = .033$, $d = .73$, 95% CI [−5.38, −.23], and TR-IM-AD, $p = .007$, $d = .86$, 95% CI [−6.35, −1.04], groups performed significantly worse than GER-AD patients on the MMSE, whereas TR-AD and TR-IM-AD groups had a comparable performance on the MMSE, $p = .488$, $d = .19$, 95% CI [−1.65, 3.43]. The subsequent ANOVA on the total RUDAS score revealed no main effect of the group, indicating that three groups had a similar performance on the RUDAS, $F(2, 62) = .542$, $p = .584$, $\eta_p^2 = .017$.

After adjustment for years of education, differences in MMSE performance were no longer significant between groups, $F(2, 61) = 1.16$, $p = .319$, $\eta_p^2 = .037$, and years of education were significantly associated with the total MMSE score, $F(1, 61) = 16.80$, $p < .001$, $\eta_p^2 = .216$. The results remained unchanged in total RUDAS scores between groups after controlling for years of education, $F(2, 61) = .479$, $p = .622$, $\eta_p^2 = .015$.

Significant performance differences in individual items of the MMSE and RUDAS across groups are shown in Table 3, while additional information on these results is shown in Figure S1 and Table S1.

Table 1
Sociodemographic and Clinical Characteristics of Each AD Group

Variables	TR-AD ($n = 24$)	TR-IM-AD ($n = 21$)	GER-AD ($n = 20$)	p value	Group differences
Age, years[a]					
M (*SD*)	70.33 (7.73)	71.62 (7.41)	74.70 (7.50)		
Median (range)	70.0 (55–85)	74.0 (57–83)	77.0 (56–85)	.090	
Sex[b]					
Female (%)	14 (58%)	12 (57%)	10 (50%)	.841	
Education, years[a]					
M (*SD*)	8.13 (4.11)	7.05 (4.44)	11.00 (4.09)		
Median (range)	7.0 (0–16)	5.0 (0–17)	9.5 (8–21)	**.003**	GER-AD > TR-IM-AD
GDS score[a]					
Mean (*SD*)	3.58 (1.58)	4.14 (2.85)	2.25 (2.09)		
Median (range)	3.0 (1–6)	3.0 (0–8)	1.5 (0–7)	**.033**	NS
Global CDR score[b]					
0.5 (%)	12 (50%)	13 (62%)	10 (50%)	.667	
1 (%)	12 (50%)	8 (38%)	10 (50%)		
CDR-SoB[c]					
M (*SD*)	4.91 (2.05)	4.07 (1.85)	3.94 (1.94)	.210	
BL Com score					
M (*SD*)	Not Available	−.33 (5.96)	−5.58 (1.10)	NA	

Note. AD = Alzheimer's disease; TR-AD = Native-born Turkish participants with AD; TR-IM-AD = Turkish immigrants with AD; GER-AD = Native-born German participants with AD; GDS = Geriatric Depression Scale; CDR-SoB = Clinical Dementia Rating Scale Sum of Boxes, BL Com score = Bilingualism Composite Score; NA = Not Applicable; NS = Not Significant in Pairwise Comparisons; ANOVA = analysis of variance. Significant p values are marked in bold. CDR-SoB scores are not available for two participants in the GER-AD group.
[a] Kruskal–Wallis Test. [b] Pearson's chi-square test. [c] One-way ANOVA.

Table 2

Immigration-Related Characteristics of Turkish Immigrants Diagnosed With AD and Their Spearman Correlation Coefficients With MMSE and RUDAS Total Scores

Variables	*M* (*SD*)	Min–Max	MMSE (r_s)	RUDAS (r_s)
Age at immigration (years)	25.24 (7.3)	14–43	.07	.11
Length of residence in Germany (years)	46.38 (7.9)	31–60	−.08	−.14
CO	43.80 (8.9)	24–56	−.08	.18
HC	38.30 (10.8)	12–52	.12	−.41
AI	54.50 (17.3)	16–79	.06	−.42

Sub-group classification of Turkish immigrants with AD based on the acculturation and bilingualism scores

		MMSE *M* (*SD*)	r_s	RUDAS *M* (*SD*)	r_s
Acculturation	Low (n = 10)	17.60 (4.6)	.04	21.20 (4.2)	−.21
	High (n = 10)	18.40 (5.8)		19.60 (3.9)	
Language groups	Monolingual (n = 9)	15.78 (4.3)	.52*	21.11 (4.2)	−.15
	Bilingual (n = 6)	21.83 (5.4)		19.67 (5.1)	

Note. AD = AD = Alzheimer's disease; CO = orientation toward culture of origin; HC = orientation toward host culture; AI = assimilation index; FRAKK = Frankfurt Acculturation Scale; MMSE = Mini-Mental State Examination; RUDAS = Rowland Universal Dementia Assessment Scale. CO, HC, and AI scores are available for 20 participants as one patient did not fill in the FRAKK questionnaire.
* $p < .05$.

Factors Associated With Performance on the MMSE and RUDAS

The associations between demographic and clinical variables, and total MMSE and RUDAS scores in each group revealed that lower global-CDR scores were significantly related to higher total MMSE scores in the three groups (TR-AD: $r_s = -.76$, $p < .001$; TR-AD-IM: $r_s = -.55$, $p = .011$; GER-AD: $r_s = -.64$, $p = .002$). On the other hand, lower global-CDR scores were significantly associated with higher total RUDAS scores in TR-AD ($r_s = -.68$, $p < .001$) and GER-AD ($r_s = -.53$, $p = .015$) groups, whereas there was no correlation between these variables in TR-IM-AD group ($r_s = -.29$, $p = .197$). Furthermore, performance on the MMSE was strongly correlated with years of education in the TR-AD-IM group ($r_s = .75$, $p < .001$), whereas there was no significant relationship between years of education and total MMSE score in the TR-AD ($r_s = .31$, $p = .140$) and GER-AD ($r_s = -.054$, $p = .820$) groups. A linear regression indicated that years of education significantly predicted total MMSE score in the TR-IM-AD group, $F(1.19) = 32.46$, $p < .001$ and it accounted for 63.1% of the variation in total MMSE score with adjusted $R^2 = 61.1\%$. Correlations between years of education and total RUDAS score did not reach significance for any of the groups (TR-AD: $r_s = .15$, $p = .487$; TR-AD-IM: $r_s = .15$, $p = .526$; GER-AD: $r_s = -.05$, $p = .819$). Correlation coefficients showed no significant relationship between age, gender, or GDS-score and total MMSE and RUDAS scores.

In terms of immigration-related characteristics, the TR-IM-AD group was categorized into two groups based on AI and bilingualism composite scores. The scores suggested by Anderson et al. (2018) and the median value of ≤59.50 were used as cut-offs for the classification of language groups and low versus high acculturation, respectively. A significant association was only evident for bilingualism and total MMSE score, with bilinguals performing better than monolinguals ($p = .046$). This performance difference remained significant after adjustment for CDR-SoB scores, $F(1, 12) = 5.248$, $p = .041$, $\eta_p^2 = .304$. When bilinguals were excluded from the analyses, one-way ANOVA results revealed that the main effect of group was significant for the total MMSE score, $F(2, 50) = 7.371$, $p = .002$, $\eta_p^2 = .228$, with both TR-AD ($p = .021$, $d = .73$, 95% CI [−5.18, −.44] and TR-IM-AD ($p < .001$, $d = 1.60$, 95% CI [−8.96, −2.68] groups performing significantly worse than GER-AD patients. Furthermore, differences in the total MMSE score between TR-AD and TR-IM-AD groups approached, but did not reach statistical significance ($p = .053$, $d = .71$, 95% CI [−.04, 6.07]. After adjustment for years of education, the MMSE performance difference between the TR-IM-AD and GER-AD groups remained significant, $F(2, 49) = 3.731$, $p = .031$, $\eta_p^2 = .132$, with the GER-AD group performing better than the TR-IM-AD participants ($p = .028$, 95% CI [.381, 8.452]). The level of acculturation (MMSE: $p = .855$, RUDAS: $p = .355$) and other immigration-related variables were not related to total MMSE or RUDAS scores. Correlations between immigration-related variables and total MMSE and RUDAS scores in the TR-IM-AD group are presented in Table 2.

Discussion

The results of the study provided partial support regarding the first hypothesis with poorer performance on the MMSE for the TR-AD and TR-IM-AD groups compared to GER-AD patients. The findings also lent support to the second hypothesis, which proposed that there would be no performance differences in the RUDAS among the groups, and no effect of demographic, cultural background, and immigration-related variables on this test. However, contrary to our expectations, significant differences in the MMSE performance were no longer observed after accounting for years of education. The cultural background had no independent effect on the total MMSE score, particularly in the comparison of nonimmigrant and monolingual groups such as the TR-AD and GER-AD participants. Furthermore, variables associated with immigration experience did

Table 3

Between-Groups Comparisons of MMSE and RUDAS

Tests/items	Median (range)/*M* (*SD*) for each group: TR-AD (n = 24)	TR-IM-AD (n = 21)	GER-AD (n = 20)	*p* value	*p* value for group comparisons: TR-AD vs. TR-IM-AD	TR-AD vs. GER-AD	TR-IM-AD vs. GER-AD
MMSE total score[a]	18.79 (4.19)	17.90 (5.06)	21.60 (3.30)	**.019**	.488	**.033**	**.007**
Adjusted *M* (*SE*)[b]	19.04 (.778)	18.66 (.849)	20.50 (.891)	.319	—	—	—
Attention, calculation[c]	3.0 (0–5)	1.0 (0–5)	4.5 (1–5)	**.003**	.520	.087	**.002**
Mean (*SD*)	2.88 (1.78)	1.95 (2.04)	4.05 (1.23)				
Phrase repetition[d] (correct, %)	16 (66.7%)	9 (42.9%)	17 (85%)	**.018**	NS	NS	**<.05**
Constructional praxis[d] (correct, %)	13 (54.2%)	5 (23.8%)	13 (65%)	**.022**	**<.05**	NS	**<.05**
RUDAS total score[a]	19.63 (4.61)	20.62 (4.09)	20.75 (2.84)	.584	—	—	—
Adjusted *M* (*SE*)[b]	19.69 (.812)	20.82 (.887)	20.45 (.930)	.622	—	—	—
Judgment[c]	2.0 (0–4)	2.0 (0–4)	3.0 (0–4)	**.012**	1.00	.064	**.015**
M (*SD*)	1.92 (1.28)	1.76 (.995)	2.75 (1.12)				

Note. TR-AD = Native-born Turkish participants with AD; TR-IM-AD = Turkish immigrants with AD; GER-AD = Native-born German participants with AD; NS = Not Significant; ANOVA = analysis of variance; ANCOVA = analysis of covariance; MMSE = Mini-Mental State Examination; RUDAS = Rowland Universal Dementia Assessment Scale. Significant *p* values are marked in bold.
[a] One-way ANOVA. [b] One-way ANCOVA (education corrected). [c] Kruskal–Wallis test. [d] Pearson's chi-square test, significant test results were followed-up by pairwise comparisons using the *z*-test of two proportions with a Bonferroni correction.

not result in performance differences in the MMSE, in comparisons between the TR-AD and TR-IM groups.

The finding regarding the comparable MMSE performance between native-born cultural groups (TR-AD and GER-AD participants), after adjustment for education, is consistent with the results of a few other studies. Studies comparing different native-born cultural groups with or without clinically diagnosed dementia have demonstrated that educational level accounted for cultural-group differences (Dodge et al., 2009; Ng et al., 2007; Salmon et al., 1989). For instance, in a cross-cultural comparison of MMSE performance between Chinese and Finnish older adults, MMSE scores were similar between groups after exclusion of a subgroup of Chinese participants with no formal education from the analysis (Salmon et al., 1989).

The result regarding the equivalent MMSE performance between the TR-IM-AD and GER-AD groups, after adjustment for educational level, complements previous studies of the immigrant and host-cultural group or English and Spanish-speaking individual comparisons (Lindesay et al., 1997; Loewenstein et al., 1993; Mungas et al., 1996; Murden et al., 1991). Those studies have found no significant MMSE performance differences between groups when the educational level was accounted for. Specifically, in a study comparing community-dwelling older immigrant Gujarati adults living in the U.K. and native-born British participants, differences in total MMSE scores across groups were mainly attributed to the effects of educational level and age, instead of differences related to the cultural background (Lindesay et al., 1997). However, the issue of the extent to which cultural background impacts MMSE performance is not well-understood in the scientific literature (Leveille et al., 1998), as some other studies indicated an association between MMSE performance and cultural background in those groups, even after adjustment for education (Bohnstedt et al., 1994; Welsh et al., 1995). The discrepancy in findings can partly be explained by recruitment of immigrant participants in host-cultural groups, considerable cultural heterogeneity within groups, bilingual status, and acculturation of older adult immigrants. The interesting result from the comparison of the three groups supports one of these explanations by indicating significant group differences in the education-adjusted MMSE total scores when the immigrant patients with AD had limited proficiency in the language of the host country. With regard to acculturation, in contrast to the study conducted with nondemented Turkish older adults in Denmark (Nielsen et al., 2012), no correlation was found between acculturation and MMSE score in the TR-IM-AD group. The lack of correlation between these variables can be attributed to differences in acculturation measures employed in the studies. A bidimensional instrument that independently assessed acculturative change in each culture was used in our study, as opposed to the unidimensional measure employed by Nielsen et al. (2012). However, language proficiency as a strong component of acculturative change (Coronado et al., 2005; Thomson & Hoffman-Goetz, 2009) was associated with better MMSE performance in this group, suggesting that language-related acculturation is a prominent factor in cognitive screening test performance. This finding is in line with a study conducted with Mexican American older adults, indicating that language-related acculturation, more than identity-related acculturation (e.g., self-identity, traditions, social interactions), triggers acculturative changes in cognitive performance (Martinez-Miller et al., 2020).

In our study, no performance difference in the MMSE was observed between TR-AD and TR-IM groups with and without adjustment for years of education. This result is in contrast with the results of a study conducted with healthy native-born Greek older adults and Greek immigrants living in Australia, which indicated significant differences in MMSE performance between the groups, even after adjustment for educational level (Plitas et al., 2009). However, as suggested by Plitas et al. (2009), second language proficiency mediated the MMSE performance of the immigrant group in our study, since the exclusion of bilinguals from TR-IM-AD group increased performance differences among the groups on this test. It is likely that the small sample size in our study resulted in reduced power to detect between-group differences, as performance difference on the MMSE across groups approached, but failed to reach statistical significance. Regarding the RUDAS performance of the three groups with AD, the finding that performance on the RUDAS was not influenced by educational level is consistent with previous studies that included low-educated participant groups with and without dementia (Ayan et al., 2019; Goudsmit et al., 2018; Mateos-Álvarez et al., 2017) and multicultural populations (Basic et al., 2009; Goudsmit et al., 2018; Rowland et al., 2006). Moreover, in line with previous studies conducted with German (Nielsen et al., 2019) and Turkish (Ayan et al., 2019) healthy older adults and dementia patients, this study revealed no effect of immigration status on the RUDAS, as assessed by the comparison of the TR-AD and TR-IM-AD groups. However, the absence of correlations between global-CDR and total RUDAS scores that was only observed in the TR-IM-AD group contradicts previous studies reporting an association between these variables (Goudsmit et al., 2018; Mateos-Álvarez et al., 2017). The reason for this rather contradictory result is not entirely clear.

The lower scores in "attention and calculation," "phrase repetition," and "constructional praxis" items of the MMSE in TR-IM-AD group are consistent with previous studies investigating the pattern of performance differences in the individual items of the MMSE in Turkish (Nielsen et al., 2012), African-Caribbean (Stewart et al., 2002), non-Western older adult immigrant groups with lower levels of education (Goudsmit et al., 2018). Several studies examining differential item functioning of individual MMSE items have shown that these items were performed poorly by individuals of different ethnic groups with lower educational levels and non-English backgrounds (Escobar et al., 1986; Jones & Gallo, 2002; Teresi et al., 1995).

Taken together, the results of this study extend prior findings regarding the usefulness of the RUDAS in culturally and linguistically diverse populations and individuals with lower educational levels. This study also provides evidence that the RUDAS can be a valuable instrument in individuals with varying degrees of acculturation, bilingualism, and a better alternative to the MMSE in clinical practice and research. On the other hand, total MMSE score and individual items of the MMSE may be influenced by educational and language background, rather than cultural differences, between German and Turkish patients with AD. However, the small sample size in this study warrants caution in the interpretation of the results and further investigation on larger samples, composed of individuals from homogenous cultural backgrounds, are needed to ascertain the validity of these findings. Despite efforts in several cities in Germany, the recruitment of Turkish immigrant patients was a major challenge due to the stigma attached to mental-health research and the delay in seeking treatment resulting from perceptions of AD as being a natural part of the aging process or as a form of insanity (Nielsen & Waldemar, 2016).

The use of the MMSE in people from the diverse cultural background who have low educational attainment and/or limited proficiency in the language of the host country may create an important hurdle for the interpretation of cognitive results and may have a detrimental impact on the treatment of AD as, for instance, in Belgium, reimbursement for medication is terminated depending on the MMSE score obtained by a patient (Segers et al., 2013). Since low levels of education, particularly, six or fewer years of education, have been shown to be linked to a greater risk for AD in different cultural groups (Harmanci et al., 2003; Harwood et al., 1999; McDowell et al., 1994), lower levels of education should not be considered as a confounding variable in this study.

Although years of formal education was taken into account in this study, recent cross-sectional studies have emphasized the role of the quality of education on the neuropsychological test performance of individuals with diverse cultural and linguistic background after accounting for years of education (Manly et al., 2002; Sachs-Ericsson & Blazer, 2005; Sisco et al., 2015). Thus, quality of education, rather than the years of education, may have caused the differences in MMSE performance observed between the Turkish and German patients. Future studies may examine the educational level in older adults by measuring self-rated literacy (Kavé et al., 2012) and self-evaluated school performance (Mehta et al., 2009), which have been shown to contribute to the prediction of late-life cognition and AD. These measures may offer a practical indicator of education quality and standards (Piccinin et al., 2013) and would further enhance the studies on the role of education and cultural background in cognitive test performance.

Conclusions

The findings highlight the need for consideration of educational bias in the diagnosis of AD and the measurement of the progression of cognitive decline for the first-generation immigrants while employing cognitive screening tools that have been developed and normed in a Western/European context. Further, the results underscore the importance of characterizing the linguistic integration or bilingual status of the immigrants with early stages of AD for the objective evaluation of cognitive deficits, irrespective of the test administration in the primary language of the test taker. The present findings should be considered as preliminary and exploratory, considering that no other studies have examined the role of culture and immigration-related variables on test performance in the comparison of these three samples.

References

Acevedo, A., Loewenstein, D. A., Agrón, J., & Duara, R. (2007). Influence of sociodemographic variables on neuropsychological test performance in Spanish-speaking older adults. *Journal of Clinical and Experimental Neuropsychology*, *29*(5), 530–544. https://doi.org/10.1080/13803390600814740

Aichberger, M. C., Neuner, B., Hapke, U., Rapp, M. A., Schouler-Ocak, M., & Busch, M. A. (2012). Der Zusammenhang zwischen Migrationsstatus und depressiven Symptomen in der älteren Bevölkerung in Deutschland. *Psychiatrische Praxis*, *39*(3), 116–121. https://doi.org/10.1055/s-0031-1276936

Al-Jawahiri, F., & Nielsen, T. R. (2020). Effects of acculturation on the Cross-Cultural Neuropsychological Test Battery (CNTB) in a culturally and linguistically diverse population in Denmark. *Archives of Clinical Neuropsychology*. Advance online publication. https://doi.org/10.1093/arclin/acz083

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & 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

Ayan, G., Afacan, C., Poyraz, B. C., Bilgic, O., Avci, S., Yavuzer, H., Erdincler, D. S., Ayan, B., & Doventas, A. (2019). Reliability and validity of Rowland Universal Dementia Assessment Scale in Turkish population. *American Journal of Alzheimer's Disease and Other Dementias*, *34*(1), 34–40. https://doi.org/10.1177/1533317518802449

Basic, D., Khoo, A., Conforti, D., Rowland, J., Vrantsidis, F., Logiudice, D., Hill, K., Harry, J., Lucero, K., & Prowse, R. (2009). Rowland Universal Dementia Assessment Scale, Mini-mental state examination and general practitioner assessment of cognition in a multicultural cohort of community-dwelling older persons with early dementia. *Australian Psychologist*, *44*(1), 40–53. https://doi.org/10.1080/00050060802593593

Beutel, M. E., Jünger, C., Klein, E. M., Wild, P., Lackner, K. J., Blettner, M., Banerjee, M., Michal, M., Wiltink, J., & Brähler, E. (2016). Depression, anxiety and suicidal ideation among 1st and 2nd generation migrants—Results from the Gutenberg health study. *BMC Psychiatry*, *16*(1), Article 288. https://doi.org/10.1186/s12888-016-0995-2

Bohnstedt, M., Fox, P. J., & Kohatsu, N. D. (1994). Correlates of mini-mental status examination scores among elderly demented patients: The influence of race-ethnicity. *Journal of Clinical Epidemiology*, *47*(12), 1381–1387. https://doi.org/10.1016/0895-4356(94)90082-5

Bongard, S., Kelava, A., Sabic, M., Aazami-Gilan, D., & Kim, Y. B. (2007). Akkulturation und gesundheitliche Beschwerden bei drei Migrantenstichproben in Deutschland. In H. Eschenbeck, U. Heim-Dreger, & C. W. Kohlmann (Eds.), *Beiträge zur Gesundheitspsychologie* (Vol. 29, p. 53). Gmünder Hochschulreihe Band 29.

Bongard, S., Pogge, S. F., Arslaner, H., Rohrmann, S., & Hodapp, V. (2002). Acculturation and cardiovascular reactivity of second-generation Turkish migrants in Germany. *Journal of Psychosomatic Research*, *53*(3), 795–803. https://doi.org/10.1016/S0022-3999(02)00347-1

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, *22*(3), 355–365. https://doi.org/10.1016/j.acn.2007.01.010

Coffey, D. M., Marmol, L., Schock, L., & Adams, W. (2005). The influence of acculturation on the Wisconsin Card Sorting Test by Mexican Americans. *Archives of Clinical Neuropsychology*, *20*(6), 795–803. https://doi.org/10.1016/j.acn.2005.04.009

Coronado, G. D., Thompson, B., McLerran, D., Schwartz, S. M., & Koepsell, T. D. (2005). A short acculturation scale for Mexican-American populations. *Ethnicity & Disease*, *15*(1), 53–62.

Dodge, H. H., Meguro, K., Ishii, H., Yamaguchi, S., Saxton, J. A., & Ganguli, M. (2009). Cross-cultural comparisons of the Mini-Mental State Examination between Japanese and U.S. cohorts. *International Psychogeriatrics*, *21*(1), 113–122. https://doi.org/10.1017/S1041610208007886

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252. https://doi.org/10.1080/00401706.1964.10490181

Durmaz, B., Soysal, P., Ellidokuz, H., & Isik, A. T. (2018). Validity and reliability of geriatric depression scale-15 (short form) in Turkish older adults. *Northern Clinics of Istanbul*, *5*(3), 216–220. https://doi.org/10.14744/nci.2017.85047

Escobar, J. I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J., & Golding, J. M. (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity. Cultural and linguistic artifacts. *Journal of Nervous and Mental Disease*, *174*(10), 607–614. https://doi.org/10.1097/00005053-198610000-00005

Espino, D. V., Lichtenstein, M. J., Palmer, R. F., & Hazuda, H. P. (2001). Ethnic differences in Mini-Mental State Examination (MMSE) scores: Where you live makes a difference. *Journal of the American Geriatrics Society*, *49*(5), 538–548. https://doi.org/10.1046/j.1532-5415.2001.49111.x

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

Gasquoine, P. G. (1999). Variables moderating cultural and ethnic differences in neuropsychological assessment: The case of Hispanic Americans. *The Clinical Neuropsychologist*, *13*(3), 376–383. https://doi.org/10.1076/clin.13.3.376.1735

Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007). Language of administration and neuropsychological test performance in neurologically intact Hispanic American bilingual adults. *Archives of Clinical Neuropsychology*, *22*(8), 991–1001. https://doi.org/10.1016/j.acn.2007.08.003

Gatz, J. L., Tyas, S. L., St John, P., & Montgomery, P. (2005). Do depressive symptoms predict Alzheimer's disease and dementia? *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *60*(6), 744–747. https://doi.org/10.1093/gerona/60.6.744

Gauggel, S., & Birkner, B. (1999). Validität und Reliabilität einer deutschen Version der Geriatrischen Depressionsskala (GDS). *Zeitschrift für Klinische Psychologie und Psychotherapie*, *28*(1), 18–27. https://doi.org/10.1026//0084-5345.28.1.18

Goudsmit, M., van Campen, J., Schilt, T., Hinnen, C., Franzen, S., & Schmand, B. (2018). One size does not fit all: Comparative diagnostic accuracy of the Rowland Universal Dementia Assessment Scale and the Mini Mental State Examination in a memory clinic population with very low education. *Dementia and Geriatric Cognitive Disorders. Extra*, *8*(2), 290–305. https://doi.org/10.1159/000490174

Güngen, C., Ertan, T., Eker, E., Yaşar, R., & Engin, F. (2002). Reliability and validity of the standardized Mini Mental State Examination in the diagnosis of mild dementia in Turkish population. *Turkish Journal of Psychiatry*, *13*, 273–281.

Harmanci, H., Emre, M., Gurvit, H., Bilgic, B., Hanagasi, H., Gurol, E., Sahin, H., & Tinaz, S. (2003). Risk factors for Alzheimer disease: A population-based case-control study in Istanbul, Turkey. *Alzheimer Disease and Associated Disorders*, *17*(3), 139–145. https://doi.org/10.1097/00002093-200307000-00003

Harwood, D. G., Barker, W. W., Loewenstein, D. A., Ownby, R. L., St George-Hyslop, P., Mullan, M., & Duara, R. (1999). A cross-ethnic analysis of risk factors for AD in white Hispanics and white non-Hispanics. *Neurology*, *52*(3), 551–556. https://doi.org/10.1212/WNL.52.3.551

Holmes, C., Russ, C., Kirov, G., Aitchison, K. J., Powell, J. F., Collier, D. A., & Lovestone, S. (1998). Apolipoprotein E: Depressive illness, depressive symptoms, and Alzheimer's disease. *Biological Psychiatry*, *43*(3), 159–164. https://doi.org/10.1016/S0006-3223(97)00326-0

Holtzer, R., Scarmeas, N., Wegesin, D. J., Albert, M., Brandt, J., Dubois, B., Hadjigeorgiou, G. M., & Stern, Y. (2005). Depressive symptoms in Alzheimer's disease: Natural course and temporal relation to function and cognitive status. *Journal of the American Geriatrics Society*, *53*(12), 2083–2089. https://doi.org/10.1111/j.1532-5415.2005.00535.x

Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, *140*(6), 566–572. https://doi.org/10.1192/bjp.140.6.566

Jones, R. N., & Gallo, J. J. (2002). Education and sex differences in the Mini-Mental State Examination: Effects of differential item functioning. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *57*(6), 548–558. https://doi.org/10.1093/geronb/57.6.P548

Kavé, G., Shrira, A., Palgi, Y., Spalter, T., Ben-Ezra, M., & Shmotkin, D. (2012). Formal education level versus self-rated literacy as predictors of

cognitive aging. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *67*(6), 697–704. https://doi.org/10.1093/geronb/gbs031

Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American acculturation and neuropsychological test performance following traumatic brain injury. *Journal of the International Neuropsychological Society*, *10*(4), 566–577. https://doi.org/10.1017/S1355617704104128

Kessler, J., Markowitsch, H. J., & Denzler, P. (1990). *Mini-Mental-Status-Test (MMST)* [German Version]. Beltz Test GmbH.

Krist, L., Keller, T., Sebald, L., Yesil-Jürgens, R., Ellert, U., Reich, A., Becher, H., Heuschmann, P. U., Willich, S. N., Keil, T., & Liman, T. G. (2019). The Montreal Cognitive Assessment (MoCA) in a population-based sample of Turkish migrants living in Germany. *Aging & Mental Health*, *23*(1), 30–37. https://doi.org/10.1080/13607863.2017.1396577

Leveille, S. G., Guralnik, J. M., Ferrucci, L., Corti, M. C., Kasper, J., & Fried, L. P. (1998). Black/white differences in the relationship between MMSE scores and disability: The Women's Health and Aging Study. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *53B*(3), 201–208. https://doi.org/10.1093/geronb/53B.3.P201

Lindesay, J., Jagger, C., Mlynik-Szmid, A., Sinorwala, A., Peet, S., & Moledina, F. (1997). The Mini-Mental State Examination (MMSE) in an elderly immigrant Gujarati population in the United Kingdom. *International Journal of Geriatric Psychiatry*, *12*(12), 1155–1167. https://doi.org/10.1002/(SICI)1099-1166(199712)12:12<1155::AID-GPS705>3.0.CO;2-E

Loewenstein, D. A., Argüelles, T., Barker, W. W., & Duara, R. (1993). A comparative analysis of neuropsychological test performance of Spanish-speaking and English-speaking patients with Alzheimer's disease. *Journal of Gerontology*, *48*(3), 142–149. https://doi.org/10.1093/geronj/48.3.P142

Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*, *11*(1), 37–46. https://doi.org/10.1207/s15324826an1101_5

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, *8*(3), 341–348. https://doi.org/10.1017/S1355617702813157

Martinez-Miller, E. E., Robinson, W. R., Avery, C. L., Yang, Y. C., Haan, M. N., Prather, A. A., & Aiello, A. E. (2020). Longitudinal associations of US Acculturation with cognitive performance, cognitive impairment, and dementia. *American Journal of Epidemiology*, *189*(11), 1292–1305. https://doi.org/10.1093/aje/kwaa088

Mateos-Álvarez, R., Ramos-Ríos, R., & López-Moríñigo, J. D. (2017). Comparative analysis between the MMSE and the RUDAS for dementia screening in low educated people in a Spanish psychogeriatric clinic. *European Journal of Psychiatry*, *31*(3), 119–126. https://doi.org/10.1016/j.ejpsy.2017.06.003

McDowell, I., Hill, G., & Lindsay, J. (1994). The Canadian Study of Health and Aging: Risk factors for Alzheimer's disease in Canada. *Neurology*, *44*(11), 2073–2080. https://doi.org/10.1212/WNL.44.11.2073

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, *34*(7), 939–944. https://doi.org/10.1212/WNL.34.7.939

Mehta, K. M., Stewart, A. L., Langa, K. M., Yaffe, K., Moody-Ayers, S., Williams, B. A., & Covinsky, K. E. (2009). "Below average" self-assessed school performance and Alzheimer's disease in the Aging, Demographics, and Memory Study. *Alzheimer's & Dementia*, *5*(5), 380–387. https://doi.org/10.1016/j.jalz.2009.07.039

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

Mungas, D., Marshall, S. C., Weldon, M., Haan, M., & Reed, B. R. (1996). Age and education correction of Mini-Mental State Examination for English and Spanish-speaking elderly. *Neurology*, *46*(3), 700–706. https://doi.org/10.1212/WNL.46.3.700

Murden, R. A., McRae, T. D., Kaner, S., & Bucknam, M. E. (1991). Mini-Mental State exam scores vary with education in blacks and whites. *Journal of the American Geriatrics Society*, *39*(2), 149–155. https://doi.org/10.1111/j.1532-5415.1991.tb01617.x

Ng, T.-P., Niti, M., Chiam, P.-C., & Kua, E.-H. (2007). Ethnic and educational differences in cognitive test performance on Mini-Mental State Examination in Asians. *The American Journal of Geriatric Psychiatry*, *15*(2), 130–139. https://doi.org/10.1097/01.JGP.0000235710.17450.9a

Nielsen, T. R., Andersen, B. B., Gottrup, H., Lützhøft, J. H., Høgh, P., & Waldemar, G. (2013). Validation of the Rowland Universal Dementia Assessment Scale for multicultural screening in Danish memory clinics. *Dementia and Geriatric Cognitive Disorders*, *36*(5–6), 354–362. https://doi.org/10.1159/000354375

Nielsen, T. R., & Jørgensen, K. (2020). Cross-cultural dementia screening using the Rowland Universal Dementia Assessment Scale: A systematic review and meta-analysis. *International Psychogeriatrics*, *32*(9), 1031–1044. https://doi.org/10.1017/S1041610220000344

Nielsen, T. R., Segers, K., Vanderaspoilden, V., Bekkhus-Wetterberg, P., Bjørkløf, G. H., Beinhoff, U., Minthon, L., Pissiota, A., Tsolaki, M., Gkioka, M., & Waldemar, G. (2019). Validation of the Rowland Universal Dementia Assessment Scale (RUDAS) in a multicultural sample across five Western European countries: Diagnostic accuracy and normative data. *International Psychogeriatrics*, *31*(2), 287–296. https://doi.org/10.1017/S1041610218000832

Nielsen, T. R., Vogel, A., Gade, A., & Waldemar, G. (2012). Cognitive testing in non-demented Turkish immigrants—Comparison of the RUDAS and the MMSE. *Scandinavian Journal of Psychology*, *53*(6), 455–460. https://doi.org/10.1111/sjop.12018

Nielsen, T. R., & Waldemar, G. (2016). Knowledge and perceptions of dementia and Alzheimer's disease in four ethnic groups in Copenhagen, Denmark. *International Journal of Geriatric Psychiatry*, *31*(3), 222–230. https://doi.org/10.1002/gps.4314

Parker, C., & Philp, I. (2004). Screening for cognitive impairment among older people in black and minority ethnic groups. *Age and Ageing*, *33*(5), 447–452. https://doi.org/10.1093/ageing/afh135

Piccinin, A. M., Muniz-Terrera, G., Clouston, S., Reynolds, C. A., Thorvaldsson, V., Deary, I. J., Deeg, D. J. H., Johansson, B., Mackinnon, A., Spiro, A., Starr, J. M., Skoog, I., & Hofer, S. M. (2013). Coordinated analysis of age, sex, and education effects on change in MMSE scores. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *68*(3), 374–390. https://doi.org/10.1093/geronb/gbs077

Plitas, A., Tucker, A., Kritikos, A., Walters, I., & Bardenhagen, F. (2009). Comparative study of the cognitive performance of Greek Australian and Greek national elderly: Implications for neuropsychological practice. *Australian Psychologist*, *44*(1), 27–39. https://doi.org/10.1080/00050060802587694

Ramirez, M., Teresi, J. A., Holmes, D., Gurland, B., & Lantigua, R. (2006). Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE). Overview, sample, and issues of translation. *Medical Care*, *44*(Suppl. 3), S95–S106. https://doi.org/10.1097/01.mlr.0000245181.96133.db

Razani, J., Murcia, G., Tabares, J., & Wong, J. (2007). The effects of culture on WASI test performance in ethnically diverse individuals. *The Clinical Neuropsychologist*, *21*(5), 776–788. https://doi.org/10.1080/13854040701437481

Reiss, K., Breckenkamp, J., Borde, T., Brenne, S., David, M., & Razum, O. (2015). Smoking during pregnancy among Turkish immigrants in Germany-are there associations with acculturation? *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco*, *17*(6), 643–652. https://doi.org/10.1093/ntr/ntu237

Rowland, J. T., Basic, D., Storey, J. E., & Conforti, D. A. (2006). The Rowland Universal Dementia Assessment Scale (RUDAS) and the Folstein MMSE in a multicultural cohort of elderly persons. *International Psychogeriatrics*, *18*(1), 111–120. https://doi.org/10.1017/S1041610205003133

Sachs-Ericsson, N., & Blazer, D. G. (2005). Racial differences in cognitive decline in a sample of community-dwelling older adults: The mediating role of education and literacy. *The American Journal of Geriatric Psychiatry*, *13*(11), 968–975. https://doi.org/10.1176/appi.ajgp.13.11.968

Salmon, D. P., Riekkinen, P. J., Katzman, R., Zhang, M. Y., Jin, H., & Yu, E. (1989). Cross-cultural studies of dementia. A comparison of mini-mental state examination performance in Finland and China. *Archives of Neurology*, *46*(7), 769–772. https://doi.org/10.1001/archneur.1989.00520430063019

Segers, K., Benoit, F., Colson, C., Kovac, V., Nury, D., & Vanderaspoilden, V. (2013). Pioneers in migration, pioneering in dementia: First generation immigrants in a European metropolitan memory clinic. *Acta Neurologica Belgica*, *113*(4), 435–440. https://doi.org/10.1007/s13760-013-0245-z

Sheehan, B. (2012). Assessment scales in dementia. *Therapeutic Advances in Neurological Disorders*, *5*(6), 349–358. https://doi.org/10.1177/1756285612455733

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, *5*(1–2), 165–173. https://doi.org/10.1300/J018v05n01_09

Sisco, S., Gross, A. L., Shih, R. A., Sachs, B. C., Glymour, M. M., Bangen, K. J., Benitez, A., Skinner, J., Schneider, B. C., & Manly, J. J. (2015). The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *70*(4), 557–567. https://doi.org/10.1093/geronb/gbt133

Söhn, J., & Özcan, V. (2006). The educational attainment of Turkish Migrants in Germany. *Turkish Studies*, *7*(1), 101–124. https://doi.org/10.1080/14683840500520626

Stewart, R., Johnson, J., Richards, M., Brayne, C., & Mann, A. (2002). The distribution of Mini-Mental State Examination scores in an older UK African-Caribbean population compared to MRC CFA study norms. *International Journal of Geriatric Psychiatry*, *17*(8), 745–751. https://doi.org/10.1002/gps.698

Storey, J. E., Rowland, J. T. J., Basic, D., Conforti, D. A., & Dickson, H. G. (2004). The Rowland Universal Dementia Assessment Scale (RUDAS): A multicultural cognitive assessment scale. *International Psychogeriatrics*, *16*(1), 13–31. https://doi.org/10.1017/S1041610204000043

Teresi, J. A., Golden, R. R., Cross, P., Gurland, B., Kleinman, M., & Wilder, D. (1995). Item bias in cognitive screening measures: Comparisons of elderly white, Afro-American, Hispanic and high and low education subgroups. *Journal of Clinical Epidemiology*, *48*(4), 473–483. https://doi.org/10.1016/0895-4356(94)00159-N

Thomson, M. D., & Hoffman-Goetz, L. (2009). Defining and measuring acculturation: A systematic review of public health studies with Hispanic populations in the United States. *Social Science & Medicine*, *69*(7), 983–991. https://doi.org/10.1016/j.socscimed.2009.05.011

Welsh, K. A., Fillenbaum, G., Wilkinson, W., Heyman, A., Mohs, R. C., Stern, Y., Harrell, L. M., Edland, S. D., & Beekly, D. (1995). Neuropsychological test performance in African-American and white patients with Alzheimer's disease. *Neurology*, *45*(12), 2207–2211. https://doi.org/10.1212/WNL.45.12.2207

Yagmur, K. (2009). Language use and ethnolinguistic vitality of Turkish compared with the Dutch in the Netherlands. *Journal of Multilingual and Multicultural Development*, *30*(3), 219–233. https://doi.org/10.1080/01434630802369445

Yagmur, K., & van de Vijver, F. J. R. (2012). Acculturation and language orientations of Turkish immigrants in Australia, France, Germany, and the Netherlands. *Journal of Cross-Cultural Psychology*, *43*(7), 1110–1130. https://doi.org/10.1177/0022022111420145

Received March 2, 2021
Revision received June 11, 2021
Accepted June 22, 2021 ■

5. Chapter 5-General Discussion

The present thesis aimed to fill the gap with systematic literature reviews on language and its implications in early AD detection and objective characterization of cognitive profiles in people with dementia. This work also contributed to the neuropsychology literature by addressing the cross-cultural potential of a cognitive screening tool, RUDAS, in people from diverse cultural, linguistic, or experiential backgrounds.

The sections below summarize and synthesize the findings of the studies revealed by these research topics. Subsequently, the strengths and limitations of the studies are critically addressed. Finally, this section concludes with an assessment of how the thesis may contribute to advancing theoretical models and practical implications.

5.1. Summary and Synthesis of Findings

The goal of Study 1 was to summarize the findings of studies examining whether people with mild AD or MCI have significant deficits in macro-structural discourse comprehension compared to cognitively healthy people. Naming latencies, global synopsis, lesson, main idea, inferential clauses, and comprehension questions were characterized as measures of macro-structural comprehension in the eight studies included in the review. The findings of Study 1 revealed that people with AD and MCI experienced significant impairments in macro-structural discourse comprehension compared to healthy older adults in all measures, except for one measure; comprehension questions. More specifically, groups with AD performed poorer than healthy older adults on five of those measures, with comprehension questions showing mixed results. Furthermore, the MCI groups showed similar impairments in performance to the AD groups compared to healthy older adults.

The review findings have also shown that syntax and phonology remain retained in language production throughout the early or even early moderate stages of AD. However, the speech lacked important macro-linguistic characteristics of conversation, such as information content, coherence, and cohesion. Given the impairments in the macro-structural measures of comprehension, as revealed by the review, communication impairments observed in AD cannot be deducted to only one linguistic component. Discourse production is a partially conscious cognitive process involving constructing, allocating, and updating mental representations of events and the communication environment (van Dijk, 2006). This type of assessment paradigm, which reflects everyday interactions and experiences, offers a comprehensive insight into cognitive and language deficits in AD, adding a new dimension to neuropsychological testing and interventions. As the review findings pointed out, the correlations between standardized neuropsychological tests and discourse comprehension tasks were moderate only in verbal and working memory measures, indicating a lack of ecological validity and limited insight into language functions in most neuropsychological tests used commonly in clinical practice. Consequently, the multi-level nature of discourse comprehension may tap into a more holistic view of cognitive and linguistic deficiencies rather than tests evaluating separate or specific linguistic processes.

To summarize, Study 1 provides a consistent pattern of findings across the studies regarding the potential of discourse comprehension measures in the early detection of AD. This review's detailed description of performance differences in discourse comprehension of MCI and AD groups compared to healthy older adults add to our understanding of language deficits in the preclinical stage of AD. Additionally, it highlights several issues in our current neuropsychological measures and aids in identifying critical domains of discourse comprehension to provide a framework for the neuropsychological profile of MCI and AD.

Studies 2 and 3 also underscored the close relationship between language and cognitive changes in older adults and examined the role of bilingualism on neuropsychological assessment tools to interpret neuropsychological test results more accurately. To our knowledge, these studies are the first to elucidate the neuropsychological consequences of bilingualism and its associated factors in culturally and linguistically diverse older adult participants. More specifically, Study 2 pointed out that a more scientific approach is necessary to understand and advance the neuropsychological assessment of bilingual and monolingual older adults. The literature review of the studies investigating the neuropsychological test performance of bilingual and monolingual older adults over time found no conclusive evidence that bilingualism was associated with advantages/disadvantages in test performance. The only consistent finding in the review was that moderate advantages were observed in inhibitory control measures, whereas the disadvantages were detected in language domain measures. However, these findings were noted in cross-sectional studies, and they varied based on the language proficiency of the tested participants and the language of test administration. Another finding was that cultural and immigration-related biases (e.g. acculturation) in neuropsychological measures limit the ability to analyze performance differences in language groups as bilinguals differed substantially from monolinguals in terms of cultural background and immigrant status. Overall, the studies pose substantial methodological challenges in identifying the nature and determinants of test performance differences as a function of bilingualism. The language of test administration and participants' language proficiency, age of language acquisition, immigrant status and cultural background were all critical moderating factors for neuropsychological test performance differences in language groups. The review indicated that the identification of moderating variables and a theoretical construct of bilingualism are of critical importance in gaining insight into neuropsychological test performance differences. A novel

finding from this review is that this study offers a conceptualization in relation to different aspects of bilingualism pertinent for understanding neuropsychological test results when conducting neuropsychological tests with bilingual older adults.

In light of the findings from Study 2, Study 3 focused on several variables that were addressed in Study 2 to examine their impact on neuropsychological test performance in culturally and linguistically homogeneous groups. More specifically, the neuropsychological test performance of Turkish immigrants living in Germany was compared to their counterparts living in Turkey who had no immigration background to explore the effects of immigrant status and its associated variables (e.g. bilingualism and acculturation). Furthermore, the comparisons between non-immigrant Turkish and German patients with AD were drawn to examine whether cultural background plays a role in test performance. In addition, given the critical role of cognitive screening tools in detecting dementia, a cognitive screening tool developed in a Western/European context or for WEIRD societies, MMSE, and a cross-cultural screening measure, RUDAS, were chosen to investigate whether they were affected by those variables. The study by Nielsen et al. (2012) also examined the role of these variables on the performance of RUDAS and MMSE in non-demented Turkish immigrants in Denmark. Consistent with Hypothesis 3 (see Table 1) and the findings from Nielsen et al. (2012), the results have revealed that RUDAS can be a valuable tool for use with culturally and linguistically diverse older adults as performance differences in the RUDAS were less evident between groups with AD.

Furthermore, the performance on the RUDAS was not affected by demographic variables, educational level, bilingualism, and acculturation. The results provide additional support for the cross-cultural applicability of the RUDAS that can be useful in improving screening accuracy for dementia across all cultural and linguistic groups, not just minorities. This is important because,

even after adjusting for demographic variables, cognitively healthy minority groups are more likely to be misdiagnosed as impaired than the native-born populations (Manly & Mayeux, 2004). However, there are practical challenges with establishing separate test norms for a wide range of cultural contexts. In addition, variability of educational and cultural experiences within a particular cultural group may decrease the accuracy of the test norms (Manly & Mayeux, 2004). Therefore, the science and practice of neuropsychology are in need of assessment tools similar to the RUDAS that are responsive to the rapidly changing and diverse nature of the population.

With regard to Hypothesis 1 and 2, TR-AD and TR-IM-AD groups performed worse than the GR-AD group. However, after adjusting for the educational level, performance differences became non-significant among the groups. Furthermore, contrary to our expectations, TR-AD and TR-IM-AD groups performed similarly on the MMSE, indicating no effect of immigration status on test performance. The comparison of the non-immigrant and monolingual groups such as the TR-AD and GER-AD participants revealed no impact of cultural background on the MMSE performance, after adjusting for educational levels. Interestingly, higher levels of bilingualism were associated with better performance on the MMSE in the TR-IM-AD group.

To summarize, Study 3 indicated that performance differences were considerably greater on a test developed in a Western context than the one developed in a multi-cultural context. Instead of the cultural backgrounds of patients with AD, the educational background impacted performance on a cognitive screening test developed for WEIRD societies. The results regarding the effect of educational level on the MMSE performance are in line with many studies showing no discrepancies in test performance among different cultural groups after adjusting for educational levels (Loewenstein, Argüelles, Barker, & Duara, 1993; Mungas et al., 1996; Rasmusson, Carson, Brookmeyer, Kawas, & Brandt, 1996). However, contrary to these findings, several studies,

including culturally diverse populations, have also found that cultural factors significantly impact the MMSE performance after accounting for educational levels (Kuller et al., 1998; Welsh et al., 1995). In the following section, possible explanations for this inconclusive evidence in relation to the results of Study 3 will be argued.

5.2. Is it Culture, Education, or Bilingualism? Critical Factors in Explaining Neuropsychological Test Performance Disparities among Older Adults

The most widely recognized research criteria (McKhann et al., 1984) require evidence of decline in cognitive performance in memory and two other cognitive domains confirmed by neuropsychological tests for the diagnosis of probable or potential AD. However, the majority of tests cannot efficiently be utilized in culturally diverse, low-educated populations due to their reliance on Western culture, education, and literacy (Franzen et al., 2020). The lack of availability of cross-culturally valid tests may explain why performance differences on neuropsychological tests lead to over and under-diagnosis of cognitive impairment, especially in ethnic minority groups. This section will review studies that have evaluated neuropsychological/cognitive test performance of culturally different groups and discuss constructs that may be used to conduct more methodologically rigorous research of cultural differences in the future.

Difficulties in interpreting test scores among ethnic minority older adults have been shown in several cognitive screening instruments, with some of them conducted with the MMSE (Escobar et al., 1986; Welsh et al., 1995). For instance, applying a standard cut-off score of 23 on the MMSE in African-American and Hispanic patients living in the US caused over-diagnosis of dementia, even after accounting for years of education (Bohnstedt et al., 1994).

The impact of the cultural background was not limited to the cognitive screening tests and was also observed in neuropsychological test batteries, including Bender-Gestalt Test and the 1981 Wechsler Adult Intelligence Scale-Revised (Adams, Boake, & Crain, 1982; Reynolds, Chastain, Kaufman, & McLean, 1987). Moreover, disparities in neuropsychological test performance have persisted even when culturally different groups are matched in age and education (Jacobs et al., 1997; Kaufman, McLean, & Reynolds, 1988; Manly et al., 1998).

Contrary to these findings stated above, in line with the results of Study 3, several studies have shown no variance in scores of MMSE and other neuropsychological tests after adjusting for years of education (Loewenstein et al., 1993; Marcopulos, McLain, & Giuliano, 1997) or after matching the participant groups on education (Ford, Haley, Thrower, West, & Harrell, 1996). In some studies mentioned above, after adjusting for years of education, many of the variances in test scores became non-significant in tests mainly assessing language domains (Jacobs et al., 1997; Manly et al., 1998; Snitz et al., 2009). Education was found to be a more potent variable affecting test scores in language tests (Snitz et al., 2009). However, the cultural background still impacted scores of visuospatial skills from the Benton Visual Retention Test and category fluency (Jacobs et al., 1997; Manly et al., 1998). Interestingly, more recent studies point out that years of education are an inadequate indicator of educational experience among culturally/linguistically diverse populations (Chin, Negash, Xie, Arnold, & Hamilton, 2012; Krch et al., 2015; Manly, Jacobs, Touradji, Small, & Stern, 2002). Reading level measures and the country where education was received, namely factors associated with the quality of education, were the strongest predictor of test performance, rather than years of education or the cultural background of the participants.

Furthermore, another study employing the CERAD has pinpointed a novel influencing factor for disparities in neuropsychological test performance (Fillenbaum, Heyman, Huber, Ganguli, &

Unverzagt, 2001). In this study, the performance of older African Americans and Whites living in North Carolina was examined by comparing it to the African Americans in Indianapolis and Whites in Pennsylvania. The findings have revealed that the cultural background had no impact on the North Carolina sample's test performance after controlling for education. However, this sample had lower scores than their African American and White counterparts in Indianapolis and Pennsylvania (Fillenbaum et al., 2001). Taken together, in most studies, ethnicity or cultural background is regarded as a proxy for factors of interest, such as genetic composition, cultural experience, or environmental exposure (Manly & Mayeux, 2004).

Moreover, another facet of culture that has received scant attention in this literature is language, specifically, bilingualism. Bilinguals who acquire their second language in a bicultural context may attain some levels of biculturalism that can intensify the complexity between these variables. Biculturalism, a cultural/ethnic identity development process that arises due to participation in, and identification with two cultures (Ramírez-Esparza & García-Sierra, 2014), may also influence cognition (Tadmor, Tetlock, & Kaiping, 2009) and cognitive reserve. To our knowledge, no previous attempt has been made to provide a comprehensive account of these environmental exposures in the literature while assessing the influence of cultural factors on test performance. The studies that do not evaluate these underlying elements provide only limited insight into the disparities in neuropsychological test performance of culturally/linguistically diverse older adults and cannot differentiate the impacts of these factors on test performance.

To sum up, ethnicity or cultural background is a challenging construct to assess as it has a variety of hidden factors that may contribute to the inconclusive findings in this field. However, as in Study 3, comparing immigrants from a particular cultural background with their native-born counterparts offers an opportunity to study groups for which genetic characteristics are mostly

unchanged, but environmental and cultural factors undergo significant transformations. In addition, although adjusting for educational levels reduces the cultural bias in test performances, the amount of variance accounted for by these factors within each group can be considerably different (Romero et al., 2009). The interactions between these factors may be unique to the particular groups which were examined as there is the heterogeneity of educational and cultural experiences within a cultural group. Thus, further research, similar to Study 3, but with a measure of the quality of education in addition to years of education, is required to understand how education and cultural background interact with each other in specific cultural groups and how these variables affect test performances. Such studies may help advance the research by enabling a more in-depth examination of the complex interactions between individual and environmental factors that generally influence cognition. In addition, the continuous investigation of the individual and combined predictive power of these factors on cognitive test performance can strengthen the validity of existing instruments. It may also help the development, validation, and standardization of more cross-culturally applicable tests.

5.3. Strengths and Limitations

A major strength of this thesis is that all studies provide a much-needed exploration into language-related cognitive changes in older adults by investigating the role of language in various facets of pathological cognitive aging in culturally and linguistically heterogeneous samples. Mainly, systematic reviews critically evaluated a number of methodologies and study designs to answer the research questions covered in this thesis. In addition, the systematic reviews used a rigorous search strategy, assessed the risk of bias in the studies, characterized the literature's strengths and weaknesses, and delineated recommendations for future research.

Another strength of this thesis lies in the population of Study 3, which has so far been an underrepresented population in European health research (Dingoyan, Schulz, & Mösko, 2020). Additionally, as stated before, previous studies have been limited to only the immigrant groups or the comparisons between immigrant participants and individuals of the host country. Study 3 included native-born populations, namely, non-immigrant Turkish and German patients with AD, as well as an underrepresented immigrant population in Europe, who were recruited from several German cities, presumably more representative of the general population. Thus, this study design has offered the opportunity to distinguish between the effects of culture and immigration-related factors. Furthermore, as revealed by Study 2, given the limited research investigating the impact of bilingualism on neuropsychological tests performed in participants' native or preferred language, to our knowledge, this study is the first to indicate the effect of bilingualism on the MMSE even when performed in the native language of participants, by examiners with the native fluency in Turkish and German.

An additional strength of this study concerns the use of multi-dimensional instruments and the applicability of these tools in German and Turkish populations. In line with the concept defined by Berry (1997, 2003), acculturation in the host country was explored with a bi-dimensional acculturation model in our study. More specifically, acculturation outcomes assessed by the FRAKK were measured based on the orientation to the culture of origin and to the host culture. This model has been considered more suitable for understanding the complexity of acculturation in immigrants and ethnic minorities (Costigan & Su, 2004). In addition, the FRAKK was validated in the Turkish immigrant groups living in Germany (Bongard, Pogge, Arslaner, Rohrmann, & Hodapp, 2002).

Regarding the bilingualism questionnaire, a continuous and multifaceted approach was taken to measure bilingual experience by including several dimensions of bilingualism, namely L1/L2 proficiency, language switching frequency and language used in certain activities or with people. These variables were merged and converted into a composite score. Thus, the effects of bilingualism on test performance were investigated as a continuous variable rather than a dichotomous variable, as has been done in other studies of bilingualism. It allowed for the concurrent consideration of several factors that may be associated with test performance (Kaushanskaya & Prior, 2014).

However, the findings and contributions of the studies reported in this thesis need to be considered in light of some potential limitations. Firstly, Studies 2 and 3 lacked data regarding another potentially significant variable, socioeconomic status. Socioeconomic factors have been linked to the risk of AD (Evans et al., 1997; Qian, Schweizer, & Fischer, 2013) and have been shown to impact cognitive functions (Turrell et al., 2002). Although the educational level was used as a proxy measure of socioeconomic status in these studies, occupational and economic prospects associated with that educational level may not be alike (Manly & Mayeux, 2004). Therefore, socioeconomic status may also have contributed to the relationship between educational level and test performance. Secondly, the small sample size of 65 participants in Study 3 presents another limitation to data generalizability and analytic approaches. As noted in Chapter 4, differences in MMSE performance between Turkish immigrant and native-born Turkish groups with AD approached but failed to reach statistical significance due to the possibility of a lack of statistical power. Despite the several recruitment strategies implemented during the study process, access to health care services was limited, especially in the Turkish immigrant group, due to certain beliefs and lack of knowledge and awareness about dementia (Nielsen & Waldemar, 2016). Another

limitation of this study relates to the assessment of the educational levels of participants. Reading fluency has evolved as an alternate measure for defining educational quality. A study has provided evidence that reading level had higher correlations with neuropsychological test scores than years of education (Fernandez & Arriondo, 2021). Therefore, the use of a reading fluency assessment would help investigate more in-depth whether various measurements of educational levels affect test performance differentially in the participant groups. However, since several tests and questionnaires were employed in this study, the use of additional questionnaires was avoided in order not to cause assessment fatigue in the participants. Finally, the lack of association between cultural background and test performance is likely influenced by specific characteristics of each sample or methodology. Therefore, findings from Study 3 may not be generalizable to all cultural groups since sample characteristics and methods are not comparable across different studies. Accordingly, conclusions on the impact of cultural, educational, and linguistic background on test performances are predicted to be limited to specific characteristics of the participants in the thesis.

5.4. Implications for Future Research

The present thesis has important implications for future research. Based on the current findings, suggested research methodologies can be pursued in more detail.

Upon closer examination of the methodologies of the studies included in the systematic reviews, participants in the studies of bilingualism and macro-structural discourse comprehension were mainly located in an English-speaking environment. Therefore, studies included in the systematic reviews were limited to specific geographic locations, including Canada and the USA. Additionally, longitudinal studies investigating macro-structural discourse comprehension and the effects of bilingualism on neuropsychological test performance in older adults were scarce.

Therefore, more research that employs a longitudinal study design with a wider geographic range and more linguistically and culturally diverse populations is warranted.

With respect to the systematic review investigating the impact of bilingualism on neuropsychological test performance, it is vital for future studies to determine the cross-cultural potential of neuropsychological measures before their utilization in bilingualism research. Or, future studies should check whether neuropsychological assessment tools planned to be used for the participant groups are validated for or relevant to the specific cultural and language groups (Loewenstein, Argüelles, Argüelles, & Linn-fuentes, 1994).

As evidenced in this thesis, bilingual individuals vary in linguistic experiences such as L1/L2 age of acquisition and proficiency, language use patterns, including language switching and mixing (Kaushanskaya & Prior, 2014). Either alone or in combination, these characteristics may have distinct effects on executive functions (Kaushanskaya & Prior, 2014), CR (Guzmán-Vélez & Tranel, 2015), and brain plasticity (Sulpizio, Del Maschio, Del Mauro, Fedeli, & Abutalebi, 2020). Since controlling for all these potential factors is not possible, an important avenue to pursue in future research would be to examine the within-group variance among bilinguals to identify which aspects of bilingualism are associated with improved performance or neural changes (Kaushanskaya & Prior, 2014; Sulpizio et al., 2020). The absence of clarity in the operationalization of bilingualism is a significant problem in this field as it invalidates comparisons across studies (Adesope, Lavin, Thompson, & Ungerleider, 2010). The application of individual-variability approaches within bilingual groups may allow for simultaneous modeling of multiple experiential factors (Sulpizio et al., 2020), thereby providing a theoretical framework for determining which features of bilingualism may impact particular aspects of cognition. Eventually, a consistent operational definition and a clear identification of the explored components may help

enhance the comparability across studies and the development of a comprehensive theoretical framework.

As for the impact of education and bilingualism on the MMSE in the immigrant group, this study provided preliminary evidence that tests developed for WEIRD societies are susceptible to variability in educational and linguistic experiences of older ethnic minority adults with AD. Although this thesis found no evidence for the relationship between acculturation and performance on the MMSE in the Turkish immigrant group, in line with the study by (Tan & Burgess, 2020), it has shown that specific constructs within acculturation, namely, language adoption was more likely to associate with test performance. The four acculturation categories as proposed by Berry (2005), namely, integration, marginalization, separation, and assimilation, were measured by the subscales of the FRAKK. However, the effects of all four acculturation strategies on test performance could not be investigated separately in our study as the small sample size of the Turkish immigrant group did not allow for such a categorization. Moreover, the distinct effects of these acculturation categories on test performance have not received much attention in the current literature (Tan & Burgess, 2020). Therefore, future research should focus on four acculturative strategies defined by Berry (2005) and include the domain of language knowledge to shed light on their association with test performances. Taken together, this study should be replicated on different culturally/linguistically homogenous groups with larger sample sizes and a broader variety of socio-demographic factors (e.g., SES, specific constructs within acculturation, ethnicity, bilingualism) to validate these findings.

5.5. Implications for Neuropsychological Practice

The studies involved in this thesis provide significant implications for improving the neuropsychological evaluation of older adults in clinical practice.

As stated in Chapter 2, features of discourse comprehension offer the possibility of detecting AD in its preclinical stage and contribute toward a better understanding of early linguistic alterations in the context of cognitive decline. There is a need for clinical tools that psychologists, speech and language therapists can use more readily and help improve diagnostic and prognosis accuracy (Bucks, Singh, Cuerden, & Wilcock, 2000). Based on the methods described in the systematic review, naming latencies, global synopsis, lesson, main idea, and inferential clauses measures may be considered as a basis for developing new tests assessing language abilities. Future validation of these methods in preclinical AD samples can demonstrate that discourse comprehension may be a potential tool for improving classification criteria to distinguish between normal aging, MCI, and AD, as well as for informing appropriate treatment plans and effective interventions for adults with MCI and AD dementia.

As for bilingualism and its association with neuropsychological test performance, the findings of Chapter 3 highlight various factors that need to be taken into consideration in clinical practice during neuropsychological testing of bilingual older adults. Firstly, it would be better to conduct the neuropsychological evaluation of bilingual adults in their preferred, frequently used, and most proficient language (Rivera Mindt et al., 2008). Language proficiency assessment should start during the interview and include an evaluation of the frequency and context of language use, age of language acquisition, the level of receptive and expressive language skills, switching between languages, and acculturation issues (Romero et al., 2009). The use of acculturation measures may also help determine the language in which the patients are evaluated (Pontón, 2001). Furthermore, identifying the most proficient language of the adults is essential for ensuring that performance on the tests is not susceptible to the language of administration. As revealed by Study 2, the language of test administration showed differing performance patterns, particularly on tests assessing

language domains. Since language interference and the mixing of words between languages are a common occurrence in daily conversations, administering and scoring language-based tests in both languages of speakers may provide a more accurate and current picture of cognitive functioning (Paplikar et al., 2021). However, another critical factor in evaluating bilinguals is the examiner's linguistic proficiency (Pontón, 2001), and these services may not be available at all times or locations. Since the use of translated tests and interpreters has been criticized for affecting test scores and standardized administration (Casas et al., 2012), the field of neuropsychology should invest more effort in increasing the recruitment and retention of culturally/linguistically diverse neuropsychologists (Romero et al., 2009).

Secondly, in the majority of bilingualism and neuropsychological performance studies, it was unclear what norms were used when the tests were employed in culturally and linguistically diverse bilingual groups. The effects of bilingualism on test performance may be eliminated by using culturally or demographically appropriate norms. Based on these findings, it is still poorly understood whether there is a need for using different norms for bilinguals and monolingual adults in a clinical diagnostic environment. Therefore, one avenue for future clinical research can be to investigate if bilingualism would associate differently depending on the norms used. This would help translate empirical work to clinical practices in terms of procedures during the neuropsychological evaluation of bilingual and monolingual adults.

Lastly, due to the interactions between language and culture, it was not possible to obtain a more precise association between bilingualism and test performance differences across language groups. Much of the work on bilingualism and neuropsychological test performance in older adults has been confounded by other group differences, including ethnicity and immigrant status. These findings reinforce the need to take into account the cultural bias of neuropsychological tests used,

especially in clinical practice. One strategy to resolve these challenges in clinical practice is to identify culture-fair tests that have shown linguistic equivalence between cultures instead of avoiding verbal-nature neuropsychological tests to diminish bias (Fernandez & Abe, 2018). Several tests, namely The European Cross-Cultural Neuropsychological Test Battery (Nielsen et al., 2019), RUDAS, Cross-Linguistic Naming Test (Ardila, 2007), Five Digit Test (Sedo, 2004), or the Stick Design Test (Baiyewu et al., 2005), have been suggested as suitable for use in culturally diverse and low-educated groups (Franzen et al., 2020). Overall, future research on bilingualism should employ clinically relevant approaches to better differentiate between culture, bilingualism, and acculturation, such as using these culturally sensitive assessment tools and norms and conducting a detailed interview or questionnaires regarding the cultural, linguistic, and immigration background of the patients.

With respect to the findings of Study 3, the results are in line with a growing body of research suggesting that MMSE score modifications based merely on age and education are insufficient to properly account for performance disparities across various ethnic groups (Milman et al., 2018). As shown in Study 3, Turkish immigrants with AD scored lower than German and Turkish patients with AD on items that have been more closely linked to education and cultural background. Although items-based performance pattern is considered in culturally diverse populations when interpreting test scores (Matallana et al., 2010; Milman et al., 2018), this approach may have practical limits in clinical settings.

Determining the most sensitive assessment tools for identifying cognitive decline is critical not only for the early detection of dementia in clinical practice but also for international clinical trials. Neuropsychological assessment tools are employed to select trial participants and used as clinical outcome measures after the intervention (Ng et al., 2018). Hence, determining tools capable of

identifying early cognitive impairment that are equivalent across culturally different groups is crucial (Ng et al., 2018). To this end, Study 3 was an attempt to gain a better understanding of the variables that impact test performance since insight into these variables can aid in identifying the "real" source of variance in estimating individual and group differences (Romero et al., 2009). Furthermore, it may eventually inform norms and promote more complex and accurate clinical interpretations of neuropsychological data (Romero et al., 2009). As revealed by Study 3, RUDAS may be a potential tool to be used for culturally and linguistically diverse populations in clinical practice, as well as in multinational clinical trials. Furthermore, it may have certain advantages over the MMSE since the scores on the RUDAS were not influenced by bilingualism in the Turkish-immigrant group.

5.6. Conclusion

Going beyond previous work and adopting a holistic approach to linguistic skills in culturally varied populations, this thesis significantly contributed to the literature on the relationship between language, culture, and neuropsychological test performance in people with dementia. However, a clear association between these variables was hampered by the differences in the methodologies, sample characteristics, neuropsychological tests used in the studies, and the operationalization of macrostructural comprehension and bilingualism. Despite these challenges, this thesis identified several discourse comprehension measures to aid in the early characterization of AD and elements of bilingualism that are empirically and conceptually relevant to test performance. Furthermore, educational level was identified as a factor in ethnic minority older adults impacting performance in a test commonly used in clinical practice and research. Interestingly, the cultural background was not associated with test performance on this test which was developed for the Western/European context. On the other hand, this thesis uncovered a culturally sensitive cognitive

screening tool that can be employed in German and Turkish populations diagnosed with AD in both clinical and research settings. Based on these findings, it is believed that this thesis will serve as a base for future studies on the potential of language as an early marker of AD pathology and as a variable bringing along multiple experiential factors that influence the neuropsychological evaluation of older adults. Given the shifts in Europe's demographic profile, there is an urgent need for neuropsychological services to competently assess the increasing number of older adults from various cultural and linguistic backgrounds. There is a scarcity of empirically based, practical information aimed at serving these populations. Therefore, future studies on this topic are required to validate the recommended approaches to increase the accuracy and precision of neuropsychological testing in culturally and linguistically diverse people with dementia.

References

2020 Alzheimer's disease facts and figures. (2020). *Alzheimer's & dementia, 16*(3), 391-460. doi:https://doi.org/10.1002/alz.12068

Abrams, L., & Farrell, M. T. (2011). Language processing in normal aging *The handbook of psycholinguistic and cognitive processes: Perspectives in communication disorders.* (pp. 49-73). New York, NY, US: Psychology Press.

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

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

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.

Adams, R. L., Boake, C., & Crain, C. (1982). Bias in a neuropsychological test classification related to education, age, and ethnicity. *J Consult Clin Psychol, 50*(1), 143-145. doi:10.1037//0022-006x.50.1.143

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. doi:10.3102/0034654310368803

Ahmed, T. F., Ahmed, A., & Imtiaz, F. (2021). History in perspective: How Alzheimer's Disease came to be where it is? *Brain Res, 1758*, 147342. doi:10.1016/j.brainres.2021.147342

Aisen, P. S., Cummings, J., Jack, C. R., Jr., Morris, J. C., Sperling, R., Frölich, L., . . . Dubois, B. (2017). On the path to 2025: understanding the Alzheimer's disease continuum. *Alzheimers Res Ther, 9*(1), 60. doi:10.1186/s13195-017-0283-5

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

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

Antecol, H., & Bedard, K. (2006). Unhealthy assimilation: Why do immigrants converge to American health status levels? *Demography, 43*(2), 337-360. doi:10.1353/dem.2006.0011

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 Pt 2), 2689-2698. doi: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*. doi:10.3389/fpsyg.2017.02217

Ardila, A. (2007). Toward the development of a cross-linguistic naming test. *Archives of clinical neuropsychology, 22*(3), 297-307. doi:https://doi.org/10.1016/j.acn.2007.01.016

Arentoft, A., Byrd, D., Robbins, R. N., Monzones, J., Miranda, C., Rosario, A., . . . Rivera Mindt, M. (2012). Multidimensional effects of acculturation on English-language neuropsychological test performance among HIV+ Caribbean Latinas/os. *J Clin Exp Neuropsychol, 34*(8), 814-825. doi:10.1080/13803395.2012.683856

Baiyewu, O., Unverzagt, F. W., Lane, K. A., Gureje, O. Y. E., Ogunniyi, A., Musick, B., . . . Hendrie, H. C. (2005). The Stick Design test: A new measure of visuoconstructional ability. *Journal of the International Neuropsychological Society, 11*(5), 598-605. doi:10.1017/S135561770505071X

Bak, T. H. (2016). The impact of bilingualism on cognitive ageing and dementia: Finding a path through a forest of confounding variables. *Linguistic Approaches to Bilingualism, 6*(1-2), 205-226. doi:https://doi.org/10.1075/lab.15002.bak

Bak, T. H., & Alladi, S. (2016). Bilingualism, dementia and the tale of many variables: Why we need to move beyond the Western World. Commentary on Lawton et al. (2015) and Fuller-Thomson (2015). *Cortex, 74*, 315-317. doi:10.1016/j.cortex.2015.04.025

Balthazar, M. L., Cendes, F., & Damasceno, B. P. (2008). Semantic error patterns on the Boston Naming Test in normal aging, amnestic mild cognitive impairment, and mild Alzheimer's

disease: is there semantic disruption? *Neuropsychology, 22*(6), 703-709. doi:10.1037/a0012919

Berry, J. W. (1997). Immigration, Acculturation, and Adaptation. *Applied Psychology, 46*(1), 5-34. doi:https://doi.org/10.1111/j.1464-0597.1997.tb01087.x

Berry, J. W. (2005). Acculturation: Living successfully in two cultures. *International Journal of Intercultural Relations, 29*(6), 697-712. doi:10.1016/j.ijintrel.2005.07.013

Bertola, L., Cunha Lima, M. L., Romano-Silva, M. A., de Moraes, E. N., Diniz, B. S., & Malloy-Diniz, L. F. (2014). Impaired generation of new subcategories and switching in a semantic verbal fluency test in older adults with mild cognitive impairment. *Frontiers in aging neuroscience, 6*. doi:10.3389/fnagi.2014.00141

Bhatia, T. K., & Ritchie, W. C. (2006). Introduction *The Handbook of Bilingualism* (pp. 1-2).

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

Bialystok, E., Craik, F. I. M., Binns, M. A., Ossher, L., & Freedman, M. (2014). Effects of bilingualism on the age of onset and progression of MCI and AD: evidence from executive function tests. *Neuropsychology, 28*(2), 290-304. doi:10.1037/neu0000023

Black, S. A., Espino, D. V., Mahurin, R., Lichtenstein, M. J., Hazuda, H. P., Fabrizio, D., . . . Markides, K. S. (1999). The Influence of Noncognitive Factors on the Mini-Mental State Examination in Older Mexican-Americans: Findings from the Hispanic EPESE. *J Clin Epidemiol, 52*(11), 1095-1102. doi:https://doi.org/10.1016/S0895-4356(99)00100-6

Blesa, R., Pujol, M., Aguilar, M., Santacruz, P., Bertran-Serra, I., Hernández, G., . . . Peña-Casanova, J. (2001). Clinical validity of the 'mini-mental state' for Spanish speaking communities. *Neuropsychologia, 39*(11), 1150-1157. doi:10.1016/s0028-3932(01)00055-0

Bohnstedt, M., Fox, P. J., & Kohatsu, N. D. (1994). Correlates of Mini-Mental Status Examination scores among elderly demented patients: the influence of race-ethnicity. *J Clin Epidemiol, 47*(12), 1381-1387. doi:10.1016/0895-4356(94)90082-5

Bondi, M. W. (2002). Genetic and brain imaging contributions to neuropsychological functioning in preclinical dementia. *J Int Neuropsychol Soc, 8*(7), 915-917. doi:10.1017/s1355617702870059

Bongard, S., Pogge, S. F., Arslaner, H., Rohrmann, S., & Hodapp, V. (2002). Acculturation and cardiovascular reactivity of second-generation Turkish migrants in Germany. *Journal of psychosomatic research, 53*(3), 795-803.

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of clinical neuropsychology, 22*(3), 355-365.

Boschi, V., Catricala, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S. F. (2017). Connected speech in neurodegenerative language disorders: a review. *Frontiers in psychology, 8*, 269.

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol, 82*(4), 239-259. doi:10.1007/bf00308809

Bschor, T., Kühl, K.-P., & Reischies, F. M. (2005). Spontaneous Speech of Patients With Dementia of the Alzheimer Type and Mild Cognitive Impairment. *International Psychogeriatrics, 13*(3), 289-298. doi:10.1017/S1041610201007682

Bucks, R. S., Singh, S., Cuerden, J. M., & Wilcock, G. K. (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology, 14*(1), 71-91. doi:10.1080/026870300401603

Budson, A., & Solomon, P. (2016). Memory Loss, Alzheimer's Disease an d Dementia. *A Practical Guide for Clinicians. second ed Elsevier*, 80-91.

Burke, D. M., & College, P. (1997). Language, Aging, and Inhibitory Deficits: Evaluation of a Theory. *The Journals of Gerontology: Series B, 52B*(6), P254-P264. doi:10.1093/geronb/52B.6.P254

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 352*(1363), 1845-1856. doi:10.1098/rstb.1997.0170

Burke, D. M., MacKay, D. G., & James, L. E. (2000). Theoretical approaches to language and aging *Models of cognitive aging.* (pp. 204-237). New York, NY, US: Oxford University Press.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*(5), 542-579. doi:https://doi.org/10.1016/0749-596X(91)90026-G

Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., . . . Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat Rev Neurosci, 19*(11), 701-710. doi:10.1038/s41583-018-0068-2

Calabria, M., Costa, A., Green, D. W., & Abutalebi, J. (2018). Neural basis of bilingual language control. *Ann N Y Acad Sci*. doi:10.1111/nyas.13879

Calvo, N., García, A. M., Manoiloff, L., & Ibáñez, A. (2015). Bilingualism and Cognitive Reserve: A Critical Overview and a Plea for Methodological Innovations. *Front Aging Neurosci, 7*, 249. doi:10.3389/fnagi.2015.00249

Calvo, N., Ibáñez, A., Muñoz, E., & García, A. M. (2018). A core avenue for transcultural research on dementia: on the cross-linguistic generalization of language-related effects in Alzheimer's disease and Parkinson's disease. *Int J Geriatr Psychiatry, 33*(6), 814-823. doi:10.1002/gps.4712

Casas, R., Guzmán-Vélez, E., Cardona-Rodriguez, J., Rodriguez, N., Quiñones, G., Izaguirre, B., & Tranel, D. (2012). Interpreter-mediated neuropsychological testing of monolingual Spanish speakers. *Clin Neuropsychol, 26*(1), 88-101. doi:10.1080/13854046.2011.640641

Catricalà, E., Della Rosa, P. A., Plebani, V., Perani, D., Garrard, P., & Cappa, S. F. (2015). Semantic feature degradation and naming performance. Evidence from neurodegenerative disorders. *Brain and Language, 147*, 58-65. doi:10.1016/j.bandl.2015.05.007

Chan, A. S., Shum, D., & Cheung, R. W. (2003). Recent development of cognitive and neuropsychological assessment in Asian countries. *Psychol Assess, 15*(3), 257-267. doi:10.1037/1040-3590.15.3.257

Chapman, S. B., Sparks, G., Levin, H. S., Dennis, M., Roncadin, C., Zhang, L., & Song, J. (2004). Discourse macrolevel processing after severe pediatric traumatic brain injury. *Dev Neuropsychol, 25*(1-2), 37-60. doi:10.1080/87565641.2004.9651921

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease:

evidence from a bilingual community. *Alzheimer Dis Assoc Disord, 24*(2), 118-125. doi:10.1097/WAD.0b013e3181ca1221

Chin, A. L., Negash, S., Xie, S., Arnold, S. E., & Hamilton, R. (2012). Quality, and not Just Quantity, of Education Accounts for Differences in Psychometric Performance between African Americans and White Non-Hispanics with Alzheimer's Disease. *Journal of the International Neuropsychological Society, 18*(2), 277-285. doi:10.1017/S1355617711001688

Choi, M. H., Kim, H. S., Gim, S. Y., Kim, W. R., Mun, K. R., Tack, G. R., . . . Chung, S. C. (2016). Differences in cognitive ability and hippocampal volume between Alzheimer's disease, amnestic mild cognitive impairment, and healthy control groups, and their correlation. *Neurosci Lett, 620*, 115-120. doi:10.1016/j.neulet.2016.03.044

Collins, N., Sachs-Ericsson, N., Preacher, K. J., Sheffield, K. M., & Markides, K. (2009). Smoking increases risk for cognitive decline among community-dwelling older Mexican Americans. *Am J Geriatr Psychiatry, 17*(11), 934-942. doi:10.1097/JGP.0b013e3181b0f8df

Copland, D. A., Chenery, H. J., & Murdoch, B. E. (2002). Hemispheric contributions to lexical ambiguity resolution: evidence from individuals with complex language impairment following left-hemisphere lesions. *Brain Lang, 81*(1-3), 131-143. doi:10.1006/brln.2001.2512

Costigan, C., & Su, T. (2004). Orthogonal versus linear models of acculturation among immigrant Chinese Canadians: A comparison of mothers, fathers, and children. *International Journal of Behavioral Development, 28*(6), 518-527. doi:10.1080/01650250444000234

Craik, F. I. M., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: bilingualism as a form of cognitive reserve. *Neurology, 75*(19), 1726-1729. doi:10.1212/WNL.0b013e3181fc2a1c

Cuetos, F., Arango-Lasprilla, J. C., Uribe, C., Valencia, C., & Lopera, F. (2007). Linguistic changes in verbal expression: a preclinical marker of Alzheimer's disease. *Journal of the International Neuropsychological Society, 13*(3), 433-439.

Cummings, L. (2019). Describing the Cookie Theft picture. *Pragmatics and Society*.

Daugherty, J. C., Puente, A. E., Fasfous, A. F., Hidalgo-Ruzzante, N., & Pérez-Garcia, M. (2017). Diagnostic mistakes of culturally diverse individuals when using North American neuropsychological tests. *Appl Neuropsychol Adult, 24*(1), 16-22. doi:10.1080/23279095.2015.1036992

Davis, D. K., Alea, N., & Bluck, S. (2015). The Difference between Right and Wrong: Accuracy of Older and Younger Adults' Story Recall. *Int J Environ Res Public Health, 12*(9), 10861-10885. doi:10.3390/ijerph120910861

De Beni, R., Borella, E., & Carretti, B. (2007). Reading comprehension in aging: The role of working memory and metacomprehension. *Aging, Neuropsychology, and Cognition, 14*(2), 189-212. doi:10.1080/13825580500229213

de Lira, J. O., Ortiz, K. Z., Campanha, A. C., Bertolucci, P. H., & Minett, T. S. (2011). Microlinguistic aspects of the oral narrative in patients with Alzheimer's disease. *Int Psychogeriatr, 23*(3), 404-412. doi:10.1017/s1041610210001092

DeTure, M. A., & Dickson, D. W. (2019). The neuropathological diagnosis of Alzheimer's disease. *Molecular Neurodegeneration, 14*(1), 32. doi:10.1186/s13024-019-0333-5

Dingoyan, D., Schulz, H., & Mösko, M. (2020). The willingness to participate in health research studies of individuals with Turkish migration backgrounds: barriers and resources. *European Psychiatry, 27*(S2), S4-S9. doi:10.1016/S0924-9338(12)75701-2

Drachman, D. A. (2006). Aging of the brain, entropy, and Alzheimer disease. *Neurology, 67*(8), 1340-1352. doi:10.1212/01.wnl.0000240127.89601.83

Drummond, C., Coutinho, G., Fonseca, R. P., Assunção, N., Teldeschi, A., de Oliveira-Souza, R., . . . Mattos, P. (2015). Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. *Frontiers in aging neuroscience, 7*. doi:10.3389/fnagi.2015.00096

Drummond, C., Coutinho, G., Monteiro, M. C., Assuncao, N., Teldeschi, A., de Souza, A. S., . . . Tovar-Moll, F. (2019). Narrative impairment, white matter damage and CSF biomarkers in the Alzheimer's disease spectrum. *Aging, 11*(20), 9188-9208. doi:10.18632/aging.102391

Dubois, B., Villain, N., Frisoni, G. B., Rabinovici, G. D., Sabbagh, M., Cappa, S., . . . Feldman, H. H. (2021). Clinical diagnosis of Alzheimer's disease: recommendations of the

International Working Group. *Lancet Neurol, 20*(6), 484-496. doi:10.1016/s1474-4422(21)00066-1

Duong, A., Giroux, F., Tardif, A., & Ska, B. (2005). The heterogeneity of picture-supported narratives in Alzheimer's disease. *Brain Lang, 93*(2), 173-184. doi:10.1016/j.bandl.2004.10.007

Ehrlich, J. S. (1994). *Studies of discourse production in adults with Alzheimer's disease*.

Eichler, T., Thyrian, J. R., Hertel, J., Michalowsky, B., Wucherer, D., Dreier, A., . . . Hoffmann, W. (2015). Rates of formal diagnosis of dementia in primary care: The effect of screening. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(1), 87-93. doi:https://doi.org/10.1016/j.dadm.2014.11.007

Escobar, J. I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J., & Golding, J. M. (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity. Cultural and linguistic artifacts. *J Nerv Ment Dis, 174*(10), 607-614. doi:10.1097/00005053-198610000-00005

Estanga, A., Ecay-Torres, M., Ibañez, A., Izagirre, A., Villanua, J., Garcia-Sebastian, M., . . . Martinez-Lage, P. (2017). Beneficial effect of bilingualism on Alzheimer's disease CSF biomarkers and cognition. *Neurobiol Aging, 50*, 144-151. doi:10.1016/j.neurobiolaging.2016.10.013

Evans, D. A., Hebert, L. E., Beckett, L. A., Scherr, P. A., Albert, M. S., Chown, M. J., . . . Taylor, J. O. (1997). Education and other measures of socioeconomic status and risk of incident Alzheimer disease in a defined population of older persons. *Arch Neurol, 54*(11), 1399-1405. doi:10.1001/archneur.1997.00550230066019

Fernandez, A. L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: problems and possible solutions. *Culture and Brain, 6*, 1-35.

Fernandez, A. L., & Arriondo, G. J. (2021). Reading fluency as a measure of educational level. *Dement Neuropsychol, 15*(3), 361-365. doi:10.1590/1980-57642021dn15-030008

Ferraro, F. R. (2016). *Minority and cross-cultural aspects of neuropsychological assessment : enduring and emerging trends / edited by F. Richard Ferraro* (Second edition. ed.). New York, NY: Psychology Press.

Ferris, S. H., & Farlow, M. (2013). Language impairment in Alzheimer's disease and benefits of acetylcholinesterase inhibitors. *Clinical interventions in aging, 8*, 1007-1014. doi:10.2147/cia.s39959

Fillenbaum, G. G., Heyman, A., Huber, M. S., Ganguli, M., & Unverzagt, F. W. (2001). Performance of elderly African American and White community residents on the CERAD Neuropsychological Battery. *Journal of the International Neuropsychological Society, 7*(4), 502-509. doi:10.1017/S1355617701744062

Fisher, N. J., Rourke, B. P., & Bieliauskas, L. A. (1999). Neuropsychological subgroups of patients with Alzheimer's disease: an examination of the first 10 years of CERAD data. *J Clin Exp Neuropsychol, 21*(4), 488-518. doi:10.1076/jcen.21.4.488.887

Fleming, V. B. (2014). Early Detection of Cognitive-Linguistic Change Associated With Mild Cognitive Impairment. *Communication Disorders Quarterly, 35*(3), 146-157. doi:10.1177/1525740113520322

Fleming, V. B., & Harris, J. L. (2008). Complex discourse production in mild cognitive impairment: Detecting subtle changes. *Aphasiology, 22*(7-8), 729-740. doi:10.1080/02687030701803762

Fleming, V. B., & Harris, J. L. (2009). Test–retest discourse performance of individuals with mild cognitive impairment. *Aphasiology, 23*(7-8), 940-950. doi:10.1080/02687030802586480

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. *J Psychiatr Res, 12*(3), 189-198. doi:10.1016/0022-3956(75)90026-6

Forbes-McKay, K. E., & Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurol Sci, 26*(4), 243-254. doi:10.1007/s10072-005-0467-9

Ford, G. R., Haley, W. E., Thrower, S. L., West, C. A., & Harrell, L. E. (1996). Utility of Mini-Mental State Exam scores in predicting functional impairment among white and African American dementia patients. *J Gerontol A Biol Sci Med Sci, 51*(4), M185-188. doi:10.1093/gerona/51a.4.m185

Franzen, S., van den Berg, E., Goudsmit, M., Jurgens, C. K., van de Wiel, L., Kalkisim, Y., . . . Papma, J. M. (2020). A Systematic Review of Neuropsychological Tests for the

Assessment of Dementia in Non-Western, Low-Educated or Illiterate Populations. *J Int Neuropsychol Soc, 26*(3), 331-351. doi:10.1017/s1355617719000894

Fuller-Thomson, E., & Kuh, D. (2014). The healthy migrant effect may confound the link between bilingualism and delayed onset of Alzheimer's disease. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 52*, 128-130. doi:10.1016/j.cortex.2013.08.009

Gasquoine, P. G. (1999). Variables moderating cultural and ethnic differences in neuropsychological assessment: The case of Hispanic Americans. *The Clinical Neuropsychologist, 13*(3), 376-383.

Gasquoine, P. G. (2009). Race-norming of neuropsychological tests. *Neuropsychology review, 19*(2), 250.

Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007). Language of administration and neuropsychological test performance in neurologically intact Hispanic American bilingual adults. *Archives of clinical neuropsychology, 22*(8), 991-1001.

Gasquoine, P. G., & Gonzalez, C. D. (2012). Using monolingual neuropsychological test norms with bilingual Hispanic americans: application of an individual comparison standard. *Arch Clin Neuropsychol, 27*(3), 268-276. doi:10.1093/arclin/acs004

Georges, J., Nielsen, T., Plejert, C., Rauf, M., Lahav, D., Jaakson, S., . . . Gove, D. (2019). *The development of intercultural care and support for people with dementia from minority ethnic groups*.

Glynn, K., O'Callaghan, M., Hannigan, O., Bruce, I., Gibb, M., Coen, R., . . . Robinson, D. (2021). Clinical utility of mild cognitive impairment subtypes and number of impaired cognitive domains at predicting progression to dementia: A 20-year retrospective study. *Int J Geriatr Psychiatry, 36*(1), 31-37. doi:10.1002/gps.5385

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *J Int Neuropsychol Soc, 13*(2), 197-208. doi:10.1017/s1355617707070038

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*(3), 787-814. doi:10.1016/j.jml.2007.07.001

Goudsmit, M., van Campen, J., Schilt, T., Hinnen, C., Franzen, S., & Schmand, B. (2018). One size does not fit all: comparative diagnostic accuracy of the rowland universal dementia assessment scale and the mini mental state examination in a memory clinic population with very low education. *Dementia and geriatric cognitive disorders extra, 8*(2), 290-305.

Green, S. (2013). Germany: A Changing Country of Immigration. *German Politics, 22*(3), 333-351. doi:10.1080/09644008.2013.832757

Guzmán-Vélez, E., & Tranel, D. (2015). Does bilingualism contribute to cognitive reserve? Cognitive and neural perspectives. *Neuropsychology, 29*(1), 139-150. doi:10.1037/neu0000105

Hall, J. R., Harvey, M., Vo, H. T., & O'Bryant, S. E. (2011). Performance on a measure of category fluency in cognitively impaired elderly. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 18*(3), 353-361. doi:10.1080/13825585.2011.557495

Harris, J. L., Kiran, S., Marquardt, T. P., & Fleming, V. B. (2008). Communication wellness check-up©: Age-related changes in communicative abilities. *Aphasiology, 22*(7-8), 813-825. doi:10.1080/02687030701818034

Hasher, L., & Zacks, R. T. (1988). Working Memory, Comprehension, and Aging: A Review and a New View. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 22, pp. 193-225): Academic Press.

Hasson, R., Wu, I., & Fine, J. (2019). Clinical utility of the WASI-II and its association with acculturation levels among Arab American adolescent males. *Applied Neuropsychology: Child, 8*(4), 295-306. doi:10.1080/21622965.2018.1442219

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behav Brain Sci, 33*(2-3), 61-83; discussion 83-135. doi:10.1017/s0140525x0999152x

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia, 42*(9), 1212-1222. doi:10.1016/j.neuropsychologia.2004.02.001

Hill, T. D., Angel, J. L., Balistreri, K. S., & Herrera, A. P. (2012). Immigrant status and cognitive functioning in late-life: an examination of gender variations in the healthy immigrant effect. *Soc Sci Med, 75*(12), 2076-2084. doi:10.1016/j.socscimed.2012.04.005

Hofmann, S. G. (2017). *Clinical psychology : a global perspective* Retrieved from http://www.myilibrary.com?id=1025661

Humpel, C. (2011). Identifying and validating biomarkers for Alzheimer's disease. *Trends Biotechnol, 29*(1), 26-32. doi:10.1016/j.tibtech.2010.09.007

Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., . . . Montine, T. J. (2012). National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement, 8*(1), 1-13. doi:10.1016/j.jalz.2011.10.007

Jack, C. R., Jr., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., . . . Phelps, C. H. (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association, 7*(3), 257-262. doi:10.1016/j.jalz.2011.03.004

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Sperling, R. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Feldman, H. H., Frisoni, G. B., . . . Dubois, B. (2016). A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology, 87*(5), 539-547. doi:10.1212/wnl.0000000000002923

Jack, C. R., Jr., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., . . . Trojanowski, J. Q. (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol, 12*(2), 207-216. doi:10.1016/s1474-4422(12)70291-0

Jacobs, D. M., Sano, M., Albert, S., Schofield, P., Dooneief, G., & Stern, Y. (1997). Cross-cultural neuropsychological assessment: A comparison of randomly selected, demographically matched cohorts of English-and Spanish-speaking older adults. *Journal of clinical and experimental neuropsychology, 19*(3), 331-339. doi:10.1080/01688639708403862

Jokel, R., Seixas Lima, B., Fernandez, A., & Murphy, K. J. (2019). Language in Amnestic Mild Cognitive Impairment and Dementia of Alzheimer’s Type: Quantitatively or

Qualitatively Different? *Dementia and geriatric cognitive disorders extra*, *9*(1), 136-151. doi:10.1159/000496824

Kaestner, R., Pearson, J. A., Keene, D., & Geronimus, A. T. (2009). Stress, Allostatic Load, and Health of Mexican Immigrants*. *Social Science Quarterly*, *90*(5), 1089-1111. doi:https://doi.org/10.1111/j.1540-6237.2009.00648.x

Kamat, R., Ghate, M., Gollan, T. H., Meyer, R., Vaida, F., Heaton, R. K., . . . Marcotte, T. D. (2012). Effects of Marathi-Hindi bilingualism on neuropsychological performance. *J Int Neuropsychol Soc*, *18*(2), 305-313. doi:10.1017/s1355617711001731

Kaufman, A. S., McLean, J. E., & Reynolds, C. R. (1988). Sex, race, residence, region, and education differences on the 11 WAIS—R subtests. *Journal of Clinical Psychology*, *44*(2), 231-248. doi:10.1002/1097-4679(198803)44:2<231::AID-JCLP2270440224>3.0.CO;2-J

Kaushanskaya, M., & Prior, A. (2014). Variability in the effects of bilingualism on cognition: It is not just about cognition, it is also about bilingualism. *Bilingualism: Language and Cognition*, *18*(1), 27-28. doi:10.1017/S1366728914000510

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychol Aging*, *23*(1), 70-78. doi:10.1037/0882-7974.23.1.70

Kertesz, A. (1994). Neuropsychological evaluation of language. *J Clin Neurophysiol*, *11*(2), 205-215. doi:10.1097/00004691-199403000-00005

Kim, B. S., Kim, Y. B., & Kim, H. (2019). Discourse Measures to Differentiate Between Mild Cognitive Impairment and Healthy Aging. *Front Aging Neurosci*, *11*, 221. doi:10.3389/fnagi.2019.00221

Kim, K., Kim, M.-J., Kim, D. W., Kim, S. Y., Park, S., & Park, C. B. (2020). Clinically accurate diagnosis of Alzheimer's disease via multiplexed sensing of core biomarkers in human plasma. *Nature Communications*, *11*(1), 119. doi:10.1038/s41467-019-13901-z

Kirova, A. M., Bays, R. B., & Lagalwar, S. (2015). Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. *Biomed Res Int*, *2015*, 748212. doi:10.1155/2015/748212

Kisser, J. E., Wendell, C. R., Spencer, R. J., & Waldstein, S. R. (2012). Neuropsychological performance of native versus non-native English speakers. *Archives of clinical neuropsychology*, *27*(7), 749-755.

Klimova, B., & Kuca, K. (2016). Speech and language impairments in dementia. *Journal of Applied Biomedicine, 14*(2), 97-103.

Kopec, J. A., Williams, J. I., To, T., & Austin, P. C. (2001). Cross-cultural comparisons of health status in Canada using the Health Utilities Index. *Ethn Health, 6*(1), 41-50. doi:10.1080/13557850125061

Krch, D., Lequerica, A., Arango-Lasprilla, J. C., Rogers, H. L., DeLuca, J., & Chiaravalloti, N. D. (2015). The Multidimensional Influence of Acculturation on Digit Symbol-Coding and Wisconsin Card Sorting Test in Hispanics. *The Clinical Neuropsychologist, 29*(5), 624-638. doi:10.1080/13854046.2015.1063696

Krist, L., Keller, T., Sebald, L., Yesil-Jürgens, R., Ellert, U., Reich, A., . . . Liman, T. G. (2019). The Montreal Cognitive Assessment (MoCA) in a population-based sample of Turkish migrants living in Germany. *Aging Ment Health, 23*(1), 30-37. doi:10.1080/13607863.2017.1396577

Kroll, J. F., & Bialystok, E. (2013). Understanding the Consequences of Bilingualism for Language Processing and Cognition. *J Cogn Psychol (Hove), 25*(5). doi:10.1080/20445911.2013.799170

Kuller, L. H., Shemanski, L., Manolio, T., Haan, M., Fried, L., Bryan, N., . . . Bhadelia, R. (1998). Relationship Between ApoE, MRI Findings, and Cognitive Function in the Cardiovascular Health Study. *Stroke, 29*(2), 388-398. doi:doi:10.1161/01.STR.29.2.388

Laws, K. R., Duncan, A., & Gale, T. M. (2010). 'Normal' semantic-phonemic fluency discrepancy in Alzheimer's disease? A meta-analytic study. *Cortex, 46*(5), 595-601. doi:10.1016/j.cortex.2009.04.009

Lemstra, A. W., de Beer, M. H., Teunissen, C. E., Schreuder, C., Scheltens, P., van der Flier, W. M., & Sikkes, S. A. (2017). Concomitant AD pathology affects clinical manifestation and survival in dementia with Lewy bodies. *J Neurol Neurosurg Psychiatry, 88*(2), 113-118. doi:10.1136/jnnp-2016-313775

Light, L. L., & Burke, D. M. (1993). Patterns of language and memory in old age *Language, memory, and aging.* (pp. 244-271). New York, NY, US: Cambridge University Press.

Lischka, A. R., Mendelsohn, M., Overend, T., & Forbes, D. (2012). A Systematic Review of Screening Tools for Predicting the Development of Dementia. *Canadian Journal on*

Aging / La Revue canadienne du vieillissement, 31(3), 295-311. doi:10.1017/S0714980812000220

Llorente, A. M. (2007). *Principles of neuropsychological assessment with Hispanics: Theoretical foundations and clinical practice*: Springer Science & Business Media.

Lobo, A., Launer, L. J., Fratiglioni, L., Andersen, K., Di Carlo, A., Breteler, M. M., . . . Hofman, A. (2000). Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology, 54*(11 Suppl 5), S4-9.

Loewenstein, D. A., Argüelles, T., Argüelles, S., & Linn-fuentes, P. (1994). Potential cultural bias in the neuropsychological assessment of the older adult. *Journal of clinical and experimental neuropsychology, 16*(4), 623-629. doi:10.1080/01688639408402673

Loewenstein, D. A., Argüelles, T., Barker, W. W., & Duara, R. (1993). A Comparative Analysis of Neuropsychological Test Performance of Spanish-Speaking and English-Speaking Patients With Alzheimer's Disease. *Journal of Gerontology, 48*(3), P142-P149. doi:10.1093/geronj/48.3.P142

Lonie, J. A., Herrmann, L. L., Tierney, K. M., Donaghey, C., O'Carroll, R., Lee, A., & Ebmeier, K. P. (2009). Lexical and semantic fluency discrepancy scores in aMCI and early Alzheimer's disease. *J Neuropsychol, 3*(Pt 1), 79-92. doi:10.1348/174866408x289935

Lopez-Gonzalez, L., Aravena, V. C., & Hummer, R. A. (2005). Immigrant Acculturation, Gender and Health Behavior: A Research Note. *Social Forces, 84*(1), 581-593. doi:10.1353/sof.2005.0112

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *J Cogn Psychol (Hove), 25*(5), 605-621. doi:10.1080/20445911.2013.795574

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections *Aging and cognition: Knowledge organization and utilization.* (pp. 213-263). Oxford, England: North-Holland.

Manly, J. J. (2008). Critical issues in cultural neuropsychology: Profit from diversity. *Neuropsychology review, 18*(3), 179.

Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied neuropsychology, 11*(1), 37-46.

Manly, J. J., & Espino, D. V. (2004). Cultural influences on dementia recognition and management. *Clinics in geriatric medicine, 20*(1), 93-119.

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., & Stern, Y. (1998). Cognitive test performance among nondemented elderly African Americans and whites. *Neurology, 50*(5), 1238-1245. doi:10.1212/wnl.50.5.1238

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *J Int Neuropsychol Soc, 8*(3), 341-348. doi:10.1017/s1355617702813157

Manly, J. J., & Mayeux, R. (2004). Ethnic differences in dementia and Alzheimer's disease. *Critical Perspectives on Racial and Ethnic Differentials in Health in Late Life*, 95-141.

Mantzavinos, V., & Alexiou, A. (2017). Biomarkers for Alzheimer's Disease Diagnosis. *Curr Alzheimer Res, 14*(11), 1149-1154. doi:10.2174/1567205014666170203125942

Manuel-Dupont, S., Ardila, A., Rosselli, M., & Puente, A. E. (1992). Bilingualism. In A. E. Puente & R. J. McCaffrey (Eds.), *Handbook of Neuropsychological Assessment: A Biopsychosocial Perspective* (pp. 193-210). Boston, MA: Springer US.

Marcopulos, B. A., McLain, C. A., & Giuliano, A. J. (1997). Cognitive impairment or inadequate norms? A study of healthy, rural, older adults with limited education. *The Clinical Neuropsychologist, 11*(2), 111-131. doi:10.1080/13854049708407040

Markides, K. S., & Gerst, K. (2011). Immigration, Aging, and Health in the United States. In R. A. Settersten & J. L. Angel (Eds.), *Handbook of Sociology of Aging* (pp. 103-116). New York, NY: Springer New York.

Martin, C.-O., Pontbriand-Drolet, S., Daoust, V., Yamga, E., Amiri, M., Hübner, L. C., & Ska, B. (2018). Narrative Discourse in Young and Older Adults: Behavioral and NIRS Analyses. *Frontiers in aging neuroscience, 10*, 69-69. doi:10.3389/fnagi.2018.00069

Maseda, A., Lodeiro-Fernández, L., Lorenzo-López, L., Núñez-Naveira, L., Balo, A., & Millán-Calenti, J. C. (2014). Verbal fluency, naming and verbal comprehension: three aspects of

language as predictors of cognitive impairment. *Aging & mental health, 18*(8), 1037-1045.

Matallana, D., de Santacruz, C., Cano, C., Reyes, P., Samper-Ternent, R., Markides, K. S., . . . Reyes-Ortiz, C. A. (2010). The Relationship Between Education Level and Mini-Mental State Examination Domains Among Older Mexican Americans. *J Geriatr Psychiatry Neurol, 24*(1), 9-18. doi:10.1177/0891988710373597

Mazzon, G., Ajčević, M., Cattaruzza, T., Menichelli, A., Guerriero, M., Capitanio, S., . . . Marini, A. (2019). Connected Speech Deficit as an Early Hallmark of CSF-defined Alzheimer's Disease and Correlation with Cerebral Hypoperfusion Pattern. *Curr Alzheimer Res, 16*(6), 483-494. doi:10.2174/1567205016666190506141733

McKhann, G. M., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*(7), 939-939.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., . . . Mayeux, R. (2011). The diagnosis of dementia due to 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), 263-269.

Meng, X., & D'Arcy, C. (2012). Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. *PloS one, 7*(6), e38268-e38268. doi:10.1371/journal.pone.0038268

Messer, R. H. (2017). *Heather Harris Wright (Ed.), Cognition, Language, and Aging. Amsterdam/Philadelphia: John Benjamins Publishing, 2016. Pp. 258 ISBN 978-9-027-21232-0* (02/27 ed. Vol. 9): Cambridge University Press.

Milman, L. H., Faroqi-Shah, Y., Corcoran, C. D., & Damele, D. M. (2018). Interpreting Mini-Mental State Examination Performance in Highly Proficient Bilingual Spanish-English and Asian Indian-English Speakers: Demographic Adjustments, Item Analyses, and Supplemental Measures. *J Speech Lang Hear Res, 61*(4), 847-856. doi:10.1044/2017_jslhr-l-17-0021

Miyake, A., Just, M. A., & Carpenter, P. A. (1994). Working Memory Constraints on the Resolution of Lexical Ambiguity: Maintaining Multiple Interpretations in Neutral Contexts. *Journal of Memory and Language, 33*(2), 175-202. doi:https://doi.org/10.1006/jmla.1994.1009

Molinuevo, J. L., Ayton, S., Batrla, R., Bednar, M. M., Bittner, T., Cummings, J., . . . Blennow, K. (2018). Current state of Alzheimer's fluid biomarkers. *Acta Neuropathol, 136*(6), 821-853. doi:10.1007/s00401-018-1932-x

Monsees, J., Hoffmann, W., & Thyrian, J. R. (2019). Prävalenz von Demenz bei Menschen mit Migrationshintergrund in Deutschland. *Zeitschrift für Gerontologie und Geriatrie, 52*(7), 654-660. doi:10.1007/s00391-018-01469-0

Monsees, J., Schmachtenberg, T., Hoffmann, W., Kind, A., Gilmore-Bykovskyi, A., Kim, A. J., & Thyrian, J. R. (2020). Dementia in People with a Turkish Migration Background: Experiences and Utilization of Healthcare Services. *J Alzheimers Dis, 77*(2), 865-875. doi:10.3233/jad-200184

Mortimer, J. A. (1997). Brain reserve and the clinical expression of Alzheimer's disease. *Geriatrics, 52 Suppl 2*, S50-53.

Moyerman, D. R., & Forman, B. D. (1992). Acculturation and Adjustment: A Meta-Analytic Study. *Hispanic Journal of Behavioral Sciences, 14*(2), 163-200. doi:10.1177/07399863920142001

Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *J Clin Exp Neuropsychol, 40*(9), 917-939. doi:10.1080/13803395.2018.1446513

Mungas, D., Marshall, S. C., Weldon, M., Haan, M., & Reed, B. R. (1996). Age and education correction of Mini-Mental State Examination for English- and Spanish-speaking elderly. *Neurology, 46*(3), 700-706. doi:10.1212/wnl.46.3.700

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *J Int Neuropsychol Soc, 12*(4), 570-574. doi:10.1017/s1355617706060590

Naqvi, R. M., Haider, S., Tomlinson, G., & Alibhai, S. (2015). Cognitive assessments in multicultural populations using the Rowland Universal Dementia Assessment Scale: a systematic review and meta-analysis. *Cmaj, 187*(5), E169-175. doi:10.1503/cmaj.140802

Naylor, M. D., Karlawish, J. H., Arnold, S. E., Khachaturian, A. S., Khachaturian, Z. S., Lee, V. M., . . . Trojanowski, J. Q. (2012). Advancing Alzheimer's disease diagnosis, treatment, and care: recommendations from the Ware Invitational Summit. *Alzheimers Dement, 8*(5), 445-452. doi:10.1016/j.jalz.2012.08.001

Ng, K. P., Chiew, H. J., Lim, L., Rosa-Neto, P., Kandiah, N., & Gauthier, S. (2018). The influence of language and culture on cognitive assessment tools in the diagnosis of early cognitive impairment and dementia. *Expert Rev Neurother, 18*(11), 859-869. doi:10.1080/14737175.2018.1532792

Nielsen, T. R., Segers, K., Vanderaspoilden, V., Beinhoff, U., Minthon, L., Pissiota, A., . . . Waldemar, G. (2019). Validation of a European Cross-Cultural Neuropsychological Test Battery (CNTB) for evaluation of dementia. *Int J Geriatr Psychiatry, 34*(1), 144-152. doi:10.1002/gps.5002

Nielsen, T. R., Vogel, A., Gade, A., & Waldemar, G. (2012). Cognitive testing in non-demented Turkish immigrants--comparison of the RUDAS and the MMSE. *Scand J Psychol, 53*(6), 455-460. doi:10.1111/sjop.12018

Nielsen, T. R., Vogel, A., Phung, T. K., Gade, A., & Waldemar, G. (2011). Over- and under-diagnosis of dementia in ethnic minorities: a nationwide register-based study. *Int J Geriatr Psychiatry, 26*(11), 1128-1135. doi:10.1002/gps.2650

Nielsen, T. R., Vogel, A., Riepe, M. W., de Mendonça, A., Rodriguez, G., Nobili, F., . . . Waldemar, G. (2011). Assessment of dementia in ethnic minority patients in Europe: a European Alzheimer's Disease Consortium survey. *Int Psychogeriatr, 23*(1), 86-95. doi:10.1017/s1041610210000955

Nielsen, T. R., & Waldemar, G. (2016). Knowledge and perceptions of dementia and Alzheimer's disease in four ethnic groups in Copenhagen, Denmark. *Int J Geriatr Psychiatry, 31*(3), 222-230. doi:10.1002/gps.4314

Oltra-Cucarella, J., Ferrer-Cascales, R., Alegret, M., Gasparini, R., Díaz-Ortiz, L. M., Ríos, R., . . . Sánchez-SanSegundo, M. (2018). Risk of progression to Alzheimer's disease for different neuropsychological Mild Cognitive Impairment subtypes: A hierarchical meta-

analysis of longitudinal studies. *Psychol Aging, 33*(7), 1007-1021. doi:10.1037/pag0000294

Orange, J. B., & Kertesz, A. (2000). Discourse Analyses and Dementia. *Brain and Language, 71*(1), 172-174. doi:https://doi.org/10.1006/brln.1999.2243

Ossher, L., Bialystok, E., Craik, F. I. M., Murphy, K. J., & Troyer, A. K. (2013). The effect of bilingualism on amnestic mild cognitive impairment. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 68*(1), 8-12. doi:10.1093/geronb/gbs038

Paplikar, A., Alladi, S., Varghese, F., Mekala, S., Arshad, F., Sharma, M., . . . Consortium, I.-N. (2021). Bilingualism and Its Implications for Neuropsychological Evaluation. *Archives of clinical neuropsychology, 36*(8), 1511-1522. doi:10.1093/arclin/acab012

Parker, C., & Philp, I. (2004). Screening for cognitive impairment among older people in black and minority ethnic groups. *Age Ageing, 33*(5), 447-452. doi:10.1093/ageing/afh135

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: a concept in evolution. *Journal of internal medicine, 275*(3), 214-228.

Pistono, A., Jucla, M., Bézy, C., Lemesle, B., Le Men, J., & Pariente, J. (2019). Discourse macrolinguistic impairment as a marker of linguistic and extralinguistic functions decline in early Alzheimer's disease. *Int J Lang Commun Disord, 54*(3), 390-400. doi:10.1111/1460-6984.12444

Pontón, M. O. (2001). Research and assessment issues with Hispanic populations *Neuropsychology and the Hispanic patient: A clinical handbook.* (pp. 39-58). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement, 9*(1), 63-75.e62. doi:10.1016/j.jalz.2012.11.007

Puente, A. E., Perez-Garcia, M., Lopez, R. V., Hidalgo-Ruzzante, N. A., & Fasfous, A. F. (2013). Neuropsychological assessment of culturally and educationally dissimilar individuals *Handbook of multicultural mental health: Assessment and treatment of diverse populations, 2nd ed.* (pp. 225-241). San Diego, CA, US: Elsevier Academic Press.

Qian, W., Schweizer, T. A., & Fischer, C. E. (2013). Impact of socioeconomic status on initial clinical presentation to a memory disorders clinic. *International Psychogeriatrics, 26*(4), 597-603. doi:10.1017/S1041610213002299

Qiu, C., Kivipelto, M., & von Strauss, E. (2009). Epidemiology of Alzheimer's disease: occurrence, determinants, and strategies toward intervention. *Dialogues in clinical neuroscience, 11*(2), 111-128. doi:10.31887/DCNS.2009.11.2/cqiu

Quinteros Baumgart, C., & Billick, S. B. (2018). Positive Cognitive Effects of Bilingualism and Multilingualism on Cerebral Function: a Review. *Psychiatr Q, 89*(2), 273-283. doi:10.1007/s11126-017-9532-9

Ramírez-Esparza, N., & García-Sierra, A. (2014). The bilingual brain: Language, culture, and identity *The Oxford handbook of multicultural identity.* (pp. 35-56). New York, NY, US: Oxford University Press.

Ramirez, M., Teresi, J. A., Holmes, D., Gurland, B., & Lantigua, R. (2006). Differential Item Functioning (DIF) and the Mini-Mental State Examination (MMSE): Overview, Sample, and Issues of Translation. *Medical Care, 44*(11, Suppl 3), S95-S106. doi:10.1097/01.mlr.0000245181.96133.db

Rasmusson, D. X., Carson, K. A., Brookmeyer, R., Kawas, C., & Brandt, J. (1996). Predicting rate of cognitive decline in probable Alzheimer's disease. *Brain Cogn, 31*(2), 133-147. doi:10.1006/brcg.1996.0038

Reitz, C., & Mayeux, R. (2014). Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem Pharmacol, 88*(4), 640-651. doi:10.1016/j.bcp.2013.12.024

Rentz, D. M., Rodriguez, M. A. P., Amariglio, R., Stern, Y., Sperling, R., & Ferris, S. (2013). Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer’s disease: a selective review. *Alzheimer's research & therapy, 5*(6), 1-10.

Reynolds, C. R., Chastain, R. L., Kaufman, A. S., & McLean, J. E. (1987). Demographic characteristics and IQ among adults: Analysis of the WAIS-R standardization sample as a function of the stratification variables. *Journal of School Psychology, 25*(4), 323-342. doi:https://doi.org/10.1016/0022-4405(87)90035-5

Rivera Mindt, M., Arentoft, A., Kubo Germano, K., D'Aquila, E., Scheiner, D., Pizzirusso, M., . . . Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychol Rev, 18*(3), 255-268. doi:10.1007/s11065-008-9069-7

Rivera Mindt, M., Byrd, D., Saez, P., & Manly, J. (2010). Increasing Culturally Competent Neuropsychological Services for Ethnic Minority Populations: A Call to Action. *The Clinical Neuropsychologist, 24*(3), 429-453. doi:10.1080/13854040903058960

Romero, H. R., Lageman, S. K., Kamath, V., Irani, F., Sim, A., Suarez, P., . . . the Summit, p. (2009). Challenges in the Neuropsychological Assessment of Ethnic Minorities: Summit Proceedings. *The Clinical Neuropsychologist, 23*(5), 761-779. doi:10.1080/13854040902881958

Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Matos, L., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. *Int J Neurosci, 112*(6), 759-776. doi:10.1080/00207450290025752

Ryder, A. G., Alden, L. E., & Paulhus, D. L. (2000). Is acculturation unidimensional or bidimensional? A head-to-head comparison in the prediction of personality, self-identity, and adjustment. *Journal of Personality and Social Psychology, 79*(1), 49-65. doi:10.1037/0022-3514.79.1.49

Sajjadi, S. A., Patterson, K., Tomek, M., & Nestor, P. J. (2012). Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. *Aphasiology, 26*(6), 847-866. doi:10.1080/02687038.2012.654933

Sala, S. D., Lorenzi, L., Spinnler, H., & Zuffi, M. (1993). Components in the breakdown of verbal communication in Alzheimer's disease. *Aphasiology, 7*(3), 285-299. doi:10.1080/02687039308249511

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

Samuel, S., Roehr-Brackin, K., Pak, H., & Kim, H. (2018). Cultural Effects Rather Than a Bilingual Advantage in Cognition: A Review and an Empirical Study. *Cogn Sci, 42*(7), 2313-2341. doi:10.1111/cogs.12672

Sayegh, P. (2015). Neuropsychological language tests in dementia diagnosis in English-speaking Hispanic and non-Hispanic white outpatients. *Applied Neuropsychology: Adult*, *22*(6), 435-444.

Schmachtenberg, T., Monsees, J., & Thyrian, J. R. (2021). What elements are important in current treatment and care guidelines to provide culturally sensitive care for people with a migration background and dementia? A systematic analysis. *Journal of Public Health*. doi:10.1007/s10389-021-01531-z

Schmitter-Edgecombe, M., Vesneski, M., & Jones, D. W. R. (2000). Aging and word-finding: A comparison of spontaneous and constrained naming tests. *Archives of clinical neuropsychology, 15*(6), 479-493. doi:10.1016/S0887-6177(99)00039-6

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

Sedo, M. (2004). Sedo, M. A. (2004) The "Five Digit Test": a color-free, non-reading alternative to the Stroop. International Neuropsychological Society Liaison Committee Newsletter, 13, 6-7. *Sedo, M. A. (2004) The "Five Digit Test": a color-free, non-reading alternative to the Stroop. International Neuropsychological Society Liaison Committee Newsletter, 13, 6-7.*, 6-7.

Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., . . . Trojanowski, J. Q. (2009). Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol, 65*(4), 403-413. doi:10.1002/ana.21610

Sheffield, K. M., & Peek, M. K. (2009). Neighborhood context and cognitive decline in older Mexican Americans: results from the Hispanic Established Populations for Epidemiologic Studies of the Elderly. *Am J Epidemiol, 169*(9), 1092-1101. doi:10.1093/aje/kwp005

Snitz, B. E., Unverzagt, F. W., Chang, C. C., Bilt, J. V., Gao, S., Saxton, J., . . . Ganguli, M. (2009). Effects of age, gender, education and race on two tests of language ability in community-based older adults. *Int Psychogeriatr, 21*(6), 1051-1062. doi:10.1017/s1041610209990214

Snyder, P. J., Kahle-Wrobleski, K., Brannan, S., Miller, D. S., Schindler, R. J., DeSanti, S., . . . Carrillo, M. C. (2014). Assessing cognition and function in Alzheimer's disease clinical trials: Do we have the right tools? *Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 10*(6), 853-860. doi:10.1016/j.jalz.2014.07.158

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., . . . Montine, T. J. (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.

Sperling, R. A., Karlawish, J., & Johnson, K. A. (2013). Preclinical Alzheimer disease—the challenges ahead. *Nature Reviews Neurology, 9*(1), 54-58.

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

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol, 11*(11), 1006-1012. doi:10.1016/s1474-4422(12)70191-6

Stern, Y., Barnes, C. A., Grady, C., Jones, R. N., & Raz, N. (2019). Brain reserve, cognitive reserve, compensation, and maintenance: operationalization, validity, and mechanisms of cognitive resilience. *Neurobiol Aging, 83*, 124-129. doi:10.1016/j.neurobiolaging.2019.03.022

Storey, J. E., Rowland, J. T., Basic, D., Conforti, D. A., & Dickson, H. G. (2004). The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale. *Int Psychogeriatr, 16*(1), 13-31. doi:10.1017/s1041610204000043

Sullivan, M. D., Poarch, G. J., & Bialystok, E. (2018). Why is Lexical Retrieval Slower for Bilinguals? Evidence from Picture Naming. *Bilingualism (Cambridge, England), 21*(3), 479-488. doi:10.1017/S1366728917000694

Sulpizio, S., Del Maschio, N., Del Mauro, G., Fedeli, D., & Abutalebi, J. (2020). Bilingualism as a gradient measure modulates functional connectivity of language and control networks. *NeuroImage, 205*, 116306. doi:https://doi.org/10.1016/j.neuroimage.2019.116306

Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J., & Pakaski, M. (2015). Speaking in Alzheimer’s disease, is that an early sign? Importance of changes in language abilities in Alzheimer’s disease. *Frontiers in aging neuroscience, 7*, 195.

Tadmor, C. T., Tetlock, P. E., & Kaiping, P. (2009). Acculturation Strategies and Integrative Complexity: The Cognitive Implications of Biculturalism. *Journal of Cross-Cultural Psychology, 40*(1), 105-139. doi:10.1177/0022022108326279

Tahami Monfared, A. A., Byrnes, M. J., White, L. A., & Zhang, Q. (2022). Alzheimer's Disease: Epidemiology and Clinical Progression. *Neurology and Therapy, 11*(2), 553-569. doi:10.1007/s40120-022-00338-8

Taler, V., & Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *Journal of clinical and experimental neuropsychology, 30*(5), 501-556.

Tan, Y. W., & Burgess, G. H. (2020). Multidimensional effects of acculturation at the construct or index level of seven broad neuropsychological skills. *Culture and Brain, 8*(1), 27-45. doi:10.1007/s40167-018-0075-1

Tarawneh, R., & Holtzman, D. M. (2012). The clinical problem of symptomatic Alzheimer disease and mild cognitive impairment. *Cold Spring Harb Perspect Med, 2*(5), a006148. doi:10.1101/cshperspect.a006148

Testa, J. A., Ivnik, R. J., Boeve, B., Petersen, R. C., Pankratz, V. S., Knopman, D., . . . Smith, G. E. (2004). Confrontation naming does not add incremental diagnostic utility in MCI and Alzheimer's disease. *J Int Neuropsychol Soc, 10*(4), 504-512. doi:10.1017/s1355617704104177

Thomson, E. F., Nuru-Jeter, A., Richardson, D., Raza, F., & Minkler, M. (2013). The Hispanic Paradox and older adults' disabilities: is there a healthy migrant effect? *Int J Environ Res Public Health, 10*(5), 1786-1814. doi:10.3390/ijerph10051786

Thornton, R., & Light, L. L. (2006). *Language Comprehension and Production in Normal Aging*.

Toledo, J. B., Arnold, S. E., Raible, K., Brettschneider, J., Xie, S. X., Grossman, M., . . . Trojanowski, J. Q. (2013). Contribution of cerebrovascular disease in autopsy confirmed neurodegenerative disease cases in the National Alzheimer's Coordinating Centre. *Brain, 136*(Pt 9), 2697-2706. doi:10.1093/brain/awt188

Traykov, L., Raoux, N., Latour, F., Gallo, L., Hanon, O., Baudic, S., . . . Rigaud, A. S. (2007). Executive functions deficit in mild cognitive impairment. *Cogn Behav Neurol, 20*(4), 219-224. doi:10.1097/WNN.0b013e31815e6254

Tsantali, E., Economidis, D., & Tsolaki, M. (2013). Could language deficits really differentiate Mild Cognitive Impairment (MCI) from mild Alzheimer's disease? *Arch Gerontol Geriatr, 57*(3), 263-270. doi:10.1016/j.archger.2013.03.011

Turrell, G., Lynch, J. W., Kaplan, G. A., Everson, S. A., Helkala, E. L., Kauhanen, J., & Salonen, J. T. (2002). Socioeconomic position across the lifecourse and cognitive function in late middle age. *J Gerontol B Psychol Sci Soc Sci, 57*(1), S43-51. doi:10.1093/geronb/57.1.s43

Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and dementia: a systematic review. *Psychol Med, 36*(4), 441-454. doi:10.1017/s0033291705006264

Valian, V. (2014). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(1), 3-24. doi:10.1017/S1366728914000522

van den Noort, M., Struys, E., & Bosch, P. (2019). Individual Variation and the Bilingual Advantage-Factors that Modulate the Effect of Bilingualism on Cognitive Control and Cognitive Reserve. *Behav Sci (Basel), 9*(12). doi:10.3390/bs9120120

van Dijk, T. A. (2006). Discourse, context and cognition. *Discourse Studies, 8*(1), 159-177. doi:10.1177/1461445606059565

Vega, J. N., & Newhouse, P. A. (2014). Mild cognitive impairment: diagnosis, longitudinal course, and emerging treatments. *Curr Psychiatry Rep, 16*(10), 490. doi:10.1007/s11920-014-0490-8

Voits, T., Pliatsikas, C., Robson, H., & Rothman, J. (2020). Beyond Alzheimer's disease: Can bilingualism be a more generalized protective factor in neurodegeneration? *Neuropsychologia, 147*, 107593. doi:10.1016/j.neuropsychologia.2020.107593

Watson, C. W., Manly, J. J., & Zahodne, L. B. (2016). Does bilingualism protect against cognitive aging?: Methodological issues in research on bilingualism, cognitive reserve, and dementia incidence. *Linguist Approaches Biling, 6*(5), 590-604. doi:10.1075/lab.15043.wat

Weintraub, S., Wicklund, A. H., & Salmon, D. P. (2012). The neuropsychological profile of Alzheimer disease. *Cold Spring Harbor perspectives in medicine, 2*(4), a006171.

Weissberger, G. H., Strong, J. V., Stefanidis, K. B., Summers, M. J., Bondi, M. W., & Stricker, N. H. (2017). Diagnostic Accuracy of Memory Measures in Alzheimer's Dementia and

Mild Cognitive Impairment: a Systematic Review and Meta-Analysis. *Neuropsychol Rev, 27*(4), 354-388. doi:10.1007/s11065-017-9360-6

Welsh, K. A., Fillenbaum, G., Wilkinson, W., Heyman, A., Mohs, R. C., Stern, Y., . . . Beekly, D. (1995). Neuropsychological test performance in African-American and white patients with Alzheimer's disease. *Neurology, 45*(12), 2207-2211. doi:10.1212/wnl.45.12.2207

Willers, I. F., Feldman, M. L., & Allegri, R. F. (2008). Subclinical naming errors in mild cognitive impairment: A semantic deficit? *Dementia & Neuropsychologia, 2*(3), 217-222. doi:10.1590/S1980-57642009DN20300010

Williams, E., McAuliffe, M., & Theys, C. (2021). Language changes in Alzheimer's disease: A systematic review of verb processing. *Brain Lang, 223*, 105041. doi:10.1016/j.bandl.2021.105041

Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., Prina, A. M., Winblad, B., . . . Prince, M. (2017). The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimers Dement, 13*(1), 1-7. doi:10.1016/j.jalz.2016.07.150

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., . . . Petersen, R. C. (2004). Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med, 256*(3), 240-246. doi:10.1111/j.1365-2796.2004.01380.x

Wingfield, A., & Grossman, M. (2006). Language and the aging brain: patterns of neural compensation revealed by functional brain imaging. *J Neurophysiol, 96*(6), 2830-2839. doi:10.1152/jn.00628.2006

Woumans, E. V. Y., Santens, P., Sieben, A., Versijpt, J. A. N., Stevens, M., & Duyck, W. (2014). Bilingualism delays clinical manifestation of Alzheimer's disease. *Bilingualism: Language and Cognition, 18*(3), 568-574. doi:10.1017/S136672891400087X

Wu, H., Yu, Z., Wang, X., & Zhang, Q. (2020). Language processing in normal aging: Contributions of information-universal and information-specific factors. *Acta Psychologica Sinica, 52*(5), 541.

Xu, H., Zhang, Y., & Wu, B. (2017). Association between migration and cognitive status among middle-aged and older adults: a systematic review. *BMC Geriatrics, 17*(1), 184. doi:10.1186/s12877-017-0585-2

Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: why variations in bilingual experiences matter. *Front Psychol, 6*, 164. doi:10.3389/fpsyg.2015.00164

Zacks, R., & Hasher, L. (1997). Cognitive Gerontology and Attentional Inhibition: A Reply to Burke and McDowd. *The Journals of Gerontology: Series B, 52B*(6), P274-P283. doi:10.1093/geronb/52B.6.P274

Appendix A 1st Publication (Study 1)

S1: PRISMA 2009 Checklist			
Section/topic	**#**	**Checklist item**	**Reported on page #**
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6-7
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	--
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	S2

Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	11
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	--
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	--
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	--
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	9, Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	9-11, Tab. 1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-9, Tab. 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	13-19, Tab. 2
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	--

Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	--
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	--
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	19-23
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	24
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	25-28
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	28

S2: Search strategy				
Database	**Search string**	**Hits**	**Date**	**Years**
Web of Science	(TS=("alzheimer's disease" OR "mild cognitive impairment") AND TS=(discourse) OR TS=("global coherence") OR TS=(macrolinguistic) OR TS=("connected language") OR TS=("connected speech") OR TS=("narrative comprehension") OR TS=("narrative speech")) *AND* **LANGUAGE:** (English) *AND* **DOCUMENT TYPES:** (Article)	1597	08.03.2018	1934-2020
	Indexes=SCI-EXPANDED, SSCI Timespan=All years	253	20.01.2020	
PubMed/ MedLine	((((((("alzheimer's disease"[Title/Abstract] OR "mild cognitive impairment"[Title/Abstract]) AND discourse[Title/Abstract]) OR "global coherence"[Title/Abstract]) OR "narrative discourse"[Title/Abstract]) OR macrolinguistic[Title/Abstract]) OR "connected language"[Title/Abstract]) OR "narrative comprehension"[Title/Abstract]) OR "connected speech"[Title/Abstract]) OR "narrative speech"[Title/Abstract]	955	08.03.2018	1954-2020
		180	20.01.2020	
PsycINFO/ EBSCO	("alzheimer's disease" OR "mild cognitive impairment") AND discourse OR "global coherence" OR "narrative discourse" OR macrolinguistic OR "connected language" OR "narrative comprehension" OR "connected speech" OR "narrative speech"	1587	08.03.2018	1934-2020
	Language: English	144	20.01.2020	

Appendix B 2nd Publication (Study 2)

TABLE S1: Characteristics of Cross-sectional Studies

First Author, Year	City, Country	Population (N)	Mean Age	Immigration status	Language of testing	Diagnostic criteria used	Quality assessment rating
Anderson et al. (2017)	Toronto, Canada	**HA** MLs: NS BLs. NS **MCI** MLs: 38 BLs: 36 **AD** MLs: 35 BLs: 40	**HA** MLs= 74.9 BLs= 74.7 **MCI** MLs= 66.5 BLs= 70 **AD** MLs= 74.2 BLs= 81.4	**HA** MLs: 4 BLs: 10 **MCI** MLs: 12 BLs: 25 **AD** MLs: 8 BLs: 27	English	NINCDS ADRDA1 (MMSE2); Albert et al., 2011	0.75
Bialystok et al. (2008)	Toronto, Canada	MLs: 24 BLs: 24	MLs= 67.2 BLs= 68.3	MLs: NS BLs: 20	English	------	0.75
Bialystok et al. (2014a)	Toronto, Canada	**HA Study 1** MLs: 25 BLs: 34 **HA Study 2** MLs: 18 BLs: 18	**HA Study 1** ML=71.3 BLs=67.6 **HA Study 2** MLs=72.4 BLs=69.1	NS	English	------	0.7
Clare et al. (2016a)	Wales, UK	**AD** ML: 49 BLs: 24	**AD** MLs=78.82 BLs: 81	Non-immigrants	English, Welsh	ICD-10[3] (MMSE)	0.95
Clare et al. (2016b)	North Wales, UK	**HA** MLs: 49 BLs: 50	**HA** MLs=72.55 BLs= 74.32	Non-immigrants	English, Welsh	------	0.9
Friesen et al. (2015)	------, Canada	**HA** MLs: 20 BLs: 21	MLs= 70.9 BLs=71.1	NS	English	------	0.75

Appendix

Ihle et al. (2016)	Different cities, Switzerland	MLs:1884 BLs: 492 TLs: 281 4Ls: 115 5Ls: 31 6Ls: 8	All= 77.9	NS	German or French or Italian (in native language of the participant)	------	0.75
Kousaie et al. (2014)	Ottowa, Quebec, Canada	**HA** ENG MLs: 31 FR MLs: 30 BLs: 36	**HA** ENG MLs=72.26 FR MLs= 72.60 BLs=70.69	Non-immigrants	English, French and in either language	------	0.8
Kowoll et al. (2015)	Heidelberg, Germany	**HA** MLs: 6 BLs: 11 **MCI** MLs: 14 BLs: 8 **AD** MLs: 25 BLs: 22	**HA** MLs= 70.2 BLs= 68.2 **MCI** MLs= 77.5 BLs= 71.3 **AD** MLs= 80.3 BLs= 77.2	**HA** MLs: 0 BLs: 8 **MCI** MLs: 4 BLs: 7 **AD** MLs: 7 BLs: 18	German, in both languages of participants for tests assessing language domain	NINCDS-ADRDA; Levy, 1994	0.7
Luo et al. (2013)	Toronto, Canada	**HA** MLs: 61 BLs: 60	**HA** MLs=71.7 BLs=67.8	NS	NS	------	0.75
Massa et al. (2020)	Toulouse, France	**HA** FR MLs: 16 BLs: 16	**HA** FR MLs= 71.1 BLs= 72.3	NS	French, Italian (participants` dominant language)	------	0.85
Nielsen et al. (2019)	Copenhagen, Denmark	**HA** MLs: 24 BLs:47	**HA** MLs= 63.00 BLs= 60.17	The whole sample was immigrants to Denmark	Turkish, Danish (participants´ language of preference)	------ (RUDAS)[4]	0.85

Ossher et al. (2012)	Toronto, Canada	**SD aMCI** MLs: 49 BLs: 19 **MD aMCI** MLs: 22 BLs: 21	**SD aMCI** MLs=74.9 BLs=79.4 **MD aMCI** MLs=75.2 BLs=72.6	NS	NS	Albert et al., 2011; Petersen, 2004	0.6
Papageorgiou et al. (2019)	London, UK	**HA** MLs: 37 BLs: 37	**HA** MLs= 69.4 BLs= 70.6	MLs: NS BLs: 12	English	------	0.75
Rosselli et al. (2000)	South Florida, USA	**HA** ENG MLs: 45 SP MLs: 18 BLs: 19	**HA** ENG MLs= 63.4 SP MLs= 61.3 BLs= 60.6	SP MLs: Immigrants BLs: 18	English, Spanish (in both languages of participants)	------	0.85
Rosselli et al. (2018)	Florida, USA	**aMCI** MLs: 25 BLs: 42	**aMCI** MLs= 73.60 BLs= 72.02	**aMCI** MLs: 4 BLs: Immigrants	English, Spanish (participants´ language of preference)	NS (CDR[5] score of 0.5; HVLT-R[6]; MINT[7]; TMT-A[8]; SCW[9])	0.9
Sheppard et al. (2016)	Ottowa-Gatineau, Quebec, Canada	**HA** ENG MLs: 31 FR MLs: 30 BLs: 36	**HA** ENG MLs= 72.26 FR MLs= 72.60 BLs= 70.69	90% of participants were born in Canada with the remainder born in the US or UK	English, French and either language	------	0.85
Soltani et al. (2019)	Ahvaz, Iran	**HA** MLs: 12 BLs: 12	**HA** All= 72.8	NS	Persian, Arabic (in both languages of participants)	------	0.8

Appendix

TABLE S2: Characteristics of Longitudinal Studies

First Author, Year	City, Country	Population (N)	Mean Age	Immigration status	Language of testing	Diagnostic criteria used	Quality assessment rating
Bak et al. (2014)	Edinburgh, UK	MLs: 591 BLs: 262	All= 72.49	Non-immigrants	English	------	0.8
Bialystok et al. (2014b)	Toronto, Canada	**MCI** MLs: 38 BLs: 36 **Probable AD** MLs:35 BLs: 40	**MCI** MLs= 66.5 BLs= 70.0 **Probable AD** MLs= 74.2 BLs= 81.4	**MCI** MLs: 12 BLs: 25 **Probable AD** MLs: 8 BLs: 27	English	NINCDS-ADRDA; Albert et al., 2011	0.85
Chertkow et al. (2010)	Montreal, Canada	**Probable AD** MLs: 379 (only 92 MLs were tested) BLs: 253 (only 62 BLs were tested)	------	MLs: 23 BLs: 135	NS	NINCDS-ADRDA	0.7
Costumero et al. (2020)	Valencia, Spain	**Baseline** / **Follow-up** ***MCI*** / **MCI** MLs:60 / MLs:15 BLs: 39 / BLs:15	**Baseline** / **Follow-up** ***MCI*** / ***MCI*** MLs= 73.58 / MLs= NS BLs= 74.26 / BLs= NS	Non-immigrants	Spanish	NS (WMS-III[10]; CDR score of 0.5)	0.6
Cox et al. (2016)	Edinburgh, UK	MLs: 64 BLs: 26	MLs= 74.45 BLs= 74.54	Non-immigrants	English	MMSE	0.95

Appendix

Mungas et al. (2018)	Sacramento, USA	**Language use** SP MLs: 628 ENG MLs: 307 BL: 524 **Proficiency** SP MLs: 409 ENG MLs: 48 BLs: 342	NS for language groups	SP MLs: 546 ENG MLs: 31 BLs: 142	English, Spanish (participants´ language of preference)	McKhann et al., 1984; Chui et al., 1992[11]	0.9
Padilla et al. (2016)	Sacramento, USA	**Baseline** **Follow-up** MLs: 289 MLs: 150 BLs: 339 BLs: 225	**Baseline** **Follow-up** MLs=73.6 MLs=73.0 BLs=74.3 BLs=73.4	The participants were born in Mexico	English, Spanish (participants´ language of preference)	------	0.8
Yeung et al. (2014)	Manitoba, Canada	**Baseline** **Follow-up** ***HA*** ***HA*** MLs: 856 MLs: 492 BLs: 79* BLs: 46* BLs: 533** BLs: 285** ***CI*** ***CI*** MLs: 29 MLs: 30 BLs: 2* BLs: 2* BLs: 60** BLs: 40** ***Dementia*** ***Dementia*** MLs:28 MLs: 54 BLs: 0* BLs: 6* BLs: 29** BLs: 35**	All= 77.9	------	English	DSM-III[12]; (3MS[13])	0.95

Appendix

Zahodne et al. (2014)	Manhattan, USA	MLs :637 BLs: 430 (Both groups consisted of different subtypes of dementia)	MLs=75.66 BLs=74.78	The whole sample was Spanish speaking immigrants to the US	Spanish	DSM-III; McKhann et al., 1984; Roman et al., 1993; McKeith et al., 1996	0.8

Notes: MLs= Monolinguals; BLs= Bilinguals; HA= Healthy older adults; MCI= Participants with mild cognitive impairment; AD= Participants with Alzheimer´s disease; TLs= Trilinguals; 4Ls= speakers of four languages; 5Ls=speakers of five languages; 6Ls= speakers of six languages; SD aMCI= Single domain amnestic MCI, MD aMCI= Multiple domain amnestic MCI; [1]National Institute of Neurological and Communicative Diseases-Alzheimer's Disease and Related Disorders Association (McKhann et al., 1984); [2]Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975); [3] Tenth Revision of the International Classification of Diseases and Related Health Problems (World Health Organization, 1992); [4] The Rowland Universal Dementia Assessment Scale (Storey, Rowland, Conforti & Dickson, 2004); [5]Clinical Dementia Rating Scale (Hughes, Berg, Danziger, Coben & Martin, 1982); [6] Hopkins Verbal Learning Test–Revised (Benedict, Schretlen, Groninger & Brandt, 1998); [7] Multilingual Naming Test (Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012); [8] Trail Making Test (Reitan, 1955); [9] Stroop Color Word Test (Stroop, 1935); [10] Wechsler Memory Scale, Third edition (Wechsler, 1997); [11] Criteria for the diagnosis of ischemic vascular dementia proposed by the State of California Alzheimer's Disease Diagnostic and Treatment Centers; [12] Diagnostic and Statistical Manual of Mental Disorders Third Edition (1980); [13] Modified Mini-Mental State Examination (Teng, Chui, 1987)

------indicates no information available or not applicable.

NS= Not specified *FR= French*

ENG= English *= English as a first language*

SP= Spanish **= English as a second language

References:

1. Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to 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), 270–279. https://doi.org/10.1016/j.jalz.2011.03.008
2. Levy, R. (1994). Aging-associated cognitive decline. *International Psychogeriatrics*, *6*(1), 63–68. https://doi.org/10.1017/S1041610294001626
3. McKeith, I. G., Galasko, D., Kosaka, K., Perry, E. K., Dickson, D. W., Hansen, L. A., Salmon, D. P., Lowe, J., Mirra, S. S., Byrne, E. J., Lennox, G., Quinn, N. P., Edwardson, J. A., Ince, P. G., Bergeron, C., Burns, A., Miller, B. L., Lovestone, S., Collerton, D., … Perry, R. H. (1996). Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on DLB international workshop. *Neurology*, *47*(5), 1113–1124. https://doi.org/10.1212/wnl.47.5.1113
4. McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*(7), 939–939. https://doi.org/10.1212/WNL.34.7.939
5. Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x
6. Román, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., Amaducci, L., Orgogozo, J. M., Brun, A., & Hofman, A. (1993). Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*, *43*(2), 250–250. https://doi.org/10.1212/wnl.43.2.250 .
7. Spitzer, R. L., Md, K. K., & Williams, J. B. (1980). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Association.

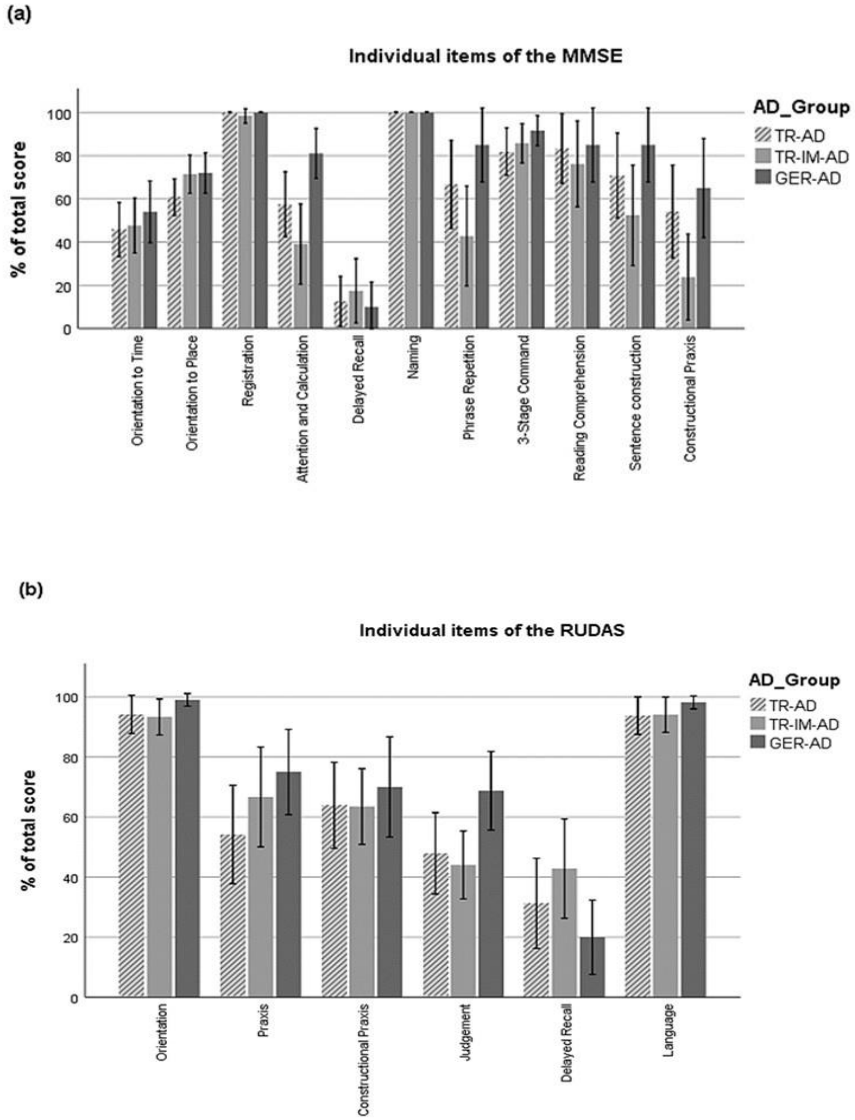
Appendix C 3rd Publication (Study 3)

Table S1. Between-Groups Comparison of the Individual Items of RUDAS and MMSE

Test Items *Median (Range)/Mean (SD)*	**TR-AD** **(n=24)**	**TR-IM-AD** **(n=21)**	**GER-AD** **(n=20)**	***p* value**
MMSE				
Orientation to Time[H]	2.5 (0-5)	2.0 (0-5)	3.0 (0-5)	.578
Mean (SD)	2.29 (1.49)	2.38 (1.40)	2.70 (1.53)	
Orientation to Place[H]	3.0 (1-5)	4.0 (1-5)	4.0 (2-5)	.075
Mean (SD)	3.04 (1.00)	3.57 (0.98)	3.60 (1.00)	
Registration[H]	3.0 (3)	3.0 (2-3)	3.0 (3)	.351
Mean (SD)	3.00 (.000)	2.95 (.218)	3.00 (.000)	
Delayed Recall[H]	.0 (0-3)	.0 (0-3)	.0 (0-3)	.740
Mean (SD)	.38 (.824)	.52 (.981)	.30 (.733)	
Naming[H]	2.0 (2)	2.0 (2)	2.0 (2)	1.000
Mean (SD)	2.00 (.000)	2.00 (.000)	2.00 (.000)	
3-Stage-Command[H]	3.0 (1-3)	3.0 (1-3)	3.0 (2-3)	.421
Mean (SD)	2.46 (.779)	2.57 (.598)	2.75 (.444)	
Reading Comprehension[‖] (correct, %)	20 (83.3%)	16 (76.2%)	17 (85%)	.784
Sentence construction[§] (correct, %)	17 (70.8%)	11 (52.4%)	17 (85%)	.076
RUDAS				
Orientation[H]	5.0 (2-5)	5.0 (3-5)	5.0 (4-5)	.241
Mean (SD)	4.71 (.751)	4.67 (.658)	4.95 (.224)	
Praxis[H]	1.0 (0-2)	1.0 (0-2)	2.0 (0-2)	.182
Mean (SD)	1.08 (.776)	1.33 (.730)	1.50 (.607)	
Constructional Praxis[H]	2.0 (0-3)	2.0 (0-3)	2.5 (0-3)	.648
Mean (SD)	1.92 (1.02)	1.90 (.831)	2.10 (1.07)	
Delayed Recall[H]	2.0 (0-8)	2.0 (0-8)	.0 (0-6)	.110
Mean (SD)	2.50 (2.84)	3.43 (2.91)	1.60 (2.13)	
Language[H]	8.0 (4-8)	8.0 (4-8)	8.0 (7-8)	.683
Mean (SD)	7.50 (1.18)	7.52 (1.03)	7.85 (.366)	

Note. TR-AD= Native-born Turkish Participants with AD, TR-IM-AD= Turkish Immigrants with AD, GER-AD= Native-born German Participants with AD. Significant *p* values are marked in bold.
H Kruskal-Wallis test
§ Pearson`s chi-square test
‖ Fisher´s exact test

Figure S2. Mean percentage of total score obtained from the individual items of the MMSE and RUDAS (Error bars represent 95% confidence interval)



Personal Contribution to the Publications of the Thesis

I. Publication

Kokje, E., Celik, S., Wahl, H. W., & von Stutterheim, C. (2021). Can discourse processing performance serve as an early marker of Alzheimer's disease and mild cognitive impairment? A systematic review of text comprehension. *European Journal of Ageing*. Advance online publication. https://doi.org/10.1007/s10433-021-00619-5.

Eesha Kokje developed the research questions, interpreted the findings of studies, and wrote the article. Eesha Kokje and I screened the article's eligibility for the review and assessed and rated the quality of the studies included in the study. The work and manuscript preparation were supervised by Prof. Christiane von Stutterheim and Prof. H.-W. Wahl.

II. Publication

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: Adult*, 1–19. https://doi.org/10.1080/23279095.2020.1788032.

I developed the research questions, evaluated the study findings, and wrote the article. Eesha Kokje and I screened the article's eligibility for the systematic review. Eesha Kokje also supported me in manuscript preparations by providing comments. Prof. Lutz Frölich, Prof. Patric Meyer, and Dr. Birgit Teichmann supervised the work and manuscript preparation.

III. Publication

Celik, S., Onur, O., Yener, G., Kessler, J., Özbek, Y., Meyer, P., Frölich, L., & Teichmann, B. (2022). Cross-cultural comparison of MMSE and RUDAS in German and Turkish patients with Alzheimer's disease.*Neuropsychology, 36*(3), 195–205. https://doi.org/10.1037/neu0000764.

I developed the study design and research questions. Additionally, I collected and analyzed the data and wrote the article. I trained the research assistants to administer test procedures for and collect data from German patients. Dr. Özgür Onur, Prof. Josef Kessler, Prof. Görsev Yener, and Yagmur Özbek provided support in finding eligible participants and obtaining ethics approval for the study. All these authors also commented on the manuscript. Prof. Lutz Frölich, Prof. Patric Meyer, and Dr. Birgit Teichmann supervised the work and manuscript preparation.

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	Simge Celik
Datum / Date	30.08.2022
Unterschrift / Signature	(Dem Dekanat der Fakultät für Verhaltens- und Empirische Kulturwissenschaften liegt eine unterschriebene Version dieser Erklärung vom 30.08.2022 vor.)