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

Title of the thesis

Macro-Level Cognitive and Linguistic Function in Early Stage Alzheimer's Disease and Mild Cognitive Impairment

presented by
Eesha Kokje

year of submission
2022

Dean: Prof. Dr. Guido Sprenger
Advisors: Prof. Dr. Christiane von Stutterheim
Prof. Dr. Hans-Werner Wahl

Acknowledgements

I would like to start by expressing my sincere gratitude to my supervisor, Prof. Dr. Christiane von Stutterheim, for sharing her knowledge and expertise, for guiding me through every stage of this dissertation project, and for extending her support and encouragement to me on every occasion; all of which has made this work possible. I am happy to have had such a positive working experience, and grateful for the opportunity to learn all that I have. Her enthusiasm for this project, has largely made this work possible.

I would like to thank my second supervisor, Prof. Dr. Hans-Werner Wahl, for his continued guidance over the last few years. Discussions with him and his in-depth feedback have helped in improving this work. I am grateful to him for sharing his resources and expertise, for his time, and for his willingness to help.

I am also grateful to Dr. Johannes Gerwien for sharing his knowledge and ideas, for giving his time to this project, and for all the stimulating conversations.

I would like to sincerely thank Prof. Dr. Konrad Beyreuther, for his constant encouragement, particularly his very encouraging feedback, and his words of wisdom. I would also thank Dr. Birgit Teichmann, for her continued help and support over the years. I am thankful to all my colleagues and friends at the Network Aging Research (NAR) for their support, their feedback, and the discussions, which have contributed to the completion of this dissertation.

I would like to express my gratitude to the Klaus Tschira Foundation for their financial support, without which this project or its completion would not have been possible.

I would sincerely like to thank the people at the Gedächtnisambulanz of Uni-Klink Heidelberg for their co-operation and support in the recruitment and testing of patients. In particular, I want to thank Prof. Dr. Daniela Rösch-Ely, Anna Jähn, and Christine Kreft for going out of their way

to make patient recruitment possible. I would also like to thank the people at the Gedächtnisambulanz of the Zentral Institut für Seeliche Gesundheit, particularly Prof. Dr. Lutz Frölich and Dr. Lucrezia Hausner for supporting me in patient recruitment. A special thanks goes out to all the people who took part in this study and their families for giving their time, without which this work would not have been possible.

I would like to extend a special thanks to my previous mentor, Prof. Markus Bindemann, for everything that he taught me in my limited time working with him, but particularly for encouraging me to pursue a PhD.

Finally, I would like to thank my friends and family. I am grateful to all my friends for their support, directly or indirectly, throughout this endeavor, and for their friendship, which made this time more enjoyable. I want to particularly thank Prerna and Rand, for their friendship and support, especially through rougher times, and for their companionship in learning new things, which kept me motivated. I am grateful to my brother, Rutwij, for his constant support, encouragement, and entertainment. And lastly, I am particularly grateful to my parents, Ketki and Milind, for their unconditional support in all forms, their belief, and their encouragement to follow my own path and pursue my dreams, over all these years. Without their support this truly would not have been possible.

Table of Contents

Abstract

Alzheimer's disease (AD) is a global health concern, particularly as there is currently no cure for the disease. Interventions to slow progression of disease, pharmacological or non-pharmacological, need to be targeted early on before any significant neurodegeneration has occurred, as these changes are irreversible, and lost cognitive function cannot be recovered. This makes it imperative to detect pathological cognitive decline as early as possible. Although biomarkers have received a lot of attention in this regard, they have several limitations, particularly outside of research settings, such as cost and availability. Cognitive markers, other than traditional neuropsychological test measures, on the other hand, have received comparatively less attention with regards to early detection; and, particularly cognitive markers that are rooted in real-world, everyday cognition, have been lacking. Due to the disease being incurable, interventions are aimed at maintaining independent living and good quality of life for as long as possible. This necessitates outcomes that can measure meaningful change in cognition and everyday functioning. The goal of the present dissertation was to identify gaps in the current literature on cognitive and linguistic assessments that are embedded in aspects of everyday cognition in AD, and work towards developing paradigms to address the gaps. Due to the emphasis on early detection, the work focused on patients in the very early stage of AD and on its preceding stage of Mild Cognitive Impairment (MCI). In light of evidence reporting the inability of AD patients to follow narratives, be it verbal or non-verbal, a systematic review of text comprehension studies was conducted to characterize and evaluate macro-level measures of discourse comprehension in their sensitivity to early stage AD, and their ability to distinguish pathological ageing due to AD or MCI from cognitive ageing. Results showed that, not only AD patients, but also MCI patients were significantly more impaired on macro-level measures of

comprehension compared to cognitively healthy older adults. These findings were consistent across all eight studies included in the review, indicating a robust effect, though there were minor differences in the sensitivity of different measures. Next, moving towards non-verbal narratives, a novel picture-based paradigm assessing event cognition, with a focus on event integration and macro-event recognition, was introduced. This study aimed to examine the macro-level processing of events by using a format requiring integration of micro-events, depicted in pictures, into a larger macro-event. AD and MCI patients' ability to connect the micro-events temporally and causally to identify the depicted macro-event was assessed. As hypothesized, the findings showed that patient groups had significant difficulties in determining temporal order of micro-events, even when provided with a verbal cue, as well as in conceptualizing the macro-event from the presented micro-events, when compared to healthy older adults. Finally, using traditional neuropsychological tests, the cognitive processes involved in performing the macro-event recognition task were determined by examining correlations. Primarily, semantic memory and executive functioning appear to play a role. However, the strength of correlations was fairly moderate, indicating added value of event recognition task in cognitive assessment. Taken together, these findings show the sensitivity of macro-level cognitive and linguistic markers based in everyday cognition in the early stages of AD, and highlight the positive role of such cognitive assessment methods in bringing together objective assessment methods and everyday cognition.

List of Tables

List of Figures

Abbreviations

Aβ	Amyloid beta
AARR	Alzheimer's Association Research Roundtable
ACE	Addenbrooke's Cognitive Examination
AD	Alzheimer's Disease
ADL	Activities of Daily Living
aMCI	amnestic Mild Cognitive Impairment
ANOVA	Analysis of Variance
APOE	Apolipoprotein E
APP	Amyloid Precursor Protein
BADL	Basic Activities of Daily Living
BNT	Boston Naming Test
CAMCOG	Cambridge Cognitive Examination
CANTAB	Cambridge Neuropsychological Test Automated Battery
CDR	Clinical Dementia Rating
CERAD-NB	Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery
CI	Confidence Interval
CN	Cognitively Normal
CR	Cognitive Reserve
CSF	Cerebrospinal Fluid
DLB	Dementia with Lewy Bodies
DSM	Diagnostic and Statistical Manual of Mental Disorders

FTD	Frontotemporal Dementia
GDS	Geriatric Depression Scale
GPCOG	General Practitioner's Assessment of Cognition
IADL	Instrumental Activities of Daily Living
ICC	Intraclass Correlation Coefficient
ICD	International Classification of Diseases
IWG	International Working Group
MCI	Mild Cognitive Impairment
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
naMCI	non-amnestic Mild Cognitive Impairment
NC	Normal Controls
NIA-AA	National Institute on Aging- Alzheimer's Association
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association
OA	Older Adults
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RBMT	Rivermead Behavioural Memory Test
RUDAS	Rowland Universal Dementia Assessment Scale
SD	Standard Deviation
SE	Standard Error
TEA	Test of Everyday Attention
TBI	Traumatic Brain Injury
ToL	Tower of London test

TMT	Trail Making Test
VaD	Vascular Dementia
VR	Virtual Reality
WAIS	Wechsler Adult Intelligence Scale
WMS	Wechsler Memory Scale

Chapter 1

General Introduction

1.1 Introduction

Ageing and issues associated with ageing have come to the forefront in recent decades due to the changing global demographic. With a continued increase in the average life expectancy, the ageing population has shown a steady growth in numbers in the last few decades (Dicker et al., 2018), and, this number is expected to continue to increase in coming decades (World Population Ageing, 2017). However, increase in life expectancy does not necessarily equate good quality of life in later age, as age itself becomes the primary risk factor for several diseases, which also has become a major public health concern. As a result of population ageing, there is increasing scientific interest in studying the ageing population, age-related disorders, and healthy ageing, in order to improve the quality of life in older age, by addressing issues associated with ageing.

Dementia, one of the major age-related disorders that significantly affects quality of life, is a syndrome characterized by progressive loss of cognitive function in areas such as memory, language, attention, visuospatial ability, and reasoning, which is significant enough to interfere with individuals' functional ability. As of 2019, over 50 million people are estimated to be living with dementia; a number which is projected to increase threefold by the year 2050 (World Alzheimer Report, 2019). Increasing life expectancy contributes to the higher estimates, as the rate of incidence of dementia increases exponentially with increasing age (von Strauss, Viitanen, De Ronchi, Winblad, & Fratiglioni, 1999). Dementia encompasses several diseases that differ in etiology, and consequently, in their presentation. Of some of the common causes of dementia, vascular dementia (VaD) accounts for about 20% of the cases, and is caused by reduced blood flow to the brain. The symptoms vary greatly depending on the area of the brain that is affected. Dementia with Lewy bodies (DLB) is characterized primarily by motor deficits similar to

Parkinson's disease. Frontotemporal dementia (FTD) encompasses a broad range of symptoms resulting from progressive neuronal loss in the frontal and temporal lobes of the brain. Primarily, it results in behavioral abnormalities, personality changes, and difficulties with language production and comprehension. There are several other reversible and irreversible causes of dementia. The most common cause, however, is Alzheimer's disease (AD), which accounts for about 50-60% of all dementia cases.

AD is a progressive, neurodegenerative disease that is characterized by a gradual decline in cognitive, linguistic, and functional abilities of individuals over a period of several years. The clinical manifestation of AD is preceded by a long preclinical phase, estimated to be around 10-20 years, during which neuropathological changes occur (Dubois et al., 2014). The hallmark of AD pathology is the deposition of amyloid beta (Aβ) and tau proteins forming senile plaques and neurofibrillary tangles, which eventually result in neuronal loss and cortical atrophy. These changes are irreversible in nature, thereby limiting the opportunities and scope for intervention.

Currently, limited success has been achieved in managing symptoms using pharmacological interventions. Non-pharmacological interventions rooted in modifiable risk and protective factors, such as cognitive training or physical activity, have had relatively greater success in slowing progression of the disease and in maintaining quality of life for a longer period of time. These interventions, however, require early detection of AD pathology, as, once significant neurodegeneration has occurred, interventions are no longer effective. Therefore, the present work chooses to focus on individuals in the early stage of AD, as well as in the preceding stage of Mild Cognitive Impairment (MCI).

At the assessment level, neuropsychological tests that were developed decades ago continue to be used for cognitive assessment of AD and MCI. Some of these tests were

developed before the construct of MCI even existed, and have not been updated. Moreover, these tests have been developed with the aim of diagnosis. The promise of interventions also demands the availability of cognitive and functional outcomes that can measure meaningful change i.e. change that has some bearing on patients' everyday life. This calls into question the appropriateness of the use of neuropsychological tests designed for the purposes of diagnosis, and calls for the broadening of the goal of cognitive tests and development of paradigms that can fulfill broader purposes, such as in tracking cognitive change over time or in intervention change assessment. A critical aspect of diagnosing and tracking disease progression is everyday functional ability. In recent years, there has been some effort in the direction of developing measures that focus on everyday cognition i.e. cognitively complex activities or tasks using naturalistic stimuli embedded with social and contextual cues that resemble daily life activity, in contrast to the controlled, decontextualized environment of laboratory-based tasks (Allaire & Marsiske, 1999; Henning, 2004; Willis, 1996). These methods, however, have not received much acceptance in clinical settings. The present dissertation aims to focus on experimental paradigms that are rooted in everyday cognition for studying cognitive and linguistic deficits in AD and MCI.

1.2 Outline of the Dissertation

The overarching goal of the present work is to study linguistic and cognitive processes using contextually-rich paradigms that are rooted in everyday, real-world cognition, and to assess the potential of such paradigms as indicators of AD in the very early stages of the disease and in measuring meaningful cognitive outcomes; and, to attempt to look beyond traditional, laboratory-based neuropsychological testing methods that lack ecological validity (Chaytor & Schmitter-Edgecombe, 2003; Spooner & Pachana, 2006).

Chapter 2 of this dissertation presents a review of the clinical features of AD and MCI- pathological features, diagnostic procedure, risk factors, and neuropsychological profile. Further, current knowledge of cognitive, linguistic, and functional impairments reported in early stage AD and MCI is reviewed. This is followed by a critical evaluation of the current state of the neuropsychological testing component of AD and MCI diagnosis, including linguistic, cognitive, and functional assessment methods, and their limitations. Finally, this chapter argues for a change in perspective when looking at neuropsychological testing, and for a broadening of goals associated with testing methods. An argument is made for focusing on outcomes that are meaningful in clinical and practical terms, and for the development and empirical testing of new cognitive assessment tools rooted in everyday cognition.

An important aspect of everyday functioning is language comprehension; not only at a word- and sentence-level, as is primarily studied, but at a macro-linguistic level, which includes understanding the global meaning, inferring unstated or indirectly stated information, and understanding the goals and intentions of the speaker, among others. It is due to the complex processing required to understand discourse at a macro-level, that it is likely to show indications of decline early in the course of the disease compared to single sentence or word level processing. **Chapter 3** presents a systematic review of discourse comprehension studies with a focus on these macro-linguistic features, conducted with early stage AD and MCI patients in comparison to healthy older adults. The review evaluates the potential usefulness of assessing macro-linguistic features in early AD, and impairment in macro-level comprehension as an early indicator of AD. Further, the advantage of employing a holistic paradigm i.e. one encompassing multiple levels of representation, such as text comprehension, over more traditional linguistic

paradigms using a piecemeal approach, such as word list learning or fluency tests, is discussed. The shortcomings of the text comprehension paradigm, and the way forward, are discussed.

In **Chapter 4**, a study investigating macro-event recognition is presented, which continues forward the theme of macro-level comprehension from Chapter 3. A novel picture-based paradigm is introduced, which addresses some issues of the text comprehension paradigm. Across two experiments, the study examines the ability of individuals in the early stage of AD and MCI, compared to healthy older adults, to recognize macro-events, when the micro-events that make up the larger macro-event are presented. The paradigm involves pictures of four smaller events within a larger event, presented in a random order, which are to be arranged in a temporally and causally appropriate sequence, followed by identification of depicted macro-event (Experiment 1). Additionally, whether a positive effect of linguistic cueing can be observed in the patient groups, is examined. Experiment 2 uses a modified, cognitively less demanding version of the paradigm in Experiment 1.

Chapter 5 further examines the macro-event recognition paradigm in relation to existing assessment procedures for AD. The aim of this chapter is twofold. One, is to determine the cognitive processes involved in the macro-event recognition task. This was done by examining associations between outcome measures from the macro-event recognition study described in Chapter 4 and traditional neuropsychological test measures that are typically used in clinical settings. Two, the additional value added by the event recognition paradigm to existing testing methods was assessed by examining number and strength of the associations between the event recognition measures and neuropsychological tests. The implications of these findings are discussed.

Finally, **Chapter 6** provides a summary and discussion of the overall findings in the context of the current state of the art, the clinical implications of the findings, the strengths and limitations of the present dissertation, and recommendations for future research.

Chapter 2

Theoretical and Empirical Background

2.1 Alzheimer's Disease: Clinical Features

AD is most commonly characterized by memory problems, which start slowly, and over time, a progressive worsening of cognitive and linguistic functions is observed. The beginning sign of AD, typically, is forgetfulness, particularly of recent events. This is also observed in cognitive ageing, which may make it difficult to distinguish them at first. AD typically begins after the age of 60, and is known as *late-onset* AD in such instances. With increasing age, the risk of developing AD increases exponentially. A more uncommon form of AD is the *early-onset* type which may begin between the ages of 40-60. It is also known as *familial AD*, as it is typically inherited, and seen in individuals with a family history of AD, usually in a first-degree relative. Both forms of AD are characterized by a progressive decline in cognitive and linguistic functions, eventually resulting in a decline in functional ability.

The etiology of AD is complex, and while the most widely held theory is the amyloid cascade hypothesis, the definite cause is not yet fully understood. This makes developing treatments and therapies complicated, and the disease is, as of now, incurable. The decline in cognitive functions occurs as a result of progressive neurodegeneration, with cortical atrophy observed largely in medial temporal region of the brain, resulting in deficits primarily in episodic memory and memory formation. This pattern of atrophy and deficits is referred to as *typical AD*. In around 25-30% of AD cases, atrophy is observed primarily in the parietal or frontal lobe, resulting in non-memory deficits, such as visuospatial, language, executive or behavioral deficits, as the primary symptoms. This is referred to as *atypical AD* (Boon et al., 2018; Graff-Radford et al., 2021).

2.1.1 Stages of Alzheimer's Disease

AD is a slowly progressing disease, and is generally divided into three distinct stages post clinical manifestation– *mild, moderate,* and *severe.* The mild stage of the disease is characterized by difficulties in decision making, planning and organizing, spatial orientation, problem solving, retaining new information, word-finding difficulties, and some decline in Instrumental Activities of Daily Living (IADLs). The moderate stage is a relatively longer lasting stage, during which individuals experience increasing decline in IADLs, loss of temporal and spatial orientation in familiar environments too, irritability, confusion, significant memory loss, and poor judgement. Additionally, communication difficulties become more pronounced in the form of repetitions, inability to express thoughts, as well as in comprehension of speech. Behavioral problems become more evident too, characterized by apathy, emotional instability, and sundowning, which is an increase in confusion and agitation in the later parts of the day. Finally, in the severe stage of the disease, a continuing decline in cognitive function results in individuals losing their ability to communicate, with their speech consisting mostly of empty phrases or words, and eventually a complete loss of speech. There is a complete loss of ability to live independently, and patients are largely dependent on round-the-clock care.

AD-related pathological changes in the brain begin ten to twenty years before AD becomes clinically evident (Bateman et al., 2012). Although decline during this stage is not significant enough to warrant a clinical diagnosis of AD, i.e. it is not clearly evident in behavior, or in clinical examination and neuropsychological tests, it is marked by subtle cognitive decline. With the increasing emphasis on early detection of AD, in order to improve prognosis, MCI was introduced as a transitional stage between healthy ageing and dementia. The increasing use of biomarkers in research settings has also led to a focus on the preclinical stages of AD, which are

characterized primarily by neuropathological changes. These changes, although not entirely overtly evident, are reflected to some extent in subtle cognitive decline.

There is increasing emphasis on treating AD as a continuum rather than discrete stages, and acknowledging the preclinical stages of the disease (Dubois, 2018). All diagnostic criteria recognize three stages of AD: (i) Preclinical stage, (ii) Prodromal stage, and (iii) Definite AD. The distinction between the preclinical and prodromal stages is the presence or absence of symptoms. The preclinical stage is an asymptomatic stage, which may be further divided into different types. For example, the International Working Group (IWG-2) classifies two types of preclinical AD– *Asymptomatic at-risk state for AD* and *Presymptomatic AD* (Dubois et al., 2014). Presymptomatic AD refers to the state of individuals who have a genetic predisposition to AD, the most prevalent being apolipoprotein E (APOE) ε4 allele, who will most likely go on to develop AD. Asymptomatic at-risk state for AD, on the other hand, refers to individuals in whom the neuropathological markers of AD are present, and are therefore at risk for developing AD, but may or may not develop AD in their lifetime. The prodromal stage is the subsequent early symptomatic stage, which encompasses individuals with MCI in whom evidence of AD pathology is also present, making them very likely to convert to AD.

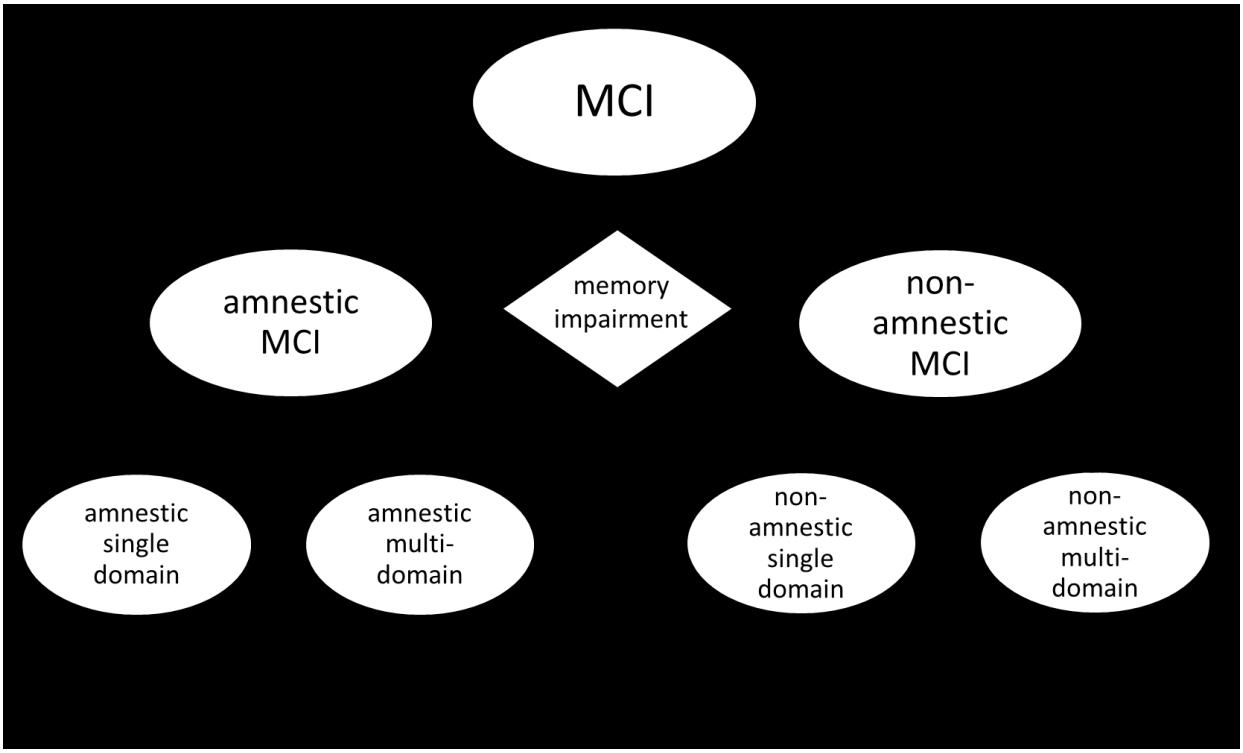
2.1.2 Mild Cognitive Impairment

MCI is the intermediate stage between normal cognitive ageing and dementia, which is marked by impairment in one or two cognitive domains, but without significant functional decline. Newer diagnostic criteria, however, acknowledge the possibility of some functional decline in cognitively more complex activities (Jack et al., 2018). MCI has heterogeneous etiologies (DeCarli, 2003), and may be reversible when the cause is reversible, such as in cases of depression, side effects from medications or vitamin deficiencies; or, it may remain stable and

never progress to dementia (Koepsell & Monsell, 2012). When caused by dementia-related pathology, as determined through biomarker evidence, it is referred to as the prodromal stage of the specific dementia subtype. The MCI stage has frequently been the target stage for interventions as, at this stage, cognitive deficits are minimal, but may still be detectable via neuropsychological testing.

In recent years, there have been efforts to classify MCI into different subtypes based on the presentation (Figure 2.1). One type of classification is *single-domain vs. multi-domain* MCI, depending on whether significant impairment is observed in one cognitive domain or more than one domain. The second type of classification is *amnestic vs. non-amnestic* sub-types, depending on whether impairment is observed in memory (aMCI) or non-memory domains (naMCI). The amnestic variant of MCI is most likely to convert to AD or VaD, whereas non-amnestic type of MCI increases risk of developing other types of dementia, such as FTD or DLB (Petersen & Negash, 2014). The National Institute on Aging and the Alzheimer's Association's (NIA-AA) recent recommendations also include etiology-based classification of MCI. This approach combines biomarker evidence with clinical presentation. When MCI is present along with evidence of AD pathology, such as amyloid and tau deposition, it is termed as MCI due to AD.

Figure 2.1: Subtypes of MCI and the type of dementia they pose a risk for developing



AD = Alzheimer's Disease; VaD = Vascular Dementia; FTD = Frontotemporal Dementia; DLB= Dementia with Lewy Bodies

The reported rate of conversion from MCI to dementia varies widely. One systematic review looking at conversion rates specifically for AD, reported rates ranging from 10.2 to 33.6% in studies after a 1-year follow up (Ward, Tardiff, Dye, & Arrighi, 2013). For studies with a follow-up period of 5 or more years, reported conversion rates ranged from 10.6 to 47.5%, and one study with a 10-year follow-up period reported a 55.5% rate of conversion. These rates vary depending on a number of variables. Conversion rates tend to be higher in clinic-based populations than in community settings (Y. Chen et al., 2017). The follow-up period tends to be a major factor influencing rates of conversion reported in studies. Many studies also do not

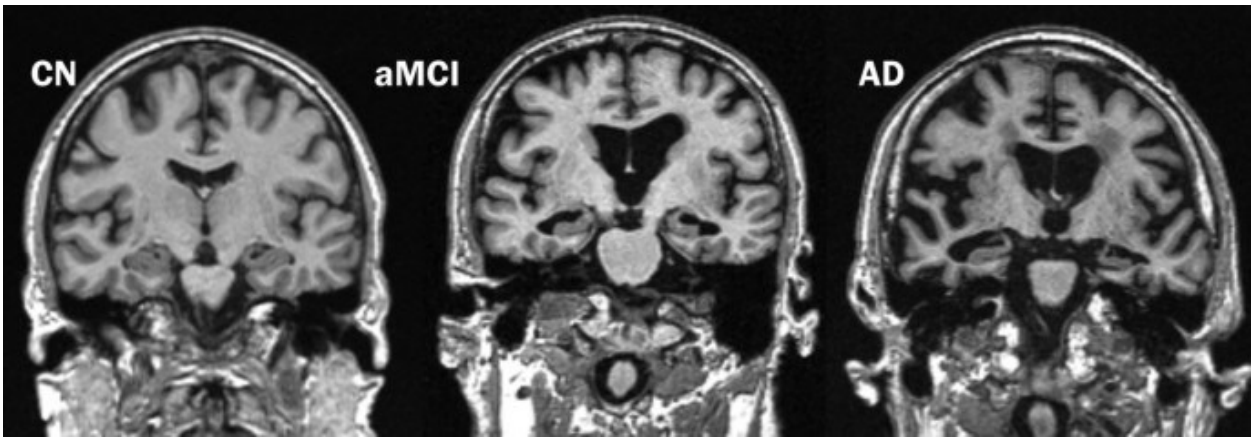
account for loss of participants to follow-up assessments, the rate of which tends to be particularly high among this group. The reasons for this attrition could be the higher mortality rates due to age and the disease, or higher drop-out rates due to cognitive and functional impairment; all of which possibly result in an underestimation of conversion rate (Michaud, Su, Siahpush, & Murman, 2017; Ward et al., 2013). Conversion rates also vary depending on the MCI-subtype. Rates are reported to be higher for the amnestic subtype compared to the non-amnestic subtype (Schmidtke & Hermeneit, 2008), as well as for multi-domain MCI compared to single-domain MCI. Diagnostic criteria acts as another variable appearing to influence conversion rates (Marcos et al., 2016). Importantly, however, baseline neuropsychological test performance appears to be a fairly good predictor of conversion from MCI to dementia. Several studies followed individuals diagnosed with MCI longitudinally, although follow-up periods varied (Bruscoli & Lovestone, 2004). Individuals with MCI, who went on to convert to dementia, had poorer performance on cognitive tests during the initial baseline assessment compared to individuals with MCI who were non-converters.

2.1.3 Pathophysiology of Alzheimer's Disease

The first pathological changes observed in AD, according to the amyloid cascade hypothesis, are formation of amyloid plaques and neurofibrillary tangles (Hardy & Higgins, 1992). Amyloid plaques are protein deposits that are formed by the breakdown of amyloid precursor protein (APP) when cleaved by the enzymes beta-secretase and gamma-secretase. This leads to the formation of Aβ peptides, which are the main component of amyloid plaques. This is followed by the formation of neurofibrillary tangles due to the hyperphosphorylation of tau protein. These accumulated hyperphosphorylated tau proteins aggregate to form paired helical filaments, known as tangles (Ballatore, Lee, & Trojanowski, 2007). The Aβ and tau proteins are

detectable in Cerebrospinal Fluid (CSF), although their degree of concentration may vary among patients (Blennow & Hampel, 2003). The accumulation of amyloid plaques and neurofibrillary tangles in the brain results in interference in communication between neurons, which causes synaptic loss and neuronal death resulting in progressive cortical atrophy (Figure 2.2), primarily in the medial temporal lobe (Pini et al., 2016). This progressive loss of neurons and resulting atrophy makes the disease irreversible. The plaques and tangles are also present in older adults without AD, however, the number is far less than in individuals with AD (Bouras, Hof, Giannakopoulos, Michel, & Morrison, 1994).

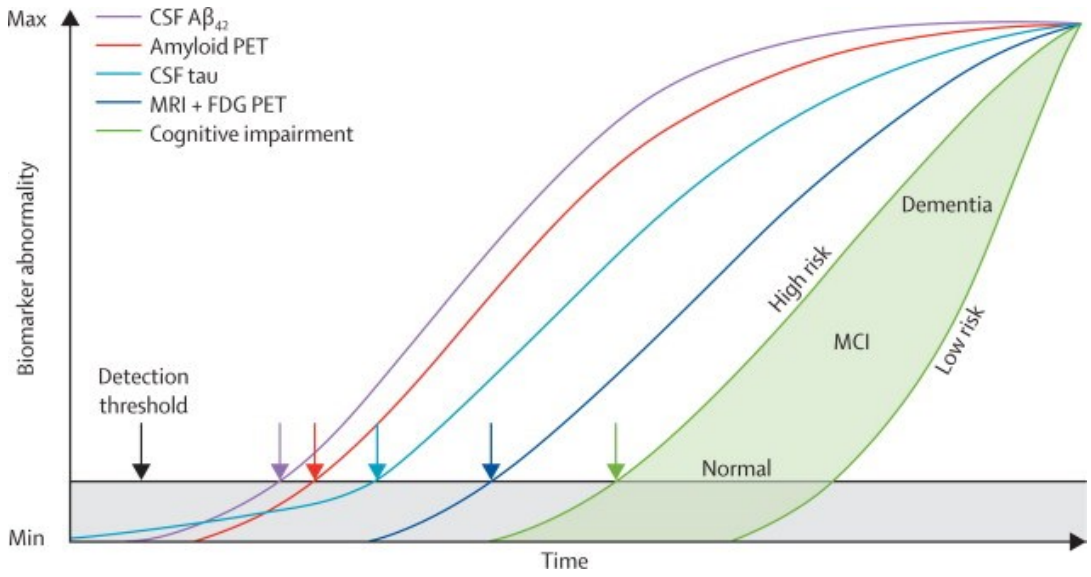
Figure 2.2: Structural MRI showing atrophy patterns in the brain of a cognitively normal (CN) individual, an individual with amnestic MCI, and an individual with AD (from Vemuri & Jack, 2010)



Although there is still some debate surrounding the sequence and timeline of neuropathological changes, the widely held theory is that the first change to occur is deposition of amyloid beta, which occurs as early as 20 years before disease onset (Jack et al., 2010). This is followed by emergence of tau tangles, which is then followed by indicators of neurodegeneration

(Figure 2.3). The exact cause and sequence of events is still poorly understood, and several other factors have emerged in recent years. For example, recent studies suggest that chronic neuroinflammation resulting from oxidative stress plays a critical role, suggesting that it is a precursor to the amyloid cascade (Z. Chen & Zhong, 2014; Zhao & Zhao, 2013). Importantly, however, although amyloid and tau depositions do appear to be the core neuropathological features of AD, the magnitude of amyloid and tau burden does not necessarily correlate with the magnitude of cognitive decline (Coughlan, Zhukovsky, Voineskos, & Grady, 2021).

Figure 2.3: Alzheimer's disease pathological and clinical cascade as a function of time (from Jack Jr et al., 2013)



2.1.4 Risk and Protective Factors

The most common, unmodifiable risk factor for AD is ageing. Beyond the age of 65 the incidence of AD doubles every five years, with the prevalence of AD being well above 30% in

those older than 85 years (Castellani, Rolston, & Smith, 2010; C. Qiu, Kivipelto, & von Strauss, 2009). The other unmodifiable risk factor for AD is genetics, with the ε4 allele of APOE being the most commonly implicated gene in higher risk for developing AD, as well as in earlier onset of disease (Verghese, Castellano, & Holtzman, 2011). The presence of one ε4 allele increases the risk of developing AD three-fold compared to individuals with both ε3 alleles; and, inheriting a pair of ε4 alleles increase the risk eight- to twelve-fold (Verghese et al., 2011). However, the presence of an APOE ε4 allele does not necessarily mean the individual will go on to develop AD, as the risk is attenuated by interaction with other genetic and environmental factors. There is also some variation in the relative increase in risk reported in different studies, and also some variation depending on ethnicity (Farrer et al., 1997). The APOE ε2 allele, on the other hand, has a protective effect. It decreases the risk of developing AD considerably and delays age of onset of AD (Li, Shue, Zhao, Shinohara, & Bu, 2020).

As AD has a complex etiology, several factors beyond ageing and genetics play a major role in determining whether individuals go on to develop AD. Individuals with a history of cardiovascular diseases, Traumatic Brain Injury (TBI), and late-life depression have a higher risk of developing AD (de Bruijn & Ikram, 2014; Diniz, Butters, Albert, Dew, & Reynolds, 2013; Sivanandam & Thakur, 2012). Alcohol consumption, smoking, diet, physical activity, cognitive and social engagement, education, and occupational complexity, are all lifestyle factors that positively or negatively influence risk of developing AD and age of onset (Weih, Wiltfang, & Kornhuber, 2007). Mediterranean diet, greater engagement in cognitive, social and physical activities, especially in mid-life, greater occupational complexity, and higher level of education are protective against AD (Crous-Bou, Minguillón, Gramunt, & Molinuevo, 2017; Sattler, Toro, Schönknecht, & Schröder, 2012). Smoking, on the other hand, increases the risk of AD, although

the precise mechanism for it is yet unknown (Durazzo, Mattsson, Weiner, & Initiative, 2014). The evidence on alcohol consumption is mixed (Piazza-Gardner, Gaffud, & Barry, 2013), with some studies showing a protective effect of moderate alcohol consumption, and others showing a detrimental effect. Overall, AD risk is modifiable (Ngandu et al., 2015), and about a third of AD cases are preventable with interventions targeted towards these modifiable risk and protective factors (Norton, Matthews, Barnes, Yaffe, & Brayne, 2014). In cases where delaying onset of AD by targeting modifiable risk factors is not possible, and due to a long preclinical phase, secondary management via early detection is the way forward (Brookmeyer, Abdalla, Kawas, & Corrada, 2018).

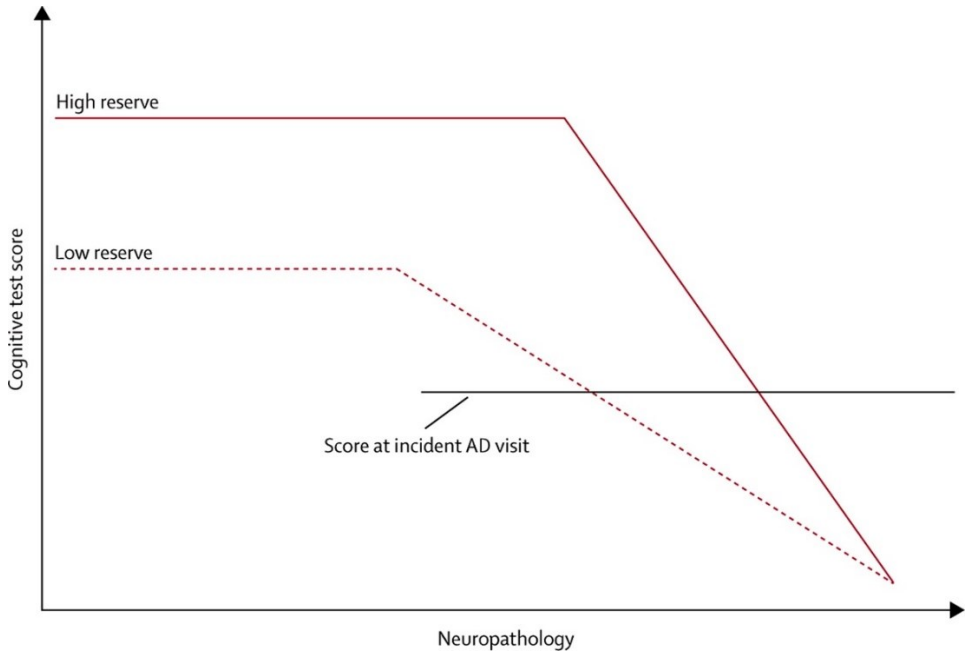
2.1.5 Cognitive Reserve

The newer, revised diagnostic criteria increasingly emphasize the use of biomarkers in the diagnosis of AD (Jack et al., 2018). While biomarkers are, without a doubt, helpful tools for detecting AD pathology, particularly in the preclinical stage of AD, they have limited utility without taking into account clinical expression of the disease. This is because not all individuals with AD pathology will go on to develop clinical symptoms in their lifetime. Even in all individuals that do develop clinical symptoms, several studies have shown that the degree of neurodegeneration often does not correspond to the level of cognitive and functional impairment observed; an incongruity attributed to cognitive reserve (Stern, 2012). According to the cognitive reserve hypothesis, brain reserve and cognitive reserve act as moderating variables accounting for the differences observed in individuals' susceptibility to neuropathology of the same magnitude.

Although it is generally referred to as *cognitive reserve* as a whole, two models of reserve have been proposed– *brain reserve* and *cognitive reserve* (Stern et al., 2020). Brain reserve refers

to a *passive* model of reserve, encompassing more physiological aspects, such as a larger brain structures due to a greater number of neurons and synapses, grey matter volume, or, cortical surface area. This means that the brain can tolerate more pathology, simply because there is more of it. Cognitive reserve, on the other hand, is an *active* form of reserve, wherein the brain actively tries to compensate for damaged networks by recruiting alternate networks to take over their function or by increasing efficiency of available resources by using alternate cognitive strategies. It is dependent on brain function rather than brain size.

Figure 2.4: Change in cognitive function in individuals with high and low cognitive reserve (from Stern, 2012)



A number of factors, directly or indirectly, contribute to cognitive reserve, including but not limited to, education level (Hoenig et al., 2017), occupational complexity (Boots et al., 2015), physical activity, cognitive engagement or leisure activities, particularly in mid-life

(Scarmeas & Stern, 2003). The aforementioned factors also tend to be inter-related, therefore making it difficult to isolate their individual contributions (Lövdén, Fratiglioni, Glymour, Lindenberger, & Tucker-Drob, 2020). For example, level of educational attainment tends to correlate with occupational complexity. So, level of education may, not only directly attenuate the functional impact of neuropathology, but also indirectly in late-life via greater cognitive engagement in the form of occupational complexity (Dekhtyar et al., 2015).

Overall, individuals with higher cognitive reserve have a higher threshold for tolerating neuropathology before AD becomes clinically evident, and some may never develop AD symptomatology in their lifetime. However, cognitive reserve is paradoxical in nature, as, once the disease manifests clinically, the decline observed in individuals with high cognitive reserve is much more rapid compared to individuals with low cognitive reserve (Soldan et al., 2017; van Loenhoud et al., 2019). This is because, individuals with high cognitive reserve start showing symptoms when coping mechanisms have been exhausted, and at that point, the degree of neuropathology is greater (Figure 2.4). This incongruity between neuropathology and clinical manifestation of AD highlights the importance of cognitive and neuropsychological testing in the diagnosis of AD, rather than relying solely on biomarkers.

2.1.6 Diagnosis

A histopathologic confirmation via autopsy used to be the only way to obtain a definite diagnosis of AD. However, with the development of valid biomarkers, a combination of these biomarkers and neuropsychological assessment have become a reliable way for a definitive diagnosis. For many decades, all diagnostic criteria emphasized the amnestic nature of AD, and necessitated memory deficits as a primary symptom to make a diagnosis of AD. This is also evident in the literature, which focused primarily on quantifying memory deficits above all else.

However, in recent years, the scientific community has acknowledged that the presentation of AD can be varied, with individuals exhibiting deficits in other cognitive domains without significant memory deficits. Cognitive domains, in this context, are delineated as the following– learning and memory, language, executive function, complex attention, perceptual-motor, social cognition. The change in approach to defining AD is reflected in the revised versions of the different diagnostic criteria, which now allow for deficits in any two cognitive domains for an AD diagnosis. Instead, AD is now classified into different subtypes depending on the presentation– *typical AD, atypical AD,* and *mixed AD*. Typical AD refers to the classic form of AD presenting with memory complaints, whereas in atypical AD, individuals present primarily with non-amnestic deficits, such as linguistic or visuospatial deficits. This type of diagnosis, however, is always made only when biomarker evidence for AD pathology is also present. Finally, when individuals present with typical AD symptomatology, but show biomarker evidence of other diseases in addition to AD pathology, it is referred to as mixed AD.

Currently, the most commonly used diagnostic criteria in clinical and research settings are the NIA-AA criteria (McKhann et al., 2011), the Diagnostic and Statistical Manual of Mental Disorder, Fifth Edition (DSM-5), the International Statistical Classification of Disease and Related Health Problems (ICD-10), and the IWG-2. The following are the core features in the diagnosis of AD: (i) insidious onset and gradual progression, (ii) amnestic or non-amnestic deficits, (iii) objective cognitive impairment (defined as 1 or 2 SDs below normal) in two or more cognitive domains, (iv) functional impairment, (v) no other neurological or psychiatric explanation for impairment. In addition to these clinical features, newer criteria are increasingly re-directing focus on inclusion of biomarkers in the diagnosis of AD. The 2018 version of the NIA-AA criteria and the IWG-2 criteria, particularly, are geared towards use in research settings,

and focus almost exclusively on biomarker evidence. As a result of biomarker availability, and therefore the possibility of detecting neuropathology before clinical disease onset, there is potential to track the temporal evolution of AD in the preclinical stages. Consequently, there is emphasis on treating clinical and preclinical stages of AD as a continuum, rather than as discrete stages, by using biomarker and clinical symptomatology as complementary evidence. Considering the move towards early detection, the literature is still found lacking in cognitive testing measures that are complex enough to be sensitive in the preclinical stage of AD, and that are specific to AD pathology, though efforts are being made in recent years. This is an issue that the current dissertation aims to address.

MCI, although clinically useful, is a relatively new concept. Therefore, there is still some controversy surrounding the diagnostic criteria. The primary feature which distinguishes MCI from dementia is that the ability to function independently is maintained in MCI. Due to a lack of formal consensus on what constitutes functional impairment, it is largely subject to individual clinicians' judgement, making MCI diagnosis challenging. Similar to AD, the initial diagnostic criteria for MCI required memory impairment, which have since evolved to include any kind of cognitive impairment, and rather there is emphasis on classification into different subtypes. The most common diagnostic criteria currently in use in clinical settings are the Petersen et al. (2001), Winblad et al. (2004), and the DSM-5 criteria for Mild Neurocognitive Disorder. Generally, core clinical criteria include the following: (i) subjective cognitive complaint from individual or family, (ii) objective cognitive impairment in one or two cognitive domains, (iii) preserved instrumental activities of daily living, and (iv) absence of dementia. The NIA-AA expanded upon the Petersen et al. (2001) and Winblad et al. (2004) criteria in 2011 to include biomarkers, for use in research settings, though core clinical criteria remain the same.

2.1.7 Neuropsychological Testing

The most commonly used tests for quick screening for cognitive impairment are the *Mini Mental State Examination* (MMSE; Folstein, Folstein, & McHugh, 1975), *Montreal Cognitive Assessment* (MoCA; Nasreddine et al., 2005), and Addenbrooke's Cognitive Examination (ACE-III; Hsieh, Schubert, Hoon, Mioshi, & Hodges, 2013). The ACE-III is becoming increasingly popular as a more comprehensive and more sensitive screening tool compared to MMSE. The sensitivity of the MMSE has, in recent years, been called into question by the scientific community, particularly as more sensitive tools have been developed. However, it continues to be extensively used in clinical practice (World Alzheimer Report, 2021). By itself, these tests are used strictly as screening tools only, and not for a comprehensive assessment. Some other brief screening tools that are frequently used include Mini-Cog (Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000), Cambridge Cognitive Examination (CAMCOG; Koning et al., 1998), and General Practitioner Assessment of Cognition (GPCOG; Brodaty et al., 2002). The primary problem with all of the above screening tools is that they were developed in primarily English-speaking, white-majority populations. Although they have been validated in other languages, simply translated versions of the tests are used rather than culturally sensitive adaptations. This results in reduced validity of the test in other languages (Ng et al., 2018). This is because of use of words or phrases in the original test which cannot easily be translated to other languages, or have little meaning when translated, as they may not conceptually exist in the language or particular culture. Additionally, performance on MMSE is also influenced by level of education. The Rowland Universal Dementia Assessment (RUDAS) is a screening tool that has been developed specifically to overcome these issues, and is recommended for use in linguistically and culturally diverse populations (Rowland, Basic, Storey, & Conforti, 2006; Storey, Rowland,

Basic, Conforti, & Dickson, 2004). This test was developed in a culturally heterogeneous population, items that may potentially be culturally biased were excluded, as well as items requiring reading and writing skills are not included, to reduce educational bias. Therefore, the test does not appear to be affected by language of administration or level of education.

The above screening tools are frequently used in clinical and research settings as a first step towards determining cognitive status of individuals, and form a part of neuropsychological assessment batteries that are used in AD diagnosis. For a more comprehensive assessment of cognitive functions, a battery of tests targeting different cognitive domains and functional abilities is used. These include tests measuring episodic memory, executive functions, working memory, language, attention, visuospatial ability, and ability to function independently. Commonly used neuropsychological test batteries include the Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery (CERAD-NB; Mirra et al., 1991) or Cambridge Neuropsychological Test Automated Battery (CANTAB; Fray, Robbins, & Sahakian, 1996).

The CERAD-NB is the most commonly used test battery in the diagnosis of AD in Germany. It includes the following tests: (i) MMSE, (ii) Semantic fluency– Animal, (iii) Boston Naming test (BNT), (iv) Word list learning– immediate and delayed recall, and (v) Constructional praxis. Optionally, the following tests are also included: (i) Logical memory, (ii) Digit span– forward and backward, (iii) Trail Making Test (TMT) A&B, and (iv) Clock drawing test.

For determining functional status, an outcome measure largely used in distinguishing AD from MCI, currently, questionnaires are primarily used in clinical practice. The commonly used tools are variations of the Activities of Daily Living (ADL) scale– Basic Activities of Daily

Living (BADL) and IADL (Giebel et al., 2014; Lawton & Brody, 1969). These scales rely on self-report or caregiver report of individual's ability to perform daily tasks independently. The BADL refers to relatively simpler daily tasks, such as, feeding, bathing, dressing, whereas the IADL consists of more cognitively complex daily activities such as, preparing food, managing finances, doing household chores, and so on. The measures, however, are not objective, as their reliance on self-report or caregiver report make them vulnerable to bias (Loewenstein & Acevedo, 2010). Additionally, they are not capable of measuring smaller changes in daily functioning longitudinally.

2.2 Cognition and Language in Early and Prodromal Alzheimer's Disease

AD is characterized by progressive decline in cognitive and linguistic functions, beginning in the preclinical stage, eventually resulting in functional impairment. Episodic memory deficit was considered the hallmark for AD, for a long time. Cognitive research in AD, since the beginning, has primarily focused on episodic memory deficits. These are commonly reported in the early stage of AD, and even during MCI or the preclinical stage. Episodic memory degradation, however, is also observed in cognitive ageing, and is also the first to decline, making it difficult to distinguish healthy ageing from pathological ageing (Tromp, Dufour, Lithfous, Pebayle, & Després, 2015). As stated briefly above, up until the late 2000s, all diagnostic criteria for AD, such as the National Institute of Neurological and Communicative Disorders and Stroke- Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) and DSM-3, required episodic memory impairment for AD diagnosis. This is, however, not the case anymore. It is now acknowledged that AD is a clinically heterogeneous disease, with cognitive profiles of individuals varying substantially (Martorelli, Sudo, & Charchat-Fichman, 2019). Episodic memory is still the most commonly reported and one of the

earliest symptoms to appear in typically developing AD. However, varied presentations of AD are now recognized. These sub-groups of patients with atypically developing AD may present with deficits in other domains, such as, language use or visuospatial abilities (Dickerson et al., 2017; Martorelli et al., 2019; Y. Qiu, Jacobs, Messer, Salmon, & Feldman, 2019). In individuals with typical AD too, there is considerable variability in the degree to which different cognitive functions may be affected.

2.2.1 Cognitive Functioning in Alzheimer's Disease

Episodic memory impairment is considered to arise from the inability of AD patients to learn new information, as is evident in word list learning task, which is commonly used in neuropsychological testing. In addition to episodic memory, a decline in semantic memory is also is also widely reported in AD patients in the early stage, as well as in MCI and the preclinical stage (Henry, Crawford, & Phillips, 2004; Joubert et al., 2010; Martorelli et al., 2019; Vogel, Gade, Stokholm, & Waldemar, 2005). It is typically assessed using category fluency task or naming tests. Some studies suggest that the deficits observed in semantic memory function result from a selective impairment in semantic memory networks, and are not simply a consequence of global cognitive decline (Clark et al., 2009; Verma & Howard, 2012). The evidence on the root of the selective impairment is mixed, with some studies suggesting degradation of knowledge structures as the cause rather than an inability to access information (Chan, Butters, & Salmon, 1997; Mårdh, Nägga, & Samuelsson, 2013). Other studies suggest the cause to be an inability to access information appropriately due to executive dysfunction, which is also commonly observed in AD (Baudic et al., 2006).

When compared to healthy older adults, individuals with AD show an overall impairment in both components of verbal fluency– semantic and phonemic. However, semantic fluency is

more markedly impaired in AD compared to phonemic fluency (Weakley & Schmitter-Edgecombe, 2014). Semantic fluency relies more on semantic memory, and phonemic fluency relies more on executive functions. In MCI population, a mixed pattern has been observed, with some showing more impairment on semantic fluency, and others on phonemic fluency (Rinehardt et al., 2014). This is likely owing to the heterogeneous nature of MCI. However, recent research examining semantic fluency, specifically in aMCI population, shows that impairment in executive control affects semantic fluency performance as much as impairment in semantic knowledge does, in this population (Tröger et al., 2021). Some studies suggest that impairments in other cognitive domains generally, such as episodic memory, visuospatial ability, or language, are attributable to executive dysfunction, rather than domain-specific impairment (Baudic et al., 2006; Franceschi et al., 2007; Grober et al., 2008), and is difficult to disentangle.

In clinical practice, Trail Making Test-B is commonly used to assess executive functions. Other tests used to assess executive function include Wisconsin Card Sorting Test (WCST) and Tower of London test (ToL). A major component of these deficits stems from working memory impairment. Studies report that the working memory deficit observed during early or preclinical stage of AD results primarily from a disruption of central executive and episodic buffer functioning (Huntley & Howard, 2010; Weintraub, Wicklund, & Salmon, 2012). This causes impairment in tasks involving complex attention, which require attention switching and inhibition, and tasks which require individuals to hold, manipulate and update information in memory. On the other hand, phonological loop and visuospatial sketchpad are fairly preserved during the early stages, and only impaired in the late stages of the disease. Working memory deficits also occurs in cognitively healthy ageing. However, a marked impairment is not observed when compared to MCI or AD patients, as healthy ageing individuals use

compensatory mechanisms by recruiting alternate networks. This was observed in studies using MRI and Positron Emission Tomography (PET), which showed greater activation, particularly in the prefrontal cortex, in cognitively healthy older adults while performing working memory tasks (Kirova, Bays, & Lagalwar, 2015).

Studies generally report visuospatial ability impairment during the early stage of AD (Quental, Brucki, & Bueno, 2009, 2013). There may be some differences in when deficits in this domain manifest depending on the AD subtype (Martorelli et al., 2019; Ossenkoppele et al., 2016). It is reported that about 25% of AD patients do not conform to the typical presentation of AD, and instead present with what is known as non-amnestic or atypical AD, of which the most common presentation is primarily with visuospatial dysfunction (Murray et al., 2011). Commonly used tests to assess visuospatial function include Corsi block test, clock drawing test, Rey complex figure. Spatial and temporal disorientation is also a common feature in early AD stage and MCI (Giannakopoulos et al., 2000).

2.2.2 Linguistic Functioning in Alzheimer's Disease

Linguistic ability is seen to progressively deteriorate during the course of AD, beginning with minor comprehension and word finding deficits, followed by phonological errors, articulation difficulties and empty speech, which eventually progresses to almost complete mutism (Forbes-McKay, Shanks, & Venneri, 2013). On a linguistic surface level, one of the most commonly reported and studied linguistic deficit in AD patients is lexico-semantic impairment, which is observed in the early stage (de Lira, Ortiz, Campanha, Bertolucci, & Minett, 2010). It is commonly reported as word-finding difficulties or tip-of-the-tongue phenomenon, wherein patients are either entirely unable to recall words or tend to substitute with conceptually-related words or pronouns (Kavé & Goral, 2018). This is, however, also a common feature of cognitive

ageing, albeit less severe (Goral, Spiro III, Albert, Obler, & Connor, 2007). A general reduction in lexical diversity and increased word frequency is reported in AD (Kavé & Dassa, 2018). It has been a matter of debate in the literature, whether the resulting lexical impairment occurring in the early stages of AD is due to a semantic dysfunction or retrieval failure, or due to a decline in general cognitive processes, such as attention and executive functions (Kempler & Goral, 2008).

Interestingly, studies overwhelmingly report a dissociation in language functions, wherein there is a preservation of syntactic structure in the early stages of AD, despite the early lexical impairment (Kempler, Curtiss, & Jackson, 1987). It is only in the late-moderate to late stages of the disease that syntax construction is reported as being affected. However, although syntactic structure may be preserved initially, the complexity of the syntax produced reduces drastically even during the early and preclinical stages of the AD (Kemper et al., 1993; Lyons et al., 1994). This has been evidenced in studies investigating written and spoken language. A well-known example for written language is that of author Iris Murdoch, who went on to develop AD. An analysis of her novels showed a marked reduction in syntactic complexity in novels written in later years, albeit before she showed clinical indication of AD, compared to the novels written in earlier years (Le, Lancashire, Hirst, & Jokel, 2011; Pakhomov, Chacon, Wicklund, & Gundel, 2011).

Studies of spoken language, using discourse production paradigms consisting of verbal descriptions elicited by pictures, have also shown reduction in syntactic complexity in the early stage of AD and MCI (Mueller, Hermann, Mecollari, & Turkstra, 2018; Slegers, Filiou, Montembeault, & Brambati, 2018). Oral discourse studies reveal deficits beyond the syntactic and lexical level, including low connectedness in speech content, excessive use of pronouns in referring to objects or people, and particularly in maintaining coherence, very early in the course

of the disease, including in the preclinical stage (Bittner, Frankenberg, & Schröder, 2022; Malcorra et al., 2021). Due to the early, albeit subtle, manifestation of deficits in language production, be it written or oral, discourse production paradigms have shown great value in identifying presence of and predicting onset of AD, particularly using machine learning for automated analysis to identify speech patterns unique to AD (Asgari, Kaye, & Dodge, 2017; Sanz et al., 2022; Weiner, Frankenberg, Schröder, & Schultz, 2019). The later stage of the disease is generally characterized by a general reduction in speech, interaction, and communicative ability, and usually consists of isolated words or empty speech (Ahmed, Haigh, de Jager, & Garrard, 2013; Forbes-McKay et al., 2013).

A marked impairment can also be observed in early AD and MCI when it comes to language comprehension. Comprehension appears to be impaired at a sentence-level in early AD (Grossman & White-Devine, 1998; van Boxtel & Lawyer, 2021), although not necessarily uniformly across all individuals (Croot, Hodges, & Patterson, 1999). While syntactic deficits play some role (Grober & Bang, 1995), working memory capacity largely appears to account for the comprehension deficits (Kempler, Almor, Tyler, Andersen, & MacDonald, 1998; Rochon, Waters Gloria, & Caplan, 2000). In addition to direct comprehension, studies investigating non-literal language comprehension at a sentential level (e.g., metaphors, proverbs, idioms, sarcasm) also report significant impairment in MCI and AD patients (Cardoso, Silva, Maroco, de Mendonça, & Guerreiro, 2014; Maki, Yamaguchi, Koeda, & Yamaguchi, 2013; Rapp & Wild, 2011). Although, other studies that made a distinction between familiar and unfamiliar proverbs or metaphors, found that familiarity plays a deciding role in non-literal comprehension and interpretation ability (Amanzio, Geminiani, Leotta, & Cappa, 2008; Chapman et al., 1997; Rapp & Wild, 2011).

Discourse

Language, beyond the sentential level, i.e. discourse, reflects a more holistic approach towards understanding language functioning. It shows great potential for detecting impairment, as AD is characterized by difficulty in maintaining overall coherence, in establishing connections between fragments of information, and subsequently, in drawing inferences. It is promising on two levels. One, in distinguishing impairment due to pathology as opposed to cognitive ageing, and in differentiating the possible underlying pathology. For example, individuals with AD report word-finding difficulties, but so do normally ageing individuals. Although the degree of impairment may differ, individual differences also come into play, as there are inter-individual differences in baseline ability. Paradigms such as discourse processing, that require more complex and simultaneous processing on multiple levels, and therefore, have the capacity to measure multiple variables simultaneously, may better be able to differentiate cognitive ageing from pathological ageing. Two, the very capacity of such paradigms to capture multiple facets gives us the opportunity to identify patterns of impairment that paradigms measuring single variables may not be able to do. This enables us to differentiate underlying pathology. For example, patients with aphasia and patients with AD may both show impairment on a discourse processing paradigm, when compared to healthy older adults; but, the level at which impairment is observed differs (Chapman, Highley, & Thompson, 1998). While aphasia patients experience difficulties at a micro-linguistic level of sentence formulation, AD patients experience difficulty with the macro-linguistic aspects, such as, in maintaining coherence or grasping the global meaning, due to problems with connecting information and drawing inferences.

Studying more complex linguistic functions is a more wholesome approach over traditional approaches, as they target multiple elements and may be more sensitive as they

require more complex processing, and therefore, engage more extensive networks. As would be expected, individuals with AD show decline on such paradigms early on in the course of the disease, when other facets such as, syntax or naming may appear to be unimpaired. Finally, discourse level language is more reflective of real-world language usage and gives us insight into impairments on a more practical, day-to-day functional level. This is opposed to frequently employed testing methods created for the laboratory, such as, word-list learning or random word-list generation, which have little bearing on everyday communication. Due to the promise that discourse-level language shows for understanding everyday language use among AD patients, discourse comprehension is examined in detail via a systematic review in the present work in the next chapter.

2.2.3 New Directions

A number of studies indicate that subtle cognitive deficits are evident during the preclinical, asymptomatic stages of AD (Blacker et al., 2007; Han, Nguyen, Stricker, & Nation, 2017; Hassenstab et al., 2015; Rentz et al., 2011). Cognitive and functional deficits, at this stage, may not be evident at a clinically significant level in routinely used neuropsychological tests. However, studies have shown that individuals exhibit cognitive decline during the preclinical stage, which may be more evident when employing paradigms requiring more complex cognitive processing. The following section further elucidates the necessity of such a renewed approach, not just in the context of neuropsychological testing, but in the larger context of the disease mechanisms, and reviews efforts in this direction so far.

2.3 The Need for Novel Cognitive and Linguistic Outcomes

The biggest challenge in Alzheimer's disease is its irreversible nature. Pharmacological treatments so far have mostly focused on alleviating symptoms, with very limited success, and

only when targeted during the very early stages (Bachurin, Gavrilova, Samsonova, Barreto, & Aliev, 2018; Bazzari, Abdallah, & El-Abhar, 2019). More recently, a drug targeting the dissolution of amyloid plaques has been developed, but its efficacy has been widely questioned as it does not appear to improve cognitive outcomes (Knopman, Jones, & Greicius, 2021). Similarly, with cognitive interventions, success has been limited to slowing progression of disease, but there has not been much success with regards to regaining lost cognitive functions (Wang et al., 2021). Therefore, the priority is early detection of disease, before significant impairment occurs. As a result, the focus of AD research has been on identifying early indicators of pathology.

2.3.1 Limitations of Biomarkers

As briefly outlined earlier, in recent years, a number of biomarkers have been developed to identify AD pathology. Currently, the AT(N) classification system is recommended to identify and classify biomarkers (Jack et al., 2018). The 'A' group of biomarkers refers to Aβ plaques-based indicators including $A\beta_{42}$ levels in CSF obtained via spinal tap and amyloid PET. Tau-based biomarkers or 'T' encompass phosphorylated tau (p-tau) levels in CSF and tau PET. Finally, 'N' refers to indicators of neurodegeneration measured by the total tau (t-tau) levels in CSF, hypometabolism, and cortical atrophy observed via MRI scans. Biomarker testing is becoming more widely available, and is thus gaining increasing importance in the diagnosis of AD, especially in distinguishing it from other types of dementia. While biomarkers have received a lot of attention in terms of their early detection potential, they are not necessarily a feasible option in the real world, as they are invasive, expensive, not easily accessible to all, and are often lacking in specificity (Hoefeijzers, Calia, & Parra, 2016). The defining biomarkers of AD are neither unique to AD, nor to pathological functioning in general. Plaques and tangles are present

in every adult beyond the age of thirty to varying degrees. However, not every individual goes on to develop AD. Moreover, they are only indicative of the presence and extent of pathology, and not of the degree of cognitive or functional impairment, or in predicting the trajectory of cognitive decline, as different individuals are able to tolerate different degrees of pathology (Dumurgier et al., 2017; Fagan et al., 2007; Palmqvist et al., 2014). As discussed earlier, overt expression of AD pathology is largely moderated by cognitive reserve (Vemuri et al., 2011). Due to large individual differences and the influence of cognitive reserve, biomarkers values such as hippocampal volume are only meaningful in relation to previous values for the same individual, i.e. in measuring change in values longitudinally (Schröder & Pantel, 2016). These issues make it difficult to define standardized biomarker values (Frisoni et al., 2017). Cognitive markers, on the other hand, have greater utility in indicating cognitive status of individuals, are not invasive, are cost-effective, and relatively easier to administer.

Current evidence on biomarker sensitivity is also conflicting. Studies report evidence of significant amyloid deposition with no indication of cognitive impairment or little to no correlation between the amyloid burden and degree of cognitive impairment (Aizenstein et al., 2008; Hedden, Oh, Younger, & Patel, 2013). Similarly, hippocampal atrophy patterns appear to have very little bearing on AD diagnosis (Falgàs et al., 2019), though several studies report tau-burden to be relatively more indicative of cognitive status (Brier et al., 2016; Cicognola et al., 2019). Cognitive and neuropsychological assessments complement biomarker evidence. A study by Gomar et al. (2011) investigated the utility of combining biomarkers and cognitive markers in predicting conversion from MCI to AD. The study found that cognitive markers were more consistent and stronger predictors of conversion to AD compared to biomarkers. Additionally, the study showed that conversion to AD was not associated with a change in neuropathological

characteristics of individuals, rather it was associated with decline in functional ability (Gomar et al., 2011). Other studies too have shown that cognitive and neuropsychological tests are better indicators or add to biomarkers, and often models combining several of these markers are better in predicting AD onset and prognosis (Devanand et al., 2008; Fleisher et al., 2007; Nation et al., 2019).

2.3.2 Shortcomings of Neuropsychological Tests

Currently, there appears to be a stagnation in the development, and even more so in the use of newly developed neuropsychological tests in clinical settings. A majority of the tests currently in use have been developed decades ago, and oftentimes have not been updated. When newer, updated versions of tests are created, they are not easily accepted by clinicians. For example, the MMSE, which was developed in 1975, before the introduction of the construct of MCI, has been proven to be far less sensitive in the earlier stage of AD and in MCI patients when compared to the relatively newer MoCA (Aggarwal & Kean, 2010; Pinto et al., 2018; Roalf et al., 2013; Siqueira, Hagemann, Coelho, Santos, & Bertolucci, 2018). This difference appears to be more pronounced in individuals with a higher education level or high cognitive reserve (Markwick, Zamboni, & de Jager, 2012). Despite overwhelming evidence confirming the lower sensitivity of MMSE in comparison to MoCA, MMSE continues to be the more popular test in clinical settings (Judge, Roberts, Khandker, Ambegaonkar, & Black, 2019). In the earlier stages of the disease, traditional neuropsychological tests may yield a ceiling effect, particularly in people with high IQ or higher education level. Established norms for neuropsychological tests are rarely updated, whereas emerging evidence points to changing norms in the population owing to the *Flynn effect* (Degen, Frankenberg, Toro, & Schröder, 2022; Munukka et al., 2021), which is the sustained increase in average intelligence in the population over time (Flynn, 1984, 1987).

One major shortcoming of current neuropsychological tests is a lack of ecological validity (Chaytor & Schmitter-Edgecombe, 2003; Ruff, 2003; Spooner & Pachana, 2006), frequently reported by psychologists as a major challenge in test selection (Rabin, Paolillo, & Barr, 2016). In the context of neuropsychological assessment, ecological validity is defined as "functional and predictive relationship between the patient's performance on a set of neuropsychological tests and the patient's behavior in a variety of real-world settings" (Sbordone, 1996), which is an extension of the concept of external validity of a test (Diehl, Wahl, & Freund, 2017). A major aspect of ecological validity is contextual-embeddedness. The tests currently employed involve paradigms that are artificial in nature, lacking any context; or, what one may refer to as laboratory tests. These paradigms attempt to test individual cognitive functions in isolation from other cognitive functions, which simply does not translate to everyday experiences of individuals, and may as a result, underestimate or overestimate the level of cognitive and functional impairment. Performance on a task requiring generation of a list of words or learning a list of words may give us insight into the cognitive processes that are impaired in the individual, but do not relay much information about how the individual processes everyday interactions or what tasks they may or may not be able to carry out in the real-world. This is because, when such tasks are conducted in a contextual vacuum, they may not reflect the true ability of individuals, given that contextual cues aid cognition, and real-world cognition is situated (Roth & Jornet, 2013). This appears more so to be the case with older adults, who display a gap in performance on laboratory tasks as opposed to naturalistic, contextually-embedded tasks (H. R. Bailey, Zacks, et al., 2013; P. E. Bailey, Henry, Rendell, Phillips, & Kliegel, 2010; Henry, MacLeod, Phillips, & Crawford, 2004; Zimerman, Hasher, & Goldstein, 2011). Additionally, it may be that impaired performance observed on a particular test may not

be caused by deficits in the particular cognitive domain that the test purports to measure, but may result from deficits in another cognitive domain (e.g. attention) that carry forward.

Further, studies assessing ecological validity of established neuropsychological assessment tools have reported only low to moderate ability of these neuropsychological tests to predict everyday functioning (Chaytor & Schmitter-Edgecombe, 2003). Established neuropsychological tests give us information on the individual cognitive abilities they purport to measure independently, but not on a real-world functional level. MMSE, the most commonly employed screening tool, has been shown to have only a moderate correlation with measures of daily functioning (Bouwens et al., 2008; Doble, Fisk, MacPherson, Fisher, & Rockwood, 2005). Groth-Marnat and Baker (2003) report that the Digit Span, which is very commonly used as a measure of attention, was a weak predictor of everyday attention. Similar findings were observed for tests of executive function such as TMT-B and WCST, with no to low predictive ability for everyday executive functioning (Chaytor, Schmitter-Edgecombe, & Burr, 2006). Despite these findings, these traditional tests continue to be used in clinical assessment over more ecologically valid measures (Rabin, Burton, & Barr, 2007). Current practice lacks cognitive outcome measures that would actually be helpful in assessing everyday cognition or day-to-day functioning in individuals with AD.

Neuropsychological tests were initially developed with the aim of assessment for the purpose of diagnosis. With the advent of newer biomarkers and relatively easier access to neuroimaging today, clinicians are increasingly relying on these techniques for the purposes of diagnosis. Recent revisions of diagnostic criteria, too, emphasize biomarker-based diagnosis, at least in research settings (Jack et al., 2018). And while these techniques are still quite expensive and not uniformly accessible everywhere, this gap is going to continue to narrow in the coming

years, as the technology becomes cheaper, and therefore, more widely available and easily accessible. However, the approach to diagnosis purely based on pathology without considering symptoms is problematic. As, elucidated above, in the context of AD, presence of pathology in and of itself is not very meaningful, as mere presence of pathology does not necessitate disordered functioning.

Currently, even with an increase in reliance on neuroimaging, existing neuropsychological assessment procedures still have value in the diagnostic procedure for AD and MCI in measuring cognitive function. However, there is a need to adapt the neuropsychological assessment procedures to realize a wider goal, to assess cognitive outcomes in the context of their functional and practical implications, such as, to assess patients' everyday functional ability or track longitudinal change in cognitive function. Laboratory-based testing measures and more ecologically valid measures of everyday cognition, each serve their unique purpose. Existing evidence indicates that measures of everyday cognition have added value in predicting real-life functioning beyond laboratory-based assessments (for a review, see Bielak, Hatt, & Diehl, 2017), and may open up avenues for cognitive training.

Two factors have been proposed to be considered in the issue of ecological validity of neuropsychological tests– *verisimilitude* and *veridicality* (Chaytor & Schmitter-Edgecombe, 2003; Franzen & Wilhelm, 1996; Spooner & Pachana, 2006). Verisimilitude, in this context, is the similarity in cognitive demands of the test and cognitive demands of the everyday environment, whereas veridicality is the degree to which performance on the test is related to measures of real-world functioning. Verisimilitude, essentially, calls for a change in approach to neuropsychological testing, with a focus on moving away from the "context-free" approach to cognitive testing, as real-life cognition occurs in and is aided by context. Similarly, Snyder and

colleagues reported the Alzheimer's Association Research Roundtable's (AARR) call for focusing on cognitive and functional decline during the early, pre-symptomatic stage of AD, and therefore, the need for improved cognitive and functional measures that are also clinically meaningful, to be used in these early stages (Snyder et al., 2014). They propose that novel approaches and alternate strategies are warranted. The current dissertation aims to take this view forward in focusing on approaches to cognitive testing that have bearing on real-world functioning.

2.3.3 Changing Role of Cognitive Assessment

As highlighted by the recent introduction of the drug Aducanumab, a reduction in pathological burden of the disease does not necessitate an improvement in cognitive outcomes (Haddad et al., 2022). The effect of pharmacological intervention at a pathophysiological level, but a lack thereof at a cognitive level, underscores the limitation of biomarker evidence in predicting cognitive trajectory or in measuring cognitive change, and calls attention to the role of cognitive tests in such scenarios. In studying the effect of interventions, be it pharmacological or non-pharmacological, the purpose of neuropsychological tests is not limited to diagnosis, but extends to measuring meaningful change in cognitive and functional outcomes. The question that then arises is, are traditional neuropsychological tests built to, and capable of measuring these changes in outcomes, considering that the testing paradigms and environment are far removed from the real-world environment and the challenges that come with it?

Subtle cognitive and linguistic deficits that may not yet be clinically evident or clinically significant have been widely reported in the preclinical stage of AD. Some studies report deficits primarily in one or two domains, such as verbal memory (Howieson et al., 1997), or episodic memory (Blacker et al., 2007), and others reporting more widespread impairment (Fox,

Warrington, Seiffer, Agnew, & Rossor, 1998; Hassenstab et al., 2015) at baseline. Several studies also claim the integrity, or lack thereof, of certain linguistic or cognitive domains (for example, executive function, visuospatial function) to be reliable predictors of conversion from MCI to AD or from preclinical stage to MCI (Belleville, Gauthier, Lepage, Kergoat, & Gilbert, 2014; Junquera, García-Zamora, Olazarán, Parra, & Fernández-Guinea, 2020; Levine et al., 2020).

As elucidated above, cognitive and linguistic assessment tools that are currently widely in use, have been developed several decades ago, and do not necessarily accommodate newer insights, such as the MCI stage or preclinical deficits. As a result, they may not be particularly sensitive to cognitive decline in the early and preclinical stages of AD (Logie, Parra, & Della Sala, 2015). An improvement in the approach to current cognitive and linguistic assessment practices in AD is warranted. One of the shortcomings of neuropsychological tests is a lack of verisimilitude. Although perfect verisimilitude is not possible in standardized testing, there is potential to increase verisimilitude. To that end, the following are some issues in testing that need to be addressed:

1. *Decontextualized testing*: Everyday cognition does not occur in a vacuum. Existing neuropsychological tests are devoid of supportive task-related context, which shapes our daily experiences. In the absence of supportive context, learning and recalling a list of random words, or generating a list of words belonging to a particular category is a difficult task in itself, one which older adults are particularly susceptible to, as they are far removed from the academic environment where such strategies may still be used. This may already put them at a disadvantage compared to younger adults who have more recent experience with academic learning. The artificial testing environment results in added stress, which is not conducive for assessing cognitive ability. Further, everyday

learning is enhanced by context. Therefore, laboratory neuropsychological tests have little bearing on real-life cognition, which is always situated. If the goal is to estimate and enhance everyday cognition in this population, the paradigms must reflect the contextual nature of cognition.

2. *Isolation of functions*: Traditional tests approach cognitive and linguistic functions as singular, compartmentalized processes working in isolation from each other. In reality, the processes are highly interdependent, and several processes work in tandem while performing any activity. While an isolationist approach may be useful in theoretical understanding of cognitive processes, cognitive assessment should reflect a practical approach, in that, cognitive functioning should be understood holistically.
3. *Task demands*: Following on from the isolationist approach towards neuropsychological assessment, isolationary neuropsychological tests do not simulate the task demands placed on an individual in the real world, and would have a tendency to underestimate or overestimate them. Although it is very challenging to fully simulate the task demands of the real world in a laboratory setting, as all the distractions and interferences experienced in the real world cannot be introduced in the lab, the use of cognitively more complex paradigms, which can match the demands of real life cognition, would go a long way towards narrowing that gap.

A renewed approach to cognitive and linguistic assessment in AD should aim to address these issues. As stated above, the AARR, 2019 emphasized the need to fundamentally change assessment tools that are used during the early stage of the disease (Rentz, Wessels, Bain, Weber, & Carrillo, 2020). It underscored the need for assessment tools that are clinically meaningful. Although there is still some ongoing debate with regards to what ‘clinically meaningful’ means,

the AARR calls for assessment tools that are sensitive to cognitive changes in the early stage of the disease and are able to predict everyday functional ability, as cognition underlies functional ability.

In recent years, the literature is increasingly shifting focus to tackle these issues and developing novel paradigms that are more complex and more challenging in nature, that attempt to simulate cognitive demands of real-life functioning, that are contextually-embedded, resistant to practice effects, and are more sensitive to impairments resulting from pathology, as opposed to cognitive ageing (Loewenstein, Curiel, Duara, & Buschke, 2017; Logie et al., 2015; Rentz et al., 2013). Over the years, paradigms with higher verisimilitude to daily life conditions have been developed. For example, Test of Everyday Attention (TEA; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996) is a better indicator of everyday attention compared to digit span test (Groth-Marnat & Baker, 2003); or, Rivermead Behavioural Memory Test (RBMT; Bolló-Gasol, Piñol-Ripoll, Cejudo-Bolivar, Llorente-Vizcaino, & Peraita-Adrados, 2014) for examining episodic memory. Distinguishing AD from effects of cognitive ageing can also be achieved by targeting deficits that are unique to AD, by implementing more complex paradigms, and particularly by using naturalistic stimuli that can help overcome test anxiety-related performance issues in healthy older adults. Taken together, the above evidence, once again, points towards focusing cognitive assessment in AD towards real-world, everyday cognition, which is the approach taken in the present dissertation.

2.4 Forming and Understanding Events

An important aspect of everyday functioning is event cognition. Events are the mental representation of the world around us and our everyday experiences, or, as defined by Zacks and Tversky (2001), “a segment of time at a given location that is conceived by an observer to have a

beginning and an end". How we perceive ongoing activity, form mental representations of them, store information about them, make predictions of what will happen, and accordingly take action, may be referred to as event cognition. These events are organized hierarchically, wherein, a coarser-grained event is constituted of several finer-grained events. The coarser-grained event can be referred to as the macro-event, and the finer-grained events within it are sub-events or micro-events. In theory, every macro-event is potentially a sub-event for a larger macro-event.

Previous work surrounding event cognition in AD has focused on two areas– script representation and event segmentation. Scripts, in this context, are defined as knowledge networks of schema-like representations of a typical sequence of events, which have been built using prior knowledge (Abbott, Black, & Smith, 1985). For example, the event of 'going grocery shopping', would typically involve making a shopping list, going to the supermarket, picking up groceries, and making the payment. In such a manner, individuals tend to have representations of actions that would typically be involved in commonly experienced events. How strong or weak the representations are, depends on factors such as, the age at which representations of a particular event are formed and stored, and how frequently or infrequently we encounter said event hierarchy.

One method for assessment involved generation of the sequential steps that would typically be involved in an event (Grafman et al., 1991; Roll, Giovannetti, Libon, & Eppig, 2019). These studies found AD patients produced fewer steps compared to cognitively healthy adults; AD patients also had difficulty following the temporal order in generating steps. Further, AD patients experienced interference from unrelated event actions, i.e. they were more likely to produce steps or actions that did not belong in the event. These issues with maintaining temporal order and with inhibition of interferences were further confirmed using a script sequencing

paradigm. AD patients, similarly, experienced difficulty in sequencing given verbal scripts for events in the correct temporal order (Allain et al., 2008; Grafman et al., 1991). Additionally, in the Allain et al. (2008) study, AD patients experienced difficulty in distinguishing scripts belonging to distinct events, and experienced interference from actions that were unrelated to represented event. These problems in sequencing and sorting persisted when the verbal scripts were replaced by picture representations of the scripts (Roll et al., 2019), indicating the issue was not necessarily connected to linguistic comprehension deficits, rather, it was an issue in the conceptualization and formation of event representations. Overall, degradation of semantic knowledge was observed, and considered to be contributing to impairment in everyday function.

Two studies have examined event cognition in AD by focusing on the segmentation behavior of subjects for everyday events. Event segmentation studies involve segmentation of a continuous stream of activity into smaller, meaningful units. The points at which subjects segment an event are the event boundaries. In keeping with our knowledge of script representations, there is generally a consensus on where event boundaries should lie. This is however, not the case in individuals with AD, whose segmentation behavior was found to be idiosyncratic i.e. event boundaries were placed more randomly rather than the more typical pattern, and there was also little consensus among the group about where event boundaries should lie (H. R. Bailey, Zacks, et al., 2013; Zacks, Speer, Vettel, & Jacoby, 2006).

Our encoding of future events is supported by event schema, which are semantic representations of common features of the event aggregated from our previous experiences of the event. A degradation of these semantic representations results in faulty encoding behavior. Segmentation behavior is also linked to subsequent memory for the event. Zacks et al. (2006) found that subsequent memory for the events was impaired in the AD patients, and was

associated with their event segmentation ability. So, more normative event segmentation resulted in better memory for the event later. These findings indicate a failure not just in storage and retrieval of memories in AD, but also during the encoding stage. It is possible that failure to retrieve memories results from faulty encoding strategies. The study by Zacks et al. (2006) also, from a different perspective, further confirmed the issues in event sequencing in AD, which were observed in the script sequencing paradigm.

Current tools that are used to measure functional ability are generally questionnaires filled out by caretakers, which measure BADLs and IADLs. Goldberg and colleagues conducted a study which included a performance-based measure for assessing functional ability, in addition to a routinely used questionnaire measuring ADLs (Goldberg et al., 2010). The performance-based measure included activities such as, making a telephone call, planning an activity, planning a route, or financial transactions. The study included individuals with mild to moderate AD and individuals with MCI, a group that is considered to be functionally unimpaired. They found that the MCI group was significantly impaired on the performance-based measures, but not on the questionnaire measuring ADLs, which was an informant-based measure. This hints at impairment in the MCI stage not only at a cognitive level, but also at a functional level. It is perhaps not to the degree where it would interfere with their ability to live independently. This, however, indicates that currently used tools for assessing everyday function may be overly simplistic, and may not entirely capture the full spectrum of functional deficits across early and preclinical stages of AD. Taking into account not only the implications for clinical diagnosis, but also in terms of developing objective and sensitive endpoints for measuring effects of interventions, novel approaches that are able to measure cognition and function of individuals with AD and MCI in a way that is meaningful for individuals' everyday life, are warranted.

Overall, cognitive assessment paradigms which focus on real-world cognition, by examining event understanding in AD patients, such as script sequencing or event segmentation, show a promising way forward to understanding deficits in AD patients in everyday life and their functional ability. A major focus of this current work is addressing the gaps in the literature on event perception and understanding in AD, by examining the interaction between micro- and macro-events. To complement previous work on event segmentation, this work focuses on event integration and macro-event formation.

2.5 The Present Dissertation

Research on novel, cognitively complex measures of language and cognition that are contextually-driven, ecologically more valid than traditional neuropsychological measures, and offer more meaningful outcomes for real-world functioning, has been gaining traction over the last few years. The current work argues for and aims to contribute to the development and use of such measures in identifying indicators of AD in the early stages of the disease, endeavors to reconcile objective lab-based measures of cognitive assessment with cognitive demands of everyday cognition, and aims to contribute towards identification and preliminary development of such measures. Extending work from previous studies on discourse comprehension, event segmentation, and script representation, the present work focuses on examining macro-level comprehension in the early stage of AD and in MCI in relation to cognitively unimpaired older adults. Processing information at a macro-level imposes a higher cognitive load as simultaneous processing on multiple levels and an interaction between these levels is required. Micro-level units of information need to be integrated to construct the macrostructure, which also requires the flexibility to move up and down the hierarchical levels. The complexity of the processing required, as opposed to recalling a list of words, makes it a good target for detecting deficits

early on. Therefore, macro-level comprehension in discourse was evaluated in a systematic review of text comprehension studies; and, in event cognition via a picture-based novel paradigm. The specific aims of the present dissertation are presented in Table 2.1.

Table 2.1: Overview of research aims in the present dissertation

Chapter	Research Aims
Chapter 3	• To review and synthesize evidence on macrostructural features in text comprehension, and evaluate their potential as an early marker of AD, and in distinguishing AD and MCI from cognitive ageing • To characterize macro-level measures of text comprehension and identify measures that are most sensitive to AD-associated cognitive decline • To examine associations of macrostructural features in discourse comprehension with existing neuropsychological tests
Chapter 4	• To examine naturalistic event cognition using a novel paradigm, and assess its potential to differentiate early stage AD and MCI patients from cognitively healthy older adults • To assess macro-level event understanding ability of AD and MCI patients by examining event sequencing, event integration, and event recognition ability • To examine the effect of verbal cueing on event integration and sequencing ability
Chapter 5	• To examine the cognitive processes involved in performing the novel macro-event recognition task, using neuropsychological tests used in AD diagnosis • To evaluate whether the event recognition task may contribute to cognitive assessment in AD beyond that of traditional neuropsychological measures

Chapter 3

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[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,* 3–18. https://doi.org/10.1007/s10433-021-00619-5

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.

3.1 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, Jones, Berger, Laukka, & Small, 2005; DeTure & 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, Tolppanen, Kivipelto, & Soininen, 2014; Livingston et al., 2017; Xu et al., 2015) can only be successful when targeted before significant pathological changes and cognitive decline have 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; P. 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 of (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, Patterson, Xuereb, & Hodges, 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 & 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; although, 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 & Reuter-Lorenz, 2009). Classic cognitive testing, so far, taps into language-related functions only marginally (Cummings, Darkins, Mendez, Hill, & Benson, 1988; Taler & Phillips, 2008; Verma & Howard, 2012; Vuorinen, Laine, & Rinne, 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, Crawford, et al., 2004; Reilly, Troche, & Grossman, 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, Waters, & Caplan, 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 & 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; Weiner et al., 2019).

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 & Chrysler, 1990; Graesser, Millis, & Zwaan, 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 & 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, macro linguistic features of discourse production were the most susceptible to decline in the early and prodromal stages of AD (Brandão, Lima, Parente, & Peña-Casanova, 2013; Pistono et al., 2019). The patterns of deficits observed in micro- and macro-structural processing have been shown to have utility in distinguishing clinical populations (Ulatowska, Chapman, Johnson, & Branch, 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 & Dijkstra, 2007; Ulatowska, Chapman, Highley, & Prince, 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 to 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.

3.2 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).

---

[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.

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: (i) the study included a group of participants who had a formal diagnosis of Alzheimer's disease or Mild Cognitive Impairment, using well-established criteria; (ii) the study included a healthy control group for comparison; (iii) mean age of the healthy group was ≥ 60 years, or population was age-matched to the patient group; (iv) 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: (i) Studies with other types of dementia population; (ii) studies measuring verbatim recall of discourse texts or only memory for details within the text; (iii) studies measuring spontaneous or picture-elicited discourse production; (iv) 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, Cook, & Lee, 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 fulfillment 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 = .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 3.1. All studies were deemed to be of a fairly good quality (≥0.75).

3.3 Results

Search Results and Study Characteristics

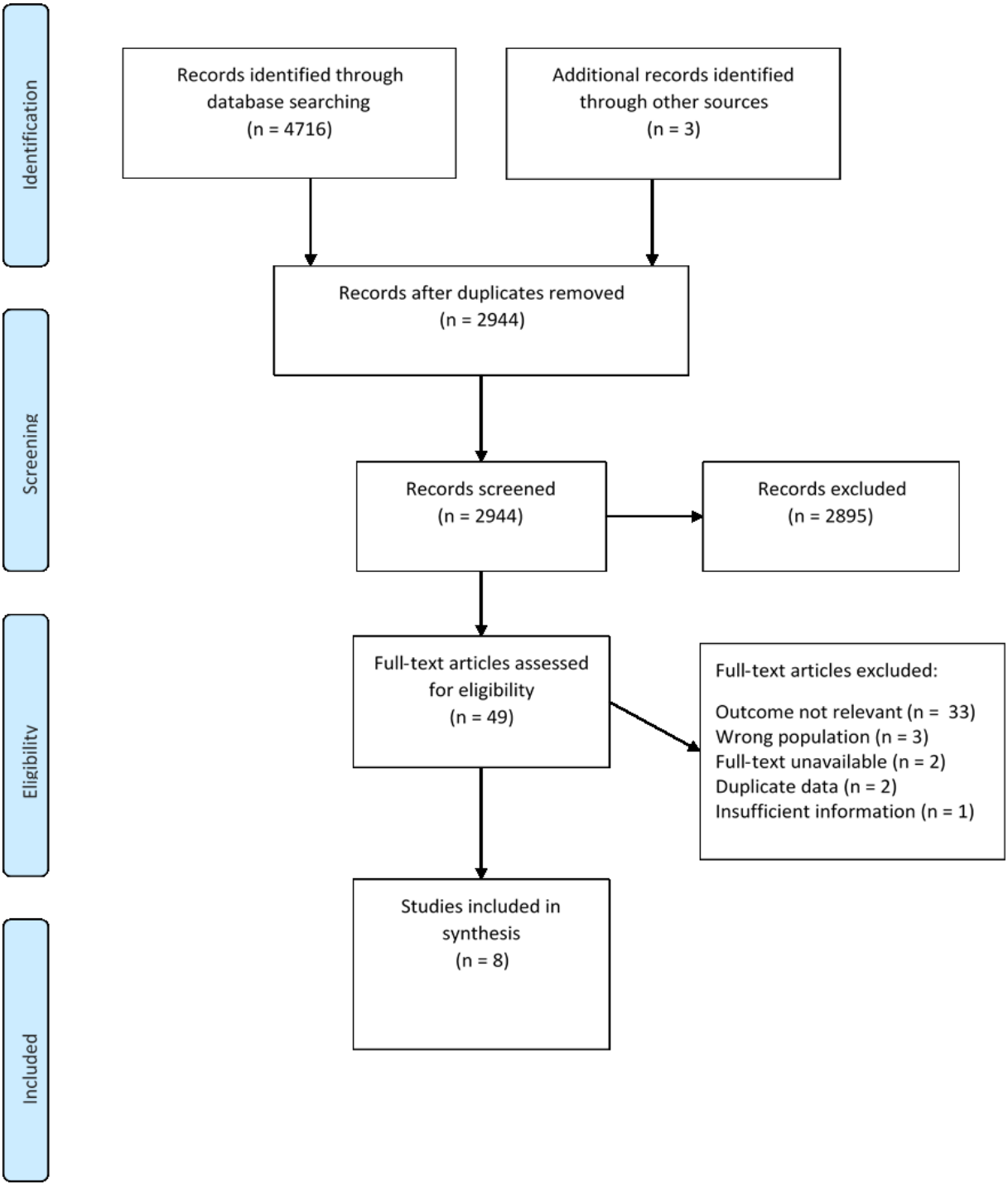
The search yielded a total of 4716 articles combined from PubMed (1954-2020), Web of Science (1934-2020), and PsycINFO/EBSCO (1934-2020). After removing duplicates 2941 articles remained, for which title and abstract were screened. Additionally, references of included articles were screened, and three additional articles, which met the inclusion criteria, were identified (Chapman, Anand, Sparks, & Cullum, 2006; Graville & Rau, 1991; MacDonald, Almor, Henderson, Kempler, & Andersen, 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 (κ = .81). The reasons for exclusion are highlighted in Figure 3.1. 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 macro-level, in adults with Alzheimer's disease or MCI.

An overview of the study characteristics is presented in Table 3.1. 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, Lubinski, & Higginbotham, 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; Chapman et al., 2002; Chapman et al., 2006; Creamer & Schmitter-Edgecombe, 2010; Drummond et al., 2019; Schmitter-Edgecombe & 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).

Figure 3.1: Flowchart of literature search and study selection process



Only one study (Drummond et al., 2019) used a test from a standardized battery (MAC battery) (Fonseca, Parente, Côté, Ska, & Joanette, 2008), and one (Welland et al., 2002) used a modified form of the Discourse Comprehension Test (DCT) battery (Brookshire & 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, 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 & Schmitter-Edgecombe, 2010; Schmitter-Edgecombe & Creamer, 2010), one study measuring naming latencies for congruent and incongruent pronouns (Almor, MacDonald, Kempler, Andersen, & Tyler, 2001), and others measuring gist-level retelling in the form of summary, lesson, main ideas (Chapman et al., 1998; Chapman et al., 2002; Chapman et al., 2006; Welland et al., 2002). Due to this heterogeneity in tasks and reported outcome measures, a meta-analysis was not performed.

Table 3.1: Characteristics of included studies

First Author, Year, REF	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 & 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 & 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., 200l; 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

Notes: [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, Berg, Danziger, Coben, & Martin, 1982); [d]The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.
** Entered as covariates*

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, Wagenpfeil, 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 & 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 etiologies 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 3.2.

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 within a macro structure, 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; Chapman et al., 1998; Chapman et al., 2002; Drummond et al., 2019), participants were presented with a short story. Following this, participants were asked to re-tell 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 analyzed 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; Chapman et al., 1998; Chapman et al., 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.

Table 3.2: Comparison of group performance on discourse comprehension measures

First Author, Year	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 & Schmitter-Edgecombe, 2010	----	----	----	----	AD<NC*	AD<NC***
Drummond et al., 2019	----	AD<MCI<NC*	AD=MCI<NC*	----	----	AD<MCI<NC*
Schmitter-Edgecombe & Creamer, 2010	----	----	----	----	MCI<NC**	MCI<NC*
Welland et al., 2002	----	----	----	----	----	MDAT=EDAT<NC**

**p≤.05, **p≤.01, ***p≤.001, †p-value not reported*

Main Idea

This probe, also administered following a short story in three of the studies (Chapman et al., 2006; Chapman et al., 1998; Chapman et al., 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 & Schmitter-Edgecombe, 2010; Schmitter-Edgecombe & 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 & Schmitter-Edgecombe, 2010; Drummond et al., 2019; Schmitter-Edgecombe & Creamer, 2010) included comprehension questions following the other re-telling 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) pose 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; one, 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. 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 = .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 = .64, r = -.83) with discourse comprehension in two studies (Almor et al., 2001; Welland et al., 2002), but two other studies (Creamer & Schmitter-Edgecombe, 2010; Schmitter-Edgecombe & 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 = .50 to r = .64) correlation with discourse comprehension measures. Only one study (Welland et al., 2002) reported a positive association with episodic memory (r = .91*)*. Additionally, one study (Creamer & Schmitter-Edgecombe, 2010) found significant correlations with TMT-A (r = .58) and D-KEFS (r = .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.

3.4 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 to 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 to 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, macro-level 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 & 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 & Schmitter-Edgecombe, 2010; Drummond et al., 2019; Schmitter-Edgecombe & 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 & Levy, 2003). Evidence from studies examining spontaneous or picture-elicited discourse

production show 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 & Murdoch, 1994; Laine, Laakso, Vuorinen, & Rinne, 1998; Toledo et al., 2018), critical macro linguistic 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 & 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 macro-level abstraction and comprehension deficits in this population (Vas, Spence, & Chapman, 2015). They showed tasks employing macro-level processing to have high sensitivity and specificity in TBI due to the complexity of processing required (Vas, Spence, Eschler, & Chapman, 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 macro-level processing benefits not only abstraction ability, but also extends to other general cognitive functions like attention and executive functions (Chapman & 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 & Merikle, 1996; Just & 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, Chertkow, & Gauthier, 2007; Huntley & 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 analyzing 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, Hubbard, & Richardson, 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, Cunha, Toledo, & Scarton, 2016; Clarke, Foltz, & Garrard, 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 & 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, Eling, Kastrup, & Hildebrandt, 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 & 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 don't convert (Celsis, 2000). Another study also showed disparate profiles of MCI patients in a text comprehension task (Chesneau, Lepage, Giroux, & Belleville, 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 multidimensional 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 multilevel 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.

Chapter 4

Macro-Event Recognition in Healthy Ageing, Alzheimer's Disease, and Mild Cognitive Impairment

Eesha Kokje[a], Johannes Gerwien[b], and Christiane von Stutterheim[b]

[a]Network Aging Research, Heidelberg University, Germany; [b]Institut für Deutsch als Fremdsprachenphilologie, Heidelberg University, Germany

Kokje, E., Gerwien, J., & von Stutterheim, C. (2021). Macro-event recognition in healthy ageing, Alzheimer's disease, and mild cognitive impairment. *Journal of Neuropsychology.* https://doi.org/10.1111/jnp.12271

Abstract

Event perception and cognition is integral to our everyday experience and functional ability. A commonly reported complaint in Alzheimer's disease (AD) is the inability to follow narratives– be it textual, conversational, video, or pictures. This phenomenon has received little systematic research so far. In the current study, we developed a novel paradigm to examine macro-event recognition in individuals with AD in the early stage and its preceding stage of Mild Cognitive Impairment (MCI) in comparison to cognitively healthy older adults, using pictures depicting events. In Experiment 1, we examined participants' ability to integrate pictorially-depicted sub-events into macro-events. The pictures were presented in a scrambled order and participants were expected to arrange them in the temporally and causally appropriate sequence, as dictated by the macro-event schema. Additionally, we investigated the effect of cueing the appropriate event schema by providing a word cue (verb). In Experiment 2, macro-event recognition was examined again using a cognitively less taxing paradigm, where pictures depicting sub-events were presented in correct order, but staggered, and recognition speed was measured. We observed significant deficits in the AD and MCI groups' performance compared to the cognitively healthy older adults, across both experiments, suggesting event perception and cognition is impaired early in the course of AD. There was no effect of cueing on the performance of any of the groups. The theoretical and clinical implications of these findings are discussed.

4.1 Introduction

Alzheimer's disease (AD), the most common cause of dementia, and Mild Cognitive Impairment (MCI), its preceding stage, are characterized by progressive decline in cognitive function. Although the most commonly reported symptom in AD is episodic memory deficits, objective impairment can be observed in several linguistic and cognitive domains during the course of AD and MCI, including but not limited to attentional control, semantic memory, and executive functions (Perry & Hodges, 1999; Verma & Howard, 2012). While memory complaints have been well documented, other areas of cognition have received relatively less attention. Neuropsychological tests that are currently in use measure attention, episodic memory, language, but employ artificial, laboratory tasks (e.g., word list learning, object naming). These typically do not target combinatorial or inferencing abilities. Moreover, existing tests are devoid of any context, and do not reflect naturally occurring experiences or situations. Activities of daily living is an important component of diagnosing AD and determining the ability of individuals to function independently (Lawton & Brody, 1969). Current assessments of functional ability comprise of self-report or caregiver-report. More objective assessments, that identify the root of the deficits, and can point to appropriate interventions, are warranted.

Previous research has shown that individuals with AD have difficulty integrating smaller units of information into a whole. This has been observed in visual and spatial perception (Duffy, Cushman, & Kavcic, 2004; Paxton et al., 2007). However, identifying a whole given smaller units plays a role in many other cognitive domains, as for example in discourse comprehension. On a more abstract level, the ability to construe larger conceptual units is fundamental to all kinds of prediction phenomena (Kuperberg, 2021).

Event recognition and prediction are closely related to the ability to perform actions (H. R. Bailey, Kurby, Giovannetti, & Zacks, 2013; Cooper, 2021). In order to perform an activity, one needs to have a relevant mental representation of the activity, often referred to as *event model*, which is activated based on an abstraction of the common features of past experiences called *event schema* (Zacks & Tversky, 2001). These schemata are created by drawing upon pre-existing experiences from episodic memory and semantic knowledge (Gerwien & von Stutterheim, 2018; Sargent et al., 2013; Zacks, Speer, Swallow, Braver, & Reynolds, 2007), and provide a framework for processing and encoding ongoing events (Zacks et al., 2007). Further, events can be segmented into smaller event units. Every event can in principle be a macro-event for a number of micro-events which are causally and temporally related to each other (Kuperberg, 2021). From this perspective, event knowledge is organized hierarchically. Making connections between micro-events, and predicting subsequent events are key components involved in holistic processing of information. This is accomplished via inference generation, referred to as *bridging inferences* and *predictive inferences* respectively (Cohn, 2019; Magliano, Dijkstra, & Zwaan, 1996; Singer & Halldorson, 1996).

What is the role of language in activating event schemata? Previous research has shown that if participants are prevented from using inner speech while performing cognitive tasks that involve - among others - information categorization, action planning, or switching between different tasks - performance typically goes down (Alderson-Day & Fernyhough, 2015). This points to an involvement of language. However, it is not yet entirely clear whether it is language use *per se* that supports performance because it helps to focus attention on relevant aspects of the task (Miyake, Emerson, Padilla, & Ahn, 2004), or whether it is really the specific semantic content that is transported via language, which supports performance. Since verbal tests have

been used to assess cognitive decline previously and some correlation between verbal behavior and cognitive decline has been attested, one aspect of the current study will be about the role of language in event schema activation as measured in a non-verbal sorting task with and without verbal cueing.

Event Cognition in Ageing and Alzheimer's Disease

Studies investigating the effects of ageing and age-related neurodegenerative diseases on event cognition have been rather limited. One area of focus has been event segmentation ability and subsequent memory for these events (H. R. Bailey, Zacks, et al., 2013; Kurby & Zacks, 2018; Zacks et al., 2006). The evidence from these studies suggests that cognitively healthy older adults maintain their event segmentation abilities when compared to AD patients. In these studies, participants were presented with 1-minute films, and asked to segment them into smaller, meaningful events. Individuals with AD segmented the stimulus in a more idiosyncratic way, i.e. agreement about where event boundaries should lie was low. This was in contrast to the younger adults group, who had very high segmentation agreement, as well as the older adults group, who also had higher segmentation agreement compared to the AD group, but less than the younger adults. Moreover, several studies found that individuals' later memory was highly correlated with their segmentation ability (Flores, Bailey, Eisenberg, & Zacks, 2017; Kurby & Zacks, 2018; Zacks et al., 2006). These findings indicate that memory problems in AD occur not only due to issues related with storage and retrieval, but they actually begin with failure to encode information correctly.

Event cognition in AD patients has also been studied using a paradigm involving verbal scripts. The tasks involved script generation, wherein participants produced the steps in a given macro-event (Grafman et al., 1991; Roll et al., 2019); or, verbal script sequencing, in which

participants either arrange the steps in event scripts in the correct sequence or determine whether the provided script sequences were correct (Allain et al., 2008; Grafman et al., 1991). The results showed that AD patients' performance was worse on script generation and sequencing compared to cognitively healthy older adults. One recent study further eliminated the verbal aspect, and replaced the verbal scripts with pictures depicting a sequence of actions in an event, which were to be arranged in the appropriate order (Roll et al., 2019). The findings remained consistent with verbal sequencing studies.

Studies using such tasks go beyond traditional episodic memory measures and psychometric tests (Sargent et al., 2013; Zacks et al., 2006) due to the richer and more realistic context they provide. In the current study, we wanted to explore further aspects of event cognition in AD patients, and additionally in MCI patients. Instead of *segmenting* a whole into smaller units, we focused on the *integration* of smaller units into a whole (West & Holcomb, 2002). Using a paradigm similar to the sequencing paradigm in the Roll et al. study, we were interested in how AD and MCI patients in comparison to older adults can activate and use schema knowledge to identify macro-events from their sub-component micro-events, and use schema knowledge for establishing temporal and causal sequences. Event segmentation and event integration both require knowledge of, and the ability to activate event schemata. However, during segmentation, the cognitive system is alert to changes in the perceptual input in order to establish event boundaries. Integration, on the other hand, requires abstraction over several event units that are each separated by an event boundary. Segmentation and integration are linked via the theoretical concept of "granularity". For example, if we consider the event of writing a paper; depending on different levels of granularity, either boundaries are established between hitting the letter 't', hitting the letter 'y' and so on, or between typing and looking up references, or between

working on a paper and taking a coffee break. Event integration is a cognitive activity that comprises of switching to a coarser granularity level, which allows integrating micro-events into macro-events by activating appropriate event schemas.

We conducted two experiments using picture narratives that examined participants ability to integrate parts of the events, draw upon event schema, and constitute the parts into their whole (Experiment 1); and, with a cognitively less demanding paradigm (Experiment 2). In both experiments, we studied patients diagnosed with AD, patients diagnosed with MCI, and cognitively healthy older adults.

4.2 Experiment 1

The aim of Experiment 1 was to investigate AD and MCI patients' ability to integrate micro-events to form macro-events and compare their performance to cognitively healthy older adults (OA group). A second goal was to examine whether cueing the appropriate macro-event schema via language improved performance. We developed a novel procedure in which participants were presented with four pictures that depicted sub-events representing different temporal stages within a single, larger macro-event, presented in a scrambled order. We assessed participants' sequencing and naming accuracy.

4.2.1 Method

Participants

A total of 47 participants completed the study. The cognitively healthy older adults were recruited from a database at the Network Aging Research, Heidelberg University. Potentially eligible participants were contacted via telephone, and a brief telephone screening was conducted to rule out visual or auditory impairment, any existing diagnoses of neurological or psychological conditions (e.g., stroke, depression, epilepsy, Parkinson's disease), or self-reported

cognitive complaints. Eligible individuals were then invited to take part in the study. Before beginning the experiment, participants in the cognitively healthy group were administered the Mini-Mental State Examination (MMSE) to screen for possible cognitive impairment. Individuals who scored ≤ 26 points on the MMSE were later excluded from the analysis. 25 cognitively healthy older adults completed the study. However, data from only 20 participants were included in the analyses, as three participants were excluded based on their MMSE score, and a further two participants' testing session was interrupted due to a technical error.

The AD and MCI patients were recruited from the memory clinics of University Hospital Heidelberg and Central Institute of Mental health, Mannheim. A total of 22 people were recruited, of which, 10 had a diagnosis of AD, and 10 were diagnosed with MCI; two were excluded due to an uncertain diagnosis. The patients were formally diagnosed by a physician following neurological and neuropsychological evaluations. These included blood tests, clinical history, CT/MRI scans, cerebrospinal fluid testing, and CERAD test battery (John C Morris et al., 1989). The National Institute on Aging and the Alzheimer's Association (NIA-AA) criteria were applied for AD diagnosis (McKhann et al., 2011). All individuals received a diagnosis of probable AD. Only individuals in the mild stage of AD (MMSE score: >19) were included in the study. The individuals in the MCI group were diagnosed according to the NIA-AA criteria for MCI (Albert et al., 2011) or the ICD-10 criteria (World Health Organization, 1992). The two criteria are comparable enough for participants to be grouped together. The classification of sub-type of MCI was only available for three of the ten participants, all of whom had amnestic MCI. For demographic characteristics, refer to Table 4.1.

All participants were native speakers of German, and reported normal or corrected-to-normal vision. The group with AD had significantly fewer years of education compared to the

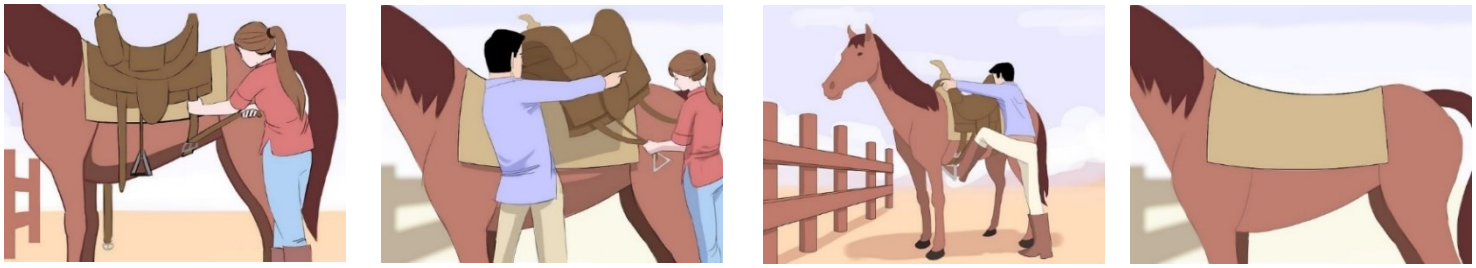
cognitively healthy group, but MCI group's education did not differ from either group. All participants provided informed consent. The study received ethical approval from the Ethics Commission of the Medical Faculty of Heidelberg University, Germany.

Stimuli

The stimuli consisted of 14 sets of events, each comprising four pictures depicting sequential stages within a macro-event. For example, in a trial depicting a 'grocery shopping' event, the pictures illustrate preparation of a shopping list, going to the supermarket, picking groceries, and paying. Each picture was 300 pixels in width and height. For example stimuli, see Figure 4.1. The stimuli were created from step-by-step tutorials demonstrating how to perform different activities, which were obtained from 'WikiHow' (http://wikihow.com). Initially, 40 such trials were created, which were tested in two groups – university students and university professors. During these test sessions, participants were presented the individual pictures of a set in a scrambled order and were asked to re-order them to resemble a temporally correct sequence, followed by naming the macro-event being depicted. The 28 events on which participants performed best in terms of naming agreement were selected– 14 in experiment 1 and 14 in experiment 2 (see below). The temporally and causally appropriate sequences were determined by the responses that maximum number of participants agreed upon during the pilot test. All, but one, events included in the final experiment had a sequence agreement of at least 95%, and one had an agreement of 86%. Macro event names for verbal cueing were also determined by the results from the pilot test (for a full list, see Supplementary materials).

Figure 4.1: Example of stimuli A) scrambled presentation, B) correct sequence

A



B

On half of the trials in the actual experiment with patients and controls, participants were given the macro-event name as a cue, and only had to re-order the pictures. On the other half, they had to name the event after re-ordering the pictures. For the purposes of counterbalancing, two versions of the experiment were created. In one version, one half of the trials were presented along with the word cue, and the other half had to be named, whereas this was reversed on the second version. Each participant performed only one of the two version, but these two versions were counterbalanced among the participants.

Procedure

Participants provided informed consent for the study, following which they filled out a sociodemographic questionnaire. In the cognitively healthy older adults group, this was followed by administration of the MMSE, and finally, the main experimental task was administered.

The experiment was programmed using Javascript, and presented on a 10.1” screen tablet. At the start of the experiment, participants received detailed instructions for the task on the screen, and additionally were given an explanation verbally when the task was not clear. They were given five practice trials at the beginning to ensure that they understood the task, were able to perform it correctly, and to familiarize them to using the tablet. Instructions were repeated as many times as required. Only after ensuring that participants understood the task, the actual experiment began.

Each participant was presented with fourteen trials – seven cued and seven non-cued – in an order that was randomized for each participant. The presented picture order within each trial was also randomized for each participant. The participants sorted the pictures using the touchscreen feature of the tablet. For naming the event, the participants typed in their response. We recorded every movement of pictures, their final sequence responses, and their naming responses. The task was self-paced, and participants could move around the pictures as many times as they wished, until they were satisfied with the sequence. The task took between 15-35 minutes, depending on individual participants’ pace.

Statistical Analyses

The recorded macro-event names were scored by two independent assessors. One of them was blinded to the diagnosis of the participants. Any discrepancies between assessors were discussed until a conclusion was reached. Sequence accuracy was measured using edit distance (for details, see Results).

All analyses were conducted in R, version 4.0.4 (R Core Team, 2021). The Shapiro-Wilk test, along with visual examination of density plots were used to determine normality of distribution, and Levene’s test was conducted to check for heteroscedasticity. For outcome

variables that were not normally distributed, the non-parametric Kruskal-Wallis test was conducted, followed by Dunn's multiple comparison test, and a Benjamin-Hochberg correction for multiple comparisons. For normally distributed variables with unequal variances, a Welch's ANOVA was conducted (Delacre, Leys, Mora, & Lakens, 2019), followed by Games-Howell post-hoc test. A robust factorial ANOVA using 10% trimmed means was conducted using the 'WRS2' package (Mair & Wilcox, 2019) to analyze the effect of cueing and diagnosis on sequence accuracy. Correlations were examined using Spearman's rho correlation coefficient. The Levenshtein distance was calculated using the 'stringdist' package (Van der Loo, 2014).

4.2.2 Results

Naming Accuracy

Naming accuracy was measured by scoring participants' responses, as correct or incorrect i.e. '1' or '0', for each trial. Acceptable responses were the pre-determined event terms, synonyms, or alternate terms that encapsulated the macro-event (for example, for the event 'eating at a restaurant', 'restaurant visit' or 'going to a restaurant' would be acceptable responses). Naming or descriptions of individual micro-events, or specific objects within the pictures were scored as incorrect (e.g., for the event 'grocery shopping', 'paying' or 'making a list' would be unacceptable as they are not indicative of the event as a whole, but single pictures/actions within the event). These general categories of acceptable or unacceptable responses were pre-determined, but not specific responses. Percentage accuracy was calculated for each participant. The means are provided in Table 4.1.

Table 4.1: Demographic characteristics and performance on measures in Experiment 1 (Means and SD)

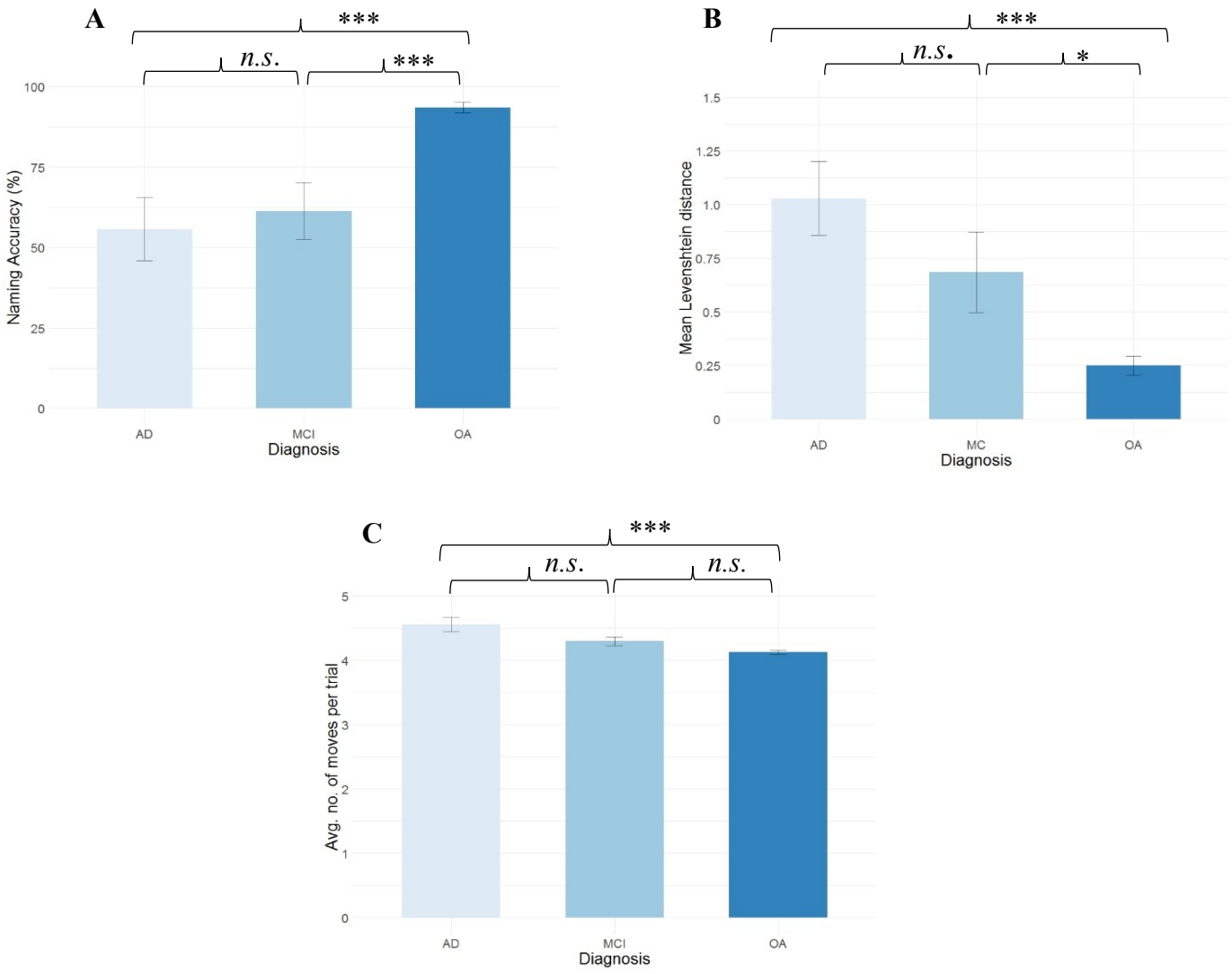
Measure	OA (n=20)	MCI (n=10)	AD (n=10)	*p*[a]	*η2*
Age	70.9 (5.5)	72.6 (6.2)	74.2 (7.5)	.17	.04
Sex (F/M)	16/4	3/7	7/3	.02[b]	.43[c]
Education (years)	11.8 (1.7)	10.4 (1.8)	9.9 (1.7)	.01	.18
MMSE score (max. 30)	28.9 (0.9)	24.7 (1.3)	23 (1.9)	<.001	.76
Naming Accuracy (%)	93.6 (7.3)	61.4 (27.8)	55.7 (31.2)	<.001	.51
Sequence Accuracy (LD[d])	0.25 (0.2)	0.69 (0.59)	1.03 (0.55)	<.001	.35
Average no. of moves per trial	4.13 (0.1)	4.3 (0.2)	4.56 (0.4)	<.001	.34

[a]Kruskal-Wallis p-value, [b]Pearson's chi sq. p-value, [c]Cramer's V, [d]Levenshtein distance

Commonly occurring errors in naming included an inability to come up with a response, describing each micro-event instead of naming the macro-event, naming just an object within the pictures, or naming unrelated event terms. One commonly observed error in the patients was their preoccupation with one particular presented event, and use of that event term repeatedly for subsequent event trials, even when the events were unrelated.

To analyze the differences between the groups formally, the non-parametric Kruskal-Wallis was conducted. The test revealed a significant effect of diagnostic category on naming accuracy, $H(2) = 21$, $p < .001$, $\eta^2 = .51$. Further examination showed that AD and MCI groups performed significantly worse than the OA group, but did not differ significantly from each other (see Figure 4.2).

Figure 4.2: Group-wise means on Experiment 1 measures of A) naming accuracy; B) sequence accuracy; C) average no. of moves per trial (error bars represent SE)



*$p \leq .05$, ***$p \leq .001$, n.s.= not significant*

Sequence Accuracy

Sequence accuracy was calculated using the edit distance ("Levenshtein distance"). In this, observed and expected picture sequences were represented as four-letter strings. Expected sequences always had the format 'ABCD' (normed order of the four sub-events). Observed

sequences were coded using the same four letters. The order of letters corresponded to the order of the sub-events, as arranged by the participants. Observed and expected sequences from each participant on every trial were then compared by calculating how many edits of the observed sequence were necessary to derive the expected sequence. For example, the response ‘CADB’ would require four edits, and, therefore, is more different than the response ‘BACD’, which requires only two edits. A higher number of edits indicated a greater distance from the original string. The lowest distance value was '0' (observed sequence was identical to the expected sequence). The highest distance value was '4' (all pictures in the wrong position). The distance values were averaged across trials (for means, see Table 4.1).

A Kruskal-Wallis test, conducted to analyze group differences, revealed a significant effect of diagnostic category on sequence accuracy, $H(2) = 14.98$, $p < .001$, $\eta^2 = .35$. The AD and MCI groups had lower sequence accuracy (i.e. higher Levenshtein distance) compared to the OA group, but did not differ significantly from each other. Additionally, we observed a strong correlation between naming and sequence accuracy, $\rho = -.61$, $p < .001$, such that naming accuracy was lower when Levenshtein distance was greater.

Moves

We recorded how many moves participants made on each trial before finalizing their response, and then calculated the average per trial for each participant (see Table 4.1). A Kruskal-Wallis test showed that the number of moves was influenced by diagnostic category, $H(2) = 14.7$, $p < .001$, $\eta^2 = .34$. The OA group made fewer moves compared to the AD group. The MCI group did not differ from either AD or OA groups. Further, the number of moves correlated negatively with naming accuracy, $\rho = -.49$, $p = .001$, but the correlation with sequence accuracy was unclear, $\rho = -.31$, $p = .05$.

Effect of Cueing

Sequence accuracy was examined separately on cued and non-cued trials. The 10% trimmed means for each group on cued and non-cued trials are provided in Table 4.2. A robust 3 (diagnosis: OA vs. MCI vs. AD) x 2 (trial type: cued vs. non-cued) mixed-factor ANOVA revealed a main effect of diagnosis, $F_t(2, 14.3) = 8.93$, $p = .005$. The main effect of trial type and the interaction between diagnosis and trial type were not significant F_ts < 1, ps > .85. Further, sequence accuracy, specifically on non-cued trials, was strongly correlated with naming accuracy, $\rho = -.70$, $p < .001$, with higher naming accuracy associated with a smaller Levenshtein distance.

Table 4.2: Group-wise performance on cued and non-cued trials in Experiment 1 (10% trimmed means and 95% confidence intervals)

		OA		**MCI**		**AD**	
		M	*CI*	*M*	*CI*	*M*	*CI*
Sequence Accuracy (Levenshtein distance)	Cued	0.2	[0.11, 0.3]	0.64	[0.29, 1.05]	0.98	[0.61, 1.38]
	Non-cued	0.25	[0.14, 0.41]	0.59	[0.25, 1.16]	1.0	[0.57, 1.5]
Average moves/trial	Cued	4.1	[4.0, 4.2]	4.3	[4.1, 4.5]	4.3	[4.1, 4.8]
	Non-cued	4.1	[4.0, 4.3]	4.2	[4.1, 4.4]	4.7	[4.4, 5.0]

Effect of cueing was also examined for the average number of moves per trial. The group-wise 10% trimmed means on cued and non-cued trials are provided in Table 4.2. A robust 3 (diagnosis: OA vs. MCI vs. AD) x 2 (trial type: cued vs. non-cued) mixed-factor ANOVA was conducted. Similar to sequence accuracy, we only found a main effect of diagnosis, $F_t(2, 11.3) = 10.8$, $p = .002$; all other effects were not significant, F_ts ≤ 3.21, ps ≥ .07.

4.2.3 Discussion

This experiment demonstrated that AD and MCI populations have difficulty in establishing a temporal and causal order of visually depicted four-event sequences. These findings reflect deficits in the activation and the use of an appropriate macro-event schema which serves integration of sub-events. This was evident in the low sequence accuracy rates displayed by the two groups, as well as in naming the depicted macro-events appropriately. These findings are in line with previous studies that demonstrated deficits experienced in binding related pieces of information in AD patients but not in healthy older adults (Parra et al., 2009; Parra, Abrahams, Logie, & Della Sala, 2010) or non-AD dementias (Della Sala, Parra, Fabi, Luzzi, & Abrahams, 2012). Additionally, MCI and AD groups also displayed more uncertainty in finalizing the sequence, as evidenced by the higher number of moves on average per trial.

The low naming accuracy among the patient groups, may, to some extent, be attributed to anomia experienced during AD. However, as is suggested by a lower sequence accuracy in addition to lower naming accuracy, the difficulties experienced by AD and MCI patients appear to emerge in part by deficits in non-verbal aspects of event cognition. Our findings indicate that individual pictures are not readily recognized as depicting interrelated sub-events, and therefore, do not serve as cues to activate an appropriate macro-event schema, on the basis of which the appropriate temporal sequence could be established and an appropriate name could be retrieved. As the individual sub-events can be considered as segments of what is a continuous flow of information, one explanation for the results may be that subjects are not able to infer what is not depicted, and therefore cannot find the links between sub-events.

Cueing with the event name did not improve sequence accuracy. In the cognitively healthy older adults, this may be because their performance was at ceiling. There may be several

reasons for the lack of an effect in the two patient populations. One, patients might not be able to access sub-events of the macro-event, i.e., they may not be able to access a more fine-grained level of event representation. Two, the issue might be an inability to associate the individual pictures with the schema that was activated by the verbal cue.

Working memory deficits which are commonly attested for in AD and MCI individuals cannot explain the effects of ordering, because the items to be arranged, as well as the cue, were visible throughout the trial and did not need to be held in memory. In sum, the performance of the AD and MCI group points to impaired bottom-up processing – retrieving the macro event schema from the sub-events –, as well as top-down processing – activating sub events from the macro event schema (verbal cue).

4.3 Experiment 2

In Experiment 1, both patient groups were evidently impaired in picture sequencing and subsequent macro-event recognition. Experiment 2 aimed to extend these findings to investigate whether macro-event recognition is impaired in AD and MCI when the additional cognitive load of unscrambling event sequences was reduced. Similar to Experiment 1, participants were presented four pictures depicting sub-events. The pictures were presented in a temporally and causally appropriate sequence, but in a staggered form. Participants were to stop the trial when they thought they could identify the macro-event, and then name it. The goal was the activation of the macro-event being depicted in as few pictures as possible, partly by drawing causal inferences between the micro-events that are available, and partly by predicting the micro-events that would follow.

4.3.1 Method

Participants

The same individuals who participated in Experiment 1, also participated in Experiment 2. In the cognitively healthy group, in addition to the twenty participants from Experiment 1, data from two additional participants, who were excluded from Experiment 1 due to a technical error, was also included. A total of twenty-two cognitively healthy older adults were included in this experiment. Demographic characteristics are reported in Table 4.3. The MCI and AD groups remained the same as in Experiment 1.

Stimuli and Procedure

The stimuli consisted of fourteen sets of events that were not used in Experiment 1. Each set consisted of four pictures, each showing different sub-events of one macro event. Each picture was 300 pixels in width and height. The experiment was programmed using Javascript, and was presented on a 10.1" tablet with an external keyboard. The pictures were presented in the correct sequence, and each trial began with the presentation of the first two pictures in the sequence. After a five-second interval, the third picture was presented. This was followed by another five-second interval before the final picture appeared. Participants were instructed to press 'space' or touch the screen as soon as they were able to identify the macro-event being depicted, to end the trial, and then write down the name of the event. It was emphasized that participants were to name the macro-event, and not the individual micro-events. The trials were presented in a randomized order. Participants were given five practice trials.

We measured the number of pictures seen by the participants before identifying the macro-event, as well as the naming accuracy. Naming accuracy was scored by two assessors

independently, one of whom was blinded to the diagnosis of the participants. Any discrepancies in scoring were resolved with discussion.

4.3.2 Results

Naming Accuracy

The naming responses were scored as '1' or '0', for a correct and incorrect response respectively, and the percentage of correct responses was calculated (for means, see Table 4.3). A Kruskal-Wallis test revealed a significant effect of diagnostic category on naming accuracy, $H(2) = 21$, $p < .001$, $\eta^2 = .49$. Further multiple comparisons indicated significantly lower naming accuracy in the AD and MCI groups compared to the OA group, but no difference was observed between AD and MCI groups.

Table 4.3: Demographic characteristics and performance on measures in Experiment 2 (Means and SD)

Measure	OA (n=22)	MCI (n=10)	AD (n=10)	*p*[a]	*η2*
Age	71 (5.3)	72.6 (6.2)	74.2 (7.5)	.16	.04
Sex (F/M)	17/5	3/7	7/3	.03[b]	.40[c]
Education (years)	11.9 (1.6)	10.4 (1.8)	9.9 (1.7)	.006	.21
MMSE score (max. 30)	28.9 (0.9)	24.7 (1.3)	23 (1.9)	<.001	.75
Naming Accuracy (%)	97.0 (5.3)	70.7 (28.1)	81.4 (8.4)	<.001	.49
Avg. no of pictures viewed per trial	2.24 (0.2)	2.79 (0.5)	2.83 (0.5)	.003[d]	.28

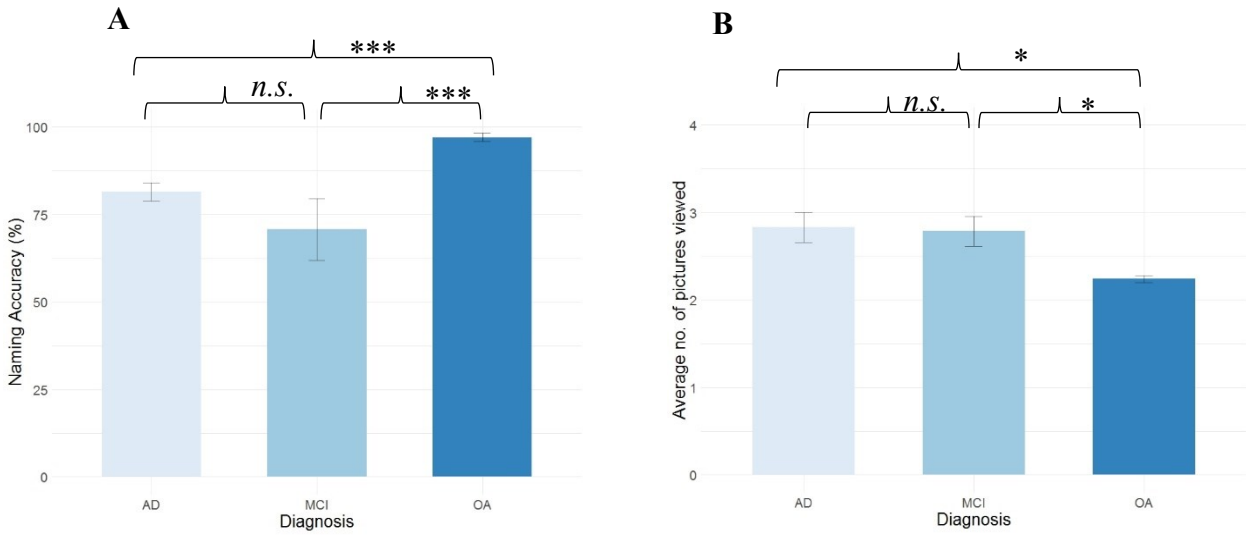
[a]Kruskal-Wallis p-value, [b]Pearson's chi sq. p-value, [c]Cramer's V, [d]Welch's ANOVA

Pictures

We calculated the average number of pictures that each participant viewed per trial before registering their response. The group-wise means are presented in Table 4.3. A Welch's ANOVA revealed a significant effect of diagnostic category on the number of pictures viewed, $F_W(2, 13.2)$

= 9.65, $p = .003$, $\eta^2 = .28$. Games-Howell post-hoc test indicated that AD and MCI groups viewed more pictures on an average compared to the OA group. AD and MCI groups did not differ significantly from each other (see Figure 4.3). Additionally, we observed a significant correlation between naming accuracy and number of pictures viewed, $\rho = -.53$, $p = .003$, indicating higher accuracy when fewer pictures were viewed.

Figure 4.3: Group-wise means on Experiment 2 measures of A) naming accuracy; B) average no. of pictures viewed per trial (error bars represent SE)



*$p \leq .05$, ***$p \leq .001$, n.s.= not significant*

4.3.3 Discussion

This experiment sought to extend findings from Experiment 1 to investigate macro-event recognition, when, in contrast to Experiment 1, recognition was facilitated by sequential sub-events, thereby reducing cognitive load. Naming accuracy was significantly lower in MCI and AD groups compared to the cognitively healthy adults, but the MCI and AD groups did not differ from each other. Naming accuracy was lower in patient groups despite the fact that they, on

average, viewed significantly more pictures per trial than the cognitively healthy group, pointing to an inability to integrate information even when it is presented without distortion. In Experiment 1, there was an added challenge of unscrambling the pictures before participants named the macro-event. Here, even when this hurdle is not present, the patient groups still display difficulty in macro-event identification. That correct event sequences did not facilitate recognition indicates that macro-event recognition may not necessarily be dependent on correct picture sequencing, rather correct sequencing relies on event recognition.

The task also involves some predictive inferencing. In order to identify the macro-event being depicted, while not viewing all sub-events, would require predictive processing. Prediction is implicit to visual perception (Cohn, 2019; Enns & Lleras, 2008), and is based on existing schemata for events (Zacks et al., 2007). It is essential for higher efficiency in information processing, but is impaired in neurodegenerative diseases, such as AD (for a review, see Kocagoncu et al., 2020). Our results correspond to findings from these studies as AD and MCI patients viewed more sub-events in the sequence before naming the event. Despite more information intake, adequate identification for naming of the macro event was impaired for the two groups.

4.4 General Discussion

In two experiments, we measured AD and MCI patients' ability to integrate smaller sub-events into larger macro-events, and to identify depicted macro-events. Five outcomes were measured. A significant deficit in performance was observed on all measures for the AD group in comparison to the cognitively healthy older adults. The AD group named and ordered fewer events correctly, required more moves to reach a decision about the sequence (Experiment 1), and viewed more pictures before naming the event (Experiment 2). For the MCI group, when

compared to the cognitively healthy group, a deficit in performance was observed on four of five measures, whereas their performance was comparable on one measure (moves). When compared to the AD group, the MCI group's performance did not differ significantly on any of the five measures across the two studies. The cognitively healthy older adults performed at ceiling on all measures.

Previous studies investigating event cognition in AD have reported poorer segmentation ability and subsequent memory for events (H. R. Bailey, Zacks, et al., 2013; Zacks et al., 2006), poorer verbal script generation (Grafman et al., 1991; Roll et al., 2019), poorer verbal script sequencing (Allain et al., 2008; Grafman et al., 1991), and poorer picture-sequencing (Roll et al., 2019). To our knowledge, the present study is the first to examine macro-event recognition in AD and MCI populations using a visual narrative paradigm that minimizes memory load. We extend findings from the aforementioned studies to show impairment in event integration, in addition to segmentation and sequencing, not only in AD patients, but in MCI too. Further, previous studies have established that AD patients experience visuospatial disorientation that is not associated with memory deficits, rather with an inability to link and integrate spatial and temporal cues (Duffy et al., 2004). This has also been observed in linguistic narrative comprehension, wherein AD and MCI patients display impairment in macro-level comprehension of textual narratives, which relies largely on inferencing, integration, and abstraction of information (Kokje, Celik, Wahl, & von Stutterheim, 2021). We extend these findings to show impairment in processing visual event scenes in AD and MCI. This indicates that the emerging deficits are not restricted to a particular domain, e.g., linguistic processing, but are likely, at least in part, attributable to a general problem in information organization, be the input verbal or visual, and likely extends to

other types of information processing, such as in spatial cognition (Lithfous, Dufour, & Després, 2013).

One question that arises is, whether this paradigm taps into declarative memory processes or procedural memory. Considering that the experimental paradigm consists of steps involved in different events or activities, one could assume that it relies primarily on procedural memory. However, this does not appear to be the case. While declarative memory is impaired early on in AD (Libon et al., 1998), procedural memory remains more or less intact until the later stages of the disease (De Wit et al., 2020). When we consider that participants do not actually perform any actions and see the events from a third-person point of view, combined with the fact that impairment can be observed in the very early stages of the disease, it is likely that this paradigm taps largely into declarative memory rather than procedural.

One goal of Experiment 1 was to examine whether a verbal cue would help in activating an event schema that can guide the sequencing task, as it may reduce the load of having to search for the appropriate event schema on the basis of the scrambled micro-events alone. Our results, however, did not indicate any improvement following cueing. Previous studies have shown mixed results, and some have suggested cueing as a possible method for intervention (for a review, see El Haj & Kessels, 2013). One possible reason for the lack of a cueing effect may simply be a failure to retrieve and activate the appropriate event schema, despite the cue, due to a degradation of semantic networks (Hodges, Salmon, & Butters, 1992). Another possibility is the inability of AD and MCI patients to appropriately use event schemata to draw bridging inferences, and to connect the scenes on the pictures (Cohn & Kutas, 2015), instead seeing them as discrete events. In the current study, it is likely a combination of both.

In both experiments, the AD group's performance was clearly distinguishable from cognitively healthy older adults. The MCI group's performance was also significantly worse than the cognitively healthy group on all but one measure, and was comparable to the AD group on all measures. This may be due to the ambiguous and heterogeneous nature of MCI, and consequently, the resulting heterogeneity in performance, as is evident in the rather wide range in their performance. The lack of sensitivity in distinguishing performance of MCI and AD groups may result from a combination of this heterogeneity and a relatively small sample size. However, the results of the current study suggest that deficits in event cognition have an early onset in degenerative brain functionality, as the MCI group's performance was more similar to that of the AD group than that of the control group. Therefore, the testing method introduced here may be a valuable addition to the currently available test battery.

The findings in the current study have important clinical implications. A testing paradigm involving indicators of everyday cognition can give us insight into individuals' functional ability. This is evidenced in the study by Roll et al. (2019), wherein a sequencing task assessing everyday task knowledge was directly associated with functional ability. In clinical assessment, functional ability is an important determinant of the degree of disease progression and the diagnosis received. The tools currently in use to measure functional ability are not objective; they rely on self-report or caregiver-report (Lawton & Brody, 1969). More objective tools are warranted, considering the vital role they play in clinical assessments. Additionally, successful event cognition is an important aspect of maintaining functional independence. In older adults, and particularly in those with neurodegenerative diseases such as AD, functional ability is a determining factor in maintaining quality of life and independence. Interventions aimed at

improving/maintaining event cognition ability can go a long way in achieving that goal, particularly when pharmacological treatments have shown limited results.

Limitations and Future Directions

One limitation of our study is the lack of classification of sub-types in the MCI population. This classification was only available for three of ten participants. The availability of classification for the whole group would have possibly further helped interpretation of results. Another limitation is the sample size. While the differences between the cognitively healthy group and patient groups were clearly evident, indicating a high sensitivity to distinguish patients from cognitively healthy adults, this was not the case for the MCI and AD groups. A study with larger samples is required to evaluate whether there is a difference between the two groups. Experiment 2 is limited in its measurement of prediction, as there was no direct measure. Future studies should aim to use a more direct measure of prediction, such as asking participants to elucidate the next steps in the sequence. Finally, the analysis was also limited by the differences in education level between the healthy and AD groups. This is, however, not unexpected, as education level plays a major role in the age of clinical onset of AD, and, as such would make it difficult to find AD and non-AD groups that are matched on age and education. But, seeing as, in our study, differences were evident on event cognition measures between the cognitively healthy and MCI groups too, who did not differ on education level, we can conclude that, while education level accounts for some variance, the effect of diagnostic category on event cognition measures persisted beyond the role of education.

Despite these limitations, our findings indicate a robust method to study event cognition in AD and MCI populations, as evidenced by the large effect sizes. The current study introduces a novel paradigm that makes further research with a wider array of stimuli possible. In this,

different dimensions of complexity can be systematically varied. Further, with the increasing emphasis on characterizing cognitive deficits in preclinical AD population and in prediction of conversion from MCI to AD, we recommend future studies to apply a longitudinal design to study the trajectories of older adults who go on to develop AD. How event cognition ability correlates with biomarker evidence would also be of interest, particularly in the context of tracking how functional ability relates to neurodegeneration. Additionally, future studies should examine in-depth the cause for a lack of effect of cueing. Finally, it would be of interest to see how the paradigm introduced in this study compares to neuropsychological assessments routinely used in clinical practice. Binding deficits appear to be specific to AD dementia. It would be of interest to compare AD and non-AD patients' performance on such a paradigm, and compare sensitivity with neuropsychological assessments currently used. Generally, it has been observed that complex cognitive testing paradigms are more sensitive during the early and preclinical stages of AD compared to routine neuropsychological assessments (Rentz et al., 2013); even more so in certain populations such as those with high cognitive reserve (Stern, 2013).

Chapter 5

Macro-Event Recognition in the Context of Traditional Neuropsychological Tests

Abstract

For several decades, clinicians have continued to use a battery of neuropsychological tests in the diagnosis of AD and MCI, which have been developed several decades prior. These tests assess overall cognitive functioning by isolating the different cognitive and linguistic functions and studying them independently of the others. Such an approach has value in direct inter-individual comparisons of individual cognitive functions, but do not necessarily provide an accurate picture of overall cognitive functioning, particularly in the real world. In the present study, the macro-event recognition paradigm was further evaluated to look at cognitive processes that may be involved in performing the task. A battery of established neuropsychological tests that are routinely used in clinical practice was administered to AD and MCI patients as part of their diagnostic procedure, in addition to the macro-event recognition task. Correlations between the neuropsychological test performance and performance on the macro-event recognition task were examined for the AD and MCI groups separately. Moderate correlations between some measures in the event recognition task and neuropsychological tests examining semantic memory, executive functioning, and episodic memory, among others, were observed. Overall findings suggest that, in AD patients, semantic memory and executive functioning (task switching) contribute to identification of events and event sequencing in Experiment 1, whereas episodic memory appeared to contribute to event identification in Experiment 2. The implications of the moderate strength of correlations and recommendations for future studies are discussed.

5.1 Introduction

Alzheimer's disease is characterized by gradual, progressive decline in cognitive functions and in the daily functional ability. Functional impairment is a commonly observed feature, as well as a critical criterion in diagnosing AD, particularly in distinguishing it from MCI. As the disease progresses, individuals eventually lose the ability to function independently. The root of this functional impairment in AD is primarily cognitive in nature, and can be predicted by level of cognitive function (Liu-Seifert et al., 2015), as opposed to, for example, Parkinson's disease or DLB, where the root of functional impairment is primarily motor dysfunction (Jankovic & Kapadia, 2001; McKeith et al., 2006). Despite this, current neuropsychological measures have only a moderate relationship with measures of functional ability (Chaytor & Schmitter-Edgecombe, 2003).

The novel paradigm introduced in Chapter 4, which measures event recognition and event sequencing shows promising results in distinguishing early stage AD and MCI patients from cognitively healthy older adults, using a task approximating real-world cognition. Although script representation and event segmentation studies have given some insight into the event cognition process in AD, this study addressed two major gaps in the literature: (1) examining event integration ability and subsequent macro-event recognition ability, and (2) inclusion of an MCI group along with AD patients. The promising findings of the study encourage further inquiry into this novel method in the context of established neuropsychological assessment methods. This study examines the measures from the event recognition study, which are detailed in Chapter 4, along with the neuropsychological test battery that formed the neuropsychological component in the diagnostic process of the AD and MCI patients.

The aim of this study was to assess the cognitive processes that are associated with event sequencing and event recognition. Previous studies examining script generation and script sequencing found that AD patients uniquely experience impairment in maintaining temporal order of events, be it a sequence of events occurring in a short time-span (Allain et al., 2008; Grafman et al., 1991) or spanning several decades (Rasmussen & Berntsen, 2022). These deficits were associated with level of impairment in executive function. Similarly, in the macro-event recognition paradigm, the sequencing measure is expected to be associated with executive function. Further, semantic knowledge guides event processing. The Zacks et al. (2006) study, reporting on event segmentation in AD, observed segmentation ability to be associated with semantic knowledge. Similarly, the aforementioned script representation studies also report an association with semantic memory. Therefore, it is expected that event naming and sequencing would be correlated with semantic memory measures.

5.2 Method

5.2.1 Participants

Twenty-two individuals with a probable diagnosis of AD or MCI were recruited from the memory clinics of Heidelberg University Hospital and Central Institute of Mental Health, Mannheim, and tested. Of these, ten participants received a diagnosis of AD (M_{Age} = 74.2, *SD* = 7.5), and 10 were diagnosed with MCI (M_{Age} = 72.6, *SD* = 6.2). Two other participants' data was excluded from the analyses, as, after comprehensive assessment was completed, they received a different diagnosis. Three of the 10 patients with MCI were classified as having amnestic MCI, whereas no classification was made for the others. The participants underwent clinical, neurological and neuropsychological testing, including clinical history, blood tests, CSF testing, CT/MRI, and the CERAD neuropsychological test battery (J. C. Morris et al., 1989). They

received a formal diagnosis from a physician based on the NIA-AA criteria, in case of AD (McKhann et al., 2011), and NIA-AA criteria for MCI (Albert et al., 2011) or the ICD-10 criteria (World Health Organization, 1992) in case of MCI.

5.2.2 Stimuli and procedure

Neuropsychological test battery

As part of the diagnostic procedure, the AD and MCI groups underwent neuropsychological testing. The German version of the CERAD-NB was administered, which consisted of the following tests: (i) Verbal fluency (semantic); (ii) Boston Naming Test (BNT); (iii) MMSE; (iv) Word list learning; (v) Constructional Praxis; and, (vi) Clock drawing test. In addition to the CERAD-NB, the following tests were also administered: (i) TMT-A ; (ii) TMT-B; (iii) Wechsler Memory Scale (WMS)- Logical Memory I&II; (iv) WMS- digit span forward; (v) WMS- digit span backward. These tests were administered to each patient in the clinic, in a single session, by a trained psychologist. An overview of the tests and the cognitive traits they purport to measure are detailed in Table 5.1.

The *Mini-Mental State Examination* (Folstein et al., 1975) is a brief cognitive screening instrument that assesses several cognitive functions, including orientation, attention, and recall, among others. It is used to get an overall indication of cognitive function, and measures possible cognitive impairment and the severity of the impairment. Of a total possible score of 30, a score < 26 is indicative of cognitive impairment.

The semantic component of the *verbal fluency* test entails verbal generation of a list of words belonging to a particular category (in this case, animals), in a given time frame, which is generally 60 seconds (Sebaldt et al., 2014). It assesses semantic memory, and to some extent executive functions.

The *Boston Naming Test* is a test of confrontation naming, which consists of line drawings of objects of varying difficulties, ranked according to their frequency, which are to be named (Kaplan, 1983). The original test consists of 60 such drawings. However, shorter versions of 30-items and 15-items are also available. CERAD-NB generally consists of the 15-item version, which is the version used in this study.

Word List Learning Test (Gaonac'h, 1976) is a test of verbal episodic memory. In this test, a list of 10 words is presented to the participant. Following this, the participant is immediately asked to recall as many of the presented words as possible. Three such learning trials are conducted using the same list of words. After a delay of 10 minutes, participants are asked to recall the list of words presented earlier. The total possible immediate recall score is 30, 10 for each of the 3 trials. For the delayed recall condition, the possible score is 10.

Constructional Praxis is a test evaluating visuospatial ability (Rosen, Mohs, & Davis, 1984). In the test, participants are presented with four geometric figures of increasing complexity. In the first condition, participants are asked to copy the figures i.e. they can view the figures while drawing them. In the second condition, which is the delayed recall condition, participants are to draw the figures previously presented to them from memory. This follows an interval of 2-2.5 minutes. Each of the conditions receives a score ranging from 0 to 11, depending on the number of features that were correctly depicted for each figure.

The *Clock Drawing Test* is used as a general cognitive screening tool, as well as to assess visuo-constructive ability specifically (Agrell & Dehlin, 1998). It involves asking subjects to draw a complete clock, including all the numbers and the clock hands, with the clock hands pointing to a specified time. Participants are not provided a picture, and have to draw the clock from memory. There are several scoring systems for this test. In CERAD-NB, the test is scored

using the Shulman et al. (1993) criteria. The score ranges from 1 to 6, with a score of 1 (perfectly drawn clock) indicating no impairment, a score of 6 (no recognizable clock) indicating severe impairment, and the scores in between indicating different levels of impairment– minor or moderate visuospatial impairment, or severe level of disorganization.

Logical Memory Test I & II is a subtest of the Wechsler Memory Scale (Wechsler, 1945), and are measures of episodic memory. The test entails presentation of two short passages, followed by an immediate free recall (part I) of the story. After a 20-minute delay, in the delayed recall condition (part-II) another free recall of the story follows, along with delayed recognition, which consists of Yes/No questions about the story. The test is scored on the basis of the number of details recalled correctly, and the number of recognition questions answered correctly. The total possible score is 53 and 39, respectively, for Parts I & II.

Digit Span Forward and Backward Tests are also part of the Wechsler Memory Scale (Wechsler, 1945). While the *forward* part of the test evaluates attention and verbal working memory, the *backward* test measures executive functions. It entails presentation of a set of digits which are to be repeated, either in the same order (forward) or in the reverse order (backward). There are several trials, with a new set of digits on each new trial, and the number of digits in a set continue to increase in every subsequent trial. The test ends when subject is unable to correctly repeat the digits in a maximum of two attempts. The test is scored as the number of trials completed successfully, and ranges from 0 to 16.

The *Trail Making Tests A & B* (Reitan, 1958) evaluate attention and executive function, respectively. The A part of the test consists of numbers from 1 to 25 which are presented scattered. Subjects are instructed to connect the numbers sequentially. In Part B, a variation of Part A is used, wherein a combination of numbers (1-10) and letters (A-J) are presented, and

subjects have to alternately connect numbers and letters sequentially (1-A-2-B-3-C, and so on). The test is scored in terms of the time taken to complete each of the tests.

Table 5.1: List of neuropsychological tests and cognitive functions assessed

Neuropsychological Test	Cognitive Functions
MMSE	– Short screening– attention, orientation, language, recall, constructional praxis
Verbal Fluency (semantic)	– Semantic memory – Executive function
BNT	– Confrontation naming
World list learning	– Verbal episodic memory
Constructional praxis	– Visuospatial ability
Logical memory I & II	– Episodic memory
Digit span forward	– Verbal working memory – Attention
Digit span backward	– Executive function
TMT-A	– Processing speed – Visual attention
TMT-B	– Executive function
Clock drawing test	– Cognitive screening – Visuoconstructive ability

Macro-event recognition

Two experiments measuring macro-event recognition were conducted. These were conducted in a separate session within a few days following the neuropsychological testing. The materials and procedure for these experiments are detailed in the methods section in Chapter 3. Five measures, also described and analyzed in Chapter 3, were analyzed together with the

neuropsychological test measure. These were the following: Experiment 1– naming accuracy, sequence accuracy, and average number of moves per trial; Experiment 2– naming accuracy and average number of pictures viewed per trial.

5.3 Results

The data was analyzed using RStudio, version 4.0.4 (R Core Team, 2021). The demographic characteristics of AD and MCI participants are summarized in Table 4.1 of Chapter 4. Each variable from the event recognition paradigm and the neuropsychological test battery was assessed for normality of distribution using the Shapiro-Wilk test. When both variables were normally distributed the Pearson's correlation coefficient was reported. If one of the two variables was not normally distributed, Spearman's rank correlation coefficient (*rho*) was reported.

The means and SDs on each individual neuropsychological test for the AD and MCI groups are reported in Table 5.2. A t-test was performed to analyze group level differences on each neuropsychological test, which is also reported in Table 5.2. Independent samples t-test revealed a marginally significant difference in test scores between the AD and MCI patients on MMSE, $t(17) = -2.21$, $p = .04$, $d = 1.02$. Additionally, a Welch's t-test revealed a significant difference between the two groups on constructional praxis, $t_W(10.1) = -2.71$, $p = .02$, $d = 1.27$. In both cases, the MCI group performed better than the AD group. Further, the relationships between neuropsychological measures employed in the diagnosis of AD and MCI, and the measures of macro-event recognition obtained from the novel paradigm employed in Chapter 4 were examined. Correlations were performed separately for each group.

Table 5.2: AD and MCI groups' performance on neuropsychological tests

Neuropsychological Test	**AD**			**MCI**			**t-test**
	Mean	*SD*	*n*	*Mean*	*SD*	*n*	*p*
Education	9.9	1.7	10	10.5	1.8	10	0.46
MMSE (30)	23	1.9	10	24.7	1.3	9	0.04
Verbal Fluency (animals)	14.2	4.7	9	13.3	3.2	10	0.62
BNT (15)	11.7	2.6	9	13.7	1.6	10	0.05
Word List Learning Immediate Recall (30)	11	5.2	9	10.8	2.6	10	0.91
Word List Delayed Recall (10)	1.22	2.2	9	1.9	1.6	10	0.45
Constructional Praxis (11)	8.67	1.8	9	10.4	0.7	10	0.02
Constructional Praxis Delayed Recall (11)	2.67	3	9	4.7	2.6	10	0.13
WMS: Logical Memory I (53)	18.5	7.8	8	14.5	5.5	10	0.22
WMS: Logical Memory II (39)	4.88	6.8	8	3.3	3.6	10	0.53
WMS: Digit Span Forward (16)	5.7	1.8	10	6.57	1.3	7	0.29
WMS: Digit Span Backward (16)	4.4	1.5	10	5	0.8	7	0.35
TMT-A (secs)	84.4	39.9	9	66.1	49.6	10	0.39
TMT-B (secs)	262.8	126.5	8	233.4	169.1	9	0.69
Clock Drawing Test (6)	3.1	1	10	2.5	1	10	0.19

Associations in the AD group

The correlations between neuropsychological tests and measures from the event recognition paradigm, for the AD group, are summarized in Table 5.3. The most notable correlation observed was between TMT-B and sequence accuracy (r = .81), pointing to fewer sequencing errors when time taken to complete TMT-B was lesser. These are both measures which involve sequencing, with TMT-B purporting to measure executive function ability. Additionally, sequence accuracy was also negatively associated with semantic fluency i.e. there were fewer sequencing errors with a higher fluency score (ρ = -.68), suggesting involvement of semantic component in sequencing. Further, naming accuracy in Experiment 1 was positively correlated with semantic fluency (ρ = .71), as well as with digit span backward test (r = 0.72), again indicating a combination of semantic component and executive function ability in recognizing and naming the macro-events. Overall, naming and sequence accuracy in Experiment 1 appear to tap into common cognitive processes, and the strong correlation observed between these variables (Chapter 4), suggest these are highly integrative actions with interdependent processing.

Interestingly, however, naming accuracy in Experiment 2 was not associated with verbal fluency, but was positively associated with Logical memory II (ρ = .87), suggesting that naming ability in Experiment 1 and in Experiment 2 tapped into distinct cognitive mechanisms. Further, in Experiment 2, the number of pictures viewed was associated negatively with semantic fluency (ρ = -.72), word list learning- delayed recall (ρ = -.66), and with digit span forward test (r = -.74), indicating verbal memory involvement.

Table 5.3: Correlations between neuropsychological measures and event cognition measures in AD patients

Neuropsychological Test	**Naming accuracy (Exp 1)**	**Sequence accuracy**	**Avg. moves/trial**	**Naming accuracy (Exp 2)**	**Avg. pictures viewed**
MMSE	0.48	-0.44	-0.21	0.35	-0.26
Verbal Fluency	0.71*	-0.68*	0.03	0.33	-0.72*
Boston Naming Test	0.53	-0.58	0.05	0.64	-0.33
Word List Learning Immediate	-0.01	-0.13	0.21	-0.02	0.17
Word List Delayed Recall	0.23	-0.27	-0.005	0.34	-0.66*
Constructional Praxis	0.15	-0.06	0.1	-0.05	-0.42
Constructional Praxis Delayed Recall	0.28	-0.44	0.44	0.31	-0.43
WMS: Logical Memory I	0.22	-0.58	-0.33	0.63	-0.21
WMS: Logical Memory II	0.18	-0.17	-0.02	0.87**	-0.37
WMS: Digit Span Forward	0.16	-0.32	0.17	0.11	-0.74**
WMS: Digit Span Backward	0.72*	-0.58	-0.14	-0.23	-0.31
Trail Making Test-A (secs)	-0.52	0.53	-0.16	0.52	0.6
Trail Making Test-B (secs)	-0.65	0.81**	0.63	0.21	0.37
Clock Drawing Test	-0.13	0.04	0.41	-0.11	0.04

****p < .001; ** p <.01; *p < .05*

Associations in the MCI group

A summary of correlations between event recognition paradigm and neuropsychological measures is given in Table 5.4. As with the AD group, there was a significant positive association between the sequence accuracy measure in Experiment 1 and TMT-B (ρ = .76), a measure of executive function. Additionally, sequence accuracy also correlated negatively with word list learning- immediate recall (r = -.70), suggesting involvement of verbal memory. Contrary to the AD group, no association was observed between semantic fluency and naming accuracy in Experiment 1. However, an unexpected negative association was observed between naming accuracy in Experiment 2 and Logical memory II (ρ = -.68), an association which was also observed in AD patients, but in the opposite direction. Additionally, naming in Experiment 2 also correlated positively with digit span forward (ρ = 0.77) and negatively with Clock Drawing Test (ρ = -0.70).

5.4 Discussion

The aim of this study was to identify the cognitive processes implicated in the processing and recognition of events, and assess the value added by novel paradigms, such as the one introduced in the previous chapter, to traditional neuropsychological tests. As was hypothesized, sequencing ability in the macro-event recognition task was strongly correlated, in both groups, with TMT-B, a measure of executive function, specifically task switching. TMT-B and the event sequencing measure, both involve arranging presented stimuli in a particular sequence. The stimuli in the TMT-B is at a very basic level of information– letters and numbers–, whereas, the stimuli in the event recognition paradigm is conceptually more complex, in accounting for temporal, causal and spatial relations, and drawing upon existing event schema. But, while the level of information to be processed in the event sequencing paradigm and TMT-B is different,

Table 5.4: Correlations between neuropsychological measures and event cognition measures in MCI patients

Neuropsychological Test	Naming accuracy (Exp 1)	Sequence accuracy	Avg. moves/trial	Naming accuracy (Exp 2)	Avg. pictures viewed
MMSE	0.54	-0.04	-0.17	0.3	-0.44
Verbal Fluency	-0.02	-0.36	-0.2	0.04	-0.28
Boston Naming Test	0.13	-0.34	0.28	0.3	-0.3
Word List Learning Immediate	-0.08	-0.70*	0.18	0.23	0.1
Word List Delayed Recall	-0.17	-0.44	0.006	0.1	0.5
Constructional Praxis	-0.13	0.07	-0.48	0.2	-0.2
Constructional Praxis Delayed Recall	-0.05	-0.25	-0.31	0.03	0.28
WMS: Logical Memory I	-0.23	-0.09	-0.02	-0.24	-0.05
WMS: Logical Memory II	-0.38	-0.14	-0.25	-0.68*	0.25
WMS: Digit Span Forward	0.69	-0.11	0.07	0.77*	-0.33
WMS: Digit Span Backward	0.09	0.25	0.58	-0.09	-0.29
Trail Making Test-A (secs)	-0.11	0.49	0.38	0.04	0.21
Trail Making Test-B (secs)	-0.59	0.76*	-0.1	-0.48	0.66
Clock Drawing Test	-0.18	-0.08	0.3	-0.70*	0.18

*$p < .05$

there appears to be some commonality in the cognitive mechanisms involved. This is in line with findings from previous script sequencing studies, which also observed a correlation between the sequencing measures and measures of executive function (Allain et al., 2008; Cosentino, Chute, Libon, Moore, & Grossman, 2006; Rasmussen & Berntsen, 2022; Roll et al., 2019). Sequencing errors can be attributed partly to executive dysfunction and partly to semantic memory deficits, as evidenced in the association with semantic fluency. Naming accuracy in Experiment 1 also largely involved semantic memory and executive function in AD patients; and, the correlation observed between sequencing and naming in Chapter 4, indicate they are interdependent, i.e. both processes occur simultaneously, involve common cognitive processes, and successful event sequencing is dependent on successful event identification to draw upon schema, and similarly, successful macro-event recognition is dependent on successful identification of the micro-events and establishment of a connection between them.

Interestingly, naming ability in the two experiments appeared to rely on distinct cognitive mechanisms. While in Experiment 1, semantic memory and executive functions seem to play a role; in Experiment 2, episodic memory appears to be involved. Unexpectedly, this relationship seems to be inverse in the MCI group. A possible reason for this could be the heterogeneity in the MCI group, as the sub-classifications of MCI were unknown for most patients. It is likely that, as a result of heterogeneity, there was a variability in performance among the MCI group across different tests. A large standard deviation was observed for the naming measure in Experiment 2. Additionally, although the difference was not significant, the MCI group's mean on both measures of the Logical Memory test was lower than that of the AD group, which is not anticipated. It is possible that a smaller sample size meant that the results were biased by the performance of a few participants. A combination of the heterogeneity of the group and a bias

due to the sample size is likely to have contributed to the unexpected direction of association. Further, the overall level of cognitive function and attention also contributed to naming ability in Experiment 2.

Finally, in AD patients, the number of pictures viewed was found to be inversely associated with a number of cognitive processes– semantic memory, verbal memory, and attention. It is interesting to note that while the number of pictures viewed was inversely correlated with naming accuracy (Chapter 4), the two measures appear to be associated with distinct cognitive processes, indicating that naming accuracy was higher despite viewing fewer pictures, not because of it. This was perhaps because cognitively healthy older adults were able to activate the event schema instantaneously, whereas AD and MCI patients experienced difficulties in activating the schema even within a longer period of time and with more information available. While episodic memory deficits are widely discussed in the context of AD, the above results along with evidence from previous script representation studies and event segmentation studies indicate a consistently observed deficit in semantic memory in the early stage of AD, and, in this study, in the MCI stage too. The studies by Rasmussen and Bernsten (2022) and Johnson and Smith (1998) show that encoding and retrieval of episodic memory is supported by semantic memory, which is something to be taken into consideration in designing interventions.

It is noteworthy, that, on the neuropsychological tests, the AD and MCI patients' performance differed significantly only on constructional praxis and marginally on MMSE, considering that these tests were involved in the process of conferring diagnoses. Such a difference was also not observed on any measure of event recognition. However, it is notable that MCI patients, who generally do not experience functional impairment or only experience

impairment on complex functional activities, as opposed to AD patients, were not distinguishable from AD patients on the event recognition task, but their performance was distinguishable from the healthy group. Previous studies using script representation paradigms have not included MCI patients, therefore, it is difficult to draw a conclusion about whether the lack of a difference observed between AD and MCI patients is because the paradigm is genuinely not capable of distinguishing between the two groups, or whether factors such as sample size and MCI group heterogeneity contributed to the lack of sensitivity. The study by Rasmussen and Bernsten (2022) examined events on a larger scale, in the form of cultural life scripts in AD patients. While they observed problems in sequence of events, they also observed that events in the life-script remain relatively stable, and closely follow cultural norms. This gives some confidence in using event-based paradigms, but also to be cautious of accounting for cultural differences.

Finally, in the current study, while we observe some correlations between measures from the event recognition paradigm and existing neuropsychological test measures, which give us some insight into the underlying cognitive components in the event recognition paradigm, the event recognition paradigm tells us far more than individual neuropsychological tests do. Chaytor and Schmitter-Edgecombe (2003) highlighted the fairly moderate correlations between traditional neuropsychological tests and measures of everyday functioning. Similarly, in this study, associations observed were moderate, as was also the case with the event segmentation paradigm (Zacks et al., 2006). In the Zacks et al. (2006) study, a linear regression revealed that neuropsychological test performance accounted for < 50% of the variance in measures of event memory in AD patients, indicating that such paradigms account for additional variables.

While individual neuropsychological tests inform us about discrete cognitive processes, they do not give us a unified picture of cognitive function, or an indication of everyday

functional ability. Considering that there is only a moderate correlation between traditional neuropsychological tests and measures of everyday functional ability (Chaytor & Schmitter-Edgecombe, 2003), newer paradigms resembling real-world situations are warranted to predict ability of patients to function in the real world. Therefore, an argument can be made that while traditional neuropsychological tests have their uses, they do not necessarily provide a holistic picture of cognitive functioning, particularly in the context of everyday cognition, where paradigms such as event recognition have added value.

5.4.1 Limitations and future directions

A major limitation in the current study is the small sample size. A linear regression analysis was not possible due to the small sample, and the study was limited to correlation analysis. The correlations, too, need to be interpreted with caution due to the small sample size. This is particularly the case with the MCI group. As sub-type classification for all MCI patients was not available, they group may possibly have been heterogeneous, with rather large variability in the group in a relatively small sample. A replication of the study with a larger sample, and in the MCI group, specifically with patients of the amnestic subtype is warranted to confirm the findings.

Another shortcoming is the lack of availability of IADL scores for the patients. It would be of interest to examine to what extent performance on measures of the macro-event recognition paradigm is associated with IADL, as this is currently the popular measure in clinical settings to assess daily functional ability. Further, it would be interesting to examine performance on the neuropsychological test battery and macro-event recognition paradigm, longitudinally, to track change in trajectory of outcomes on both, and compare these, particularly beginning in the preclinical stage of AD. During this stage, cognitive impairment is reported to be not evident on

routine neuropsychological tests, so, it would be of interest to assess the predictive value of the event recognition paradigm.

Finally, neuropsychological tests evidently have low sensitivity in distinguishing between MCI and early stage AD patients, as well as in predicting conversion from MCI to AD. In the current sample, differences were only observed on two of the neuropsychological test measures, even though these tests scores formed part of the diagnostic process. With the MCI stage being an ambiguous stage, neuropsychological test performance lies between AD and cognitively normal, and is sometimes not distinguishable from the performance of cognitively healthy older adults. The event recognition paradigm shows promise in distinguishing MCI patients from cognitively healthy older adults. These results are also noteworthy in light of the fact that functional impairment is not observed in MCI. A step forward with the event recognition paradigm would be to conduct a study with a larger sample and a homogenous aMCI group, in order to assess its potential to distinguish them from AD patients, and in predicting converters and non-converters in the MCI group.

Chapter 6

General Discussion

6.1 Summary and Integration of Findings

The current work broadly aimed to highlight the paucity of meaningful cognitive measures that are contextually-rich and echoing everyday cognitive and linguistic functioning, in the assessment of AD in the early stages of disease; and, to contribute towards developing such outcomes. Within this larger framework, specifically, this work focused on examining macro-level understanding from micro-level information, in patients in the early stage of AD and MCI, compared to cognitively unimpaired older adults. To this end, macro-level processing of narratives was assessed via a systematic review of discourse comprehension studies (Chapter 3), a novel paradigm was developed to assess event sequencing, integration, and macro-event recognition (Chapter 4), and the relationship between the novel macro-event recognition paradigm and existing neuropsychological tests was examined (Chapter 5).

The results from studies reported in this dissertation overwhelmingly support the idea that, at least part of the difficulty experienced in everyday cognition in AD, results from an inability to integrate related chunks of information into a whole. This finding remained consistent, whether it was examined using textual narratives (Chapter 3) or picture-based event recognition (Chapter 4). Moreover, this impairment was observed not only in AD patients in the very early stage, but also in individuals with MCI. One of the shortcomings of the text comprehension studies was their over-reliance on language production measures to assess macro-level comprehension of narratives, as there is a possibility that performance of AD and MCI patients observed on comprehension measures did not really reflect comprehension deficits, but may have stemmed from impaired language production ability. This limitation was overcome in the event sequencing paradigm. Despite employing a more facilitative paradigm, both patient groups displayed difficulty in the non-verbal task of event sequencing. Further, in the event

sequencing paradigm, a possible facilitative effect of language in the form of a verbal cue, was also examined. Linguistic cueing did not appear to improve event sequencing in the current study. However, this requires further examination, as this was a preliminary investigation, and modified paradigms may reveal different results.

Information is organized hierarchically, i.e. it is organized at multiple levels simultaneously wherein information at a lower level is integrated into information at a higher level. Segments of events can always be broken down further into even smaller segments, or integrated to form larger segments. Impairment observed in macro-level processing, in the patient groups, appears to stem from an inability to establish temporal and causal connections between events within a narrative, and by extension, in linking objects with actions, and objects with other objects within that narrative. Previous studies have observed deficits among AD patients in binding features (objects-colors) in verbal and visual short-term memory (Parra et al., 2009; Parra et al., 2010), which appear to be unique to AD dementia, and are not evident in other types of dementia (Della Sala et al., 2012). These studies reported significant impairment in performance when AD patients were asked to recall multiple features together as bound entities (color + object) compared to when they recalled the same features as independent entities (only color or only object). This was not the case with cognitively healthy older adults, who performed similarly whether recalling bound features or independent features.

While these above studies highlight AD patients' difficulties in binding unrelated information, the evidence from the discourse comprehension and macro-event recognition studies further illustrate problems in binding and integrating related chunks of information too. This is not surprising, considering the medial temporal lobe is largely involved in binding features (Hannula & Ranganath, 2008), and the medial temporal lobe is known to show

considerably higher rates of atrophy in the early stage of AD (McDonald et al., 2009). Further, as a consequence of their inability to relate entities, they appear to be unable to fill in missing information, or draw inferences, that would bridge the given information.

Studies examining connected speech in AD report issues in referential processing, including pronouns, an inability to maintain coherence, and an overall low level of connectedness in speech (Malcorra et al., 2021). Evidence from the macro-event recognition study and the discourse comprehension studies indicates that the root of these impairments is patients' inability to draw inferences between propositions, which is the information that is required to connect the propositions. Early stage AD and MCI patients evidently experience difficulty in inferencing irrespective of whether the medium of information is verbal or non-verbal, as seen from the paradigms highlighted in the present work, indicating that it is not a language-specific deficit. Their problems with referential pronouns, with maintaining coherence in speech, with macro-level features of comprehension and production of discourse, appear to emerge from deficits in cognitive capabilities of the higher-order i.e. in this case, processing of information at multiple hierarchical levels and interacting up and down these levels, as opposed to a simple recall of information. And, higher-order cognition or complex cognitive processing is affected very early on in AD (Bondi et al., 2008; Hansson et al., 2017).

Considering that AD and MCI patients had difficulty integrating parts into a whole across both, discourse comprehension and event recognition paradigms, raises the possibility that this problem may be generalizable to information processing in general in AD. This may, at least in part, account for other deficits commonly observed in AD patients. For example, impairment in spatial navigation, particularly in unfamiliar environments, but also to an extent in familiar environments, is a common feature in the very early and preclinical stages of AD (Allison,

Fagan, Morris, & Head, 2016; Coughlan, Laczó, Hort, Minihane, & Hornberger, 2018). This dysfunction may be attributable to the inability to construct a cognitive map for unfamiliar environments from the fragments of information available (visual feed, landmark cues, directional cues), similar to the inability to construct a mental framework of the macro-event from micro-events and the temporal sequence of these micro-events, in the event recognition task. It is an approach to consider while characterizing functional deficits in everyday life in AD and MCI.

An important implication of the overall findings is the impairment beyond long-term episodic memory. Episodic memory impairment is the most widely highlighted and most widely researched issue in the context of AD, and general decline in functional ability is widely attributed to episodic memory impairment. However, the paradigms presented here were designed so as to minimize memory load. Participants had all relevant information available to them throughout each trial. It, then, follows that the results from the above studies provide strong evidence that functional impairment in AD and MCI cannot be accounted for by episodic memory impairment alone.

Further, the impairment observed in the patients on both paradigms in this work, despite efforts to minimize memory load, combined with evidence from event segmentation studies (H. R. Bailey, Zacks, et al., 2013; Zacks et al., 2006), indicates that memory deficits in AD cannot be attributed solely to degradation of memory networks. The event segmentation studies highlight the idiosyncratic encoding pattern for information in AD patients, and its effect on storage of that information long-term memory. So, it is evident that memory problems in AD begin with issues in encoding information appropriately, and these faulty encoding patterns perhaps further multiply the problems in storage and retrieval.

Finally, it is noteworthy that the measures in both, the macro-event recognition paradigm and the discourse comprehension paradigm, were only moderately associated with existing neuropsychological test measures. This implies that the paradigms highlighted here tap into additional facets of cognition that are not accounted for by traditional testing methods, which is for future studies to examine further. It is clear, however, that objective assessment tools that are rooted in everyday cognition have the potential to contribute uniquely to the assessment process, and should receive more attention.

6.2 Clinical Implications

The major finding in the context of clinical settings is the need for review of cognitive and functional testing outcomes. Both paradigms presented here attempt to closely mirror everyday functioning. MCI patients, by definition, have intact activities of daily living. In the above findings, however, it is evident that there is considerable variation in performance in MCI patients. Moreover, with cognitively more complex tasks such as the discourse comprehension and event recognition tasks presented above, performance of MCI patients was closer to the early stage AD patients than it was to the cognitively unimpaired adults. Previous studies assessing ecological validity of common neuropsychological assessment tools have reported low to moderate predictive ability of these neuropsychological tests (Chaytor & Schmitter-Edgecombe, 2003; Chaytor et al., 2006; Groth-Marnat & Baker, 2003). Perhaps it is time to move beyond the overly reductionist testing methods and embrace cognitively complex paradigms that would be better indicators of the true degree of cognitive decline and functional capabilities.

Further, it is evident from the macro-event recognition study that the MCI group's performance falls between that of the cognitively healthy older adults and mild AD patients, and actually edges closer to AD patients' performance. Similar findings were observed in the

discourse comprehension studies, wherein macrostructural comprehension was significantly impaired in MCI patients when compared to the healthy adults. However, when compared to AD patients, their performance was distinguishable on some measures, but comparable on other measures. Considering these findings, it may be concluded that MCI patients are in fact impaired, albeit to a lesser degree, on everyday functioning; particularly on complex activities, as is also evidenced in previous studies measuring complex activities of daily living (Farias et al., 2006; Perneczky, Pohl, et al., 2006). Event recognition and discourse comprehension paradigms are evidently sensitive to these subtle deficits in MCI. Clinical assessment of MCI should account for these deficits rather than relying simply on informant-reported questionnaires for functional assessment, as previous evidence indicates that routinely used functional assessment measures are not sensitive enough to everyday functional deficits during the MCI stage (Burton, Strauss, Bunce, Hunter, & Hultsch, 2009; Jefferson et al., 2008; Wadley, Okonkwo, Crowe, & Ross-Meadows, 2008).

Finally, more complex, contextually-rich measures of language and cognition, as presented in this work, which include an interplay between multiple cognitive and linguistic functions are not only more naturalistic methods for gauging daily functioning, but also appear to be sensitive in the early stages of disease. Considering that early detection of indications of cognitive impairment is currently key in tackling AD, such methods need to be adopted in clinical practice. Further, the nature of the paradigms presented above make them less susceptible to practice effects, compared to a verbal fluency test or a digit span test, as there is scope for introducing variations in the stimuli presented. Such formats make repeated testing possible, and therefore, lend themselves to broader purposes, such as, intervention change assessments or tracking cognitive decline longitudinally.

6.3 Strengths and Limitations

Strengths

The present work addressed a major gap in research on the holistic processing of events. So far, in the context of AD, studies have looked at spontaneous segmentation of a continuous flow of events into smaller units, or in spontaneous generation of these smaller events. This work addressed AD patients' ability to integrate smaller events to form a macro-representation. Additionally, the study also included MCI patients, which previous studies in the area of event cognition have not included. The findings from the study gave us an indication that, in terms of event cognition, MCI patients ability tends to lie closer to the early stage AD patients on the AD continuum than it does to the cognitively unimpaired individuals.

Another major strength of the work is its focus on contextually-rich paradigms which attempt to echo real-world cognition. A change in approach to neuropsychological testing has been to increase verisimilitude of cognitive assessment tools. Three issues concerned with low verisimilitude were highlighted in Chapter 2– decontextualized testing, isolation of functions, and underestimation of task demands. The novel macro-event recognition paradigm and the text comprehension studies attempt to address these issues. Both, paradigms use a contextually-immersed approach to assessment, and do not attempt to test individual cognitive processes in isolation. Although it is difficult to replicate the task demands of the real-world in a research setting, adding context and not isolating cognitive functions goes a long way towards imitating the task demands. Previous studies have shown that older adults, in particular, respond better to stimuli that are contextually-rich (P. E. Bailey et al., 2010; Zimerman et al., 2011). In considering the *gestalt* view, the whole is greater than the sum of its parts. Therefore, paradigms assessing cognition holistically may be more informative than a piecemeal approach with several tests

assessing individual cognitive functions independently of the others. The large effect sizes observed in the macro-event recognition study in distinguishing healthy older adults from AD patients demonstrate that it is possible to combine everyday cognition with objective measures.

In light of the evolving role of neuropsychological testing, highlighted in Chapter 2– from diagnostic purposes to purpose of assessing change in cognition and function over time-, the cognitive testing paradigms highlighted here have great value in terms of repeated testing. Such paradigms can incorporate a large variety of stimuli. This, along with their more complex nature make them resistant to practice effects, and therefore, more feasible for repeated testing.

Limitations

One of the general limitations is the use of MMSE. This has been discussed in previous chapters in the context of individual studies. However, this is a shortcoming that is not limited to this work, but to the field in general. MMSE is still the primarily used cognitive screening tool in clinical settings, largely because of a reluctance on clinicians' part to change familiar practices (Rabin et al., 2007). MMSE also forms part of some of the commonly employed neuropsychological test batteries, such as CERAD-NB, which is very commonly used in clinical settings in countries like Germany and USA. It is for this reason that it was not possible to use the MoCA or ACE in the present work. The MMSE is not as sensitive as MoCA or ACE-III in detecting cognitive impairment during the MCI stage (Pinto et al., 2018; Roalf et al., 2013; Senda et al., 2020), or generally in distinguishing MCI from early AD or from healthy ageing (Mitchell, 2009), as it was developed before the construct of MCI was developed.

Another general limitation was the sample size. Dementia patients are generally hard to recruit for a number of reasons– dependency on caregiver for mobility, inability to give informed consent, frequent occurrences of co-morbidities due to age which result in exclusion, and a

reluctance to participate due to a lack of existing treatment options or benefits for patients, among others. With further restrictions in criteria, such as in this study, which was restricted not only to one type of dementia, but also to a specific stage of disease, it becomes an increasingly challenging task. In this study, the sample size was adequate in distinguishing AD patients from healthy older adults, and, to an extent, MCI patients could be distinguished from the healthy group too. However, with a population like MCI, which occupies a rather broad range on the AD continuum, a larger sample may be required to detect differences. A similar pattern is observed in the text comprehension studies, which have similar sample sizes as the event recognition study. This problem was further compounded when assessing neuropsychological test data. It is difficult for AD patients to complete a whole battery of tests, and some may find certain individual tests difficult to complete due to problems with attention. Therefore, all AD and MCI patients did not complete the full battery of tests, resulting in missing values for several of the tests.

An additional limitation, which is again not restricted to this work, is the lack of classification of MCI into sub-types. The classification of MCI sub-type has become increasingly common when diagnosing MCI for research-specific purposes, but it is not so common in clinical settings yet, except in specialized memory clinics, due to a lack of knowledge and training on the subject. As a result, a classification of aMCI was only available for a few of the participants in the MCI group, with no classification available for the others, making it unclear whether the group was homogenous in that respect or not. A heterogeneous MCI group would mean more variations in the cognitive and linguistic deficits in individuals, and as a consequence, more variation in performance on the task.

6.4 Future Directions

A general direction for future research using novel paradigms aimed at understanding everyday functioning would be to conduct longitudinal studies, and track the trajectory of decline in functioning. It would also be of interest to contrast this with existing traditional neuropsychological measures. Additionally, as previously discussed, there is conflicting evidence on the ability of biomarkers to predict cognitive and functional decline, as clinical expression of AD is mediated by a number of factors associated with cognitive reserve. It would be of interest to longitudinally track association of such everyday cognition oriented measures with biomarkers. With the increasing emphasis on intervention at the earliest possible time point, the preclinical stage of AD i.e. the stage with detectable biomarker abnormalities, but no clinical indication, is a ripe stage for employing paradigms highlighted in this work. This is because, such paradigms require more cognitively complex processing than do straightforward tests of attention and memory. It is a possibility that subtle impairments are evident on more complex paradigms, while performance on traditional neuropsychological tests still remains within the normal range.

Another avenue for future research is to use immersive paradigms such as Virtual Reality (VR). There have already been a few studies in recent years which have used VR for assessment and for cognitive intervention and training purposes. VR-based cognitive training paradigms, so far, have not reported a particularly high success rate in terms of immediate improvement in cognitive function, as is also the case with existing pharmacological and other non-pharmacological treatment options. However, studies employing VR-based paradigms report high rates of engagement and enjoyment in patients (Clay et al., 2020), which contributes towards higher motivation and lower drop-out rates from the studies. In studying real-world

functioning with event cognition-based paradigms, immersive VR technology would be a good method for assessment and training using an egocentric reference frame. Allocentric frame of reference is known to deteriorate in AD, and also in healthy ageing, as opposed to egocentric frame. Using immersive VR may be a way to overcome the limitations that allocentric reference frame-based paradigms are subject to. A modification of the event recognition paradigm described in Chapter 4 to include immersive VR may be a possibility for developing the paradigm further.

Further, VR-based paradigms would be advantageous in terms of possibly offering greater ecological validity, due to the possibility of simulating naturalistic, real-world environments using immersive VR. In recent years, several studies using immersive and non-immersive VR have shown great promise as cognitive assessment tools with greater ecological validity (Kourtesis, Collina, Doumas, & MacPherson, 2020; Parsons, Carlew, Magtoto, & Stonecipher, 2017; Tarnanas et al., 2013). These studies investigated episodic memory, ADLs, executive functioning, and reported better predictive ability for everyday functioning, addressing the issue of veridicality that is a challenge with current tests. With VR's ability to create virtual environments, designing virtual environments that incorporate the conditions and demands of the real-world is a promising step toward resolving the issue of verisimilitude too.

Future studies should also aim to assess the ecological validity of the macro-event recognition paradigm. However, current tools to assess ecological validity are lacking, as they rely largely on subjective measures such as, self- or caregiver-report, or clinician-ratings, which is something to consider for future work. Finally, considering the promise that discourse comprehension studies and the event recognition study show, it would be of interest to replicate

these paradigms in different cultures and different languages, as well as to develop them further in the direction of standardization.

6.5 Conclusion

The present work demonstrated the viability of employing contextually-rich, objective, experimental paradigms in the assessment of everyday linguistic and cognitive deficits in the course of AD; and, the utility of such methods in distinguishing early stage AD and MCI from the effects of cognitive ageing. By means of two different paradigms– discourse comprehension and macro-event recognition– , this work focused on macro-level cognitive processing. Discourse comprehension, at a macro-level, was found to be impaired in early stage AD and in MCI patients, across all studies examined in the systematic review. A novel paradigm assessing event cognition was designed, and similarly, impairment in processing events at a macro-level in early stage AD and in MCI was evident. Even with a relatively smaller sample, the findings were robust, and reliable differences were observed between cognitively healthy older adults and older adults with early stage AD or MCI. Further, the findings also hinted that the reach of these paradigms extends to facets of cognition beyond those captured by traditional neuropsychological tests, though this requires further confirmation.

Two findings are underscored here. One, the presence of impairments in macro-level event cognition and discourse processing at the very early AD stage as well as during MCI highlights that these tasks are sensitive in the very early stages of the disease, and are specific enough to distinguish pathological ageing from cognitive ageing, which is currently a critical need in the management of AD. This also gives further impetus for examining the possibility and degree of impairment in macro-level comprehension in the preclinical stage of AD. Two, the above studies are further evidence that measurement of everyday cognition can be reconciled

with objective assessment. This facilitates the use of cognitive assessment tools for more than just diagnostic purposes, in extending their utility to tracking meaningful change in cognition over time and in assessing effects of interventions. Overall, the above findings highlight that novel assessment techniques involving everyday cognition can play a complementary role to the currently available roster of biomarkers and cognitive assessment tools.

References

Abbott, V., Black, J. B., & Smith, E. E. (1985). The representation of scripts in memory. *Journal of Memory and Language, 24*(2), 179-199. doi:https://doi.org/10.1016/0749-596X(85)90023-3

Aggarwal, A., & Kean, E. (2010). Comparison of the Folstein Mini Mental State Examination (MMSE) to the Montreal Cognitive Assessment (MoCA) as a cognitive screening tool in an inpatient rehabilitation setting. *Neuroscience and Medicine, 1*(2), 39.

Agrell, B., & Dehlin, O. (1998). The clock-drawing test. *Age and ageing, 27*(3), 399-404.

Ahmed, S., Haigh, A.-M. F., de Jager, C. A., & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain, 136*(12), 3727-3737. doi:10.1093/brain/awt269

Aizenstein, H. J., Nebes, R. D., Saxton, J. A., Price, J. C., Mathis, C. A., Tsopelas, N. D., . . . Klunk, W. E. (2008). Frequent Amyloid Deposition Without Significant Cognitive Impairment Among the Elderly. *Archives of Neurology, 65*(11), 1509-1517. doi:10.1001/archneur.65.11.1509

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

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: development, cognitive functions, phenomenology, and neurobiology. *Psychological bulletin, 141*(5), 931.

Allain, P., Le Gall, D., Foucher, C., Etcharry-Bouyx, F., Barré, J., Dubas, F., & Berrut, G. (2008). Script representation in patients with Alzheimer's disease. *Cortex, 44*(3), 294-304.

Allaire, J. C., & Marsiske, M. (1999). Everyday cognition: age and intellectual ability correlates. *Psychology and aging, 14*(4), 627-644. doi:10.1037//0882-7974.14.4.627

Allison, S. L., Fagan, A. M., Morris, J. C., & Head, D. (2016). Spatial Navigation in Preclinical Alzheimer's Disease. *Journal of Alzheimer's Disease, 52*, 77-90. doi:10.3233/JAD-150855

Almor, A., MacDonald, M. C., Kempler, D., Andersen, E. S., & Tyler, L. K. (2001). Comprehension of long distance number agreement in probable Alzheimer's disease. *Language and Cognitive Processes, 16*(1), 35-63.

Aluisio, S., Cunha, A., Toledo, C., & Scarton, C. (2016). *Computational tool for automated language production analysis aimed at dementia diagnosis.* Paper presented at the International Conference on Computational Processing of the Portuguese Language, Demonstration Session.

Alzheimer's Disease International. (2019). *World Alzheimer Report 2019: Attitudes to dementia*. Retrieved from London, England:

Amanzio, M., Geminiani, G., Leotta, D., & Cappa, S. (2008). Metaphor comprehension in Alzheimer's disease: Novelty matters. *Brain and Language, 107*(1), 1-10. doi:https://doi.org/10.1016/j.bandl.2007.08.003

Amieva, H., Le Goff, M., Millet, X., Orgogozo, J. M., Pérès, K., Barberger‐Gateau, P., . . . Dartigues, J. F. (2008). Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society, 64*(5), 492-498.

Asgari, M., Kaye, J., & Dodge, H. (2017). Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimer's & Dementia: Translational Research & Clinical Interventions, 3*(2), 219-228. doi:https://doi.org/10.1016/j.trci.2017.01.006

Bachurin, S. O., Gavrilova, S. I., Samsonova, A., Barreto, G. E., & Aliev, G. (2018). Mild cognitive impairment due to Alzheimer disease: Contemporary approaches to diagnostics and pharmacological intervention. *Pharmacological Research, 129*, 216-226. doi:https://doi.org/10.1016/j.phrs.2017.11.021

Bäckman, L., Jones, S., Berger, A.-K., Laukka, E. J., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. *Neuropsychology, 19*(4), 520.

Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia, 51*(11), 2294-2304.

Bailey, H. R., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Head, D., Kurby, C. A., & Sargent, J. Q. (2013). Medial temporal lobe volume predicts elders' everyday memory. *Psychological science, 24*(7), 1113-1122.

Bailey, P. E., Henry, J. D., Rendell, P. G., Phillips, L. H., & Kliegel, M. (2010). Dismantling the "age–prospective memory paradox": The classic laboratory paradigm simulated in a naturalistic setting. *Quarterly Journal of Experimental Psychology, 63*(4), 646-652. doi:10.1080/17470210903521797

Ballatore, C., Lee, V. M.-Y., & Trojanowski, J. Q. (2007). Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nature reviews neuroscience, 8*(9), 663-672.

Bateman, R. J., Xiong, C., Benzinger, T. L., Fagan, A. M., Goate, A., Fox, N. C., . . . Blazey, T. M. (2012). Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N Engl J Med, 367*, 795-804.

Baudic, S., Barba, G. D., Thibaudet, M. C., Smagghe, A., Remy, P., & Traykov, L. (2006). Executive function deficits in early Alzheimer's disease and their relations with episodic memory. *Archives of Clinical Neuropsychology, 21*(1), 15-21.

Bazzari, F. H., Abdallah, D. M., & El-Abhar, H. S. (2019). Pharmacological interventions to attenuate Alzheimer’s disease progression: the story so far. *Current Alzheimer research, 16*(3), 261-277.

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*(4), 458.

Belleville, S., Gauthier, S., Lepage, É., Kergoat, M.-J., & Gilbert, B. (2014). Predicting decline in mild cognitive impairment: A prospective cognitive study. *Neuropsychology, 28*(4), 643.

Bielak, A. A. M., Hatt, C. R., & Diehl, M. (2017). Cognitive Performance in Adults’ Daily Lives: Is There a Lab-Life Gap? *Research in Human Development, 14*(3), 219-233. doi:10.1080/15427609.2017.1340050

Bittner, D., Frankenberg, C., & Schröder, J. (2022). Changes in Pronoun Use a Decade before Clinical Diagnosis of Alzheimer’s Dementia—Linguistic Contexts Suggest Problems in Perspective-Taking. *Brain Sciences, 12*(1), 121.

Blacker, D., Lee, H., Muzikansky, A., Martin, E. C., Tanzi, R., McArdle, J. J., . . . Albert, M. (2007). Neuropsychological Measures in Normal Individuals That Predict Subsequent Cognitive Decline. *Archives of Neurology, 64*(6), 862-871. doi:10.1001/archneur.64.6.862

Blennow, K., & Hampel, H. (2003). CSF markers for incipient Alzheimer's disease. *The Lancet Neurology, 2*(10), 605-613. doi:https://doi.org/10.1016/S1474-4422(03)00530-1

Bolló-Gasol, S., Piñol-Ripoll, G., Cejudo-Bolivar, J. C., Llorente-Vizcaino, A., & Peraita-Adrados, H. (2014). Ecological assessment of mild cognitive impairment and Alzheimer disease using the Rivermead Behavioural Memory Test. *Neurología (English Edition), 29*(6), 339-345. doi:https://doi.org/10.1016/j.nrleng.2013.07.001

Bondi, M. W., Jak, A. J., Delano-Wood, L., Jacobson, M. W., Delis, D. C., & Salmon, D. P. (2008). Neuropsychological Contributions to the Early Identification of Alzheimer's Disease. *Neuropsychology review, 18*(1), 73-90. doi:10.1007/s11065-008-9054-1

Bondi, M. W., Monsch, A. U., Galasko, D., Butters, N., Salmon, D. P., & Delis, D. C. (1994). Preclinical cognitive markers of dementia of the Alzheimer type. *Neuropsychology, 8*(3), 374.

Bondi, M. W., & Smith, G. E. (2014). Mild cognitive impairment: a concept and diagnostic entity in need of input from neuropsychology. *Journal of the International Neuropsychological Society : JINS, 20*(2), 129-134. doi:10.1017/S1355617714000010

Boon, B. D. C., Hoozemans, J. J. M., Lopuhaä, B., Eigenhuis, K. N., Scheltens, P., Kamphorst, W., . . . Bouwman, F. H. (2018). Neuroinflammation is increased in the parietal cortex of atypical Alzheimer's disease. *Journal of Neuroinflammation, 15*(1), 170. doi:10.1186/s12974-018-1180-y

Boots, E. A., Schultz, S. A., Almeida, R. P., Oh, J. M., Koscik, R. L., Dowling, M. N., . . . Bendlin, B. B. (2015). Occupational complexity and cognitive reserve in a middle-aged cohort at risk for Alzheimer's disease. *Archives of Clinical Neuropsychology, 30*(7), 634-642.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *Int J Geriatr Psychiatry, 15*(11), 1021-1027. doi:10.1002/1099-1166(200011)15:11<1021::aid-gps234>3.0.co;2-6

Bouras, C., Hof, P. R., Giannakopoulos, P., Michel, J.-P., & Morrison, J. H. (1994). Regional distribution of neurofibrillary tangles and senile plaques in the cerebral cortex of elderly patients: a quantitative evaluation of a one-year autopsy population from a geriatric hospital. *Cerebral cortex, 4*(2), 138-150.

Bouwens, S. F., Van Heugten, C. M., Aalten, P., Wolfs, C. A., Baarends, E. M., Van Menxel, D. A., & Verhey, F. R. (2008). Relationship between measures of dementia severity and observation of daily life functioning as measured with the Assessment of Motor and Process Skills (AMPS). *Dementia and geriatric cognitive disorders, 25*(1), 81-87. doi:https://doi.org/10.1159/000111694

Brandão, L., Lima, T. M., Parente, M. A. d. M. P., & Peña-Casanova, J. (2013). Discourse coherence and its relation with cognition in Alzheimer's disease. *Revista Psicologia em Pesquisa, 7*(1).

Brier, M. R., Gordon, B., Friedrichsen, K., McCarthy, J., Stern, A., Christensen, J., . . . Ances, B. M. (2016). Tau and Aβ imaging, CSF measures, and cognition in

Alzheimer’s disease. *Science Translational Medicine, 8*(338), 338ra366-338ra366. doi:doi:10.1126/scitranslmed.aaf2362

Brodaty, H., Pond, D., Kemp, N. M., Luscombe, G., Harding, L., Berman, K., & Huppert, F. A. (2002). The GPCOG: a new screening test for dementia designed for general practice. *J Am Geriatr Soc, 50*(3), 530-534. doi:10.1046/j.1532-5415.2002.50122.x

Brookmeyer, R., Abdalla, N., Kawas, C. H., & Corrada, M. M. (2018). Forecasting the prevalence of preclinical and clinical Alzheimer's disease in the United States. *Alzheimer's & Dementia, 14*(2), 121-129.

Brookshire, R. H., & Nicholas, L. E. (1993). *Discourse comprehension test*: Communication Skill Builders.

Bruscoli, M., & Lovestone, S. (2004). Is MCI really just early dementia? A systematic review of conversion studies. *International Psychogeriatrics, 16*(2), 129-140.

Burton, C. L., Strauss, E., Bunce, D., Hunter, M. A., & Hultsch, D. F. (2009). Functional abilities in older adults with mild cognitive impairment. *Gerontology, 55*(5), 570-581. doi:10.1159/000228918

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory, 9*(4-6), 365-381. doi:10.1080/09658210143000083

Cardoso, S., Silva, D., Maroco, J., de Mendonça, A., & Guerreiro, M. (2014). Non-literal language deficits in mild cognitive impairment. *Psychogeriatrics, 14*(4), 222-228. doi:https://doi.org/10.1111/psyg.12101

Castellani, R. J., Rolston, R. K., & Smith, M. A. (2010). Alzheimer disease. *Disease-a-month : DM, 56*(9), 484-546. doi:10.1016/j.disamonth.2010.06.001

Celsis, P. (2000). Age-related cognitive decline, mild cognitive impairment or preclinical Alzheimer's disease? *Annals of Medicine, 32*(1), 6-14. doi:10.3109/07853890008995904

Chan, A. S., Butters, N., & Salmon, D. P. (1997). The deterioration of semantic networks in patients with Alzheimer's disease: A cross-sectional study. *Neuropsychologia, 35*(3), 241-248. doi:https://doi.org/10.1016/S0028-3932(96)00067-X

Chapman, S. B., Anand, R., Sparks, G., & Cullum, C. M. (2006). Gist distinctions in healthy cognitive aging versus mild Alzheimer's disease. *Brain Impairment, 7*(3), 223-233.

Chapman, S. B., Highley, A. P., & Thompson, J. L. (1998). Discourse in fluent aphasia and Alzheimer's disease: Linguistic and pragmatic considerations. *Journal of Neurolinguistics, 11*(1-2), 55-78.

Chapman, S. B., & Mudar, R. A. (2014). Enhancement of cognitive and neural functions through complex reasoning training: evidence from normal and clinical populations. *Frontiers in systems neuroscience, 8*, 69.

Chapman, S. B., Ulatowska, H. K., Franklin, L. R., Shobe, A. E., Thompson, J. L., & McIntire, D. (1997). Proverb interpretation in fluent aphasia and Alzheimer's disease: implications beyond abstract thinking. *Aphasiology, 11*(4-5), 337-350.

Chapman, S. B., Zientz, J., Weiner, M., Rosenberg, R., Frawley, W., & Burns, M. H. (2002). Discourse changes in early Alzheimer disease, mild cognitive impairment, and normal aging. *Alzheimer Disease & Associated Disorders, 16*(3), 177-186.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology review, 13*(4), 181-197.

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology, 21*(3), 217-227. doi:10.1016/j.acn.2005.12.002

Chen, P., Ratcliff, G., Belle, S. H., Cauley, J. A., DeKosky, S. T., & Ganguli, M. (2001). Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study. *Archives of general psychiatry, 58*(9), 853-858.

Chen, Y., Denny, K. G., Harvey, D., Farias, S. T., Mungas, D., DeCarli, C., & Beckett, L. (2017). Progression from normal cognition to mild cognitive impairment in a diverse clinic-based and community-based elderly cohort. *Alzheimer's & Dementia, 13*(4), 399-405.

Chen, Z., & Zhong, C. (2014). Oxidative stress in Alzheimer's disease. *Neuroscience bulletin, 30*(2), 271-281.

Chenery, H. J., & Murdoch, B. E. (1994). The production of narrative discourse in response to animations in persons with dementia of the Alzheimer's type: Preliminary findings. *Aphasiology, 8*(2), 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= Canadian Journal of Speech-Language Pathology and Audiology, 40*(1), 51-65.

Cicognola, C., Brinkmalm, G., Wahlgren, J., Portelius, E., Gobom, J., Cullen, N. C., . . . Höglund, K. (2019). Novel tau fragments in cerebrospinal fluid: relation to tangle pathology and cognitive decline in Alzheimer's disease. *Acta neuropathologica, 137*(2), 279-296. doi:10.1007/s00401-018-1948-2

Clark, L. J., Gatz, M., Zheng, L., Chen, Y.-L., McCleary, C., & Mack, W. J. (2009). Longitudinal Verbal Fluency in Normal Aging, Preclinical, and Prevalent Alzheimer's Disease. *American Journal of Alzheimer's Disease & Other Dementias®, 24*(6), 461-468. doi:10.1177/1533317509345154

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*.

Clay, F., Howett, D., FitzGerald, J., Fletcher, P., Chan, D., & Price, A. (2020). Use of Immersive Virtual Reality in the Assessment and Treatment of Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's Disease, 75*, 23-43. doi:10.3233/JAD-191218

Cohen, G. (1979). Language comprehension in old age. *Cognitive psychology, 11*(4), 412-429.

Cohn, N. (2019). Being explicit about the implicit: inference generating techniques in visual narrative. *Language and Cognition, 11*(1), 66-97.

Cohn, N., & Kutas, M. (2015). Getting a cue before getting a clue: Event-related potentials to inference in visual narrative comprehension. *Neuropsychologia, 77*, 267-278.

Cooper, R. P. (2021). Action production and event perception as routine sequential behaviors. *Topics in cognitive science, 13*(1), 63-78.

Cosentino, S., Chute, D., Libon, D., Moore, P., & Grossman, M. (2006). How does the brain support script comprehension? A study of executive processes and semantic knowledge in dementia. *Neuropsychology, 20*(3), 307-318. doi:10.1037/0894-4105.20.3.307

Coughlan, G., Laczó, J., Hort, J., Minihane, A.-M., & Hornberger, M. (2018). Spatial navigation deficits — overlooked cognitive marker for preclinical Alzheimer disease? *Nature Reviews Neurology, 14*(8), 496-506. doi:10.1038/s41582-018-0031-x

Coughlan, G., Zhukovsky, P., Voineskos, A., & Grady, C. (2021). A profile of brain reserve in adults at genetic risk of Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 13*(1), e12208.

Creamer, S., & Schmitter-Edgecombe, M. (2010). Narrative comprehension in Alzheimer's disease: Assessing inferences and memory operations with a think-aloud procedure. *Neuropsychology, 24*(3), 279.

Croot, K., Hodges, J. R., & Patterson, K. (1999). Evidence for impaired sentence comprehension in early Alzheimer's disease. *Journal of the International Neuropsychological Society, 5*(5), 393-404.

Crous-Bou, M., Minguillón, C., Gramunt, N., & Molinuevo, J. L. (2017). Alzheimer's disease prevention: from risk factors to early intervention. *Alzheimer's research & therapy, 9*(1), 71. doi:10.1186/s13195-017-0297-z

Cummings, J. L. (2000). Cognitive and behavioral heterogeneity in Alzheimer's disease: seeking the neurobiological basis. *Neurobiology of aging, 21*(6), 845-861. doi:https://doi.org/10.1016/S0197-4580(00)00183-4

Cummings, J. L., Darkins, A., Mendez, M., Hill, M. A., & Benson, D. (1988). Alzheimer's disease and Parkinson's disease: Comparison of speech and language alterations. *Neurology, 38*(5), 680-680.

Dalton, S. G., Hubbard, H. I., & Richardson, J. D. (2019). *Moving toward non-transcription based discourse analysis in stable and progressive aphasia.* Paper presented at the Semin Speech Lang.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic bulletin & review, 3*(4), 422-433.

Das, N., Spence, J. S., Aslan, S., Vanneste, S., Mudar, R., Rackley, A., . . . Chapman, S. B. (2019). Cognitive training and transcranial direct current stimulation in mild cognitive impairment: A randomized pilot trial. *Frontiers in neuroscience, 13*, 307.

de Bruijn, R. F. A. G., & Ikram, M. A. (2014). Cardiovascular risk factors and future risk of Alzheimer's disease. *BMC Medicine, 12*(1), 130. doi:10.1186/s12916-014-0130-5

de Lira, J. O., Ortiz, K. Z., Campanha, A. C., Bertolucci, P. H. F., & Minett, T. S. C. (2010). Microlinguistic aspects of the oral narrative in patients with Alzheimer's disease. *International Psychogeriatrics, 23*(3), 404-412. doi:10.1017/S1041610210001092

De Wit, L., Marsiske, M., O'Shea, D., Kessels, R. P., Kurasz, A. M., DeFeis, B., . . . Smith, G. E. (2020). Procedural learning in individuals with amnestic Mild Cognitive Impairment and Alzheimer's dementia: A systematic review and meta-analysis. *Neuropsychology review*, 1-12.

DeCarli, C. (2003). Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment. *The Lancet Neurology, 2*(1), 15-21. doi:https://doi.org/10.1016/S1474-4422(03)00262-X

Degen, C., Frankenberg, C., Toro, P., & Schröder, J. (2022). Differences in Cognitive Functioning in Two Birth Cohorts Born 20 Years Apart: Data from the Interdisciplinary Longitudinal Study of Ageing. *Brain Sciences, 12*(2), 271.

Dekhtyar, S., Wang, H.-X., Scott, K., Goodman, A., Koupil, I., & Herlitz, A. (2015). A Life-Course Study of Cognitive Reserve in Dementia—From Childhood to Old Age. *The American Journal of Geriatric Psychiatry, 23*(9), 885-896. doi:https://doi.org/10.1016/j.jagp.2015.02.002

DeKosky, S. (2003). Early intervention is key to successful management of Alzheimer disease. *Alzheimer Disease & Associated Disorders, 17*, S99-S104.

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2019). Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology, 32*(1).

Della Sala, S., Parra, M. A., Fabi, K., Luzzi, S., & Abrahams, S. (2012). Short-term memory binding is impaired in AD but not in non-AD dementias. *Neuropsychologia, 50*(5), 833-840.

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

Devanand, D. P., Liu, X., Tabert, M. H., Pradhaban, G., Cuasay, K., Bell, K., . . . Pelton, G. H. (2008). Combining Early Markers Strongly Predicts Conversion from Mild Cognitive Impairment to Alzheimer's Disease. *Biological Psychiatry, 64*(10), 871-879. doi:https://doi.org/10.1016/j.biopsych.2008.06.020

Dicker, D., Nguyen, G., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., . . . Murray, C. J. L. (2018). Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet, 392*(10159), 1684-1735. doi:https://doi.org/10.1016/S0140-6736(18)31891-9

Dickerson, B. C., McGinnis, S. M., Xia, C., Price, B. H., Atri, A., Murray, M. E., . . . Wolk, D. A. (2017). Approach to atypical Alzheimer's disease and case studies of the major subtypes. *CNS Spectrums, 22*(6), 439-449. doi:10.1017/S109285291600047X

Diehl, M., Wahl, H.-W., & Freund, A. (2017). Ecological Validity as a Key Feature of External Validity in Research on Human Development. *Research in Human Development, 14*(3), 177-181. doi:10.1080/15427609.2017.1340053

Diniz, B. S., Butters, M. A., Albert, S. M., Dew, M. A., & Reynolds, C. F. (2013). Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *The British Journal of Psychiatry, 202*(5), 329-335.

Doble, S. E., Fisk, J. D., MacPherson, K. M., Fisher, A. G., & Rockwood, K. (2005). Measuring Functional Competence in Older Persons With Alzheimer's Disease. *International Psychogeriatrics, 9*(1), 25-38. doi:10.1017/S1041610297004171

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 (Albany NY), 11*(20), 9188-9208. doi:10.18632/aging.102391

Dubois, B. (2018). The Emergence of a New Conceptual Framework for Alzheimer’s Disease. *Journal of Alzheimer's Disease, 62*, 1059-1066. doi:10.3233/JAD-170536

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., . . . Bateman, R. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *The Lancet Neurology, 13*(6), 614-629.

Duffy, J., Cushman, L., & Kavcic, V. (2004). Visuospatial disorientation in Alzheimer's disease: Impaired spatiotemporal integration in visual information processing. *INTERDISCIPLINARY TOPICS IN GERONTOLOGY, 34*, 155-172.

Dumurgier, J., Hanseeuw, B. J., Hatling, F. B., Judge, K. A., Schultz, A. P., Chhatwal, J. P., . . . Hyman, B. T. (2017). Alzheimer’s disease biomarkers and future decline in cognitive normal older adults. *Journal of Alzheimer's Disease, 60*(4), 1451-1459.

Durazzo, T. C., Mattsson, N., Weiner, M. W., & Initiative, A. s. D. N. (2014). Smoking and increased Alzheimer's disease risk: a review of potential mechanisms. *Alzheimer's & Dementia, 10*, S122-S145.

El Haj, M., & Kessels, R. P. (2013). Context memory in Alzheimer's disease. *Dementia and geriatric cognitive disorders extra, 3*(1), 342-350.

Emery, V. O. B. (2000). Language Impairment in Dementia of the Alzheimer Type: A Hierarchical Decline? *The International Journal of Psychiatry in Medicine, 30*(2), 145-164. doi:10.2190/x09p-n7au-ucha-vw08

Enns, J. T., & Lleras, A. (2008). What's next? New evidence for prediction in human vision. *Trends in cognitive sciences, 12*(9), 327-333.

Fagan, A. M., Roe, C. M., Xiong, C., Mintun, M. A., Morris, J. C., & Holtzman, D. M. (2007). Cerebrospinal fluid tau/β-amyloid42 ratio as a prediction of cognitive decline in nondemented older adults. *Archives of Neurology, 64*(3), 343-349.

Falgàs, N., Sánchez-Valle, R., Bargalló, N., Balasa, M., Fernández-Villullas, G., Bosch, B., . . . Lladó, A. (2019). Hippocampal atrophy has limited usefulness as a diagnostic biomarker on the early onset Alzheimer's disease patients: A comparison between visual and quantitative assessment. *NeuroImage: Clinical, 23*, 101927. doi:https://doi.org/10.1016/j.nicl.2019.101927

Farias, S. T., Mungas, D., Reed, B. R., Harvey, D., Cahn-Weiner, D., & Decarli, C. (2006). MCI is associated with deficits in everyday functioning. *Alzheimer disease and associated disorders, 20*(4), 217-223. doi:10.1097/01.wad.0000213849.51495.d9

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., . . . van Duijn, C. M. (1997). Effects of Age, Sex, and Ethnicity on the Association Between

Apolipoprotein E Genotype and Alzheimer Disease: A Meta-analysis. *JAMA, 278*(16), 1349-1356. doi:10.1001/jama.1997.03550160069041

Fleisher, A. S., Sowell, B. B., Taylor, C., Gamst, A. C., Petersen, R. C., & Thal, L. J. (2007). Clinical predictors of progression to Alzheimer disease in amnestic mild cognitive impairment. *Neurology, 68*(19), 1588. doi:10.1212/01.wnl.0000258542.58725.4c

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: Recognition memory for three types of textual information. *Discourse processes, 13*(2), 175-190.

Flores, S., Bailey, H. R., Eisenberg, M. L., & Zacks, J. M. (2017). Event segmentation improves event memory up to one month later. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(8), 1183.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychol Bull, 95*(1), 29. doi:https://doi.org/10.1037/0033-2909.95.1.29

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychol Bull, 101*(2), 171. doi:10.1037/0033-2909.101.2.171

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.

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: revista de atualizacao cientifica, 20*(4), 285-291.

Forbes-McKay, K., Shanks, M. F., & Venneri, A. (2013). Profiling spontaneous speech decline in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatrica, 25*(6), 320-327. doi:10.1017/neu.2013.16

Fox, N. C., Warrington, E. K., Seiffer, A. L., Agnew, S. K., & Rossor, M. N. (1998). Presymptomatic cognitive deficits in individuals at risk of familial Alzheimer's disease. A longitudinal prospective study. *Brain, 121*(9), 1631-1639. doi:10.1093/brain/121.9.1631

Franceschi, M., Caffarra, P., De Vreese, L., Pelati, O., Pradelli, S., Savarè, R., . . . Grossi, E. (2007). Visuospatial Planning and Problem Solving in Alzheimer's Disease Patients: A Study with the Tower of London Test. *Dementia and geriatric cognitive disorders, 24*(6), 424-428. doi:10.1159/000109827

Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. *Ecological validity of neuropsychological testing*, 91-112.

Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuorpsychiatyric applications of CANTAB. *International journal of geriatric psychiatry, 11*(4), 329-336. doi:10.1002/(SICI)1099-1166(199604)11:4<329::AID-GPS453>3.0.CO;2-6

Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., . . . Winblad, B. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *The Lancet Neurology, 16*(8), 661-676. doi:https://doi.org/10.1016/S1474-4422(17)30159-X

Galton, C. J., Patterson, K., Xuereb, J. H., & Hodges, J. R. (2000). Atypical and typical presentations of Alzheimer's disease: a clinical, neuropsychological, neuroimaging and pathological study of 13 cases. *Brain, 123*(3), 484-498. doi:10.1093/brain/123.3.484

Gaonac'h, D. (1976). Word-list learning and short-term memory. *Acta Psychologica, 40*(1), 9-20. doi:https://doi.org/10.1016/0001-6918(76)90039-1

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

Gerwien, J., & von Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition, 180*, 225-237.

Giannakopoulos, P., Gold, G., Duc, M., Michel, J.-P., Hof, P. R., & Bouras, C. (2000). Neural substrates of spatial and temporal disorientation in Alzheimer's disease. *Acta neuropathologica, 100*(2), 189-195.

Giebel, C. M., Sutcliffe, C., Stolt, M., Karlsson, S., Renom-Guiteras, A., Soto, M., . . . Challis, D. (2014). Deterioration of basic activities of daily living and their impact on quality of life across different cognitive stages of dementia: a European study. *International Psychogeriatrics, 26*(8), 1283-1293. doi:10.1017/S1041610214000775

Goldberg, T. E., Koppel, J., Keehlisen, L., Christen, E., Dreses-Werringloer, U., Conejero-Goldberg, C., . . . Davies, P. (2010). Performance-based measures of everyday function in mild cognitive impairment. *American Journal of Psychiatry, 167*(7), 845-853.

Gomar, J. J., Bobes-Bascaran, M. T., Conejero-Goldberg, C., Davies, P., Goldberg, T. E., & Alzheimer's Disease Neuroimaging Initiative, f. t. (2011). Utility of Combinations of Biomarkers, Cognitive Markers, and Risk Factors to Predict Conversion From Mild Cognitive Impairment to Alzheimer Disease in Patients in the Alzheimer's Disease Neuroimaging Initiative. *Archives of General Psychiatry, 68*(9), 961-969. doi:10.1001/archgenpsychiatry.2011.96

Goral, M., Spiro III, A., Albert, M. L., Obler, L. K., & Connor, L. T. (2007). Change in lexical retrieval skills in adulthood. *The Mental Lexicon, 2*(2), 215-238. doi:https://doi.org/10.1075/ml.2.2.05gor

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual review of psychology, 48*(1), 163-189.

Graff-Radford, J., Yong, K. X. X., Apostolova, L. G., Bouwman, F. H., Carrillo, M., Dickerson, B. C., . . . Murray, M. E. (2021). New insights into atypical Alzheimer's disease in the era of biomarkers. *The Lancet Neurology, 20*(3), 222-234. doi:https://doi.org/10.1016/S1474-4422(20)30440-3

Grafman, J., Thompson, K., Weingartner, H., Martinez, R., Lawlor, B. A., & Sunderland, T. (1991). Script generation as an indicator of knowledge representation in patients with Alzheimer's disease. *Brain and Language, 40*(3), 344-358. doi:https://doi.org/10.1016/0093-934X(91)90134-M

Graville, D. J., & Rau, M. T. (1991). Reading Comprehension of Directly Stated and Inferred Information in Paragraph-Length Material by Nondemented and Demented Elderly Subjects. *Clinical Aphasiology, 19*, 77-86.

Grober, E., & Bang, S. (1995). Sentence comprehension in Alzheimer's disease. *Developmental Neuropsychology, 11*(1), 95-107.

Grober, E., Hall, C. B., Lipton, R. B., Zonderman, A. B., Resnick, S. M., & Kawas, C. (2008). Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society, 14*(2), 266-278. doi:10.1017/S1355617708080302

Grossman, M., & White-Devine, T. (1998). Sentence Comprehension in Alzheimer's Disease. *Brain and Language, 62*(2), 186-201. doi:https://doi.org/10.1006/brln.1997.1898

Groth-Marnat, G., & Baker, S. (2003). Digit Span as a Measure of Everyday Attention: A Study of Ecological Validity. *Perceptual and Motor Skills, 97*(3_suppl), 1209-1218. doi:10.2466/pms.2003.97.3f.1209

Haddad, H. W., Malone, G. W., Comardelle, N. J., Degueure, A. E., Kaye, A. M., & Kaye, A. D. (2022). Aducanumab, a Novel Anti-Amyloid Monoclonal Antibody, for the Treatment of Alzheimer's Disease: A Comprehensive Review. *Health Psychology Research, 10*(1), 31925.

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. *Journal of Alzheimer's Disease, 22*(3), 971-980.

Han, S. D., Nguyen, C. P., Stricker, N. H., & Nation, D. A. (2017). Detectable neuropsychological differences in early preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology review, 27*(4), 305-325.

Hannula, D. E., & Ranganath, C. (2008). Medial Temporal Lobe Activity Predicts Successful Relational Memory Binding. *The Journal of Neuroscience, 28*(1), 116. doi:10.1523/JNEUROSCI.3086-07.2008

Hansson, O., Grothe, M. J., Strandberg, T. O., Ohlsson, T., Hägerström, D., Jögi, J., . . . Schöll, M. (2017). Tau Pathology Distribution in Alzheimer's disease Corresponds Differentially to Cognition-Relevant Functional Brain Networks. *Frontiers in Neuroscience, 11*. doi:10.3389/fnins.2017.00167

Hardy, J. A., & Higgins, G. A. (1992). Alzheimer's disease: the amyloid cascade hypothesis. *Science, 256*(5054), 184-186.

Hassenstab, J., Monsell, S. E., Mock, C., Roe, C. M., Cairns, N. J., Morris, J. C., & Kukull, W. (2015). Neuropsychological markers of cognitive decline in persons with Alzheimer disease neuropathology. *Journal of Neuropathology & Experimental Neurology, 74*(11), 1086-1092.

Hedden, T., Oh, H., Younger, A. P., & Patel, T. A. (2013). Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology, 80*(14), 1341. doi:10.1212/WNL.0b013e31828ab35d

Henning, P. H. (2004). Everyday cognition and situated learning. *Handbook of research on educational communications and technology: A project of the association for educational communications and technology*, 829-861.

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:https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychol Aging, 19*(1), 27-39. doi:10.1037/0882-7974.19.1.27

Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? *Neuropsychologia, 30*(4), 301-314. doi:https://doi.org/10.1016/0028-3932(92)90104-T

Hoefeijzers, S., Calia, C., & Parra, M. A. (2016). Is it time to change the way we detect Alzheimer's disease and monitor its progression? Towards affordable and theory-driven

approaches from cognitive neurosciences. *JSM Alzheimer's Disease and Related Dementia, 3*(2).

Hoenig, M. C., Bischof, G. N., Hammes, J., Faber, J., Fliessbach, K., van Eimeren, T., & Drzezga, A. (2017). Tau pathology and cognitive reserve in Alzheimer's disease. *Neurobiology of aging, 57*, 1-7.

Howieson, D. B., Dame, A., Camicioli, R., Sexton, G., Payami, H., & Kaye, J. A. (1997). Cognitive markers preceding Alzheimer's dementia in the healthy oldest old. *Journal of the American Geriatrics Society, 45*(5), 584-589.

Hsieh, S., Schubert, S., Hoon, C., Mioshi, E., & Hodges, J. R. (2013). Validation of the Addenbrooke's Cognitive Examination III in Frontotemporal Dementia and Alzheimer's Disease. *Dementia and geriatric cognitive disorders, 36*(3-4), 242-250. doi:10.1159/000351671

Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British journal of psychiatry, 140*(6), 566-572.

Huntley, J., & Howard, R. (2010). Working memory in early Alzheimer's disease: a neuropsychological review. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences, 25*(2), 121-132.

Imtiaz, B., Tolppanen, A.-M., Kivipelto, M., & Soininen, H. (2014). Future directions in Alzheimer's disease from risk factors to prevention. *Biochemical pharmacology, 88*(4), 661-670.

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Karlawish, J. (2018). NIA‐AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia, 14*(4), 535-562.

Jack, C. R., 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. *The Lancet Neurology, 12*(2), 207-216. doi:https://doi.org/10.1016/S1474-4422(12)70291-0

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., . . . Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology, 9*(1), 119-128. doi:10.1016/S1474-4422(09)70299-6

Jankovic, J., & Kapadia, A. S. (2001). Functional Decline in Parkinson Disease. *Archives of Neurology, 58*(10), 1611-1615. doi:10.1001/archneur.58.10.1611

Jefferson, A. L., Byerly, L. K., Vanderhill, S., Lambe, S., Wong, S., Ozonoff, A., & Karlawish, J. H. (2008). Characterization of activities of daily living in individuals with mild cognitive impairment. *Am J Geriatr Psychiatry, 16*(5), 375-383. doi:10.1097/JGP.0b013e318162f197

Johnson, D. L., & Smith, S. D. (1998). Effects of Familiarity and Temporal Organization on Memory for Event Schemas in Aged and Alzheimer Subjects: Implications for Clinical Management. *Alzheimer Disease & Associated Disorders, 12*(1), 18-25.

Joubert, S., Brambati, S. M., Ansado, J., Barbeau, E. J., Felician, O., Didic, M., . . . Kergoat, M.-J. (2010). The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia, 48*(4), 978-988. doi:https://doi.org/10.1016/j.neuropsychologia.2009.11.019

Judge, D., Roberts, J., Khandker, R. K., Ambegaonkar, B., & Black, C. M. (2019). Physician Practice Patterns Associated with Diagnostic Evaluation of Patients with Suspected Mild

Cognitive Impairment and Alzheimer’s Disease. *International Journal of Alzheimer’s Disease, 2019*, 4942562. doi:10.1155/2019/4942562

Junquera, A., García-Zamora, E., Olazarán, J., Parra, M. A., & Fernández-Guinea, S. (2020). Role of Executive Functions in the Conversion from Mild Cognitive Impairment to Dementia. *Journal of Alzheimer's Disease, 77*, 641-653. doi:10.3233/JAD-200586

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review, 99*(1), 122.

Kaplan, E. G. H. W. S. G. H. (1983). *Boston naming test*. Philadelphia: Lea & Febiger.

Kavé, G., & Dassa, A. (2018). Severity of Alzheimer’s disease and language features in picture descriptions. *Aphasiology, 32*(1), 27-40. doi:10.1080/02687038.2017.1303441

Kavé, G., & Goral, M. (2018). Word retrieval in connected speech in Alzheimer’s disease: a review with meta-analyses. *Aphasiology, 32*(1), 4-26.

Kavé, G., & Levy, Y. (2003). Morphology in picture descriptions provided by persons with Alzheimer's disease. *Journal of Speech, Language, and Hearing Research*.

Kemper, S., LaBarge, E., Ferraro, F. R., Cheung, H., Cheung, H., & Storandt, M. (1993). On the Preservation of Syntax in Alzheimer's Disease: Evidence From Written Sentences. *Archives of Neurology, 50*(1), 81-86. doi:10.1001/archneur.1993.00540010075021

Kempler, D., Almor, A., Tyler, L. K., Andersen, E. S., & MacDonald, M. C. (1998). Sentence comprehension deficits in Alzheimer's disease: A comparison of off-line vs. on-line sentence processing. *Brain and Language, 64*(3), 297-316.

Kempler, D., Curtiss, S., & Jackson, C. (1987). Syntactic Preservation in Alzheimer's Disease. *Journal of Speech, Language, and Hearing Research, 30*(3), 343-350. doi:10.1044/jshr.3003.343

Kempler, D., & Goral, M. (2008). Language and dementia: Neuropsychological aspects. *Annual review of applied linguistics, 28*, 73.

Khoury, R., & Ghossoub, E. (2019). Diagnostic biomarkers of Alzheimer's disease: A state-of-the-art review. *Biomarkers in Neuropsychiatry, 1*, 100005. doi:https://doi.org/10.1016/j.bionps.2019.100005

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review, 95*(2), 163.

Kintsch, W., & Rawson, K. A. (2005). Comprehension *The Science of Reading: A Handbook* (pp. 211-226).

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review, 85*(5), 363.

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 research international, 2015*.

Kmet, L. M., Cook, L. S., & Lee, R. C. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields.

Knopman, D. S., Jones, D. T., & Greicius, M. D. (2021). Failure to demonstrate efficacy of aducanumab: An analysis of the EMERGE and ENGAGE trials as reported by Biogen, December 2019. *Alzheimer's & Dementia, 17*(4), 696-701. doi:https://doi.org/10.1002/alz.12213

Kocagoncu, E., Klimovich-Gray, A., Hughes, L. E., & Rowe, J. B. (2020). Evidence and implications of abnormal predictive coding in dementia. *arXiv preprint arXiv:2006.06311*.

Koepsell, T. D., & Monsell, S. E. (2012). Reversion from mild cognitive impairment to normal or near-normal cognition. *Risk factors and prognosis, 79*(15), 1591-1598. doi:10.1212/WNL.0b013e31826e26b7

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*, 1-16.

Koning, I. d., Kooten, F. v., Dippel, D. W. J., Harskamp, F. v., Grobbee, D. E., Kluft, C., & Koudstaal, P. J. (1998). The CAMCOG: A Useful Screening Instrument for Dementia in Stroke Patients. *Stroke, 29*(10), 2080-2086. doi:doi:10.1161/01.STR.29.10.2080

Kourtesis, P., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2020). Validation of the Virtual Reality Everyday Assessment Lab (VR-EAL): An Immersive Virtual Reality Neuropsychological Battery with Enhanced Ecological Validity. *Journal of the International Neuropsychological Society, 27*(2), 181-196. doi:10.1017/S1355617720000764

Kuperberg, G. R. (2021). Tea with milk? A hierarchical generative framework of sequential event comprehension. *Topics in cognitive science, 13*(1), 256-298.

Kurby, C. A., & Zacks, J. M. (2018). Preserved neural event segmentation in healthy older adults. *Psychology and Aging, 33*(2), 232.

Laine, M., Laakso, M., Vuorinen, E., & Rinne, J. (1998). Coherence and informativeness of discourse in two dementia types. *Journal of Neurolinguistics, 11*(1-2), 79-87.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *The gerontologist, 9*(3_Part_1), 179-186.

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. *Literary and Linguistic Computing, 26*(4), 435-461. doi:10.1093/llc/fqr013

Levine, T. F., Allison, S. L., Stojanovic, M., Fagan, A. M., Morris, J. C., & Head, D. (2020). Spatial navigation ability predicts progression of dementia symptomatology. *Alzheimer's & Dementia, 16*(3), 491-500. doi:https://doi.org/10.1002/alz.12031

Li, Z., Shue, F., Zhao, N., Shinohara, M., & Bu, G. (2020). APOE2: protective mechanism and therapeutic implications for Alzheimer's disease. *Molecular neurodegeneration, 15*(1), 63. doi:10.1186/s13024-020-00413-4

Libon, D. J., Bogdanoff, B., Cloud, B. S., Skalina, S., Giovannetti, T., Gitlin, H. L., & Bonavita, J. (1998). Declarative and procedural learning, quantitative measures of the hippocampus, and subcortical white alterations in Alzheimer's disease and ischaemic vascular dementia. *Journal of Clinical and Experimental Neuropsychology, 20*(1), 30-41.

Lim, Y. Y., Baker, J. E., Bruns, L., Mills, A., Fowler, C., Fripp, J., . . . Maruff, P. (2020). Association of deficits in short-term learning and Aβ and hippoampal volume in cognitively normal adults. *Neurology*.

Lithfous, S., Dufour, A., & Després, O. (2013). Spatial navigation in normal aging and the prodromal stage of Alzheimer's disease: Insights from imaging and behavioral studies. *Ageing Research Reviews, 12*(1), 201-213. doi:https://doi.org/10.1016/j.arr.2012.04.007

Liu-Seifert, H., Siemers, E., Price, K., Han, B., Selzler, K. J., Henley, D., . . . Alzheimer's Disease Neuroimaging, I. (2015). Cognitive Impairment Precedes and Predicts Functional Impairment in Mild Alzheimer's Disease. *Journal of Alzheimer's disease : JAD, 47*(1), 205-214. doi:10.3233/JAD-142508

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Cohen-Mansfield, J. (2017). Dementia prevention, intervention, and care. *The Lancet, 390*(10113), 2673-2734.

Loewenstein, D., & Acevedo, A. (2010). The relationship between instrumental activities of daily living and neuropsychological performance. *Neuropsychology of everyday functioning*, 93-112.

Loewenstein, D., Curiel, R., Duara, R., & Buschke, H. (2017). Novel Cognitive Paradigms for the Detection of Memory Impairment in Preclinical Alzheimer's Disease. *Assessment, 25*(3), 348-359. doi:10.1177/1073191117691608

Logie, R. H., Parra, M. A., & Della Sala, S. (2015). From Cognitive Science to Dementia Assessment. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 81-91. doi:10.1177/2372732215601370

Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and Cognitive Functioning Across the Life Span. *Psychological Science in the Public Interest, 21*(1), 6-41. doi:10.1177/1529100620920576

Lyons, K., Kemper, S., Labarge, E., Ferraro, F. R., Balota, D., & Storandt, M. (1994). Oral language and alzheimer's disease: A reduction in syntactic complexity. *Aging, Neuropsychology, and Cognition, 1*(4), 271-281. doi:10.1080/13825589408256581

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., & Andersen, E. S. (2001). Assessing working memory and language comprehension in Alzheimer's disease. *Brain and language, 78*(1), 17-42.

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes, 22*(3), 199-224.

Mair, P., & Wilcox, R. (2019). Robust statistical methods in R using the WRS2 package. *Behavior research methods*, 1-25.

Maki, Y., Yamaguchi, T., Koeda, T., & Yamaguchi, H. (2013). Communicative competence in Alzheimer's disease: metaphor and sarcasm comprehension. *American Journal of Alzheimer's Disease & Other Dementias®, 28*(1), 69-74.

Malcorra, B. L. C., Mota, N. B., Weissheimer, J., Schilling, L. P., Wilson, M. A., & Hübner, L. C. (2021). Low Speech Connectedness in Alzheimer's Disease is Associated with Poorer Semantic Memory Performance. *Journal of Alzheimer's Disease, 82*, 905-912. doi:10.3233/JAD-210134

Marcos, G., Santabárbara, J., Lopez‐Anton, R., De‐la‐Cámara, C., Gracia‐García, P., Lobo, E., . . . Stephan, B. (2016). Conversion to dementia in mild cognitive impairment diagnosed with DSM‐5 criteria and with Petersen's criteria. *Acta Psychiatrica Scandinavica, 133*(5), 378-385.

Mårdh, S., Nägga, K., & Samuelsson, S. (2013). A longitudinal study of semantic memory impairment in patients with Alzheimer's disease. *Cortex, 49*(2), 528-533.

Markwick, A., Zamboni, G., & de Jager, C. A. (2012). Profiles of cognitive subtest impairment in the Montreal Cognitive Assessment (MoCA) in a research cohort with normal Mini-Mental State Examination (MMSE) scores. *Journal of clinical and experimental neuropsychology, 34*(7), 750-757. doi:10.1080/13803395.2012.672966

Martorelli, M., Sudo, F. K., & Charchat-Fichman, H. (2019). This is not only about memory: A systematic review on neuropsychological heterogeneity in Alzheimer's disease. *Psychology & Neuroscience, 12*(2), 271.

Mazzon, G., Ajcevic, 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

McDonald, C. R., McEvoy, L. K., Gharapetian, L., Fennema-Notestine, C., Hagler, D. J., Holland, D., . . . Dale, A. M. (2009). Regional rates of neocortical atrophy from normal aging to early Alzheimer disease. *Neurology, 73*(6), 457. doi:10.1212/WNL.0b013e3181b16431

McKeith, I. G., Rowan, E., Askew, K., Naidu, A., Allan, L., Barnett, N., . . . O'Brien, J. T. (2006). More severe functional impairment in dementia with lewy bodies than Alzheimer disease is related to extrapyramidal motor dysfunction. *Am J Geriatr Psychiatry, 14*(7), 582-588. doi:10.1097/01.JGP.0000216177.08010.f4

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.

Michaud, T. L., Su, D., Siahpush, M., & Murman, D. L. (2017). The Risk of Incident Mild Cognitive Impairment and Progression to Dementia Considering Mild Cognitive

Impairment Subtypes. *Dementia and geriatric cognitive disorders extra, 7*(1), 15-29. doi:10.1159/000452486

Mirra, S. S., Heyman, A., McKeel, D., Sumi, S. M., Crain, B. J., Brownlee, L. M., . . . participating, C. n. (1991). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). *Neurology, 41*(4), 479. doi:10.1212/WNL.41.4.479

Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research, 43*(4), 411-431. doi:https://doi.org/10.1016/j.jpsychires.2008.04.014

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J.-c. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta psychologica, 115*(2-3), 123-142.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J., van Belle, G., Fillenbaum, G., . . . Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., . . . Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39*(9), 1159-1165. doi:10.1212/WNL.39.9.1159

Mudar, R. A., Nguyen, L. T., Eroh, J., Chiang, H.-S., Rackley, A., & Chapman, S. B. (2019). Event-related neural oscillation changes following reasoning training in individuals with mild cognitive impairment. *Brain research, 1704*, 229-240.

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

Munukka, M., Koivunen, K., von Bonsdorff, M., Sipilä, S., Portegijs, E., Ruoppila, I., & Rantanen, T. (2021). Birth cohort differences in cognitive performance in 75- and 80-year-olds: a comparison of two cohorts over 28 years. *Aging Clinical and Experimental Research, 33*(1), 57-65. doi:10.1007/s40520-020-01702-0

Murray, M. E., Graff-Radford, N. R., Ross, O. A., Petersen, R. C., Duara, R., & Dickson, D. W. (2011). Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *The Lancet Neurology, 10*(9), 785-796.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . 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. doi:https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nation, D. A., Ho, J. K., Dutt, S., Han, S. D., Lai, M. H., & Initiative, A. s. D. N. (2019). Neuropsychological decline improves prediction of dementia beyond Alzheimer's disease biomarker and mild cognitive impairment diagnoses. *Journal of Alzheimer's Disease, 69*(4), 1171-1182.

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 Review of Neurotherapeutics, 18*(11), 859-869. doi:10.1080/14737175.2018.1532792

Ngandu, T., Lehtisalo, J., Solomon, A., Levälahti, E., Ahtiluoto, S., Antikainen, R., . . . Kivipelto, M. (2015). A 2 year multidomain intervention of diet, exercise, cognitive training, and

vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *The Lancet, 385*(9984), 2255-2263. doi:https://doi.org/10.1016/S0140-6736(15)60461-5

Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K., & Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. *The Lancet Neurology, 13*(8), 788-794. doi:https://doi.org/10.1016/S1474-4422(14)70136-X

Ossenkoppele, R., Schonhaut, D. R., Schöll, M., Lockhart, S. N., Ayakta, N., Baker, S. L., . . . Rabinovici, G. D. (2016). Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. *Brain, 139*(5), 1551-1567. doi:10.1093/brain/aww027

Pakhomov, S., Chacon, D., Wicklund, M., & Gundel, J. (2011). Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing. *Behavior Research Methods, 43*(1), 136-144. doi:10.3758/s13428-010-0037-9

Palmqvist, S., Zetterberg, H., Blennow, K., Vestberg, S., Andreasson, U., Brooks, D. J., . . . Minthon, L. (2014). Accuracy of brain amyloid detection in clinical practice using cerebrospinal fluid β-amyloid 42: a cross-validation study against amyloid positron emission tomography. *JAMA neurology, 71*(10), 1282-1289.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual review of psychology, 60*, 173-196.

Parra, M. A., Abrahams, S., Fabi, K., Logie, R., Luzzi, S., & Sala, S. D. (2009). Short-term memory binding deficits in Alzheimer's disease. *Brain, 132*(4), 1057-1066.

Parra, M. A., Abrahams, S., Logie, R. H., & Della Sala, S. (2010). Visual short-term memory binding in Alzheimer's disease and depression. *Journal of neurology, 257*(7), 1160-1169.

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation, 27*(5), 777-807. doi:10.1080/09602011.2015.1109524

Patterson, C. (2018). World Alzheimer Report 2018: The State of the Art of Dementia Research: New Frontiers. *Alzheimer's Disease International (ADI): London, UK*.

Paxton, J. L., Peavy, G. M., Jenkins, C., Rice, V. A., Heindel, W. C., & Salmon, D. P. (2007). Deterioration of visual-perceptual organization ability in Alzheimer's disease. *Cortex, 43*(7), 967-975.

Perneczky, R., Pohl, C., Sorg, C., Hartmann, J., Tosic, N., Grimmer, T., . . . Kurz, A. (2006). Impairment of activities of daily living requiring memory or complex reasoning as part of the MCI syndrome. *Int J Geriatr Psychiatry, 21*(2), 158-162. doi:10.1002/gps.1444

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. *The American Journal of Geriatric Psychiatry, 14*(2), 139-144. doi:https://doi.org/10.1097/01.JGP.0000192478.82189.a8

Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease: A critical review. *Brain, 122*(3), 383-404.

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., . . . Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology, 58*(12), 1985-1992.

Petersen, R. C., & Negash, S. (2014). Mild Cognitive Impairment: An Overview. *CNS Spectrums, 13*(1), 45-53. doi:10.1017/S1092852900016151

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology, 56*(3), 303-308.

Piazza-Gardner, A. K., Gaffud, T. J. B., & Barry, A. E. (2013). The impact of alcohol on Alzheimer's disease: A systematic review. *Aging & Mental Health, 17*(2), 133-146. doi:10.1080/13607863.2012.742488

Pini, L., Pievani, M., Bocchetta, M., Altomare, D., Bosco, P., Cavedo, E., . . . Frisoni, G. B. (2016). Brain atrophy in Alzheimer's disease and aging. *Ageing research reviews, 30*, 25-48.

Pinto, T. C. C., Machado, L., Bulgacov, T. M., Rodrigues-Júnior, A. L., Costa, M. L. G., Ximenes, R. C. C., & Sougey, E. B. (2018). Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) in the elderly? *International Psychogeriatrics, 31*(4), 491-504. doi:10.1017/S1041610218001370

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*(3), 390-400. doi:10.1111/1460-6984.12444

Price, B. H., 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. *Archives of Neurology, 50*(9), 931-937.

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.

Qiu, Y., Jacobs, D. M., Messer, K., Salmon, D. P., & Feldman, H. H. (2019). Cognitive heterogeneity in probable Alzheimer disease. *Clinical and neuropathologic features, 93*(8), e778-e790. doi:10.1212/wnl.0000000000007967

Quental, N. B. M., Brucki, S. M. D., & Bueno, O. F. A. (2009). Visuospatial function in early Alzheimer's disease: Preliminary study. *Dementia & neuropsychologia, 3*(3), 234-240. doi:10.1590/S1980-57642009DN30300010

Quental, N. B. M., Brucki, S. M. D., & Bueno, O. F. A. (2013). Visuospatial Function in Early Alzheimer's Disease—The Use of the Visual Object and Space Perception (VOSP) Battery. *PLOS ONE, 8*(7), e68398. doi:10.1371/journal.pone.0068398

R Core Team. (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rabin, L. A., Burton, L. A., & Barr, W. B. (2007). Utilization Rates of Ecologically Oriented Instruments Among Clinical Neuropsychologists. *The Clinical Neuropsychologist, 21*(5), 727-743. doi:10.1080/13854040600888776

Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in Test-Usage Practices of Clinical Neuropsychologists in the United States and Canada Over a 10-Year Period: A Follow-Up Survey of INS and NAN Members. *Archives of Clinical Neuropsychology, 31*(3), 206-230. doi:10.1093/arclin/acw007

Radvansky, G. A., & Dijkstra, K. (2007). Aging and situation model processing. *Psychonomic bulletin & review, 14*(6), 1027-1042.

Rapp, A. M., & Wild, B. (2011). Nonliteral language in Alzheimer dementia: a review. *Journal of the International Neuropsychological Society, 17*(2), 207-218.

Rasmussen, K. W., & Berntsen, D. (2022). Deficient semantic knowledge of the life course—Examining the cultural life script in Alzheimer's disease. *Memory & Cognition, 50*(1), 1-15. doi:10.3758/s13421-021-01202-0

Reilly, J., Troche, J., & Grossman, M. (2011). Language processing in dementia. *The handbook of Alzheimer's disease and other dementias*, 336-368.

Reitan, R. M. (1958). Validity of the Trail Making Test as an Indicator of Organic Brain Damage. *Perceptual and Motor Skills, 8*(3), 271-276. doi:10.2466/pms.1958.8.3.271

Rentz, D. M., Amariglio, R. E., Becker, J. A., Frey, M., Olson, L. E., Frishe, K., . . . Sperling, R. A. (2011). Face-name associative memory performance is related to amyloid burden in normal elderly. *Neuropsychologia, 49*(9), 2776-2783. doi:https://doi.org/10.1016/j.neuropsychologia.2011.06.006

Rentz, D. M., Parra Rodriguez, M. A., 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), 58. doi:10.1186/alzrt222

Rentz, D. M., Wessels, A. M., Bain, L. J., Weber, C. J., & Carrillo, M. C. (2020). Clinical meaningfulness addressed at Alzheimer's Association Research Roundtable. *Alzheimer's & Dementia, 16*(5), 814-814. doi:https://doi.org/10.1002/alz.12111

Rinehardt, E., Eichstaedt, K., Schinka, J. A., Loewenstein, D. A., Mattingly, M., Fils, J., . . . Schoenberg, M. R. (2014). Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease. *Dementia and geriatric cognitive disorders, 38*(1-2), 1-9.

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

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society, 2*(6), 525-534.

Rochon, E., Waters Gloria, S., & Caplan, D. (2000). The Relationship Between Measures of Working Memory and Sentence Comprehension in Patients With Alzheimer's Disease. *Journal of Speech, Language, and Hearing Research, 43*(2), 395-413. doi:10.1044/jslhr.4302.395

Rochon, E., Waters, G. S., & Caplan, D. (1994). Sentence comprehension in patients with Alzheimer′ s disease. *Brain and language, 46*(2), 329-349.

Roll, E. E., Giovannetti, T., Libon, D. J., & Eppig, J. (2019). Everyday task knowledge and everyday function in dementia. *Journal of neuropsychology, 13*(1), 96-120.

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *The American Journal of Psychiatry, 141*(11), 1356-1364. doi:10.1176/ajp.141.11.1356

Roth, W.-M., & Jornet, A. (2013). Situated cognition. *WIREs Cognitive Science, 4*(5), 463-478. doi:https://doi.org/10.1002/wcs.1242

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. doi:10.1017/S1041610205003133

Ruff, R. M. (2003). A friendly critique of neuropsychology: facing the challenges of our future. *Archives of Clinical Neuropsychology, 18*(8), 847-864. doi:https://doi.org/10.1016/j.acn.2003.07.002

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nature Reviews Neurology, 10*(11), 634.

Sanz, C., Carrillo, F., Slachevsky, A., Forno, G., Gorno Tempini, M. L., Villagra, R., . . . García, A. M. (2022). Automated text-level semantic markers of Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 14*(1), e12276. doi:https://doi.org/10.1002/dad2.12276

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., . . . Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition, 129*(2), 241-255.

Sattler, C., Toro, P., Schönknecht, P., & Schröder, J. (2012). Cognitive activity, education and socioeconomic status as preventive factors for mild cognitive impairment and Alzheimer's disease. *Psychiatry Research, 196*(1), 90-95. doi:https://doi.org/10.1016/j.psychres.2011.11.012

Sbordone, R. J. (1996). Ecological validity: Some critical issues for the neuropsychologist.

Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of clinical and experimental neuropsychology, 25*(5), 625-633.

Schmidtke, K., & Hermeneit, S. (2008). High rate of conversion to Alzheimer's disease in a cohort of amnestic MCI patients. *International Psychogeriatrics, 20*(1), 96-108.

Schmitter-Edgecombe, M., & Creamer, S. (2010). Assessment of strategic processing during narrative comprehension in individuals with mild cognitive impairment. *Journal of the International Neuropsychological Society, 16*(4), 661-671.

Schröder, J., & Pantel, J. (2016). Neuroimaging of hippocampal atrophy in early recognition of Alzheimer´s disease – a critical appraisal after two decades of research. *Psychiatry Research: Neuroimaging, 247*, 71-78. doi:https://doi.org/10.1016/j.pscychresns.2015.08.014

Sebaldt, R., Dalziel, W., Massoud, F., Tanguay, A., Ward, R., Thabane, L., . . . Lescrauwaet, B. (2014). Detection of Cognitive Impairment and Dementia Using the Animal Fluency Test: The DECIDE Study. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques, 36*(5), 599-604. doi:10.1017/S0317167100008106

Senda, M., Terada, S., Takenoshita, S., Hayashi, S., Yabe, M., Imai, N., . . . Yamada, N. (2020). Diagnostic utility of the Addenbrooke's Cognitive Examination – III (ACE-III), Mini-ACE, Mini-Mental State Examination, Montreal Cognitive Assessment, and Hasegawa Dementia Scale-Revised for detecting mild cognitive impairment and dementia. *Psychogeriatrics, 20*(2), 156-162. doi:https://doi.org/10.1111/psyg.12480

Shulman, K. I., Pushkar Gold, D., Cohen, C. A., & Zucchero, C. A. (1993). Clock-drawing and dementia in the community: A longitudinal study. *International journal of geriatric psychiatry, 8*(6), 487-496. doi:https://doi.org/10.1002/gps.930080606

Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology, 30*(1), 1-38.

Siqueira, G. S. A., Hagemann, P. d. M. S., Coelho, D. d. S., Santos, F. H. D., & Bertolucci, P. H. F. (2018). Can MoCA and MMSE Be Interchangeable Cognitive Screening Tools? A Systematic Review. *The gerontologist, 59*(6), e743-e763. doi:10.1093/geront/gny126

Sivanandam, T. M., & Thakur, M. K. (2012). Traumatic brain injury: A risk factor for Alzheimer's disease. *Neuroscience & Biobehavioral Reviews, 36*(5), 1376-1381. doi:https://doi.org/10.1016/j.neubiorev.2012.02.013

Slegers, A., Filiou, R.-P., Montembeault, M., & Brambati, S. M. (2018). Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's Disease, 65*, 519-542. doi:10.3233/JAD-170881

Snyder, P. J., Kahle-Wrobleski, K., Brannan, S., Miller, D. S., Schindler, R. J., DeSanti, S., . . . Chandler, J. (2014). Assessing cognition and function in Alzheimer's disease clinical trials: do we have the right tools? *Alzheimer's & Dementia, 10*(6), 853-860.

Soldan, A., Pettigrew, C., Cai, Q., Wang, J., Wang, M.-C., Moghekar, A., . . . Albert, M. (2017). Cognitive reserve and long-term change in cognition in aging and preclinical Alzheimer's disease. *Neurobiology of aging, 60*, 164-172. doi:https://doi.org/10.1016/j.neurobiolaging.2017.09.002

Sparks, J. R. (2012). Language/discourse comprehension and understanding. *Encyclopedia of the Sciences of Learning*, 1713-1717.

Spooner, D. M., & Pachana, N. A. (2006). Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology, 21*(4), 327-337.

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology, 11*(11), 1006-1012.

Stern, Y. (2013). Cognitive Reserve: Implications for Assessment and Intervention. *Folia Phoniatrica et Logopaedica, 65*(2), 49-54. doi:10.1159/000353443

Stern, Y., Arenaza‐Urquijo, E. M., Bartrés‐Faz, D., Belleville, S., Cantilon, M., Chetelat, G., . . . Kremen, W. S. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia, 16*(9), 1305-1311.

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

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.

Tarnanas, I., Schlee, W., Tsolaki, M., Müri, R., Mosimann, U., & Nef, T. (2013). Ecological Validity of Virtual Reality Daily Living Activities Screening for Early Dementia: Longitudinal Study. *JMIR Serious Games, 1*(1), e1. doi:10.2196/games.2778

Thorndyke, P. W. (1976). The role of inferences in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior, 15*(4), 437-446.

Toledo, C. M., Aluísio, S. M., dos Santos, L. B., Brucki, S. M. D., Trés, E. S., de Oliveira, M. O., & Mansur, L. L. (2018). Analysis of macrolinguistic aspects of narratives from individuals with Alzheimer's disease, mild cognitive impairment, and no cognitive impairment. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 10*, 31-40.

Tort-Merino, A., Valech, N., Peñaloza, C., Grönholm-Nyman, P., León, M., Olives, J., . . . Martínez-Lage, P. (2017). Early detection of learning difficulties when confronted with

novel information in preclinical Alzheimer's disease stage 1. *Journal of Alzheimer's Disease, 58*(3), 855-870.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse processes, 21*(3), 255-287.

Tröger, J., Lindsay, H., Mina, M., Linz, N., Klöppel, S., Kray, J., & Peter, J. (2021). Patients with amnestic MCI Fail to Adapt Executive Control When Repeatedly Tested with Semantic Verbal Fluency Tasks. *Journal of the International Neuropsychological Society*, 1-8. doi:10.1017/S1355617721000849

Tromp, D., Dufour, A., Lithfous, S., Pebayle, T., & Després, O. (2015). Episodic memory in normal aging and Alzheimer disease: Insights from imaging and behavioral studies. *Ageing research reviews, 24*, 232-262. doi:https://doi.org/10.1016/j.arr.2015.08.006

Ulatowska, H. K., Chapman, S., Johnson, J., & Branch, C. (1999). Macrostructure and inferential processing in discourse of aphasic patients. *Psychology of Language and, 3*(2), 3-20.

Ulatowska, H. K., Chapman, S. B., Highley, A. P., & Prince, J. (1998). Discourse in healthy old-elderly adults: A longitudinal study. *Aphasiology, 12*(7-8), 619-633.

van Boxtel, W., & Lawyer, L. (2021). Sentence comprehension in ageing and Alzheimer's disease. *Language and Linguistics Compass, 15*(6), e12430. doi:https://doi.org/10.1111/lnc3.12430

Van der Loo, M. P. (2014). The stringdist package for approximate string matching. *R J., 6*(1), 111.

Van Dijk, T. A. (2019). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*: Routledge.

van Loenhoud, A. C., van der Flier, W. M., Wink, A. M., Dicks, E., Groot, C., Twisk, J., . . . Initiative, f. t. A. s. D. N. (2019). Cognitive reserve and clinical progression in Alzheimer disease. *A paradoxical relationship, 93*(4), e334-e346. doi:10.1212/wnl.0000000000007821

Vas, A. K., Spence, J., & Chapman, S. B. (2015). Abstracting meaning from complex information (gist reasoning) in adult traumatic brain injury. *Journal of clinical and experimental neuropsychology, 37*(2), 152-161.

Vas, A. K., Spence, J. S., Eschler, B., & Chapman, S. B. (2016). Sensitivity and specificity of abstraction using gist reasoning measure in adults with traumatic brain injury. *Journal of Applied Biobehavioral Research, 21*(4), 216-224.

Vemuri, P., & Jack, C. R. (2010). Role of structural MRI in Alzheimer's disease. *Alzheimer's research & therapy, 2*(4), 23. doi:10.1186/alzrt47

Vemuri, P., Weigand, S. D., Przybelski, S. A., Knopman, D. S., Smith, G. E., Trojanowski, J. Q., . . . Bernstein, M. A. (2011). Cognitive reserve and Alzheimer's disease biomarkers are independent determinants of cognition. *Brain, 134*(5), 1479-1492.

Verghese, P. B., Castellano, J. M., & Holtzman, D. M. (2011). Apolipoprotein E in Alzheimer's disease and other neurological disorders. *The Lancet Neurology, 10*(3), 241-252. doi:https://doi.org/10.1016/S1474-4422(10)70325-2

Verma, M., & Howard, R. J. (2012). Semantic memory and language dysfunction in early Alzheimer's disease: a review. *International journal of geriatric psychiatry, 27*(12), 1209-1217. doi:10.1002/gps.3766

Villemagne, V. L., Pike, K. E., Chételat, G., Ellis, K. A., Mulligan, R. S., Bourgeat, P., . . . Rowe, C. C. (2011). Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Annals of Neurology, 69*(1), 181-192. doi:10.1002/ana.22248

Vogel, A., Gade, A., Stokholm, J., & Waldemar, G. (2005). Semantic Memory Impairment in the Earliest Phases of Alzheimer's Disease. *Dementia and geriatric cognitive disorders, 19*(2-3), 75-81. doi:10.1159/000082352

von Strauss, E., Viitanen, M., De Ronchi, D., Winblad, B., & Fratiglioni, L. (1999). Aging and the Occurrence of Dementia: Findings From a Population-Based Cohort With a Large Sample of Nonagenarians. *Archives of Neurology, 56*(5), 587-592. doi:10.1001/archneur.56.5.587

Vuorinen, E., Laine, M., & Rinne, J. (2000). Common pattern of language impairment in vascular dementia and in Alzheimer disease. *Alzheimer Disease & Associated Disorders, 14*(2), 81-86.

Wadley, V. G., Okonkwo, O., Crowe, M., & Ross-Meadows, L. A. (2008). Mild cognitive impairment and everyday function: evidence of reduced speed in performing instrumental activities of daily living. *Am J Geriatr Psychiatry, 16*(5), 416-424. doi:10.1097/JGP.0b013e31816b7303

Wang, Y.-Y., Yang, L., Zhang, J., Zeng, X.-T., Wang, Y., & Jin, Y.-H. (2021). The Effect of Cognitive Intervention on Cognitive Function in Older Adults With Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Neuropsychology review*. doi:10.1007/s11065-021-09486-4

Ward, A., Tardiff, S., Dye, C., & Arrighi, H. M. (2013). Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a systematic review of the literature. *Dementia and geriatric cognitive disorders extra, 3*(1), 320-332. doi:10.1159/000354370

Weakley, A., & Schmitter-Edgecombe, M. (2014). Analysis of verbal fluency ability in Alzheimer's disease: the role of clustering, switching and semantic proximities. *Archives of Clinical Neuropsychology, 29*(3), 256-268.

Wechsler, D. (1945). *Wechsler memory scale*. San Antonio, TX, US: Psychological Corporation.

Weih, M., Wiltfang, J., & Kornhuber, J. (2007). Non-pharmacologic prevention of Alzheimer's disease: nutritional and life-style risk factors. *Journal of Neural Transmission, 114*(9), 1187. doi:10.1007/s00702-007-0704-x

Weiner, J., Frankenberg, C., Schröder, J., & Schultz, T. (2019, 14-18 Dec. 2019). *Speech Reveals Future Risk of Developing Dementia: Predictive Dementia Screening from Biographic Interviews.* Paper presented at the 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU).

Weintraub, S., Wicklund, A. H., & Salmon, D. P. (2012). The neuropsychological profile of Alzheimer disease. *Cold Spring Harbor perspectives in medicine, 2*(4), a006171.

Welland, R. J., Lubinski, R., & Higginbotham, D. J. (2002). Discourse comprehension test performance of elders with dementia of the Alzheimer type. *Journal of Speech, Language, and Hearing Research, 45*(6), 1175-1187.

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research, 13*(3), 363-375. doi:https://doi.org/10.1016/S0926-6410(01)00129-X

Willis, S. L. (1996). Everyday Cognitive Competence in Elderly Persons: Conceptual Issues and Empirical Findings1. *The gerontologist, 36*(5), 595-601. doi:10.1093/geront/36.5.595

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. *Journal of Internal Medicine, 256*(3), 240-246. doi:10.1111/j.1365-2796.2004.01380.x

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

World Population Ageing 2017-Highlights. (2017). Retrieved from

Xu, W., Tan, L., Wang, H.-F., Jiang, T., Tan, M.-S., Tan, L., . . . Yu, J.-T. (2015). Meta-analysis of modifiable risk factors for Alzheimer's disease. *J Neurol Neurosurg Psychiatry, 86*(12), 1299-1306.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological bulletin, 133*(2), 273.

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and aging, 21*(3), 466.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological bulletin, 127*(1), 3.

Zhao, Y., & Zhao, B. (2013). Oxidative stress and the pathogenesis of Alzheimer's disease. *Oxidative medicine and cellular longevity, 2013*.

Zimerman, S., Hasher, L., & Goldstein, D. (2011). Cognitive ageing: a positive perspective. *The paradoxical brain*, 130-150.

Appendix

mentary Table A1: PRISMA 2009 Checklist

opic	#	Checklist item	Reported on page #
	1	Identify the report as a systematic review, meta-analysis, or both.	1
ACT			
d	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
DUCTION			
	3	Describe the rationale for the review in the context of what is already known.	3-6
es	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6-7
DDS			
and on	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.	--
y	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
on	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
	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	S2
	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
n	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
ns	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8

isk of bias individual udies	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
ummary easures	13	State the principal summary measures (e.g., risk ratio, difference in means).	11
ynthesis of sults	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.	--
isk of bias ross studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	--
dditional nalyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	--
ESULTS			
tudy election	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
tudy haracteristics	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.
isk of bias ithin 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
esults of dividual udies	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
ynthesis of sults	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	--
isk of bias ross studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	--
dditional nalysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	--
ISCUSSION			
ummary of vidence	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
imitations	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
onclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	25-28

urces of funding for the systematic review and other support (e.g., supply of data); role of funders ematic review.	28

Supplementary Table A2: 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) Indexes=SCI-EXPANDED, SSCI Timespan=All years	1597 253	08.03.2018 20.01.2020	1934- 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]	955 180	08.03.2018 20.01.2020	1954- 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" Language: English	1587 144	08.03.2018 20.01.2020	1934- 2020

Supplementary Table A3: List of events in Experiment 1, along with the trial-wise sequence accuracy (edit distance) by groups, and the sequence agreement in the pilot study

Original event names	**English translation**	**Edit distance**			**Pilot agreement (%)**
		AD	**MCI**	**OA**	
Torte dekorieren	decorating a cake	0.9	0.4	0	100
Lagerfeuer machen	lighting a campfire	0.9	0.2	0.4	95
Cocktail mixen	mixing a cocktail	2.3	1.4	0.2	95
Ostern feiern	celebrating Easter	0.8	0.4	0.2	100
Blume einpflanzen	planting a flower	1.7	0.9	0.5	100
töpfern	pottery	0.8	0.6	0.3	95
im Restaurant essen	eating at a restaurant	1.3	1.6	0.5	100
Sandburg bauen	building a sandcastle	2.1	1.0	0.9	100
nähen	sewing	0.5	0.4	0.1	86
einkaufen	grocery shopping	0.6	0.8	0.2	100
fallschirmspringen	skydiving	0.8	0.5	0	100
Tee kochen und trinken	making & drinking tea	0.2	0.2	0	100
verreisen	travelling	0.8	0.8	0.2	100
Truthahn braten	roasting a turkey	0.7	0.4	0	100

Supplementary Table A4: List of events in Experiment 2 and the sequence agreement in the pilot study

Original event names	**English translation**	**Pilot sequence agreement (%)**
backen	baking	100
Getränk einschenken	pouring a drink	100
bowlen/kegeln	bowling	95
Kuh melken	milking a cow	95
ins Wasser springen	diving	100
Hund baden	Bathing a dog	100
Blumenschale arrangieren	Arranging flowers	100
angeln	fishing	95
Golf spielen	playing golf	95
grillen	grilling	95
Brief schreiben	writng a letter	100
Bild malen	painting	100
Pferd satteln	saddling a horse	100
Weihnachtsbaum schmücken	decorating a Christmas tree	100

List of Publications

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,* 3–18. https://doi.org/10.1007/s10433-021-00619-5

C. v. Stutterheim and I conceptualized the review. I conducted the database searches, article screenings, and selection. I extracted the data from articles, conducted quality assessments, and synthesized and interpreted the findings. I wrote and revised the manuscript. S. Celik conducted articles screenings, selection, and quality assessments. H.W. Wahl advised on interpretation of the findings and revision of manuscript. C.v. Stutterheim contributed to interpretation of findings and manuscript revision.

Kokje, E., Gerwien, J., & von Stutterheim, C. (2021). Macro-event recognition in healthy aging, Alzheimer's disease, and mild cognitive impairment. *Journal of Neuropsychology.* https://doi.org/10.1111/jnp.12271

C.v. Stutterheim, J. Gerwien, and I developed and designed the study. I prepared the study materials and collected the data. I processed, analyzed, and interpreted the data. I wrote and revised the manuscript. J. Gerwien advised on the analysis and interpretation of data, and in the revision of the manuscript. C.v. Stutterheim contributed to the interpretation of data and revision of manuscript.

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	Eesha Kokje
Datum / Date	28.07.2022
Unterschrift / Signature	(Dem Dekanat der Fakultät für Verhaltens- und Empirische Kulturwissenschaften liegt eine unterschriebene Version dieser Erklärung vom 28.07.2022 vor.)