Memory Precision across Space and Time in Alzheimer's disease

Candidate: Dr. Sofia Toniolo

Supervisors: Prof. Masud Husain and Prof. Clare Mackay

Dissertation for the degree of Doctor of Philosophy (DPhil) in Clinical Neurosciences

Nuffield Department of Clinical Neurosciences, University of Oxford



Table of contents

Legend:

• = Young Healthy Controls (YHC)

§ = Elderly Healthy Controls (EHC)

‡ = Subjective Cognitive Impairment patients (SCI)

¥ = Mild Cognitive Impairment patients (MCI)

= Alzheimer's Disease patients (AD)

* = Parkinson's Disease patients (PD)

General Abstract

1. Introduction

- 1.1 Brain mechanisms underlying Visual Short-Term Memory (VSTM)

-- 1.1.1 Fundamentals of working memory: the central executive, phonological loop, visuospatial sketchpad, and some magic (numbers)

-- 1.1.2 The Slot Model: a fixed number of resources stored with fixed precision

-- 1.1.3 The Mixture Model: flexible allocation of shared resources determines memory precision

-- 1.1.4 The ability of Filtering out distractors as key determinant of working memory capacity

- 1.2 Memory impairment profile in healthy ageing and patients at risk of developing Alzheimer's Disease

-- 1.2.1 The role of memory precision, target detection, ability to filter out distractors, misbinding and guessing in healthy ageing

-- 1.2.2 Traditional neuropsychological hallmarks of early AD: the role of long-term memory

-- 1.2.3 Visual Short-Term Memory impairment in patients at risk of developing AD

- 1.3 Brain Imaging correlates of Visual Short-Term Memory

-- 1.3.1 Functional neuroanatomy of WM

-- 1.3.2 Neural circuitry involved in Filtering abilities

-- 1.3.3 Brain correlates of misbinding, guessing and memory precision in healthy subjects

-- 1.3.4 Misbinding as hippocampal lesion model in patients with neurological diseases

- 1.4 Plasma biomarkers of Alzheimer's Disease

-- 1.4.1 Why are biomarkers needed?

-- 1.4.2 The 'ATN' system

-- 1.4.3 Relationship with cognitive impairment

-- 1.4.4 Plasma biomarkers breakthrough

2. Filtering at Encoding and Maintenance modelled through Mixture Model metrics•§#*

- 2.1 Introduction

- 2.2 Experiment 1: Filtering at Encoding and Maintenance in healthy ageing

- 2.3 Experiment 2: Filtering at Encoding and Maintenance in patients with AD

- 2.4 Experiment 3: Filtering at Encoding and Maintenance in patients with PD

- 2.5 Transdiagnostic analysis of cognitive measures and hippocampal volumes

- 2.6 Discussion

- 2.7 Conclusion

3. Cross-Sectional and Longitudinal assessment of Visual Short-Term Memory in patients at risk of developing Alzheimer's Disease with the "What was where?" Oxford Memory Task §‡¥#

- 3.1 Cross-sectional Visual Short-Term memory assessment in patients with SCI, MCI and AD §‡¥#

-- 3.1.1 Delay, Set Size effects and Group differences across 8 key VSTM metrics (Identification Time, Localization Time, Proportion Correct, Absolute Error, Target Detection, Misbinding, Guessing and memory precision)

-- 3.1.2 Support Vector Machine classifier using VSTM metrics for patients' stratification

-- 3.1.3 Relationship between VSTM metrics and baseline hippocampal volumes

- 3.2 Longitudinal evaluation of Visual Short-Term Memory decline after 1 year in patients with SCI, MCI and AD §‡¥#

-- 3.2.1 Differential trajectories of VSTM metrics for monitoring disease progression

-- 3.2.2 Features extraction for predicting performance decline

-- 3.2.3 Relationship between VSTM performances' decline over time and baseline hippocampal atrophy

4. Structural correlates of Visual Short-Term Memory using Voxel-Based Morphometry (VBM) in AD and EHC §#

5. Diffusion correlates of Visual Short-Term Memory using Tract-based spatial statistics (TBSS) in AD and EHC §#

6. Combining Plasma Biomarkers and digital VSTM metrics in AD and EHC §#

- 6.1 Relationship between Plasma Biomarkers and digital cognitive metrics

- 6.2 Test-retest variability of Plasma Biomarkers and digital cognitive metrics

7. Discussion and Conclusion

- 7.1 Overview of findings and contribution to existing literature

- 7.2 Potential translational value

- 7.3 Future Directions

- 7.4 Conclusion

References

Appendix

Acknowledgements

GENERAL ABSTRACT

This thesis aimed to deepen current understanding of the mechanisms underlying visual short-term memory (VSTM) impairments in Alzheimer's Disease (AD). The mixture model of working memory (WM) was deployed to dissect the differential contributions of distinct VSTM processes: ability to detect a target, swapping an item with another item in memory (misbinding), random guessing and precision of item representation (memory precision). Proficiency in filtering an unwanted item out (filtering ability), whether presented simultaneously with the to-be-remembered stimuli (encoding) or during the retention phase (maintenance) was also studied. In addition, measures of overall task performance, spatial localisation errors and reaction times were analysed.

Chapter 2 showed that healthy ageing was associated with a decline in VSTM memory precision, while in AD patients higher guessing and lower target detection were the main sources of memory errors. AD patients were also particularly affected when a distractor was presented during maintenance (filtering at maintenance), but not at encoding. As a comparator, patients with Parkinson's Disease (PD) showed higher guessing rates, but preserved filtering abilities. In **Chapter 3** the "What was where" Oxford Memory Task (OMT), was used to assess VSTM performance in people at risk of developing AD dementia, i.e., patients with subjective cognitive impairment (SCI) and mild cognitive impairment (MCI), compared to patients with established AD dementia. Apart from reaction time measures, most metrics derived from the OMT task were able to discriminate between healthy controls and patients with MCI, and between MCI and AD dementia. Mixture model metrics and spatial localization error also differed significantly between people with SCI and MCI. All metrics were correlated with hippocampal atrophy.

Chapter 4 showed that OMT metrics clustered into four specific spatial patterns of atrophy in the grey matter across a cohort of healthy controls and AD patients, with the Precuneus being associated only with metrics that carried spatial information. In **Chapter 5** a key white matter region encompassing three tracts (Optic radiation, Forceps Major and Middle Longitudinal Fasciculus) was identified to be associated with VSTM performance on OMT in the left hemisphere in AD patients, but in the right hemisphere in healthy controls.

Finally, In **Chapter 6** OMT metrics and several other cognitive measures derived from an online battery testing VSTM, executive functions, processing speed, long-term memory and visuospatial abilities, were related to plasma biomarkers of AD; phospotau (pTau)181, Glial fibrillary acidic protein (GFAP), Neurofilament Light Chain (NfL), and amyloid β 42/40 ratio (Aβ42/40 ratio). The results showed that of these biomarkers, pTau181 was most closely correlated to cognitive performance. However, cognitive measures and plasma biomarkers had different degrees of test-retest reliability, with pTau181 showing the highest degree of variability among plasma biomarkers in this sample.

Chapter 1. GENERAL INTRODUCTION

1.1 Brain mechanisms underlying Visual Short-Term Memory (VSTM)

1.1.1 Fundamentals of working memory: the central executive, phonological loop, visuospatial sketchpad, and some magic (numbers)

Working memory (WM) has historically been defined by its limited capacity, and by its impersistence. In the original model by Baddeley and Hitch of 1974, working memory is divided into 3 essential components. These are the '*central executive'*, an attentional control system, aided by two subsidiary short-term memory (STM) stores or buffers, the '*phonological loop*' and the '*visuospatial sketchpad*' [1]. In this model, the phonological loop temporarily holds verbal and acoustic information thanks to a phonological short-term storage and an articulatory rehearsal system, while the visuospatial sketchpad retains visuospatial information over a short period of time. It was only in 2000 that the model was revised by Baddeley himself, to add a fourth component, the '*episodic buffer*', which provides temporary storage of information in a multimodal code, binding information from the subsidiary stores and long-term memory (LTM) into a unitary episodic representation [2].

The term STM is most commonly used to refer to passive, temporary storage of a limited amount of information for a brief amount of time, while WM refers to the ability to manipulate the information contained within STM in order to plan and carry out intended actions. It is therefore heavily dependent on attentional processes guided by the central executive, and more tightly correlated with speed of processing and reasoning [3].

Furthermore, it can also be influenced and aided by information stored in LTM through the episodic buffer. This can help to increase WM capacity by chunking items semantically, i.e., compressing information thanks to shared commonalities between items [3].

While the distinction between STM and WM retains an important historical value, these terms are often used interchangeably, as most tasks require some forms of manipulation of STM content. The experiments described in this DPhil have different grades of required content processing, and therefore I will treat these terms as equivalent, while acknowledging they are conceptually distinct.

WM has been shown to persist up to few seconds, initially described as 18-30 seconds [4, 5], with some arguing that does not exceed 2 seconds if not aided by rehearsal [6]. Others might say that the opposite is also true: technically, with continued rehearsal it can last indefinitely, but thankfully, this not implemented in everyday life or in experimental conditions. Regarding the limit on the number of items that can be stored, George Miller approached this topic rather heuristically, in what has been regarded as one of the most influential, yet controversial, papers on the limits of WM [7]. He stated that WM can only hold only 7, plus or minus 2, items (Miller's Law, 'magic number 7'), and this capacity limits applies to a variety of unidimensional tasks that encompass different senses, such as repeating words, hearing sounds, perceiving intensity, duration and locations of a vibration, and spatial discrimination of visual stimuli. Even in his original paper, he mentioned that the channel capacity for judgments of visual position seems to be significantly larger compared to other modalities and can potentially go up to 15.

Another important corollary statement of Miller's paper is that the *number* of items might not be as important as the *bits* or *chunks* of information associated with that item, and that the number of items that can be stored varies according to the complexity of the stimulus and the ability of compressing information. Later studies by Cowan et al identified at 4 the number of chunks that can be retained in working memory, without additional contributions of rehearsal or LTM ('magic number 4') [8].

Even Miller, who was the first to hypothesize a discrete number, concluded that there are ways to expand our WM capacity, and the three most important ones are either to make *relative* rather than *absolute* judgments, to increase the number of *dimensions* along which stimuli can differ, or to make a *sequence* of several absolute judgments in a row [7]. Regarding the first method, it is interesting to note that this has evolved in time into the concept of decreasing the *precision* of memory representations, a theme which will be extensively explored in this thesis.

1.1.2 The Slot model: a fixed number of resources stored with fixed precision

Miller and Cowan's views of a fixed number of items, bits, and chunks have been incorporated in what is now known as the 'Slot model' [9]. In their ground-breaking paper of 1997, Luck and Vogel found that only four colors or orientations could be stored in visual WM at one time [10]. However, they also reported that is also possible to retain both the color and the orientation of four objects, suggesting that visual WM stores integrated objects rather than individual features, allowing therefore 16 individual features to be stored if distributed across four objects. According to the Slot model, all items are encoded with equal

precision, and if there are more items than slots, the remaining items are simply not encoded, and therefore the responses will be represented by random guesses. This is mirrored by the shape of the response curves they observed in several different STM tasks, which they modelled in terms of high precision responses, centered narrowly around the target location, and by random guesses, represented by a uniform distribution [9] (**Figure 1.1**, left).

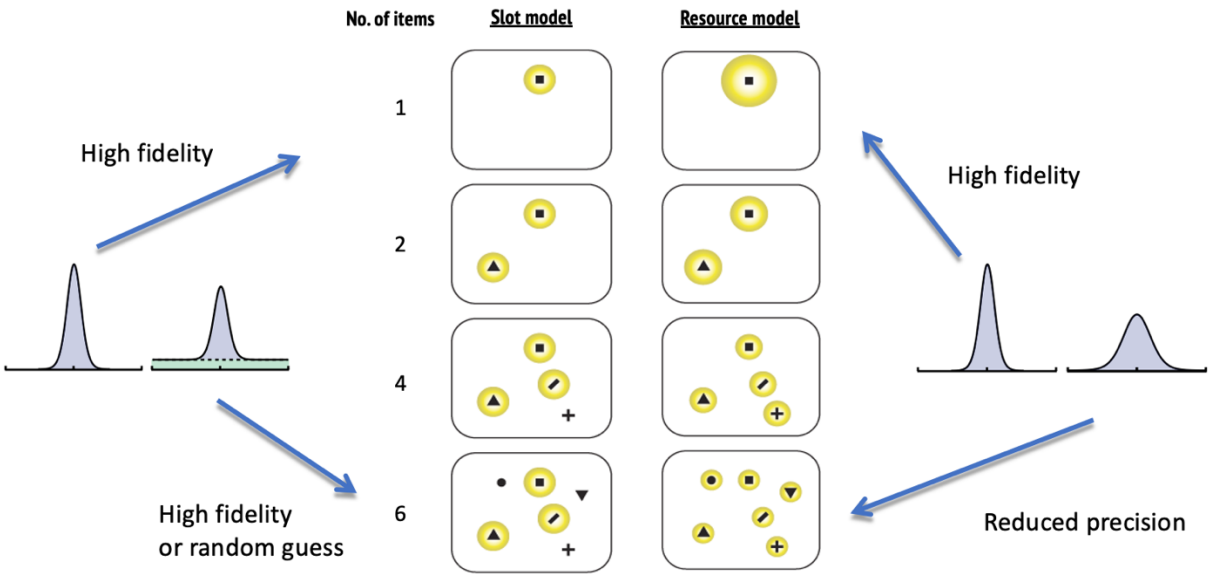


Figure 1.1 | Slot model and Resource model of WM

Adapted from Ma, Husain and Bays [11]. On the left: According to the Slot model, items can be stored with a fixed amount of precision. If the number of items exceeds a fixed number of 'slots', those items are not stored, and participants resort to random guessing. On the right: In the Resource model, items can be stored with variable degrees of precision. When the number of items is small, items are stored with high precision. With increasing number of items, the precision of the representation decreases.

However, the findings of Luck and Vogel regarding fixed capacity limits depend heavily on specific features of the items to-be-retained. Indeed, several factors can impact upon the total amount of information that can be stored: the ability to *bind* information together in a single representation (or encoding multidimensionally according to Miller), the degree of similarity that items share and covert or overt attentional shifts [12, 13, 14].

A variant of the Slot model from the same group, the 'Slots + Averaging model' [15], tries also to account for the increased guessing rates observed with increasing set sizes (higher number of items), which would not be predicted by the original Slot model. They postulated that if only one item is stored in memory, it can be represented in more than one slot, that can team up and increase the precision of the item recall. According to their model, slots act independently if there are more items to be stored, leaving a fixed number of slots that can be filled. However, a prevalent competing view claims that there is no such upper limit to the number of items that can be stored [16, 17], and instead supports new ways of modelling working memory capacity, known collectively as 'Resource models'.

1.1.2 The Mixture Model: flexible allocation of shared resources determines memory precision

In recent years, two different types of paradigms have been used to study visual STM (VSTM). The first uses *change-detection tasks*, in which participants indicate whether a change occurred or not. The second involves *delayed-reproduction tasks,* where they are asked to reproduce features of the remembered item (e.g., its location, color or orientation) in a continuous, analogue response space (**Figure 1.2**). Importantly, only the latter measures the precision of memory recall, rather than just the number of the items remembered in a binary (correct/incorrect) manner.

Resource models still consider VSTM to be limited in its capacity, but rather than measure a finite number of features or objects that can be stored, they regard VSTM as a limited pool of resource, where quantity can be traded for quality (**Figure 1.1**, right, **Figure 1.2**) [11]. These

models have been able to show that this trade-off follows a power law, where the more items that are stored, the lower the precision by which they are retained.

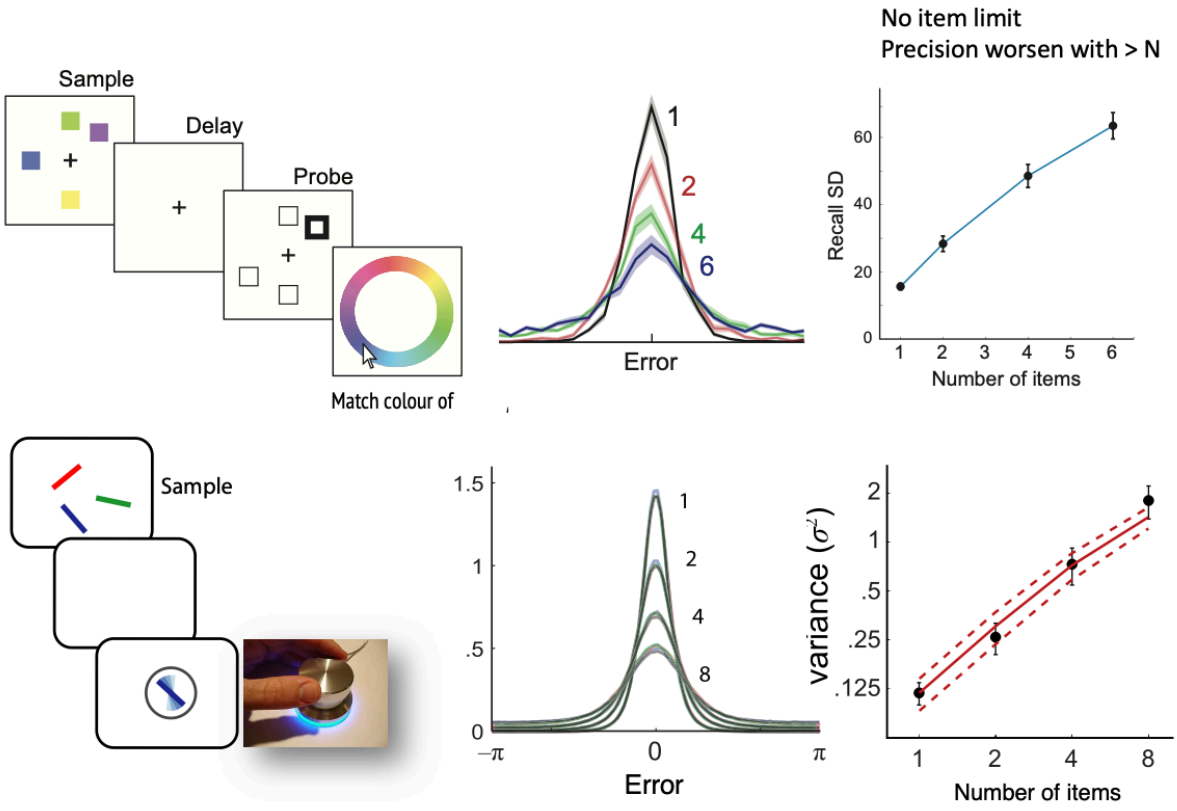


Figure 1.2 | Delayed-reproduction tasks

Adapted from Ma, Husain and Bays [11], and Bays et al [18]. In delayed reproduction tasks, participants are required to reproduce features of the remembered item (e.g., its location, color, or orientation) in a continuous fashion. The precision of the recall decreases as the number of items increases.

Precision can be determined not only by the number of items, but also by other factors, including whether the items to remember are in a privileged attentional state, often termed as the 'focus of attention' [19]. This property has been referred to as *flexibility* of resource allocation, where resources can be redirected and non-equally distributed to gain a behavioral advantage. However, this behavioral advantage comes at the expense of the lower precision of recall of the other items in memory, thus remaining faithful to the idea that VSTM is a closed neural system, where if a bigger amount of resource is consumed by an item, the other items will have less available for them [20].

The flexibility of resource allocation can be mediated by voluntary redeployment of attention, but can also be driven by unwanted attentional capture by visually salient objects [21], and can occur at encoding as well as more dynamically during the maintenance phase, as when retrocues are deployed, i.e., attentional cues presented after the stimulus [22]. Top-down processes that include expectations, memory and motivation are considered to interact with bottom-up processes, including salience, to form spatial priority navigation maps that are able to guide behavior through competing cognitive demands [23].

Precision seems to be a crucial determinant of variability across memory representations but is not the only factor at play. A very influential Resource model of VSTM is represented by the 'Mixture Model' of working memory [17]. According to this model, the probability of responding with the reported orientation ($\hat{\theta}$) is described by the following equation:

$$p(\hat{\theta}) = (1-\gamma-\beta)\Phi_\kappa(\hat{\theta}-\theta) + \gamma\frac{1}{2\pi} + \beta\frac{1}{m}\sum_{i}^{m}\Phi_\kappa(\hat{\theta}-\theta_i)$$

where θ is the true orientation of the target item, $\hat{\theta}$ the orientation reported by the participant, and φ_k is the von Mises distribution (the circular analogue of the Gaussian) with mean zero and concentration parameter κ, which is a measure of memory precision (**Figure 1.3**). The probability of reporting the correct target item is given by $1-\gamma-\beta$ (in red), while the probability of mistakenly reporting a non-target item is given by β (in light blue), where $\{\theta_1, \theta_2, \theta_m\}$ are the orientations of the m non-target items. γ represents the probability of responding randomly (in green).

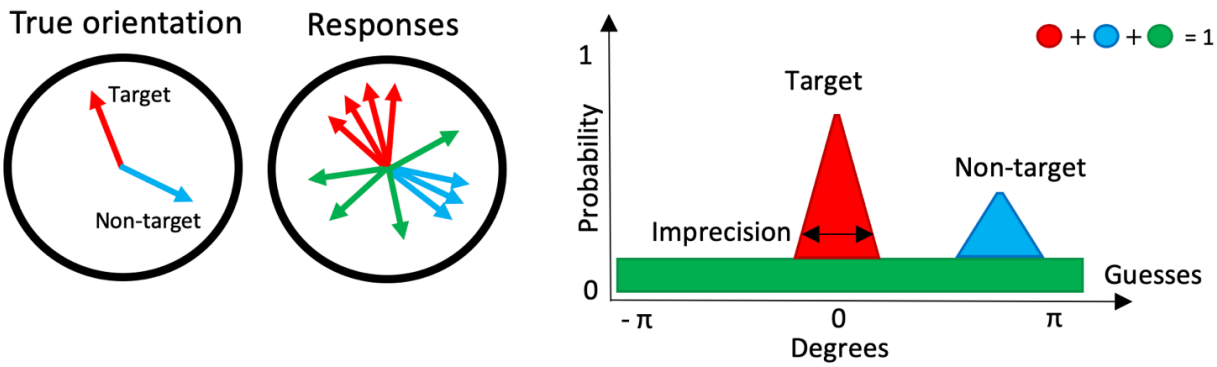


Figure 1.3 | Mixture Model of Working Memory

Adapted from Grogan et al [24]. Summary of the Mixture Model metrics (Target detection, Misbinding, Guessing, Imprecision). The probabilities of detecting a target (in red), a non-target (in light blue) and guessing (in green) add to one.

The Mixture Model will be at the basis of the analyses in my thesis. These parameters will be referred from now onwards as **Target detection** (the probability of correctly identifying a target), **Misbinding** (the probability of erroneously swapping the target with another item in memory), **Guessing** (the probability of random guessing), and **Precision** (the concentration of the responses around the target, or its opposite Imprecision (its width)).

This model, which unlike the Slot model includes misbinding, or 'swap' errors, shows that there is an inverse relationship between the number of items in VSTM and their precision, without a fixed number of items that can be stored. Thus, misbinding errors may previously have been mistaken for random guesses in traditional Slots models [17].

The Mixture Model has four different factors that account for subjects' response variability, but there is a fifth element that plays an important role in flexibility of resource allocation. This is filtering ability, i.e., the ability to filter out task-irrelevant distractors. There are other attentional and non-attentional mechanisms that can influence VSTM capacity, including chunking of information [25], which will not be my primary focus. Therefore, these five

elements are not to be considered exhaustive of all possible contributors to VSTM capacity but are nevertheless key determinants and will be the primary subject of the work reported in this thesis.

1.1.3 Filtering out distractors as key determinant of Working Memory capacity

As WM is limited in its capacity, items compete with each other for resource utilization. Therefore, optimizing the amount of task-relevant information to attend to, might be important to boost performance [26]. Filtering out unwanted distractors has the potential of freeing up space within a crowded pool of items to-be-remembered [27, 28].

How does the brain deal with distracting information? Several factors need to be considered, such as whether that item is relevant or not to the task, previous experience that boosts attention to the target item or otherwise shifts it away from the distractor, target-distractor similarity, and *when* filtering needs to be accomplished [28].

The last factor is of particular importance: intuitively, filtering out a distractor should represent a process that arises only after that distractor has been presented to the viewer, who will subsequently decide whether to discard it or not (a process also known as reactive inhibition) [29]. However, evidence shows that it is possible also to suppress a distractor even before it has been presented (i.e., preattentive inhibition) through an early top-down process [30]. This can be largely dependent on previous experience shaping spatial priority maps to prevent unwanted attentional capture and is consistent with the view that WM can be modelled as transitioning between two functional states, one in which there is maintenance

of information without a specific action plan, and one where WM prioritizes items in an action-oriented format [28]. This reconfiguration of representational state in anticipation of upcoming task demands might serve to optimize task responses and is potentially a crucial element of filtering abilities.

Besides pre-attentive and post-attentive considerations, it is also important not to forget an important characteristic of WM, which is the ability to manipulate information over time. This opens a second question of whether filtering processes not only can occur at pre and post, peri-encoding stages, but also later on during the maintenance period (**Figure 1.4**). Indeed, it has been shown that even long after encoding, filtering can improve memory performance, by re-prioritizing the content of information in WM [31]. This led researchers to propose that WM might possess multiple '*gates*', including an input gate and an output gate [32]. The input gate reflects what is allowed to enter WM, while the output gate rules the reallocation of WM representations made irrelevant by acquired task-relevant information. Filtering therefore might not simply be tied to a single event but could be involved in input and output processes, both tightly linked to WM capacity.

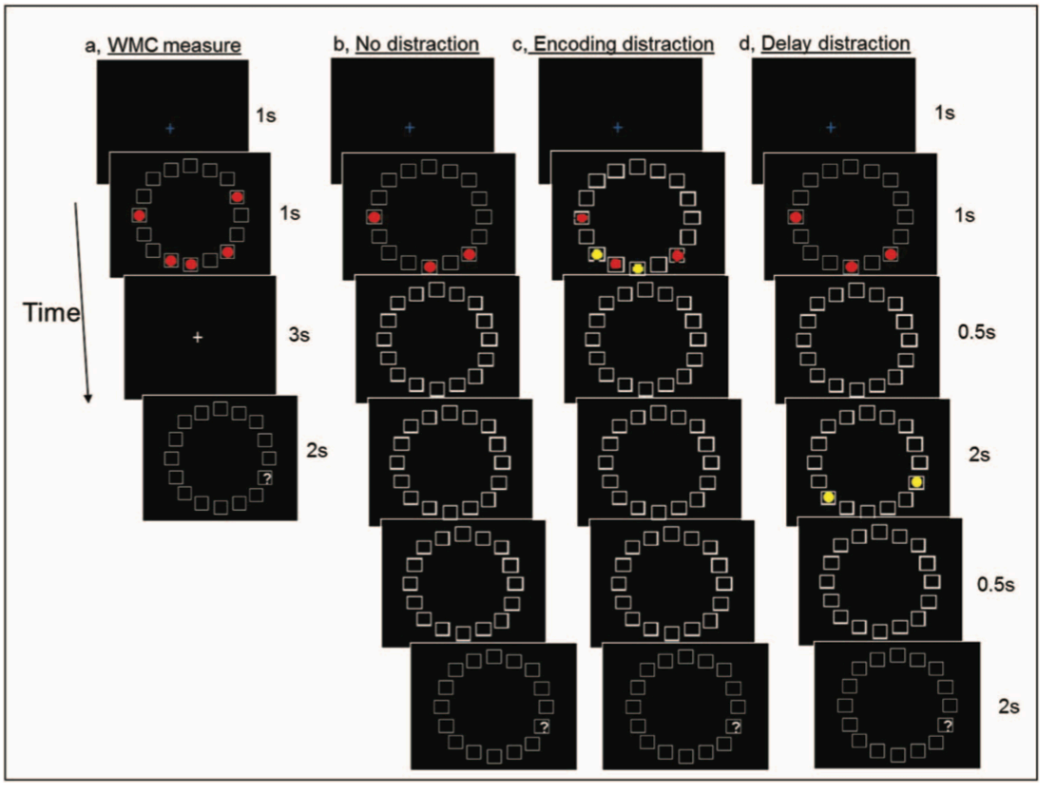


Figure 1.4 | Distractor filtering at encoding and maintenance

Adapted from McNab & Dolan [33]. An example of trials when distractor filtering is accomplished during the encoding phase or during maintenance.

It is also important to underscore that target selection and distractor suppression can be independent processes and do not necessarily rely on each other. Distractor suppression relies more heavily on maintaining a facilitated WM reconfiguration, as seen by the presence of the benefit often seen in fixed conditions where WM remodeling can be maintained, and not in mixed, flexible trials where that preparatory mechanism is disrupted [34]. Indeed, WM reconfiguration varies according to different types of filtering demands and experimental set-up.

1.2 Memory in healthy ageing and people at risk of developing Alzheimer's Disease

1.2.1 Memory precision, target detection, misbinding, guessing and filtering ability in healthy ageing

In the previous paragraphs, I introduced some major determinants of WM capacity. These are usually modelled on young, healthy participants; but while the theoretical constructs might hold true along the lifespan, WM performance declines even in normal, healthy individuals [35, 36]. Large scale studies found that WM capacity grows over time and peaks in our 20-30s, while then slowly but steadily declines over time during our adulthood following an inverted U-shape; indeed, some evidence shows that performances by age 55 are comparable to the ones of an 8-9 year-old [35, 36].

Several distinct components that contribute to worsening performance have been identified in ageing: a reduction in the number of items that can be stored, or WM capacity [35]; a deficit in maintaining the associations (bindings) between individual object features [37] resulting in a higher proportion of misbinding [37, 38]; an increase in random guessing [39]; a decrease in response precision [38, 39, 40, 41]; and an impaired filtering ability [42]. The age-related decline in WM capacity is important as it is of considerable magnitude, with changes in capacity explaining at least five times the variance in performance compared to changes in other abilities as for example binding abilities in one study [35].

However, these components are not mutually exclusive. Indeed, binding errors have been proposed to significantly contribute to the age-related decline in WM capacity, with several studies reporting higher rates of *misbinding* in healthy ageing [37, 38], especially at higher set sizes [39]. Notably, not all investigations have found such a difference in binding abilities in ageing, possibly due to the type of task used (change detection versus delayed-reproduction tasks) [43], low number of items probed which might hamper its detection at smaller set sizes [35], the control group chosen (young healthy controls or Alzheimer's Disease (AD) patients) [44], whether binding conditions were assessed in mixed versus fixed trials (the former requiring greater cognitive flexibility) [37], or the content of the items to be bound (binding deficits to a location can be more prominent for orientation rather than color) [40].

Increased *guessing* has also been reported as a cause of lower WM performance in healthy ageing, especially in conditions with high cognitive load, when more items need to be stored in memory [39]. However, other studies failed to find such an increased rate of guessing even in high memory load conditions [38].

A decline in memory *precision* has also recently been linked with physiological ageing, with load-dependent worsening of performance with higher cognitive load [39, 40, 41]. Age-related reduction in memory precision has also been associated with lower scores on established WM tasks such as digit span and spatial span, especially with backwards tasks that have higher cognitive demands [40].

The ability to efficiently *filter* irrelevant information is crucial for optimal use of WM resources, and its impairment has recently been proposed as an additional source of decline of WM performance in healthy ageing [42]. Filtering has been considered a core executive

function [45], and while executive functions might decline with ageing [46], evidence on decline of filtering abilities has been mixed [44]. A recent large meta-analysis on inhibitory deficits in elderly healthy controls showed that ageing alone does not impact all tasks requiring inhibition [47]. According to these data, some tasks such as the Stroop, Flanker and n-back task, do not seem to be particularly affected by age, while others, such as go/no- go and stop-signal tasks demonstrate an age-related deficit in inhibitory processes.

Some evidence also reveals that older adults show poorer recall performance if distractors are presented during maintenance rather than at encoding [42]. Interestingly, data on whether these filtering deficits, either at encoding or maintenance, are unique to healthy ageing or are also present in populations with known memory deficits such as AD patients, or those with more prominent deficits in executive functions, such as Parkinson's Disease (PD) patients, are very scarce.

While all these five factors (reduced WM capacity and memory precision, increased misbinding and guessing, and impaired filtering ability) have been shown in isolation to play a role in healthy ageing, in **Chapter 2** I will explore the complex relationship between them. I will use as control groups young healthy participants, patients with AD and PD, to test differences across these factors in physiological and pathological ageing. I will also investigate whether presenting a distractor during the maintenance or encoding phase plays a detrimental role specifically in healthy ageing, in AD or in PD patients.

1.2.2 Traditional neuropsychological hallmarks of early AD: the role of LTM

AD is a neurodegenerative disease characterized by accumulation of amyloid plaques and neurofibrillary tangles at neuropathology and is the most common cause of dementia worldwide [48]. The accumulation of neurofibrillary tangles is not evenly distributed across the cortex, but rather starts in the transentorhinal cortex (Braak stages I and II), to then spread to other limbic structures such as the enthorinal cortex and the hippocampus (Braak stages III and IV) and then more widely to isocortical areas, i.e., most areas of the brain, which have six layers of nerve cells (Braak stages V and VI) [49]. The spatial progression of these changes well reflects the neuropsychological hallmarks of AD across its severity continuum, in line with the theory of *selective regional vulnerability* of neurodegenerative diseases, where the molecular biology is linked to the behavioral phenotype through vulnerability of specific cell populations, with the topography of the cellular dysfunction determining the clinical phenotype [50].

Initial prominent involvement of the transenthorinal cortex is associated with early episodic memory deficits, which have traditionally been regarded as a constituent feature of AD. With disease progression, there is involvement of high-order association isocortical areas, which accounts for progressive impairment of additional cognitive domains, including visuospatial abilities (e.g., parietal lobe), executive function (e.g., prefrontal cortex), and even semantic memory (e.g., anterior temporal cortex), culminating in the multidomain dementia syndrome seen in later stages of the disease [51]. Conversely, neuropathological amyloid burden does not correlate tightly with the severity or the duration of dementia [52]. Indeed, in regions of early amyloid deposition such as the medial temporal lobe (MTL),

pathological studies reveal that the amyloid burden reaches a plateau in the preclinical phase of the disease, even if in-vivo neuroimaging with amyloid tracers is able to detect progression of amyloid load over time [53, 54].

While in the full-blown dementia stage an impairment across multiple cognitive domains is observed, a large number of investigations have tried to dissect what is the most specific test to differentiate AD from other forms of dementia, as well as the most sensitive test to detect AD in its earliest phases. Clinical criteria for AD diagnosis state that delayed recall and free and cued recall show the highest specificity and sensitivity [55, 56]. Free and cued recall procedures are based on the assumption that in MTL amnestic syndromes such as in typical AD, encoded information is lost because of a deficit in the consolidation processes of the memory trace, i.e., *storage*. Therefore, it cannot be recollected even in conditions that aid memory at the time of *retrieval*, as category cues. This can be tested by making sure effective *encoding* of the memorandum during the study phase has been achieved, so that poor encoding does not act as a confounder at time of retrieval.

This pattern is different compared to frontal lobe syndromes, where encoding is often defective due to attentional deficits, and therefore, if encoding is controlled, cues are able to aid memory at the time of retrieval. However, more recent evidence, including a large metanalysis, challenge the superiority of these tests, and found that delayed recall did not predict progression better than immediate recall, and that there was no major difference between free and cued recall [57, 58]. Overall, whilst the testing paradigm chosen depends on set-up of the specific center, a simple and relatively quick global test of cognition as the ACE-III (Addenbrooke's cognitive examination III) [59] also yields excellent overall accuracy,

has been shown to be comparable to the abovementioned tests [58, 59, 60] and was the clinical test chosen for my thesis.

It is also important to mention that most of these findings relate mostly to amnestic, typical AD phenotypes, but newer criteria of AD include a biological definition of the disease, where regardless of the neuropsychological profile, evidence of accumulation of tau and amyloid must be present [61]. Indeed, AD also encompasses rarer variants as Posterior Cortical Atrophy (where there is a greater early impairment of visuospatial functions) [62], Logopenic variant (with prominent deficits in selective language abilities such as naming and repetition) [63], and Behavioural/dysexecutive variant (with severe initial changes in behavior and executive functions) [64]. These show how an impairment in episodic LTM, despite being an important hallmark of AD, is neither necessary, nor sufficient to make a diagnosis of AD.

1.2.3 Visual Short-Term Memory impairment in patients at risk of developing AD

Early studies showed that atrophy in the MTL, and particularly in the hippocampus, is correlated with impairment in declarative episodic LTM, but not with verbal and visual STM [65, 66, 67, 68]. While the role of LTM impairment in AD has been extensively discussed, interestingly, one of the first accounts of the importance of STM comes from patients with AD. The original experiment by Baddeley and colleagues showed that STM, and particularly the central executive system, is affected early during the course of the disease [69]. In this study, patients showed prominent impairment in a task that required multitasking, and therefore engaging the central executive system, compared to when they were tested only

on one task, even if the level of difficulty was matched between patients and controls, thereby excluding an overall effect of global impairment.

Further evidence has shown that also the initially spared phonological loop (which is usually indexed by the forward Digit Span) can be impacted later on during the course of the disease [70]. Moreover, the visuospatial sketchpad (usually measured by the Corsi Block Tapping Task) is impaired in earlier stages compared to the phonological loop in AD patients [71, 72], probably because the function of the visuospatial sketchpad seems to be more dependent on an intact central executive system compared to verbal tasks [71]. Backwards verbal and spatial WM tasks, such as backwards Digit Span and backwards Corsi Block Tapping Task, also require a higher engagement of the central executive system compared to forward tasks and can already be affected when the phonological loop is still intact [72]. The episodic buffer seems also impacted in more advanced stages of the disease [73].

Regarding WM memory deficits in AD, one feature that stands out as being particularly fragile in this patient population is *memory binding* [44]. Indeed, increased misbinding has been described in patients with AD using either change-detection or delayed reproduction tasks [44, 74, 75, 76, 77, 78]. Binding deficits seem to be more specific to AD compared to other types of dementias [75], or to other neurodegenerative disorders such as PD, where an increased amount of random guessing rather than misbinding is often seen [76]. Increased misbinding seems to affect patients with young-onset, familial AD (FAD) as well as individuals with late onset, sporadic AD [76, 77, 78].

Strikingly, asymptomatic carriers of causative genes for FAD as presenilin 1 (PSEN1) or amyloid precursor protein (APP) might show normal performance on standard tests of

cognition and overall proportion of items identified in tests of VSTM, while already having increased rates of misbinding (**Figure 1.5**) [77].

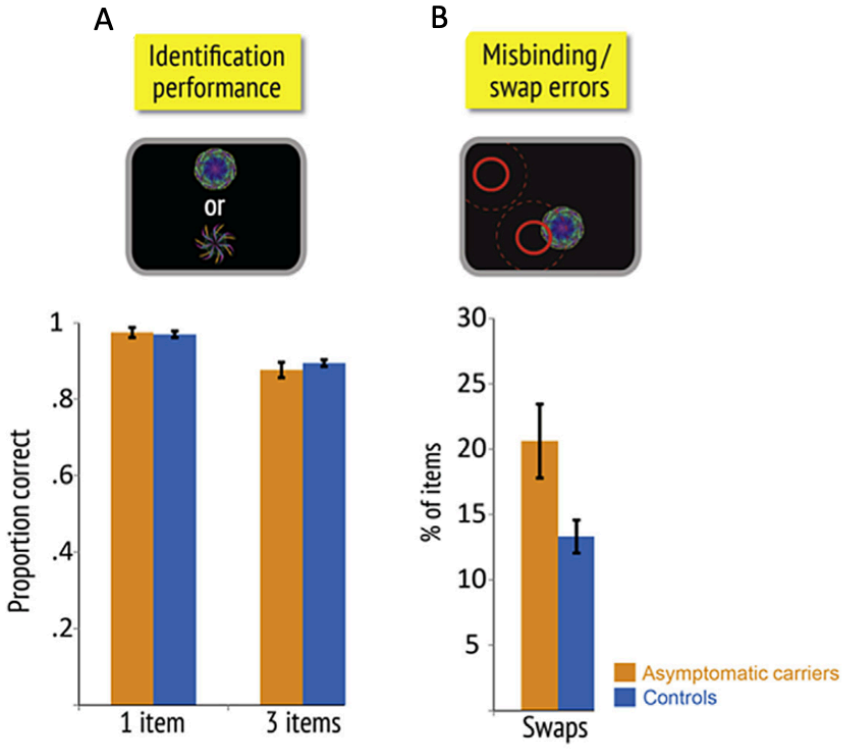


Figure 1.5 | Misbinding in asymptomatic mutation carriers for FAD genes

Adapted from Liang et al [77]. Asymptomatic carriers of FAD genes (APP and PSEN1) showed normal overall identification performance (panel A), but increased rates of misbinding (panel B).

This does not seem directly applicable to all genetic risk factors for sporadic AD, such as carrying the apolipoprotein E4 (APOE4) allele, where the evidence is more mixed [79, 80, 81]. Some evidence shows no increased misbinding in asymptomatic APOE4 carriers [79], while others highlight a paradoxical behavioral advantage in VSTM tasks, particularly at shorter delays [80], but no benefit on LTM tasks [81].

However, little is known about the selective vulnerability of binding across different clinical stages of AD, from subjective cognitive impairment (SCI), mild cognitive impairment (MCI) through to dementia. Some evidence gained from change-detection tasks shows that patients in the MCI stage already have lower performance in tasks requiring VSTM binding [82, 83], but to date this has not been investigated using the mixture model. Also, most studies compared a single cohort of AD patients, whether young or old, to healthy age-matched controls, but data looking at differential impact of age and pathology on VSTM binding is currently lacking. Additionally, most investigations are cross-sectional, and longitudinal data on patients with AD, or at risk of developing AD, are very scarce.

As discussed above, an essential component of VSTM performance is distractor filtering, which relies on an intact central executive system. Therefore, findings of impaired filtering abilities in AD patients are far from unexpected [84, 85, 86, 87, 88]. However, data on whether a failure in distractor filtering in this population is associated with increased misbinding, increased guessing, or reduced memory precision are lacking.

In **Chapter 2** I will focus on filtering abilities in AD patients, using mixture modelling. In **Chapter 3** I will investigate primarily the role of central executive system and visuospatial sketchpad in patients in different stages of AD, while also relating performance to global measures of cognition, e.g., ACE-III scores, and tasks that engage primarily the phonological loop, e.g., Digit Span. I will also look at whether mixture model metrics, including misbinding, are selectively affected in different AD stages compared to healthy controls, and how these metrics change longitudinally.

1.3 Brain Imaging correlates of Visual Short-Term Memory

1.3.1 Functional neuroanatomy of WM

The anatomical localization of working memory has been a subject of debate throughout the years due to its complex architecture, which requires a flexible interplay between attentional, mnemonic, perceptual and visuospatial systems.

Several different functional magnetic resonance imaging (fMRI) studies investigating VSTM, including a large metanalysis, showed a common activation of a fronto-parietal network including frontal regions as the premotor cortex, cingulate cortex, frontal pole, dorsolateral prefrontal cortex (DLPFC) and ventrolateral prefrontal cortex (VLPFC), parietal regions including posterior parietal cortex and especially the intraparietal sulcus (IPS), areas within the occipital cortex including lateral occipital cortex (LOC), and also the cerebellum [89, 90, 91, 92].

As a conceptual framework for understanding visual processing, both monkey and human studies support the presence of distinct anatomical pathways, an occipito-temporal pathway, or ventral stream, which is responsible for object recognition within an allocentric framework ('what'), and an occipito-parietal, or dorsal stream, which codes spatial location ('where') in egocentric coordinates in order to visually guide behavior and motor planning [93]. Both visual systems are structured hierarchically, as late anterior-rostral areas decode more complex stimuli compared to early posterior-caudal locations within the occipital cortex, both share extensive reciprocal connections, and terminate respectively in the ventrolateral prefrontal cortex and dorsolateral prefrontal cortex [94]. Electrophysiological

evidence in monkeys show that maintaining visual information throughout a short delay is associated with sustained activity in neurons in the inferior temporal cortex, posterior parietal and prefrontal cortex, which are respectively associated with maintaining information over object features, spatial localization, or are active across different stimuli [95]. Focusing selectively on lateral prefrontal cortex, it seems plausible that ventrolateral prefrontal cortex processes non-spatial, object-related information, while dorsolateral prefrontal cortex spatial information [96]. Lesion studies in humans have also supported this anatomical segregation [97]. Kleist reported that patients with focal brain lesions in different parts of posterior cortices demonstrated inability to perceive the form of an object (if the lesion was located ventrally) or its spatial relation to other items (if located dorsally) [98]. Dissociation between object and spatial location information has been replicated also in patients with anterior temporal lobe damage, where object recognition abilities are severely impaired, whilst spatial information is retained [99, 100, 101]. Conversely, dorso-parietal lesions have been shown to severely impact the ability to process and retain spatial information [102].

Human neuroimaging studies have been able to replicate the distinction between the 'what' and 'where' stream [103, 104, 105, 106]. According to this *domain-specific* VSTM segregation theory, early areas within the visual cortex respond non selectively to object identity and location. Further along the processing stream, object identity seems to localize specifically to the fusiform gyrus, inferior temporal cortex, and to ventral prefrontal areas. Conversely, spatial information activates dorsal occipital cortex, inferior and superior parietal cortex, intraparietal sulcus, precuneus and dorsal prefrontal cortex.

This view has been subsequently challenged by other groups, which failed to find such domain-specificity, and have proposed instead an alternative *processing-specific* theory of

VSTM, where ventrolateral prefrontal cortex is engaged in simpler processes such as the active encoding and retrieval of information and the dorsolateral PFC supports higher order control functions as monitoring and manipulation of information [89, 107].

Further evidence also dissected the differential contribution of sensory and WM demands by varying the delay by which the stimuli had to be held in memory, with the assumption that the longer the delay, the more the contribution of WM compared to pure perceptual activation of viewing the stimuli [108]. The authors of this study found that occipitotemporal areas including calcarine cortex, fusiform gyrus and parahippocampal gyrus were negatively correlated with delay, therefore possibly suggesting a more substantial involvement of perceptual quality of the stimuli, while prefrontal areas, including the inferior and anterior middle frontal gyrus, were more likely to be associated with increased visual working memory demands [108].

Prefrontal areas, particularly the DLPFC, have subsequently gained much attention in working memory maintenance mediated through attention [109, 110]. If persistent activity seen within DLPFC lasts throughout the whole experimental delay period, this could be consistent with the hypothesis that it is a site where active representations are stored long enough to guide appropriate responses or that they represent sustained activity driven by continuously focusing attention on relevant representations stored elsewhere. Experiments looking at the effect of increased memory load have yielded mixed results regarding the role of DLPFC, as some found increased activation with increased set size, while other failed to find such an association [111, 112], with some trying to reunite these apparently conflicting findings by postulating a role in chunking of information [113], or in integrating VSTM representations with goal-directed motor behavior [114]. This does not imply that these regions

are active and store information during passive maintenance, i.e., the delay period, as several lines of evidence failed to find evidence of decodability of stimuli in frontal cortices during the delay period [91, 115, 116]. According to these studies, it is however plausible that the task information and strategies are being stored and implemented in frontal regions.

These controversies touch upon another important aspect, which is whether different brain areas underly encoding, maintenance, and retrieval processes. Prefrontal areas seem particularly important at encoding, where deploying attentional resources is highly needed [90, 117]. Therefore, prefrontal regions, including the DLPFC, might provide top-down control over more posterior regions, where information is actually stored [110]. Several conflicting lines of evidence argued whether storage lies within early visual cortices or parietal regions. Some studies found that object information can be extracted during the delay period in the IPS, which also is increasingly activated with higher memory load, peaking at WM capacity [91, 118]. Other investigations found that even if regions like IPS show high activation at fMRI during the maintenance period, identity of a single object can be decoded during the delay period from early visual cortices (including V1-V4 and MT), without a concomitant delay-period activity [119, 120]. Therefore, activity in areas as IPS rather than explaining the passive storage of information per se, might represent directed attention towards a limited amount of information.

However, cumulative evidence gained through different studies suggest that during maintenance, visual information is segregated anatomically in the brain depending on the nature of the information, depending on whether is color, location, orientation, a bound feature or other more complex stimuli.

1.3.2 Neural circuitry involved in Filtering abilities

We saw in paragraph 1.1.3 that filtering occupies a special place within working memory, as requires the creation of a spatial map that will optimize actions according to task demands. A recent comprehensive review suggests that this process of prioritization involves several different core areas, including visual cortex and fronto-parietal networks [28]. According to the authors, at encoding visual information is processed by the visual cortex, which can also act as a site of maintenance if no other manipulation is needed. However, if an item needs to be preferentially selected amongst others, this will activate an attentional fronto-parietal network including the superior parietal lobule (SPL), IPS, the pre-supplementary motor area (pre-SMA), the anterior cingulate cortex (ACC) and the frontal eye fields (FEF). Subsequently, according to this model, frontal areas including VLPFC, insula and frontal operculum help 'reformatting' the information to create a preferential spatial map that can be directly used and activated when a new stimulus needs to be processed following prioritization.

While this gives us a good understanding of how the process of re-prioritization works, several questions on filtering processes remain unsolved. A line of complementary work suggests that while the prefrontal cortex is involved in maintaining prioritization rules and to a certain extent, the temporary storage of items as well, filtering and updating information can also involve regions within the basal ganglia, including globus pallidus, caudate and putamen [32, 121, 122, 123, 124, 125, 126]. This might be due to the role of the basal ganglia as hub for reinforcement learning, where these structures learn task-relevant rules and convey information to prefrontal cortices for selection and processing through intermediate structures such as the thalamus [127, 128].

The role of the basal ganglia as gatekeeper for preventing irrelevant information entering working memory and distractor manipulation has also been supported by evidence of impaired filtering and updating in PD patients and patients with structural lesions in the basal ganglia, which can be aided by dopaminergic supplementation [124, 129, 130]. Updating information compared to purely passive maintenance has been linked to structures in the DLPFC, as well as caudate, midbrain including the substantia nigra and ventral tegmental area [121]. Therefore, basal ganglia have been implicated in *input* as well *output* gating necessary to working memory filtering [32, 122, 126].

However, the neural underpinning of filtering abilities in healthy ageing and in AD is far less studied. It is also unclear if other brain structures that are involved in LTM as well as STM, such as the hippocampus, play a role in WM filtering. In **Chapter 2** I will investigate the brain anatomical localization of filtering abilities in healthy ageing, AD and PD, with a special focus on the hippocampus.

1.3.3 Brain correlates of misbinding, guessing and memory precision in healthy subjects

Dissecting the quality of behavioral responses through the mixture model allows us to extract different response dimensions. However, from an anatomical perspective, very little is known about the precise localization of constructs such as memory precision, misbinding and guessing. It is also unknown whether the topographic underpinning of these features changes with healthy ageing or neurodegenerative conditions that may present with impairment along one of these dimensions as AD or PD. Studies in young healthy subjects using a delayed reproduction task analyzed through mixture modelling suggested that these

constructs do map onto different brain areas [131]. A number of different studies have highlighted the importance of the hippocampus in binding features together in healthy subjects [132, 133, 134, 135]. Some evidence argues that tasks that use larger set sizes, and therefore exceed WM capacity, can cause misbinding even in healthy people [136], and that misbinding can increase with physiological ageing due to a dysfunctional hippocampus [134]. Another controversial question is whether hippocampal contribution to binding is essential at encoding or during maintenance. Some evidence shows that whilst is essential for encoding, it might not necessarily be implicated in maintenance of bound features [133].

Regarding memory precision, the hippocampus is considered to be essential in storing information with high fidelity [132, 137], which reconciles contrasting findings regarding hippocampal contribution to memory binding. Indeed, these can also be explained by a hierarchical impairment in binding abilities, whereas hippocampal damage does not necessarily cause increased misbinding in tasks that require only low-level or low-resolution bindings, but does impact performance when complex, high-resolution bindings are required [132, 138]. However, the hippocampus might not be the only brain area underlying memory precision. In a study by Richter et al, while the hippocampus was important for overall task accuracy, other areas such as the parietal cortex, particularly the angular gyrus, were implicated in remembering an item with high precision [131]. Other studies have found early visual areas to be associated with increased memory precision for single items, where the decodability of the stimuli varied according to WM load, with fewer items stored with higher memory precision [139]. Therefore, WM memory precision seems to be attributable to a large network encompassing different brain areas, at least in healthy subjects [140].

In **Chapters 2 and 3** I will investigate specifically the hippocampal contribution to VSTM metrics such as misbinding and guessing, in healthy elderly participants, patients with AD dementia, as well as patients at risk of developing AD. I will also explore in **Chapter 4 and 5** the whole-brain localization of misbinding, memory precision and guessing in healthy elderly subjects, as well as patients with AD using structural (Chapter 4) and diffusivity (Chapter 5) metrics.

1.3.4 Misbinding as hippocampal lesion model in patients with neurological diseases

We discussed in paragraph 1.2.2 how AD pathology starts from a very small region in the transentorhinal cortex, to then spread across the whole brain. In line with the concept of selective regional vulnerability, there is another neurological disease that comes to aid in the understanding of what might be happening in the earliest stages of AD, and that is limbic encephalitis (LE). Patients with LE have often clinical features including STM deficits, an acute confusional state, psychiatric symptoms and seizures [141]. Interestingly, short-term memory deficits are the most commonly reported symptom [142], and whilst considered a clinical hallmark of the disease, these are often overshadowed by other more explosive symptoms. VSTM binding seems particularly affected in these patients, while the correct representation of single features (color, orientation, location) may still be intact [143].

Beyond LE, several studies have found that in patients with hippocampal lesions, recollection of object or location might be spared, but binding of the two features might be affected [143, 144, 145, 146, 147, 148]. Findings of increased misbinding secondary to hippocampal dysfunction seem to be consistent across many different aetiologies, including as previously

mentioned, limbic encephalitis [143], but also surgical resection due to epilepsy surgery [146], infectious encephalitis [145], anoxia [147] and also atrophy secondary to neurodegeneration [77]. However, some studies have failed to find such deficits in binding abilities in patients with hippocampal damage [149], or found deficits only when the task exceeded working memory capacity, and therefore WM was insufficient to support performance [136, 150]. Interestingly, this capacity limit is similar to what was observed in patient H.M. (Henry Gustav Molaison) for verbal short-term memory, which was intact if the task did not require transcending WM capacity limits or couldn't be aided by continuous rehearsal [65, 66].

We have seen that multiple studies have investigated binding abilities in patients with hippocampal damage. However, not all of them have computed misbinding using the mixture model. Therefore, is still unclear to what extent only misbinding is affected in these patients, and whether other metrics are impacted as well. Most studies only explore overall task accuracy in binding paradigms, and focus on correct/incorrect responses, and therefore, quantification of guessing errors as conceptualized in the mixture model is less reported. There is some evidence that people with hippocampal damage might exhibit also higher levels of random guessing compared to healthy controls [78, 136], while other studies have failed to find increased guessing rates in patients with hippocampal atrophy as AD patients [76]. Moreover, while we argued that the hippocampus might be important for memory precision [132], some evidence shows comparable levels of memory precision in patients with hippocampal atrophy compared to healthy age-matched controls [76].

In **Chapters 2 and 3** I will focus on the role of the hippocampus in binding abilities in healthy ageing, AD and also in a control patient population without significant hippocampal atrophy as PD. In **Chapters 4 and 5** I will test the hypothesis of whether beyond the

hippocampus, other brain areas in the grey and white matter might be involved in VSTM binding.

1.4 Plasma biomarkers of Alzheimer's Disease

1.4.1 Why are biomarkers needed?

Biomarker, crasis for biological marker, is defined as 'a characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention' [151]. In AD, biomarkers are being used to measure specific pathophysiological changes associated with the disease, including accumulation of amyloid plaques and tau tangles, but also to gather information regarding concomitant processes including neurodegeneration, synaptic dysfunction, neuroinflammation and vascular changes [152]. A timely, etiological diagnosis of AD in clinical practice is essential to provide patients with diagnostic and prognostic information about their disease [153]. It also gives patients the opportunity to join clinical trials, which are aimed at selectively targeting biological pathways measured by specific biomarkers.

The need for inclusion of biomarkers when testing new therapeutic interventions results from the low accuracy of clinical diagnosis in predicting underlying brain pathologies. It is estimated that around 25% of patients fulfilling clinical criteria for AD enrolled in clinical trials do not have evidence of Aβ pathology [154, 155]. In APOE4 noncarriers this percentage can rise up to 36.1% of enrolled patients having a negative amyloid positron emission tomography (PET) scan, compared to 6.5% of APOE4 carriers [155]. At autopsy, one study showed that up to 37% of patients with a clinical diagnosis of AD showed minimal amyloid

plaque accumulation in the brain [156]. Therefore, supportive biological evidence of amyloid and tau pathology is warranted, to mitigate the risk of trial failure due to the inclusion of patients without the pathology of interest, and to dynamically monitor effects of the drug *in vivo* [157].

The recent advent of anti-amyloid therapies with potential disease-modifying effects, such as Aducanumab [158], and the more promising Lecanemab [159], has highlighted the need of a biomarker-based diagnosis, to stratify which patients with a clinical diagnosis of AD could benefit from early interventions aimed at reducing amyloid burden in the brain. Therefore, an etiological diagnosis of AD might be essential soon not only for research purposes or in the context of clinical trials, but also in everyday clinical practice.

1.4.2 The 'ATN' system

Since the introduction of the ATN (Aβ amyloid, tau and neurodegeneration) staging system in 2018 [61], the definition of AD has shifted from a clinical syndrome to a biological construct, where the presence of both amyloid and tau, rather than a multidomain cognitive impairment, constitutes the hallmark of AD. That consensus paper marked a watershed moment, modifying the role of biomarkers from being satellite, supportive features for etiological diagnosis of AD, to being essential, integral diagnostic features. Therefore, whilst cognitive impairment has been historically a key feature of all previous diagnostic criteria for AD [56, 160, 161, 162], according to the ATN criteria it is not necessary to support the diagnosis of AD. Conversely, AD is defined by the *in vivo* detection of the accumulation of amyloid plaques and neurofibrillary tangles, which can occur even in asymptomatic individuals (Preclinical AD), whilst the presence of mild cognitive impairment (Prodromal

AD) or dementia (AD with dementia) only marks the progression of cognitive impairment but is not necessary to make a diagnosis of AD.

Recognized biomarkers for AD diagnosis are divided into three main categories according to that framework:

- A: Amyloid markers: reduced Aβ42 or Aβ42/Aβ40 ratio in the cerebrospinal fluid (CSF), or positive amyloid PET.
- T: Tau: increased phosphorylated tau (p-tau) in the CSF or positive Tau PET.
- N: Neurodegeneration: atrophy on anatomic MRI, positive fluorodeoxyglucose (FDG) PET and increased total tau (t-tau) in the CSF.

AD is defined by positivity of markers of amyloid and of tau accumulation (A+T+), regardless of positivity on markers of neurodegeneration (N+). Positivity of only amyloid markers (A+T-N-) is termed 'Alzheimer's pathologic change', while if markers of neurodegeneration are also present (A+T-N+) that is considered to represent 'Alzheimer's and concomitant suspected non Alzheimer's pathologic change'. All these four entities (A+T+N-, A+T+N+, A+T-N-, A+T-N+), constitute the Alzheimer's Disease Continuum.

Besides the clear importance of biomarkers in the diagnosis of AD, the ATN criteria have also drawn considerable criticism, with competing expert views (as in Dubois et al) [163] arguing that cognition cannot be completely factored out from the diagnosis of AD, and that asymptomatic people with positive A+T+ biomarkers should be considered 'at risk' of developing AD rather than having AD.

1.4.3 Relationship with cognitive impairment

Indeed, there is a huge interindividual variability in the lag time between evidence of AD pathophysiological markers and the emergence of cognitive impairment [164], and not all people that have positive biomarkers will develop a multidomain cognitive impairment, which is ultimately what is clinically relevant to the individual. This is probably mediated by differences in brain resilience to pathological insults, cognitive reserve, and lifestyle risk factors [165, 166, 167].

As many previous autopsy studies have shown, it is possible to have considerable AD pathology in the brain, particularly amyloid deposition, without having developed cognitive impairment [168, 169]. The relationship between cognitive impairment and presence of amyloid plaques and tau tangles is also far from linear, with more chances of mismatch between expected cognitive impairment and underlying neuropathological burden in the early phases of the disease [168]. Some evidence suggests that most people with preclinical AD will not develop AD dementia during their lifetime [170]. Additionally, there is a low predictive accuracy in cognitively normal people (5 - 42%) with positive biomarkers in converting to MCI or AD, while this percentage rises to 84.7% if MCI is present [170]. For predicting future development of AD dementia, cognition is also the first factor in terms of importance that should be considered as add-on measure to etiological biomarkers of AD [171]. However, standard cognitive tests used in clinical settings such as the Mini Mental State Examination (MMSE) or Montreal Cognitive Assessment (MoCA) have low sensitivity in discriminating between healthy subjects and MCI patients [172], which represents the phase where therapeutic interventions are more likely to succeed.

1.4.4 Plasma biomarkers breakthrough

A recent game-changer has been the advent of blood-based biomarkers. The measurements of amyloid and tau included in the 2018 criteria are either expensive (PET imaging) or invasive, as they require a lumbar puncture to be performed (CSF Aβ42, Aβ42/Aβ40 ratio, t-tau and p-tau) [173]. After decades of failures to reliably identify AD specific biomarkers in blood, in 2016 the first amyloid immunoassay using the Simoa platform (Quanterix, Lexington, MA, USA) showed promising results in predicting CSF and amyloid PET levels [174]. Following that initial success, mass spectrometry methods have been able to detect Aβ amyloid in plasma with high precision [175, 176, 177]. Fully automated immunoassays are now available to measure not only amyloid species, but multiple plasma biomarkers of AD [178, 179].

Blood-based biomarkers offer several advantages. Firstly, they are safe and easy to measure, and cost-effective [178]. Deploying blood-based biomarkers would significantly cut diagnostic costs for the healthcare system [153, 180]. It could also significantly reduce screening cost for inclusion in clinical trials and be used to monitoring effects of the studied compound over time [153, 178]. Their reduced cost and accessibility are highly relevant not only in research centers, but also in the context of primary care, where the rate of misdiagnosis is even higher than specialized dementia centers [181]. Secondly, they have been shown to be both sensitive and specific, i.e., able to detect the disease early and to discriminate it well from other neurodegenerative diseases [180, 182, 183]. Plasma biomarkers follow similar trajectories compared to their CSF homologues [184], and they have also been extensively validated against PET imaging and autopsy evidence [174, 182, 183, 185, 186, 187. 188]

However, they are still relatively new if compared to tools such as amyloid and tau measures in CSF and PET, and evidence of confounding factors that break the linear association between levels of biomarkers in the plasma and in the brain is only now starting to emerge [189]. Data on test-retest reliability is also scarce [190, 191].

In **Chapter 6** I will explore the relationship between plasma biomarkers of AD and several digital cognitive metrics, measuring not only VSTM but also executive functions, processing speed, visuospatial abilities and LTM recall. My aim is also to test which combination of metrics best discriminates between healthy controls and patients with AD dementia. I will also evaluate the test-retest reliability of plasma biomarkers and digital cognitive metrics over a short time frame (0 - 3 - 6 months).

Chapter 2. Filtering at Encoding and Maintenance modelled through Mixture Model metrics •§#*

Abstract

The differential impact on working memory (WM) performance of distractors presented either at encoding or during maintenance was investigated in AD, PD and healthy ageing. Across three studies, we enrolled 28 AD and 28 PD patients, 28 elderly (EHC) and 28 young healthy controls (YHC). All participants performed a delayed reproduction task, where they reported the orientation of an arrow from a study set of either two or three items, with a distractor present either at encoding or at maintenance. Mean absolute error (the difference between probed and reported orientation) was calculated as an analogue measure of WM. Additionally, mixture modelling metrics, i.e., memory precision, target detection, misbinding, and guessing, were computed. MRI data was also acquired in AD, PD and EHC participants, and hippocampal volumes were extracted to test whether these WM constructs were related to hippocampal integrity. Healthy ageing was associated with relatively good filtering abilities both at encoding and during maintenance. However, AD patients exhibited significant filtering deficits specifically when the distractor appeared *during maintenance.* In contrast, PD patients showed preserved filtering abilities, both at encoding and maintenance. Healthy ageing and AD were associated with higher rates of both misbinding and guessing, as well as lower target detection, and memory precision. However, in healthy ageing there was a prominent decline in VSTM memory precision, whilst in AD lower target detection and higher guessing were the main sources of error. Conversely, PD was associated only with higher guessing rates. Hippocampal atrophy was significantly

correlated with filtering ability during maintenance – but not at encoding – as well as with overall mean absolute error, target detection, guessing and misbinding. These findings demonstrate how healthy ageing and neurodegenerative diseases exhibit distinct patterns of WM impairment, including differential effects on filtering irrelevant material presented at encoding and maintenance.

2.1 Introduction

Working memory is a limited data processing system responsible for encoding and briefly maintaining information, so that it can still be accessed in the absence of inputs from the environment [11, 192, 193]. Encoding processes are considered to transform transient perceptual representations into a more durable state, while maintenance mechanisms allow information to be retained so that it can be used to guide behavior over brief delays [194, 195]. WM performance has been observed to decline with normal ageing [196, 197, 198], and is further impaired in several neurological conditions such as AD [69, 72] and PD [199].

In *healthy ageing*, several distinct components that contribute to declining performance in visual WM have been identified: a reduction in the number of items that can be stored, or WM capacity [35]; a deficit in maintaining the associations (bindings) between individual object features [37] resulting in a higher proportion of misbinding or "swap errors" [38]; an increase in random guessing [130, 200]; as well as a decrease in response precision [40, 41]. In addition, since WM is limited in capacity, an important feature of its function is that items appear to compete for access to it. Hence, the ability to efficiently *filter* out irrelevant information is crucial for optimal use of WM resources, and impairment in filtering ability

has recently been proposed as an additional source of decline of WM performance with healthy ageing [42].

Indeed, the ability to successfully filter out distractors, is now considered to be one of the major processes that impact upon WM capacity [27, 201]. Filtering out distractors at encoding versus maintenance poses different challenges [202]. Attentional processes are thought to be crucial, both in successfully encoding information but also to avoid unwanted attentional capture during maintenance, and their integrity might impact on the ability to efficiently manipulate information and therefore filtering abilities [28, 203, 204, 205, 206]. Filtering *at encoding* has been envisaged as deployment of a flexible gating system, allowing at the same time relevant targets to enter WM while keeping distractors out [32, 33]. Filtering *during maintenance* on the other hand might rely on the ability to keep the gate shut firmly once targets are successfully encoded, in order to prevent any further information entering WM and thereby corrupting task-relevant information [28, 32, 204, 206].

The effectiveness of filtering can be tested experimentally by studying the impact of distractors at encoding and during maintenance. Some experimental evidence suggests that these two processes might reflect two separate cognitive mechanisms, with the performance of each being strongly correlated to total WM capacity [33]. The integrity of fronto-striatal networks has been commonly associated with performance in filtering paradigms [122, 207, 208], and has been linked to brain structures involved in WM storage including the parietal cortex [122, 209]. However, the contribution of the hippocampus, despite being crucial for short-term and long-term memory function [77, 210, 211], has received little attention in relation to filtering abilities in WM.

Elderly participants seem to be particularly prone to errors with increasing memory load due to a reduction in their WM capacity [35], which has been linked to a decline in their filtering abilities [42]. Some evidence also suggests older adults show poorer performance if distractors are presented during maintenance rather than at encoding [33].

Why might this be the case? Are elderly participants worse at filtering during maintenance simply because of decreased memory fidelity? Or do they fail to successfully form a bound, stable representation of an object, and therefore have a higher probability of misbinding features belonging to different items in memory? Alternatively, is their lower performance simply due to information exceeding their WM capacity so their responses include more random guessing? These questions can be addressed using delayed reproduction paradigms that ask participants to reproduce from memory, using a continuous analogue response space, specific features of the items they viewed at encoding [11]. This type of task allows for dissection of the contributions of different types of error to overall recall performance. Both an age-related decline in memory precision [40, 41] as well as increased guessing [39] have previously been reported using such paradigms. Evidence of increased misbinding has been more mixed, with some data showing higher rates in healthy ageing [38], while other have not, at least with more traditional change detection visual WM paradigms which rely on binary correct / incorrect response measures [43], rather than reproduction of item features which allows for more continuous measurement of error.

Previous evidence from change-detection tasks has suggested that maintenance of bound features, such as color and shape, might be selectively impaired in AD [44] rather than in healthy ageing. Increased misbinding in patients with both sporadic, late-onset AD and

familial AD has now been reported using both change-detection and delayed reproduction tasks [75, 76, 77, 212], and specifically linked to hippocampal dysfunction [132, 143].

An increased vulnerability to distractors may be a feature that is detectable from the earliest phases of AD [84, 85, 86, 87, 88]. Other studies report individuals with AD to be less impacted on simple perceptual filtering tasks at encoding, but struggle to maintain the task instructions over time, pointing towards a deficit of goal-maintenance processes [213, 214]. Crucially, comparative data on distractor filtering if presented at either encoding or maintenance within the same study are lacking. Moreover, it is unclear whether hippocampal integrity might be related to filtering ability, either at encoding or during maintenance.

PD is also associated with deficits in visual WM. Different paradigms have provided evidence of reduced WM capacity [215] as well as greater levels of guessing responses, rather than misbinding [76, 216]. Further, PD patients have been shown to exhibit deficits in distractor filtering when the distractor is present at encoding, which impacts negatively on their WM capacity, and might be linked to basal ganglia and prefrontal dysfunction [122, 215]. However, other studies have shown that PD patients also fail to efficiently manipulate information over the maintenance period, such as when updating information is required, which can selectively be modulated by replenishing their dopaminergic tone [130].

In summary, healthy ageing, AD and PD have been associated with reduction of filtering abilities, which impact negatively on their overall WM capacity and performance. Data directly comparing filtering at encoding and maintenance in these groups are currently lacking. Moreover, to the best of our knowledge, no study has hitherto attempted to delineate

the source of increased filtering errors, both at encoding and at maintenance, in these patient groups, or related WM filtering ability to hippocampal integrity.

Here, using a unified delayed reproduction task and computational modelling of sources of error, we investigated whether filtering deficits in healthy ageing, AD and PD are related to a decline in memory precision, a decreased ability to detect the correct target, an increased likelihood of misbinding or greater random guessing. We also examined whether these performance metrics are associated with hippocampal integrity, given the crucial role of the hippocampus in the interplay between attention, WM and high-resolution memory binding [132, 217, 218], and involvement in both encoding and maintenance processes [210, 218, 219]. First, we established normal performance in young and older healthy controls (Experiment 1). Then, we studied filtering ability in AD (Experiment 2) and PD (Experiment 3).

2.2 Experiment 1 | Filtering at Encoding and Maintenance in healthy ageing

2.2.1 Methods

Participants

28 young (YHC) and 28 elderly healthy controls (EHC) were recruited (YHC: 18-35 years and EHC: 50-90 years), through either the departmental online participant recruitment system (SONA) for YHC, or open day events for EHC. All of them performed the Filtering

at Encoding and Maintenance Task (**Figure 2.1**). We also collected a measure of global cognition – the Addenbrooke's cognitive examination (ACE-III) [27], referred to as ACE – in the elderly group as baseline screening for cognitive impairment. Participants who reported any psychiatric or neurological illness, were on psychoactive drugs or scored below the cut-off for normality (88/100 total ACE score) were excluded from the experiment. Participants' demographics and tests results are presented in **Table 2.1**. Between-group differences were assessed with Mann-Whitney U test for continuous variables and χ^2 test for categorical variables. Statistical significance was set as p-value < 0.05, and this threshold will be used for all subsequent Chapters.

	YHC	EHC	p-value
Age	24.8 (5.1)	70.36 (7.7)	*< 0.001
Gender (M/F)	10/18	16/12	0.108
Education	17.36 (4.2)	17.75 (6.5)	0.610
Handedness (R/L)	26/2	22/6	0.127
ACE	-	97.82 (1.8)	-

Table 2.1 | Demographics for young and elderly healthy controls

Age, education, and ACE are presented as mean (standard deviation). Gender: M = male, F = Female. Education: years of full-time education. Handedness: Right-handed, Left-handed.

All participants had normal or corrected-to-normal vision acuity and no color blindness. They all gave written informed consent prior to the start of the experiment. Ethical approval was granted by the University of Oxford Medical Sciences Inter-Divisional Research Ethics Committee (MS IDREC).

Filtering at Encoding and Maintenance Task

The Filtering at Encoding and Maintenance Task (**Figure 2.1**) had 5 different conditions. A written instruction cue was presented prior to each trial to explain what was required in that particular trial.

In **Set Size 2 (SS2)** and **Set Size three at Encoding (SS3E)** conditions, participants were presented with either two (SS2) or three (SS3E) different coloured arrows (red, green or blue) and were instructed to remember their colours and orientations. After a blank delay interval (2000 ms), one of the target arrows (i.e., the probe) re-appeared in black in its original location but presented in a random orientation. Participants were instructed to rotate the probe using the left and right arrow keys on a keyboard back to the remembered orientation. In the **Filter at Encoding (FE)** condition three arrows were presented simultaneously, as in the SS3E condition, but now participants were asked to ignore one of them.

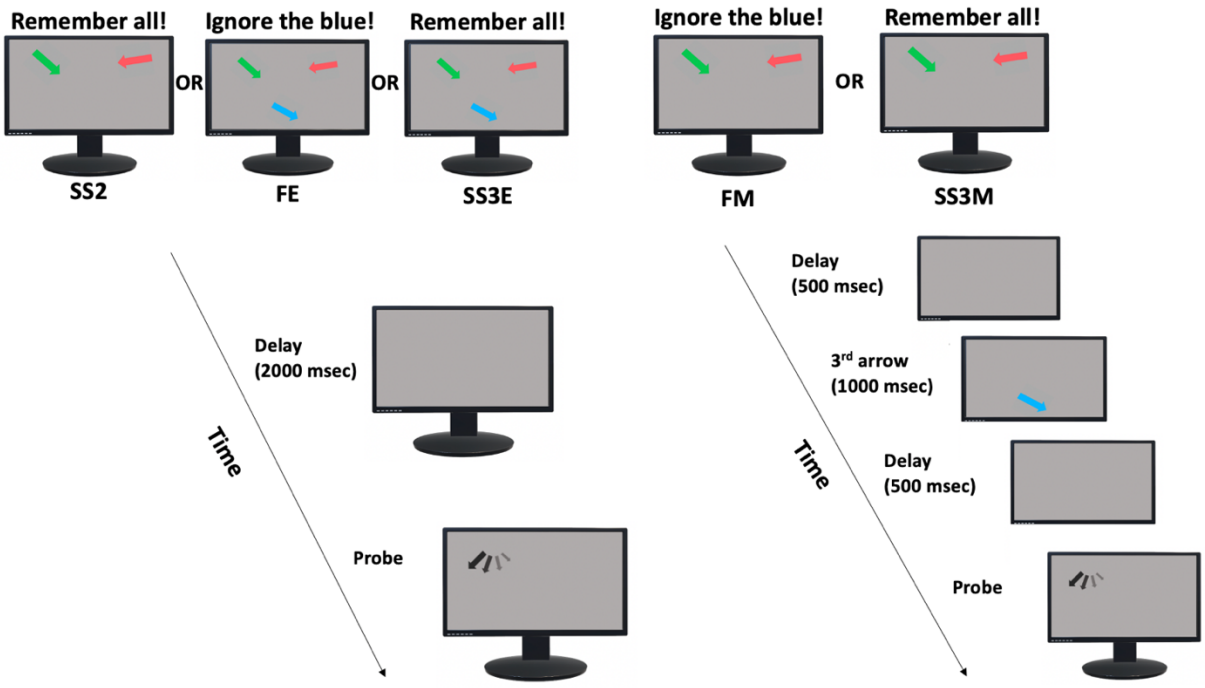


Figure 2.1 | Filtering at Encoding and Maintenance Task design

Participants were presented with either two (SS2) or three (SS3E) arrows at encoding and were instructed to remember the orientations of all arrows **(SS2 and SS3E)**, or to ignore one of them in the **Filter at Encoding (FE)** condition. After a blank interval (2000 msec), one of the target arrows (probe) re-appeared in black in its original location but in a random orientation, which the participants had to rotate to its remembered orientation. In the **Filter at Maintenance (FM**) and **Set Size three at Maintenance (SS3M)** the first two arrows appeared simultaneously and had to be remembered, and after a 500 msec delay during the maintenance period, a third arrow appeared for 1000 msec, which either had to be ignored (FM), or remembered (SS3M). After 500 msec the probe was displayed and had to be rotated to its original orientation.

In the **Filter at Maintenance (FM)** and **Set Size three at Maintenance (SS3M)** conditions, the first two arrows appeared simultaneously and had to be remembered. Subsequently, after a 500 msec delay during the maintenance period, a third arrow appeared for 1000 msec, which either had to be ignored (FM) or remembered (SS3M) and after 500 msec the probe was displayed. For all 5 conditions the instruction was presented at the start of the trial: "Remember all" in the SS2, SS3E and SS3M conditions or "Ignore the red/green/blue" in the FE and FM conditions.

Stimuli were displayed on a 24-inch LCD computer at a viewing distance of 60 cm. The orientations of the arrows were located along the boundary of an imaginary circle, with radius 8° in visual angle, centered on the cardinal center of the screen. Each arrow was 1.4° x 0.6° visual angle in size and pointing in random directions over 360 degrees. Their orientations were random but constrained to be separated by at least 20 degrees of one another. Stimulus colors were red (RGB: [255, 0, 0]), green (RGB: [0, 255, 0]) or blue (RGB: [0, 0, 255]). Which arrow had to be ignored on a given trial was chosen at random. All stimuli were presented on a grey background (RGB: [171, 171, 171]). The arrow used to probe a response was black (RGB: [0, 0, 0]). Participants performed a total of 120 trials (5 conditions x 24 trials per condition), divided into 12 blocks of 10 trials each. Trials from all 5 conditions were interleaved within each block.

Statistical analysis

All analyses were conducted using MATLAB 2019a and JASP (JASP team, 2020). As first step, we computed the ***mean absolute error*** **(MAE)** for each condition for each participant. This provides an index of raw error. We then fitted the mixture model of WM [17] to the relative error of each participant separately for each condition by fitting the following equation, as described in Chapter 1.1.2.

$$p(\hat{\theta}) = (1-\gamma-\beta)\Phi_\kappa(\hat{\theta}-\theta) + \gamma\frac{1}{2\pi} + \beta\frac{1}{m}\sum_{i}^{m}\Phi_\kappa(\hat{\theta}-\theta_i)$$

The mixture model allows modelling of responses in terms of probability of correctly identifying the target (**Target detection**, in red), random guessing (**Guessing**, in green), misbinding the features of an object with another among the probed items (**Misbinding**, in light blue), and extracts a measure of memory **Precision** (k in the equation) [17].

<u>Key parameters used for statistical analysis</u>

For MAE and all Mixture Model variables, frequentist mixed ANOVAs were performed to examine the effect of **Instruction** x **Condition** x **Group** where:

- **Instruction** - Refers to 3-item conditions with filtering (FE, FM) vs no filtering (SS3E, SS3M)
- **Condition** - All stimuli at encoding (FE, SS3E) vs one presented at maintenance (FM, SS3M)
- **Group -** YHC vs EHC

We also computed from the MAE a main effect of **Set Size**, both **at Encoding** and **at Maintenance** and their difference or **Set Size Cost**:

- **Set Size at Encoding =** SS3E - SS2
- **Set Size at Maintenance =** SS3M - SS2
- **Set Size Cost =** Set Size Encoding - Set Size Maintenance

Moreover, a main effect of seeing a to-be-filtered distractor, or **Filtering rate**, was calculated separately at Encoding and Maintenance ***from MAE*** as follows:

- **Filtering rate at Encoding =** FE - SS2
- **Filtering rate at Maintenance =** FM - SS2
- **Filtering rate Cost =** Filtering rate at Encoding - Filtering rate at Maintenance

If perfect filtering was achieved, Filtering rate would be zero. The higher the Filtering rate, the worse the performance. For the ANOVA, effect size was quantified using Eta Squared ($\eta 2$). According to the normality of the data, between-group differences for Set Size and Filtering rate metrics were assessed either with a t-test or Mann-Whitney U test, and a Bayesian independent sample analysis was performed. For the Bayesian analysis, we used the default priors for the effects (Fixed effects: r = 0.5, Random effects: r = 1).

2.2.2 Results

Mean absolute error (MAE)

With respect to the raw mean absolute error (MAE), the elderly group performed significantly worse than their younger counterparts across all conditions (main effect of Group [$F(1,48) = 40.46$, $^*p < 0.001$, $\eta 2 = 0.292$], **Figure 2.2**). Furthermore, both groups

were significantly worse in the conditions where all three items had to be remembered (SS3E and SS3M) compared to when they had to filter an item out (FE and FM, main effect of Instruction [F (1,48) = 17.75, *p < 0.001, η2 = 0.049], **Figure 2.2)**. Thus, it appears that both younger and older people were able to use the instruction to filter out an irrelevant item, either at Encoding or during Maintenance.

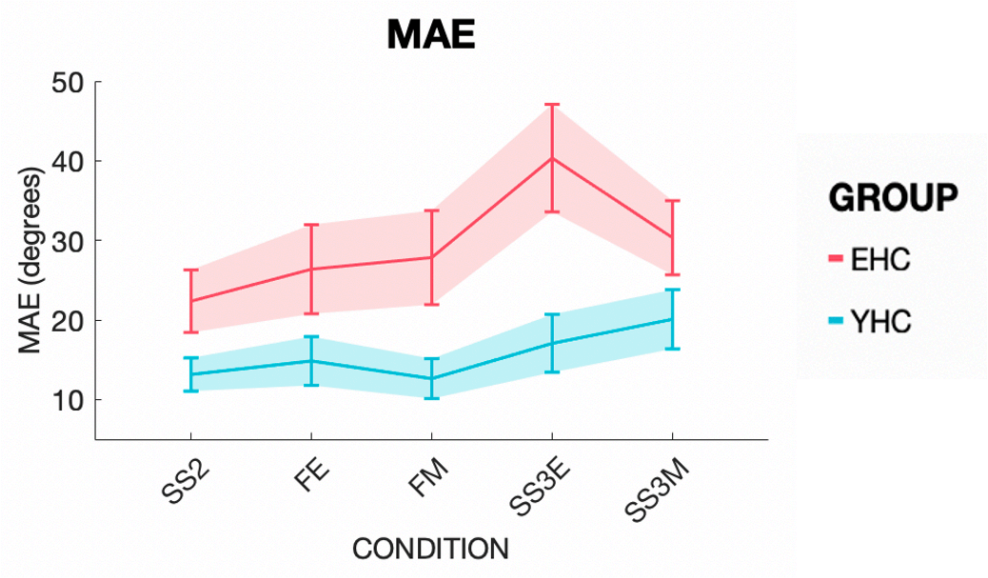


Figure 2.2 | Mean absolute error (MAE) across different working memory conditions for young (YHC) and elderly healthy controls (EHC)

Performance on the five different conditions: Set Size 2 (SS2), Filtering at Encoding (FE), Filtering at Maintenance (FM), Set Size three at Encoding (SS3E), Set Size three at Maintenance (SS3M). EHC in coral red, YHC in turquoise. Shaded areas represent confidence intervals (CI).

There was no significant main effect of Condition, i.e., comparison of performance when all 3-item stimuli were presented at Encoding (FE, SS3) versus when one was presented during Maintenance (FM, SS3M). This suggests that across the 2 groups, Encoding conditions were not more cognitively taxing than the Maintenance conditions. However, there was a significant 3-way interaction between Group x Instruction x Condition [(F (1,48) = 8.22, *p = 0.006, η2 = 0.014]. This would be consistent with the observation that EHC were significantly worse than YHC if the three items that needed to be remembered were

presented simultaneously at encoding (SS3E), compared to the if the third item was presented during maintenance (SS3M), (**Figure 2.2**).

Impact of Set Size at Encoding and Maintenance

EHC exhibited a higher Set Size effect at Encoding compared to YHC ($t(48) = -4.548$, $^*p < 0.001$), (**Figure 2.3, panel a)**. Thus, the impact of a third item was far greater in older people than younger participants when all the memoranda had to be encoded simultaneously. However, the two groups were not significantly different in Set Size at Maintenance ($t(48) = -0.936$, $p = 0.354$), i.e., when an item was added during maintenance it had similar impact on the MAE of both groups compared to their SS2 baselines (**Figure 2.3, panel b)**. To determine whether there was a greater Set Size effect at Encoding compared to Maintenance in elderly compared to younger participants we calculated the Set Size Cost. YHC had a significantly lower Set Size Cost compared to EHC ($t(47) = -3.391$, $^*p = 0.001$), (**Figure 2.3, panel c)**. Bayesian independent sample test also showed a $BF_{10} = 523.1$ for the Set Size at Encoding and $BF_{10} = 23.28$ for Set Size Cost, providing strong evidence for these Set Size effects.

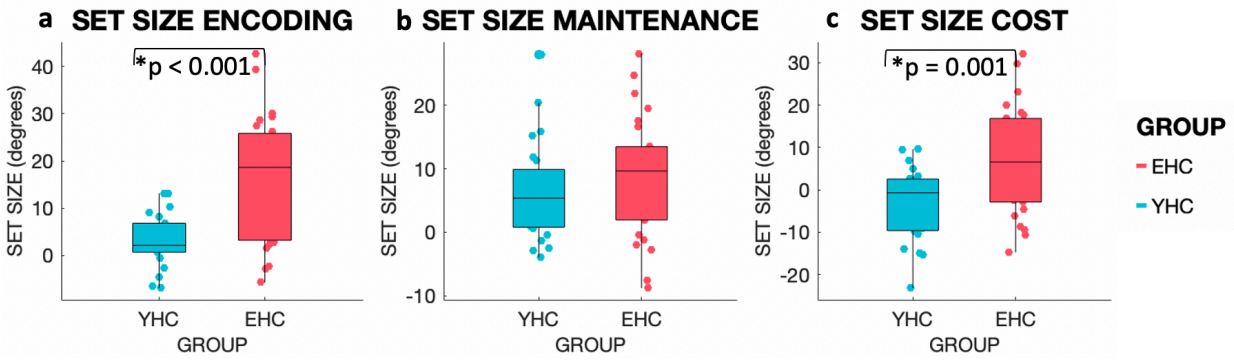


Figure 2.3 | Set Size at Encoding, Maintenance and their Cost in YHC and EHC

EHC in coral red, YHC in turquoise. Only statistically significantly p-values are shown.

Overall, these analyses support the conclusion that performance at remembering three items simultaneously was much worse for older individuals compared to younger people, while when the WM load of encoding 3 items was distributed across the initial encoding and maintenance phases elderly participants were able to cope with the increased cognitive demand as well as younger ones.

Impact of Filtering at Encoding and Maintenance

Both groups were able to efficiently filter out irrelevant items, shown by the Filtering rates at Encoding and Maintenance being not significantly different from zero in both groups (**Figure 2.4**). EHC and YHC were not significantly different in Filtering rate at Encoding (**Figure 2.4, panel a**). However, EHC showed worse performances (higher values) compared to YHC in Filtering rate at Maintenance ($U = 205$, $^*p = 0.038$), (**Figure 2.4, panel b**). Nevertheless, the Filtering rate Cost (the difference between filtering at encoding and maintenance) was not significantly different between the two groups, although performances of elderly participants varied far more than younger ones (**Figure 2.4, panel c**). Moreover, the Bayesian independent sample test showed a $BF_{10} = 1.047$, providing no definitive evidence for a significant difference in Filtering rate at Maintenance between the two groups tested here.

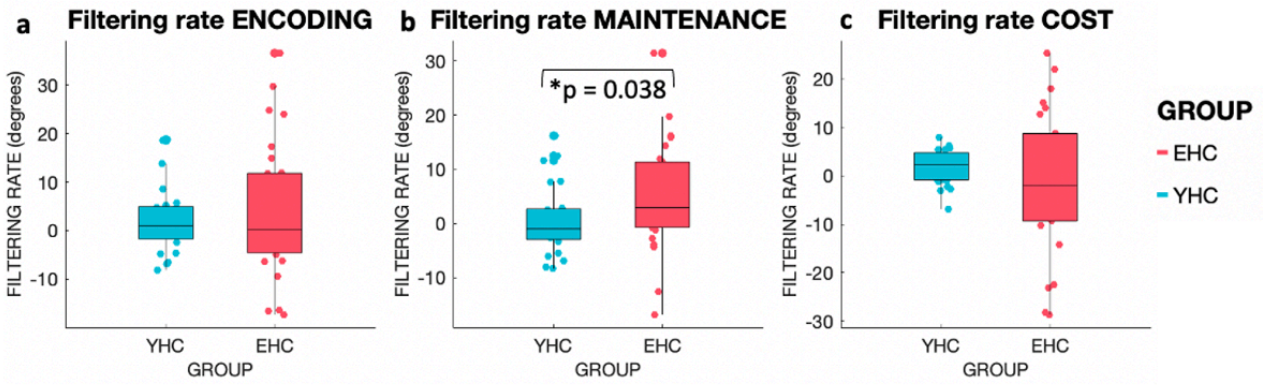


Figure 2.4 | Filtering rate at Encoding, Maintenance and their Cost in YHC and EHC

EHC in coral red, YHC in turquoise. Only statistically significantly p-values are shown.

Analysis of mixture model parameters

Precision

YHC had significantly higher Precision than EHC [$F(1, 42) = 59.06$, $^*p < 0.001$, $\eta 2 = 0.399$], but there were no significant effects of either Instruction or Condition or significant interactions, suggesting the difference in precision reflects a pure ageing effect (**Figure 2.5, panel a**). Therefore, EHC had significantly worse precision, regardless of the complexity of the WM task – with or without filtering - and regardless of the Condition (Encoding or Maintenance).

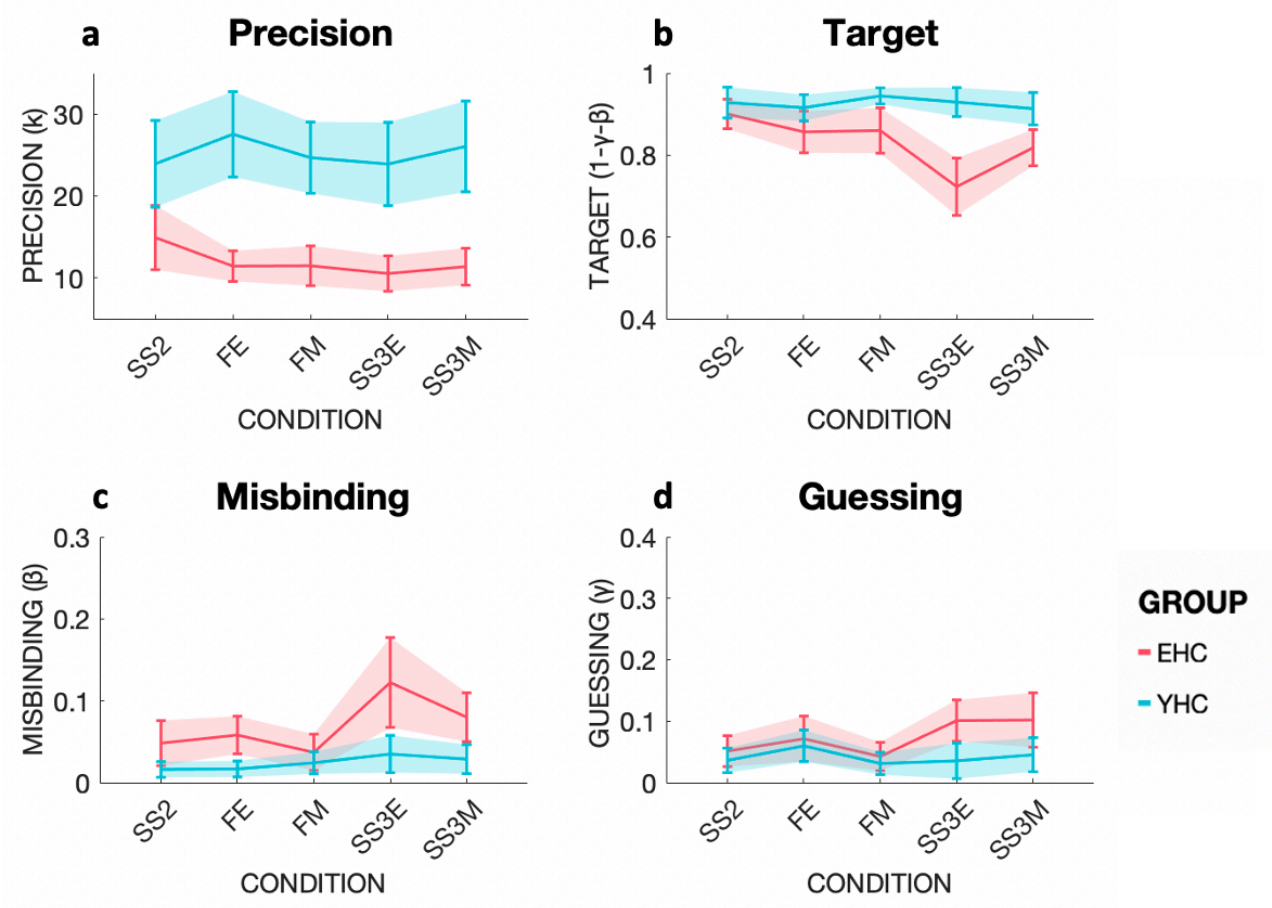


Figure 2.5 | Mixture Model metrics across conditions in YHC and EHC

EHC in coral red, YHC in turquoise.

Target detection

YHC showed significantly better target detection compared to EHC [$F(1, 48) = 25.38$, $^*p < 0.001$, $\eta 2 = 0.198$], (**Figure 2.5, panel b**). Across both groups, significantly more targets were detected in the filtering conditions (FE and FM) than in SS3 conditions (SS3E and SS3M), indexed by a main effect of Instruction [$F (1,48) = 10.01$, $^*p = 0.003$, $\eta 2 = 0.031$]. Furthermore, there was a main effect of Condition [$F (1,48) = 6.64$, $^*p = 0.013$, $\eta 2 = 0.011$], with significantly fewer targets identified in the conditions which required remembering or filtering during the Encoding phase (SS3E and FE) than during the Maintenance phase (FM and SS3M). EHC performed worse than YHC when all three items needed to be remembered compared to when an item had to be filtered out, as shown by a significant two-way interaction between Instruction and Group [$F (1,48) = 9.62$, $^*p = 0.003$, $\eta 2 = 0.030$]. Finally, EHC were worse at detecting targets when 3 items had to be remembered during the Encoding compared to the Maintenance phase, as there was a significant three-way Instruction x Condition x Group interaction [$F (1,48) = 4.39$, $^*p = 0.041$, $\eta 2 = 0.010$]. These results suggest that YHC and EHC found it more challenging to detect targets when all three items had to be encoded, and if the stimuli appeared at Encoding, with EHC paying the highest cost in terms of target detection.

Misbinding

There was a significant main effect of Group [$F(1, 43) = 15.29$, $^*p < 0.001$, $\eta 2 = 0.103$], and a main effect of Instruction [$F(1, 43) = 6.74$, $^*p = 0.013$, $\eta 2 = 0.038$], with significantly more misbinding errors committed by both groups in the SS3 conditions (SS3E and SS3M) compared to the filtering conditions (FE and FM), (**Figure 2.5, panel c**). Thus, even in healthy young individuals, misbinding rates were higher if three items had to be remembered

compared to when one could be filtered out, suggesting that filtering can reduce cognitive load and WM resources consumption.

Guessing

A main effect of Group ($F(1, 44) = 7.32$, $*p = 0.010$, $\eta 2 = 0.070$], no main effects of either Instruction or Condition, but significant 2-way Instruction x Group ($F(1, 44) = 4.33$, $*p = 0.043$, $\eta 2 = 0.016$] and Instruction x Condition [$F(1, 44) = 5.39$, $*p = 0.025$, $\eta 2 = 0.017$] interactions were seen at the ANOVA (**Figure 2.5, panel d**). This suggests that EHC guessed more than YHC, and more so in the more taxing SS3 conditions, independently of being presented during the Encoding or the Maintenance phase. Moreover, filtering during the Maintenance phase was more efficient in reducing guessing in both groups, compared to when filtering had to be carried out at Encoding. Therefore, EHC were more prone to respond with a random guess, often considered to be an index of no information about the object probed being retained, compared to YHC. Filtering helped reducing WM load in terms of reduced guessing across the two groups only if the distractor was presented during the Maintenance phase.

2.2.3 Conclusion

Overall, the results of Experiment 1 show that both EHC and YHC groups were able to filter out distractors efficiently. Elderly controls benefited from encoding three items non-simultaneously (when two items were presented first and then one later). Moreover, a weak effect of increased difficulty in Filtering during Maintenance was present in EHC compared to YHC. Overall, EHC made more errors compared to YHC, and this was due to worse performance across all the different types of error parameters: decline in memory Precision,

reduced Target detection, increased Misbinding rates and increased Guessing. Filtering out an item was clearly beneficial in term of increased rates of target detection and reduced misbinding rates across the two groups. EHC also guessed more if compared to YHC when one item could not be filtered out. Memory precision was lower in EHC across all conditions and was the metric that showed the highest effect size ($\eta 2 = 0.399$) in group comparisons if compared to other metrics, which suggests it might be a good marker of ageing, rather than reflecting effects related to the complexity of the task or underlying filtering abilities.

2.3 Experiment 2 | Filtering in Alzheimer's disease

2.3.1 Methods

Participants

Twenty-eight patients, defined as per Jack et al 2018 [61] as having Alzheimer's clinical syndrome and hereinafter referred to as the AD group, and 28 EHC (the same as in Experiment 1) were recruited from the Oxford Centre for Cognitive Disorders and through open day events respectively. At the time of the Experiment, patients in the AD group did not have CSF confirmation of amyloid and tau pathology but were diagnosed according to their clinical phenotype and supporting imaging (MRI and FDG-PET). 13/28 patients had plasma biomarkers analysis as part of the study described in Chapter 6. AD patients were enrolled in the experiment if their total ACE score was higher than 45/100. Participants underwent the Filtering at Encoding and Maintenance Task as in Experiment 1, a neuropsychological battery (**Table 2.2**), and a subset of patients (48/56) consented to perform a 3T MRI brain scan.

The neuropsychological battery comprised measures of global cognition, verbal short-term memory, depression, apathy and fatigue. These included ACE-III [59], Digit Span (DS) [220], Apathy Motivation Index (AMI) [221], Beck's Depression Inventory (BDI) [222], Hospital Anxiety and Depression Scale (HADS) [223], Snaith-Hamilton Pleasure Scale (SHAPS) [224], Mood Scale (MS), i.e., the 15-Item Geriatric Depression Scale (GDS-15) [225], Fatigue Severity Scale (FSS) [226], Visual Analogue Fatigue Scale (VAFS) [227], Pittsburgh Sleep Quality Index (PSQI) [228], World Health Organisation Five Well-Being Index (WHO-5) [229] and Cantril Ladder (CL) [230]. A summary of participants' demographics, clinical variables, and tests results is presented in **Table 2.2**. According to the normality of the data, between-group differences were assessed either with t-test or Mann-Whitney U test for continuous variables and χ^2 test for categorical variables.

All participants had normal- or corrected-to-normal vision acuity and no color blindness. Elderly healthy participants who reported any psychiatric or neurological illness, or were on psychoactive drugs, were excluded from the experiment. They all gave written informed consent prior to the start of the experiment. Ethical approval was granted by the University of Oxford MS IDREC.

EHC and AD were not significantly different with respect to age, gender and handedness, but EHC had higher years of education (**Table 2.2**). All elderly healthy controls had global cognitive scores within the normal range, and none was significantly depressed. AD patients had, as expected, lower ACE scores with respect to EHC, but comparable performances on DS, a general measure of verbal working memory. AD patients had significantly higher levels of apathy and depression (AMI, BDI and HADS scores). They also had significantly lower scores on the VAFS fatigue scale than EHC (lower scores indicating more severe

fatigue), but not if measured by the FSS. EHC and AD had comparable sleep quality levels, indexed by the PSQI scores. There was a significant difference in WHO-5 wellbeing index between EHC and AD but no difference in CL measure of life satisfaction.

	EHC	AD	p-value
Age	70.36 (7.7)	70.79 (9.5)	0.854
Gender (M/F)	16/12	15/13	0.788
Education	17.75 (6.5)	15.11 (5.5)	0.036*
Handedness (R/L/A)	22/6/0	23/3/2	0.859
ACE	97.82 (1.8)	78.75 (11.8)	*< 0.001
DS	17.36 (3.6)	16.04 (4.3)	0.223
AMI	1.08 (0.4)	1.7 (0.7)	*< 0.001
BDI	4.6 (4.5)	10.86 (9.0)	*0.002
HADS	5.1 (3.0)	8.69 (6.2)	*0.040
SHAPS	18.32 (4.7)	20.79 (3.9)	*0.021
MS	1.29 (1.5)	3.3 (3.5)	*0.005
FSS	2.8 (1.2)	3.05 (1.5)	0.591
VAFS	84.14 (21.1)	67.86 (24.2)	*0.005
PSQI	4.36 (3.2)	4.41 (3.2)	0.905
WHO-5	78.0 (11.5)	62.7 (20.8)	*0.009
CL	8.44 (1.2)	7.79 (1.6)	0.149
Disease duration	-	4.7 (3.1)	-

Table 2.2 | EHC and AD demographics and tests

Handedness: R = Right-handed, L = Left-handed, A = Ambidextrous. Disease duration: years since onset of symptoms.

MRI acquisition and analysis

T1-weighted volumetric MR brain images were acquired on a 3T Siemens Magnetom Verio syngo scanner using a magnetisation prepared rapid gradient echo (MPRAGE) protocol acquired in sagittal orientation (TR = 2000 msec, TE = 1.94 msec, TI = 880 msec, Flip angle = 8 degrees, FOV read = 256 mm, Voxel size = 1.0 x 1.0 x 1.0 mm). All images were reviewed by a trained neurologist to exclude the presence of remarkable macroscopic brain abnormalities not compatible with the original diagnosis. Hippocampal volumes (HV) were estimated using FSL-FIRST [231]. For each participant, left, and right HV were calculated, and bilateral HV was computed. We also calculated whole brain volumes for each subject. We subsequently computed the head size corrected values for whole brain volumes and HV, using the scaling factor derived from SIENAX [232]. When referring to HV and whole brain volumes throughout the chapter, only head-size corrected volumes have been used. Images were carefully visually inspected after each processing step.

Statistical analysis

Mixture model variables and cumulative Set Size and Filtering rate metrics were calculated following the same principles as in Experiment 1.

For MAE, Precision, Target detection, Misbinding, and Guessing, we performed 2 (**Instruction:** Filter and SS3) x 2 (**Condition:** Encoding and Maintenance) x 2 (**Group:** EHC, AD) mixed ANOVAs. Between-group differences in cumulative metrics (Filtering rate and Set Size) were assessed through frequentist and Bayesian ANCOVAs using age, sex, education and handedness as covariates. Correlations between WM metrics, HV, and disease duration were assessed by calculating cumulative metrics for each Mixture Model variable by averaging the performance across all five conditions (SS2, FE, FM, SS3E, SS3M).

2.3.2 Results

Mean absolute error (MAE)

On the mixed ANOVA there was a main effect of Group [F (1,52) = 43.01, *p < 0.001, η2 = 0.361], (**Figure 2.6**). Moreover, there was a main effect of Instruction [F (1,52) = 16.54, *p < 0.001, η2 = 0.022], with higher errors in the conditions where three items had to be remembered compared to when one item had to be filtered out across the two groups. There was no main effect of Condition, implying overall equal number of errors during the Encoding or the Maintenance conditions if the two groups were taken together. Nevertheless, there was a significant Condition x Group interaction [(F (1,52) = 4.26, *p = 0.044, η2 = 0.004], with AD performing worse at Maintenance (FM, SS3M), while EHC performing worse at Encoding (FE, SS3E). There was also a significant Instruction x Condition interaction [(F (1,52) = 11.79, *p = 0.001, η2 = 0.010], with Filtering at Maintenance and Set Size at Encoding being challenging conditions compared to their counterparts Filtering at Encoding and Set Size at Maintenance.

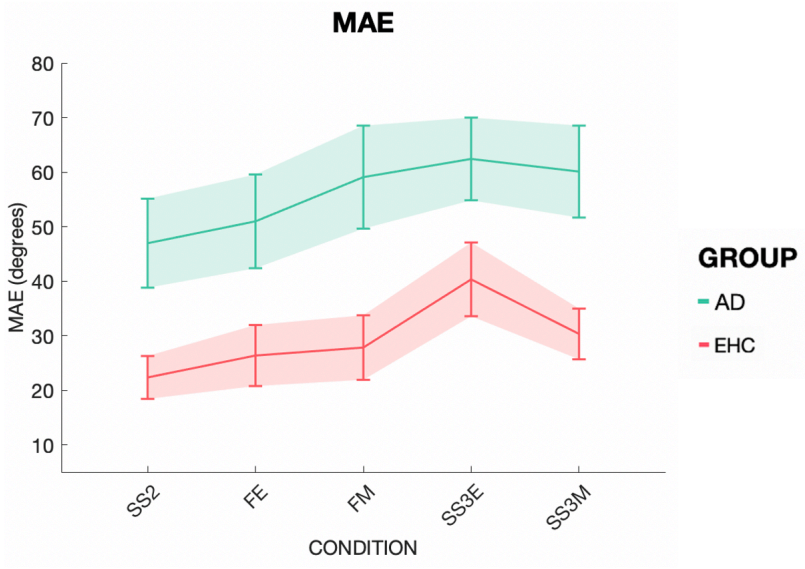


Figure 2.6 | MAE across working memory conditions for EHC and AD

EHC in coral red, AD in green. Note in particular the differential effect of Filtering at Maintenance (FM) compared to Filtering at Encoding (FE) within the AD group compared to within the EHC group. Further, in AD, FM performance was equivalent to SS3M.

Overall, AD patients performed worse than EHC, and regardless of the diagnosis, participants made more errors when three items had to be remembered compared to when they could be filtered out. Across both groups, filtering seemed to be more taxing during Maintenance, while remembering three items was more difficult if they appeared simultaneously, during the Encoding phase. AD patients experienced greater difficulty at Maintenance compared to at Encoding, whereas EHC performed worse at Encoding.

Impact of Set Size at Encoding and Maintenance

AD patients showed a greater Set Size effect at Maintenance [F (1,49) = 4.35, *p = 0.042, $\eta 2$ = 0.070] compared to EHC, but no difference in Set Size at Encoding, or their Cost (**Figure 2.7**). The Bayesian ANCOVA showed low evidence for this effect (BF_M = 3.160). This low effect is also in line with the absence of a difference in Set Size Cost.

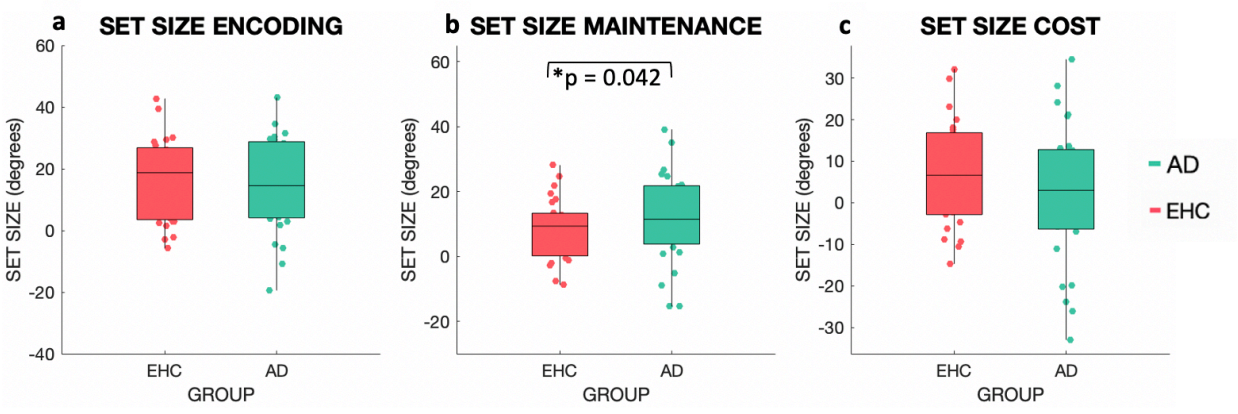


Figure 2.7 | Set Size at Encoding, Maintenance and their Cost in EHC and AD

EHC in coral red, AD in green. Only statistical significantly differences are shown.

Impact of Filtering at Encoding and Maintenance

There was no difference in Filtering rate at Encoding between the two groups, but groups were different in Filtering rate at Maintenance [F (1,48) = 8.94, *p = 0.004, η2 = 0.125] and Filtering rate Cost [F (1,48) = 4.63, *p = 0.036, η2 = 0.078], with AD having higher Filtering rates at Maintenance and lower Filtering rate Cost, indicative of worse performances at Maintenance (**Figure 2.8, panel b** and **c**). The Bayesian ANCOVA showed a $BF_M = 5.762$ for Filtering rate at Maintenance and $BF_M = 3.324$ for Filtering rate Cost, therefore providing evidence for a moderately significant difference.

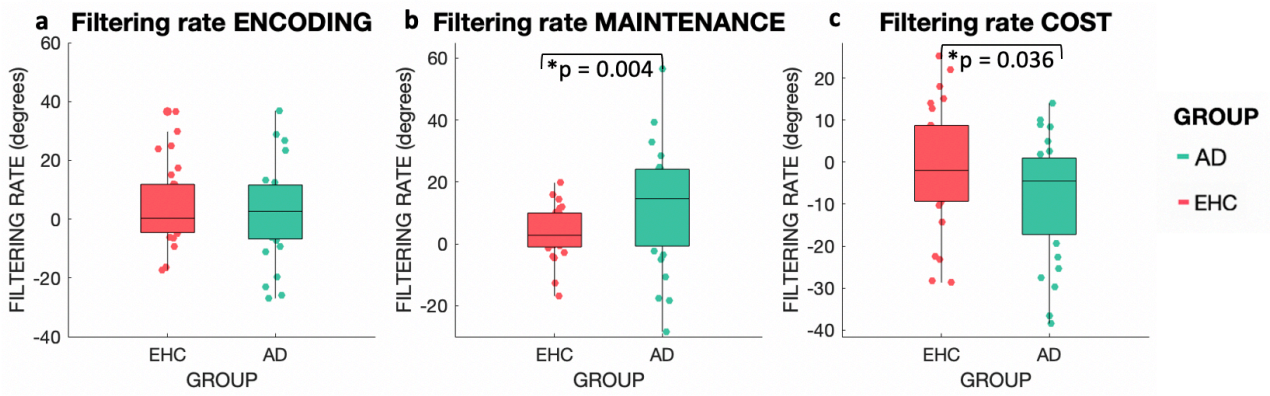


Figure 2.8 | Filtering rate at Encoding, Maintenance and Cost in EHC and AD

EHC in coral red, AD in green. Only statistical significantly differences are shown.

Analysis of mixture model parameters

Precision

The mixed ANOVA showed that AD patients had lower Precision compared to EHC [F (1,39) = 4.34, *p = 0.044, η2 = 0.039], (**Figure 2.9, panel a**). There were no main effects of Instruction or Condition, nor any significant interactions.

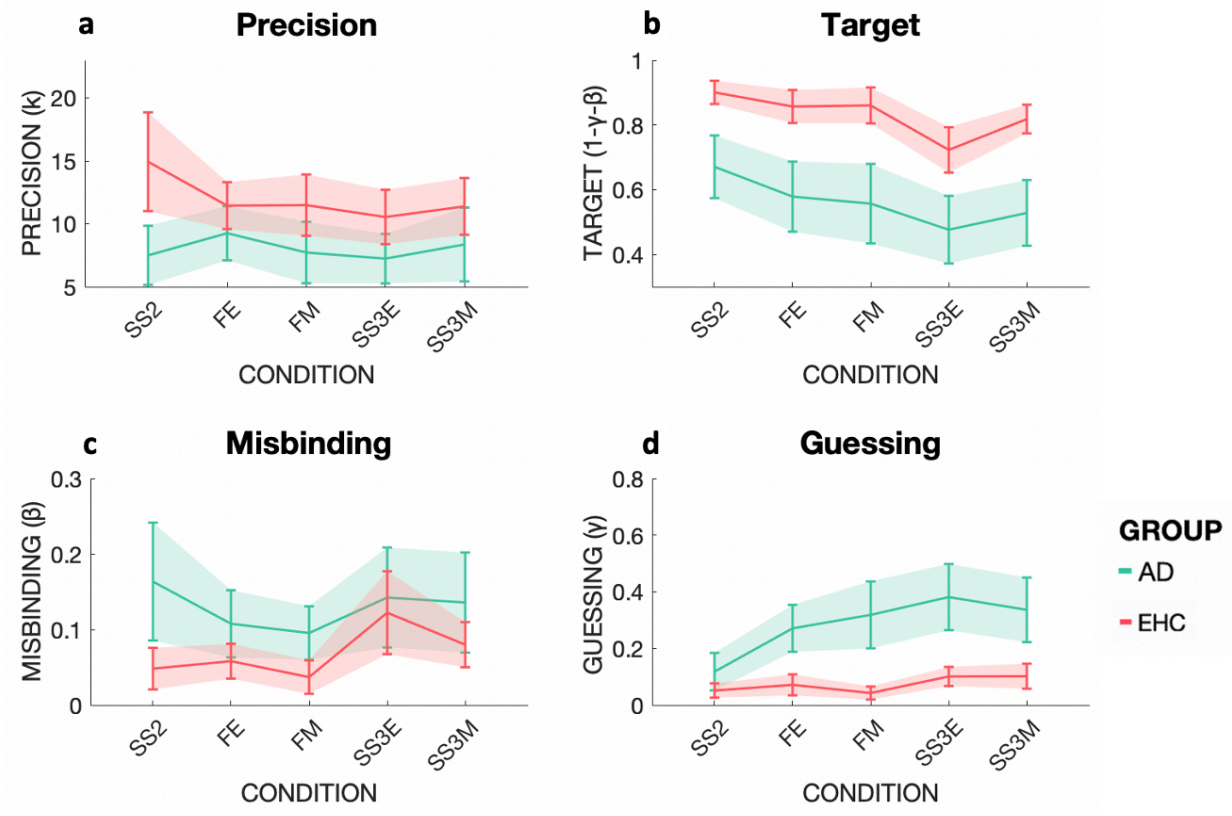


Figure 2.9 | Mixture Model metrics across conditions

EHC in coral red, AD in green.

Target detection

For Target detection, a main effect of Group [$F(1,52) = 36.66$, $^*p < 0.001$, $\eta 2 = 0.291$], with AD patients exhibiting worse performance compared to EHC, was shown by the mixed ANOVA (**Figure 2.9, panel b**). There was also a main effect of Instruction [$F(1,52) = 10.20$, $^*p = 0.002$, $\eta 2 = 0.021$], with less targets detected in SS3E and SS3M than in the filtering conditions. While AD patients were worse at detecting targets compared to EHC, both groups benefitted from filtering one item out.

Misbinding

The mixed ANOVA showed a main effect of Group [F (1,52) = 7.87, *p = 0.007, η2 = 0.047], with higher misbinding rates in AD patients compared to EHC (**Figure 9, panel c**). There was also a main effect of Instruction [F (1,52) = 6.02, *p = 0.018, η2 = 0.029], with more misbinding in SS3E and SS3M than in the filtering conditions for both groups. No other significant main effects or interactions.

Guessing

There was a main effect of Group [F (1,48) = 31.27, *p < 0.001, η2 = 0.241], with AD patients performing significantly worse than EHC (**Figure 9, panel d**). There were no other significant main effects or interactions.

In the AD group we found no correlation between disease duration and filtering abilities (Filtering rate metrics), working memory capacity (Set Size metrics), or any of the other WM metrics (Precision, Target detection, Misbinding, Guessing) computed as a continuous variable across all five conditions.

2.3.3 Conclusion

In line with Experiment 1, Experiment 2 showed that filtering out an item was clearly beneficial in term of increased Target detection and reduced Misbinding rates across the two groups. In both studies, filtering an item out was not able to increase memory Precision. Filtering also did not lead to reduced rates of Guessing overall but did so in elderly healthy controls compared to their younger counterparts in Experiment 1. One key result was that *Filtering at Maintenance* was more deficient in AD patients compared to EHC, with AD

patients showing similar performances in FM as in the SS3M condition (**Figure 2.6**). AD patients showed impairment across all multiple metrics compared to EHC: reduced memory Precision and Target detection, increased Misbinding and Guessing. However, in this case Target detection ($\eta 2 = 0.291$) and Guessing ($\eta 2 = 0.241$), and not memory Precision as in healthy ageing, showed the strongest effect size in group comparisons.

2.4 Experiment 3 | Filtering in Parkinson's disease

2.4.1 Methods

Participants

Twenty-eight patients with clinically diagnosed idiopathic Parkinson's Disease were recruited from the Cognitive Disorder Clinic and the Oxford Parkinson's Disease Centre Discovery Cohort, based at the John Radcliffe Hospital in Oxford. Their performance was compared to the EHC group described in Experiment 1 and 2. Idiopathic Parkinson's Disease patients were diagnosed in line with recent criterion evidence [233] and will be referred as the PD group. We recruited only PD patients with no evidence of cognitive impairment (ACE total score > 88/100), to avoid potential confounding factors in the results interpretation, even if further studies including a wider range of cognitively intact PD patients, PD with mild cognitive impairment and PD patients with dementia would be advisable. All participants had normal or corrected-to-normal vision acuity and no color blindness. They gave written informed consent prior to the start of the experiment. Ethical approval was granted by the University of Oxford MS IDREC.

PD patients were tested in 'ON' state, i.e., within 3 hours of their usual dopaminergic medications. Participants underwent the same Filtering at Encoding and Maintenance Task as in Experiment 1 and 2, the neuropsychological battery described in Experiment 2 (**Table 2.3**), and a subset of patients (43/56) consented to perform a 3T MRI brain scan (MRI acquisition and analysis as described in Experiment 2). In addition, we collected data on Unified Parkinson's Disease Rating Scale (UPDRS) scores and total Levodopa dose, to investigate whether differences in performance could mirror a different disease stage or could be impacted by the overall amount of dopaminergic medication. A summary of participants' demographics, clinical variables, and test scores is presented in **Table 2.3**.

EHC and PD were not significantly different with respect to age and gender but EHC had higher years of education. PD patients had lower ACE scores compared to EHC, but still within normal ranges (above 88/100 total) and the two groups had comparable performances at Digit Span. PD patients had significantly higher scores compared to EHC on AMI, indicative of higher levels of apathy, and higher levels of depression if measured by the BDI and MS, but not the HADS. PD patients also had significantly higher levels of fatigue, as seen by lower scores on the VAFS fatigue scale and higher scores at FSS. PD patients had poorer sleep quality, as seen by higher PSQI scores. Finally, there was a significant difference in CL and WHO-5 life satisfaction scores between PD and EHC, with PD having lower life satisfaction.

	EHC	**PD**	**p-value**
Age	70.36 (7.7)	70.79 (6.2)	0.831
Gender (M/F)	16/12	16/12	1

Education	17.75 (6.5)	14.82 (3.5)	*0.023
Handedness (R/L/A)	22/6/0	26/1/0	0.049
ACE	97.82 (1.8)	94.64 (3.4)	*< 0.001
DS	17.36 (3.6)	18.18 (3.9)	0.419
AMI	1.08 (0.4)	1.90 (0.9)	*< 0.001
BDI	4.6 (4.5)	9.57 (6.9)	* 0.002
HADS	5.1 (3.0)	7.14 (4.9)	0.083
SHAPS	18.32 (4.7)	19.30 (3.6)	0.391
MS	1.29 (1.5)	2.89 (3.0)	*0.025
FSS	2.8 (1.2)	3.57 (1.3)	*0.035
VAFS	84.14 (21.1)	68.6 (20.3)	*0.007
PSQI	4.36 (3.2)	6.86 (3.99)	*0.012
WHO-5	78.0 (11.5)	67.4 (15.6)	*0.006
CL	8.44 (1.2)	7.25 (1.6)	*0.004
UPDRS: total	-	51.31 (23.12)	-
part I	-	7.75 (5.79)	-
part II	-	10.39 (7.83)	-
part III	-	31.58 (1.23)	-
part IV	-	1.52 (2.64)	-
Levodopa dose (mg /daily)	-	375 (220)	-
Disease duration	-	8.2 (7.3)	-

Table 2.3 | EHC and PD demographics and tests

Statistical analysis

We followed the same principles described in Experiment 1 and 2 in computing cumulative Set Size and Filtering rate metrics. Between-group differences and correlations between

clinical variables and WM metrics were assessed following the same principles as in Experiment 2.

2.4.2 Results

Mean absolute error (MAE)

There was no significant group difference between EHC and PD patients in MAE, with EHC and PD performing overall similarly (**Figure 2.10**). However, there was a main effect of Instruction [F (1,49) = 27.07, *p < 0.001, η2 = 0.079], with less targets detected in SS3E and SS3M than in the filtering conditions and an Instruction x Condition interaction [F (1,49) = 8.12, *p = 0.006, η2 = 0.016], with filtering being more taxing during Maintenance while remembering three items was more difficult at Encoding.

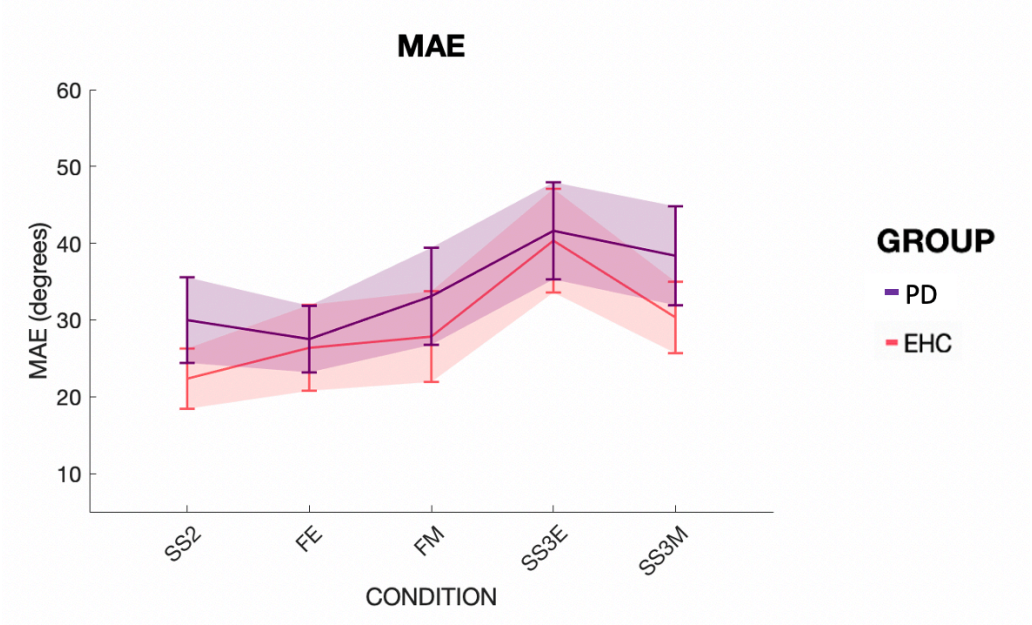


Figure 2.10 | MAE across conditions for EHC and PD

EHC in coral red, PD in violet.

Set Size and Filtering strategies

There was no significant difference in Set Size at Encoding, at Maintenance, or their Cost among the two groups. Filtering rate at Encoding, Filtering rate at Maintenance and Filtering rate Cost were also not significantly different between the two groups.

Analysis of mixture model parameters

Precision

The was no main effect of Group, Instruction or Condition, but a significant Instruction x Condition interaction [F (1,38) = 4.58, *p = 0.039, η2 = 0.022], suggesting that overall participants were more precise during Set Size three at Maintenance rather than Set Size three at Encoding (**Figure 2.11, panel a**). Therefore, unlike in healthy ageing or AD, memory Precision does not seem to be affected in PD patients compared to EHC.

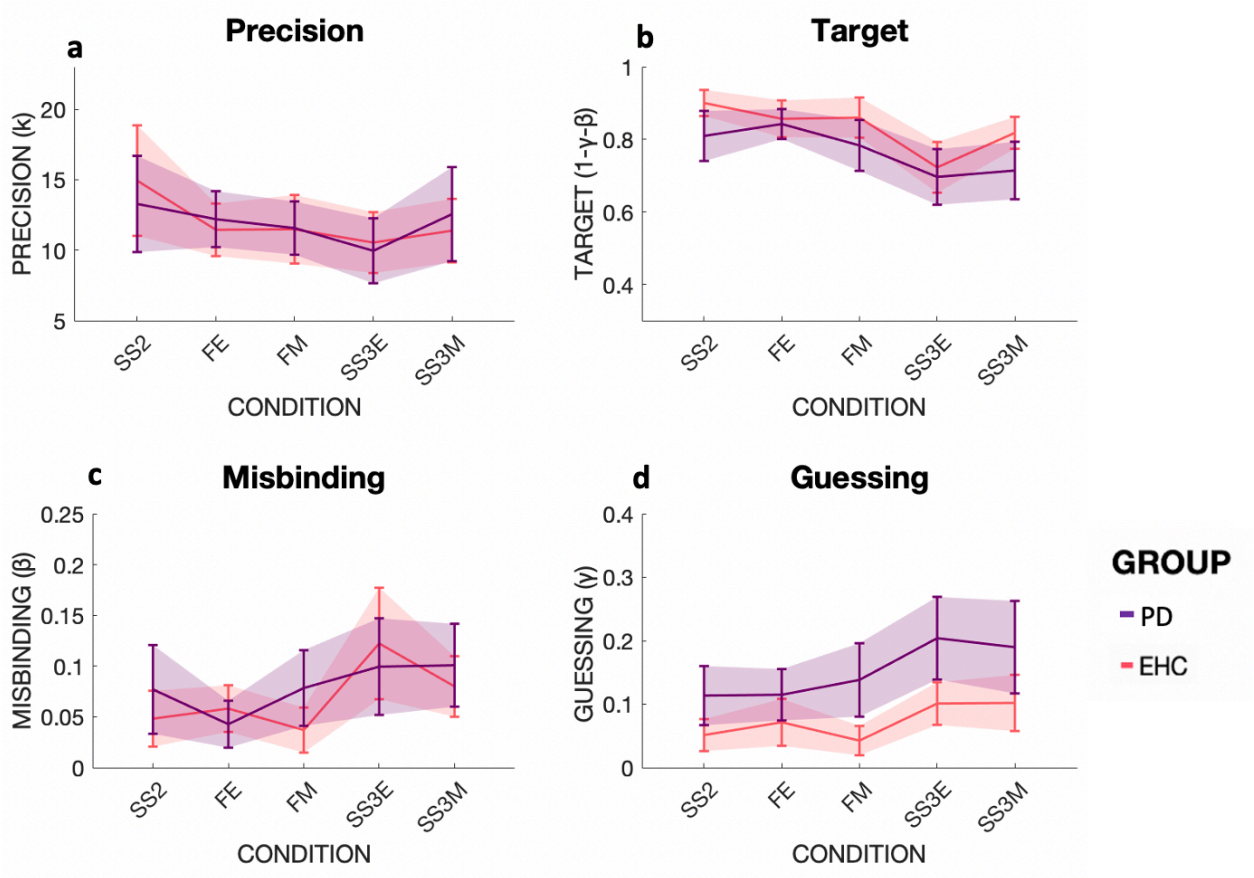


Figure 2.11 | Mixture Model metrics across conditions

EHC in coral red, PD in violet.

Target detection

There were no main effects of Group, or Condition, but there was a main effect of Instruction [$F(1,49) = 29.60$, $^*p < 0.001$, $\eta 2 = 0.090$], with less target detected in conditions where three items needed to be remembered compared to when one item could be filtered out (**Figure 2.11, panel b**). This suggests that filtering out an item could be helpful to achieve greater target detection rates in both PD and EHC.

Misbinding

The mixed ANOVA did not show any main effects of Group, or Condition, but there was a main effect of Instruction [$F(1,46) = 7.72$, $^*p = 0.008$, $\eta 2 = 0.046$], with misbinding occurring more frequently when three items need to be remembered compared to when two have to be remembered and one filtered out (**Figure 2.11, panel c**). This suggests a positive effect of filtering in helping to reduce misbinding errors in both groups.

Guessing

There was a main effect of Group [$F(1,45) = 9.43$, $^*p = 0.004$, $\eta 2 = 0.074$], with PD guessing significantly more than EHC, and a main effect of Instruction [$F(1,45) = 8.50$, $^*p = 0.006$, $\eta 2 = 0.037$], with more Guessing occurring in the conditions when all three items had to be remembered (**Figure 2.11, panel d**). This suggests two independent effects at play: the more taxing task of remembering three items and the pathological diagnosis are able to explain increased levels of guessing in this experiment. Filtering out an item seem to reduce guessing in EHC and in PD alike.

In the PD group, we found no correlation between UPDRS scores, Levodopa dose or disease duration and filtering abilities (Filtering rate metrics), working memory capacity (Set Size metrics), or any of the other WM metrics computed as a continuous variable across all five conditions.

2.4.3 Conclusion

In agreement with Experiment 1 and 2, Experiment 3 showed that filtering out an unwanted item improved Target detection, and reduced Misbinding, while having no effect on memory Precision. The key result was that PD patients exhibited a selective increase in Guessing compared to EHC, but otherwise had comparable performances in terms of WM filtering efficiency and capacity, memory Precision, Target detection and Misbinding rates. Another novel finding compared to Experiment 1 and 2, is that in Experiment 3 filtering out an item was beneficial also to reduce Guessing rates in EHC and PD if taken together.

2.5 Transdiagnostic analysis of cognitive measures and hippocampal volumes

After analysing filtering abilities across groups, we next investigated the interplay between WM capacity, filtering abilities and hippocampal volumes using a transdiagnostic approach. We also investigated whether different sources of error derived from the Mixture Model, i.e., MAE, Precision, Target detection, Misbinding and Guessing, were associated with hippocampal structural changes, as previous evidence points towards a possible involvement

of the hippocampus in feature binding, target detection and memory precision [77, 131 210, 211]. Correlations between variables were computed using a generalized linear model, with age, sex, handedness and education included as covariates of no interest.

Relationship between hippocampal volumes and filtering ability

We first examined filtering ability, the focus of this investigation. Across groups, *Filtering at Maintenance* was significantly negatively correlated with HV (r = - 0.404 *p = 0.002), (**Figure 2.12, panel a, Figure 2.13**). **Figure 2.13** displays the relationship between Filtering at Maintenance and hippocampal volumes, with participants with higher Filtering at Maintenance (on the right side), having lower hippocampal volumes (bottom of the graph). However, there was no significant correlation between hippocampal volumes with Filtering at Encoding or with Set Size, either at Encoding or Maintenance. Thus, Filtering at Maintenance was strongly related to hippocampal volume, which is often used as an important marker of AD (as age effects were regressed out in this analysis). To validate this result, we looked at whether Filtering at Maintenance was correlated with whole brain volume, and there was no such correlation (r = - 0.116, p = 0.392).

Relationship between hippocampal volume and other WM performance metrics

HV was negatively correlated with MAE (HV: r = - 0.473, *p < 0.001), Misbinding (HV: r = - 0.476, *p < 0.001) and Guessing (HV: r = - 0.578, *p < 0.001) and positively correlated with Target detection (HV: r = 0.474, *p < 0.001), (**Figure 2.12, panel b, Figure 2.13** for MAE). **Figure 2.13** shows how higher MAE is negatively correlated with hippocampal volumes, as the regression plane goes from higher hippocampal values (top if the image) to lower values (bottom) with increasing MAE values (from back to front of the image on the Y-axis). No association between hippocampal volume and Precision was found (HV: r =

0.057, p = 0.710). In this case, both MAE (r = - 0.319, *p = 0.014), Target detection (r = 0.310, *p = 0.017) and Guessing (r = - 0.335, *p = 0.015) were correlated also with whole brain volume, while Misbinding was not.

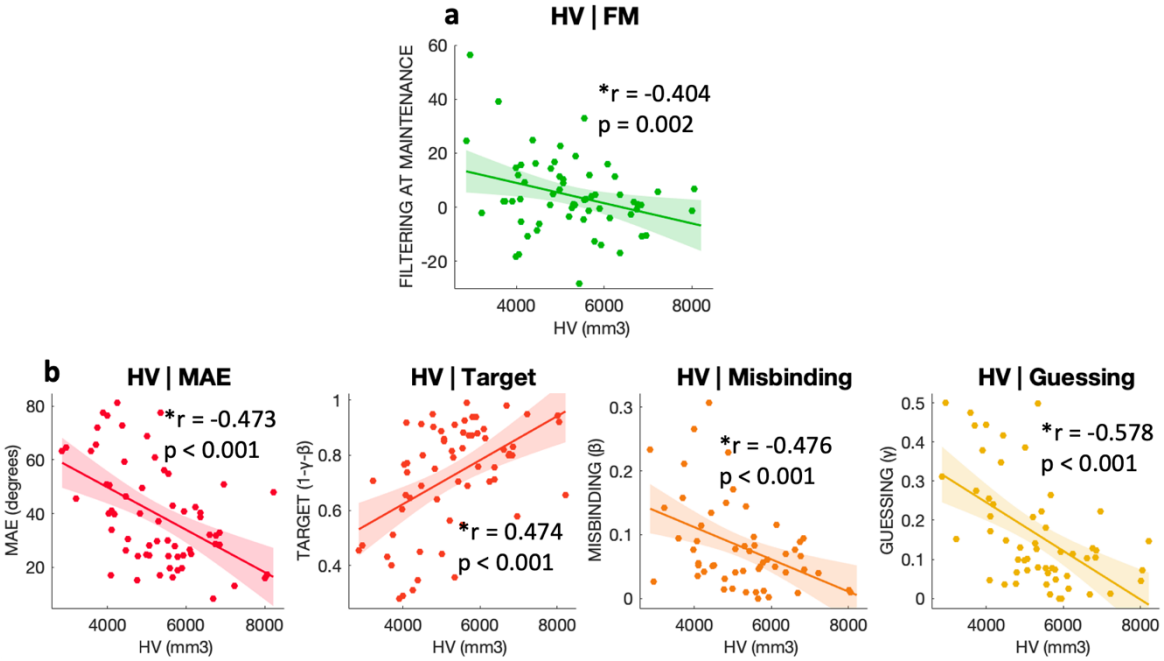


Figure 2.12 | Correlations between Filtering at maintenance and Mixture model parameters with hippocampal volumes

a | Correlation between Filtering rate at Maintenance and hippocampal volume (HV). **b** | Correlation between Mixture Model metrics and hippocampal volume. On the X-axis the volume in mm^3 of the hippocampus. In **b**, the Y-axis represent from left to right respectively MAE in degrees, probabilities of Target detection, Misbinding and random Guessing.

A stepwise regression analysis was performed to determine which of these metrics was a better predictor of hippocampal volume. HV was set as dependent variable, MAE, Misbinding, Guessing, and Filtering at Maintenance as independent variables. Target detection was not included for avoiding collinearity, which arises from how Target, Misbinding and Guessing probabilities are calculated (as they add up to 1). Overall, HV was better explained by Guessing (adjusted $R^2 = 0.419$, *p < 0.001).

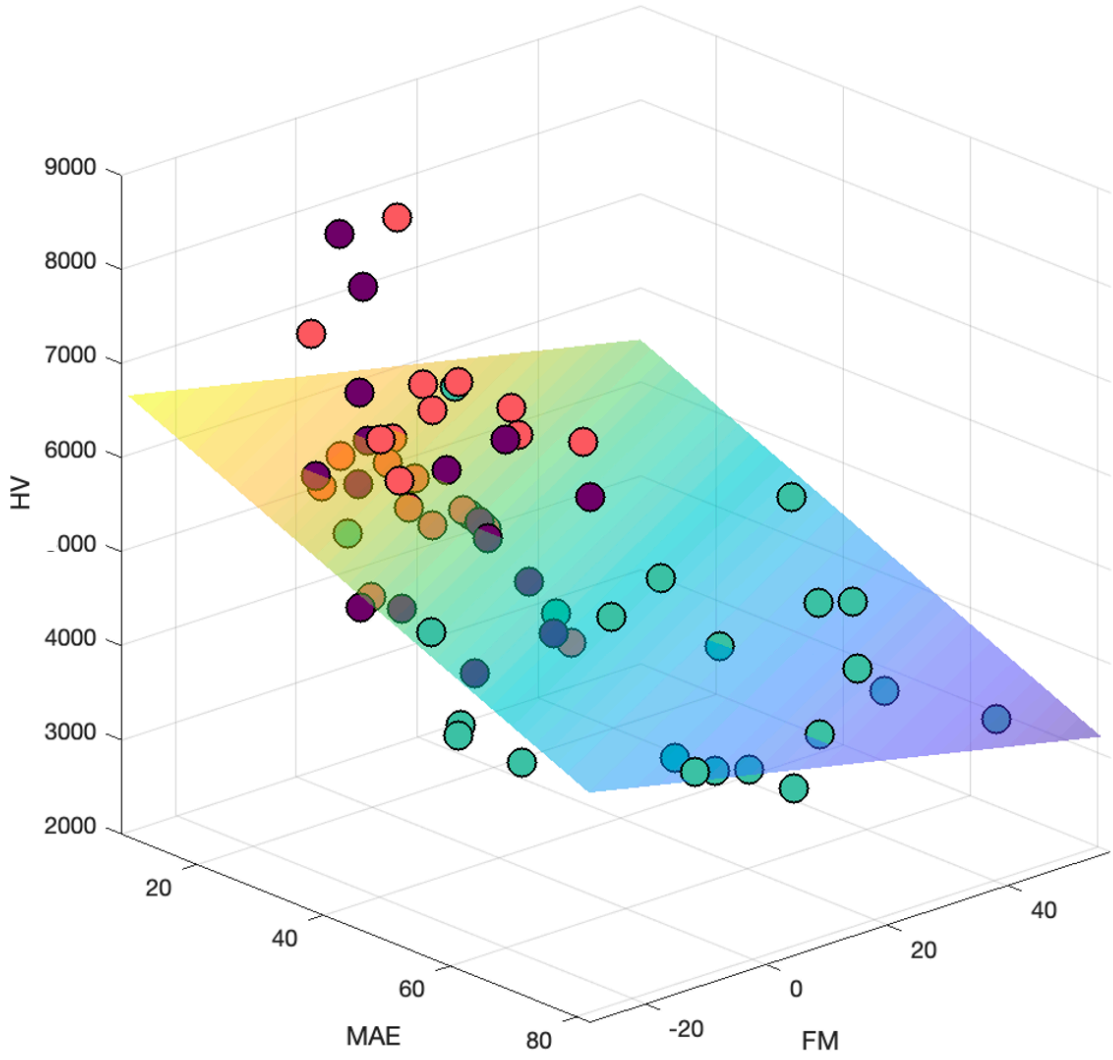


Figure 2.13 | Filtering and MAE stratified by HV and diagnosis

FM on the X-axis, MAE on the Y-axis, hippocampal volumes (HV) on the Z-axis. Subjects with higher HV (top of the image) have lower MAE (back of the image), lower Filtering at Maintenance scores (on the left) and are predominantly represented by elderly healthy subjects (in coral red), and PD patients (in violet). On the other end of the spectrum, AD patients (in green) have higher Filtering at Maintenance scores (on the right), higher MAE (front of the image), and have lower hippocampal volumes (bottom of the image).

2.6 Discussion

The studies described here examined WM performance using a novel delayed reproduction task developed specifically with conditions that required filtering out distractors either at encoding or maintenance. Elderly participants had relatively preserved filtering abilities

both at encoding and during maintenance. In AD, however, there were significant filtering deficits specifically when the distractor appeared during maintenance. In PD, by contrast, patients showed preserved filtering, both at encoding and maintenance. Healthy ageing was associated with higher rates of both misbinding and guessing, as well as lower target detection, but most importantly, with a prominent decline in memory precision. AD patients also demonstrated higher rates of misbinding and lower memory precision, but higher guessing and lower target detection were the leading sources of memory error, whereas PD patients showed only higher guessing rates. Finally, hippocampal volume were significantly correlated with filtering ability during maintenance - but not at encoding - as well as with overall mean absolute error, target detection, guessing and misbinding. Overall, these findings show that healthy ageing and two different neurodegenerative diseases (AD and PD) exhibit distinct patterns of WM impairment at Encoding and Maintenance and provide further evidence of a key role of the hippocampus in WM, including filtering distractors during maintenance.

Filtering abilities

Previous evidence suggests filtering ability is a crucial determinant of overall performance in WM tasks and good filtering abilities are tightly related to higher WM capacity [42, 234]. In this experiment, elderly controls showed preserved filtering abilities compared to younger controls if the distractor was presented at Encoding (**Figure 2.4, panel a**). Their ability to filter out a distractor during the Maintenance period might not be as efficient, as suggested by the small age-related effect in decline in Filtering at Maintenance performances in healthy ageing (**Figure 2.4, panel b**). This finding is in line with previous evidence of an age-dependent decline in performance when people are required to filter during the retention

period [42], but in our case, the Bayesian analysis did not provide strong evidence to draw a firm conclusion. This might be due to the fact that our elderly participants were thoroughly screened for depression, apathy and cognitive impairment, which might lead to a potential inflation of their true cognitive abilities in that age range compared to large, community-based studies.

AD patients on the other hand showed much clearer WM deficits if they had to filter out an irrelevant distractor during the Maintenance phase (**Figure 2.8, panel b**). This is in line with previous findings showing that AD patients struggle to prioritize relevant over irrelevant visual information where an optimal inhibitory control is required to avoid unwanted attentional capture [235, 236, 237, 238]. Our experiment was able to show not only that filtering abilities are impaired in AD patients, but that their deficit is specifically limited to the maintenance phase, as they showed good performances when asked to filter out a distractor presented at encoding (**Figure 2.8, panel a**).

In contrast, the PD patients tested here showed excellent filtering ability both at Encoding and Maintenance. All our PD patients were tested in an ON Levodopa state (when dopaminergic medications are effective). Other studies have reported an impaired filtering ability in PD patients when they were OFF Levodopa [215], so it is possible that dopaminergic therapy can have a beneficial effect of dopamine replenishment on ignoring or filtering in PD [130]. The integrity of WM maintenance processes in PD concurs with previous data on visual pattern stimuli tasks that likely require both object and spatial processing [239], and our results suggest that PD patients are able to manipulate information successfully during the maintenance period, unlike AD patients. Since we did not specifically test for a delay effect, we cannot exclude that at longer delay PD patients could also suffer from delay-dependent

declines in maintenance performances as previous evidence suggests [240], so further studies are needed to unravel a time dependent constant of possible delay-driven performance decline.

Set size effects

Overall, healthy elderly participants performed worse compared to younger participants and were specifically impaired compared to their younger counterparts when three items had to be encoded simultaneously (**Figure 2.2**). This is reflected by a significant Set Size effect at Encoding, and Set Size Cost, while Set Size at Maintenance was not affected (**Figure 2.3**) and is in line with a well-known age-related decline in WM capacity [33, 40]. AD patients showed overall lower performances compared to elderly controls. Unlike the latter, however, they seemed to cope well with increased Set Size, i.e., when the increased cognitive demand occurred at Encoding (**Figure 2.6 and 2.7, panel a**), but were mildly impaired during the Maintenance phase, as reflected by the weak Set Size effect at Maintenance (**Figures 2.6 and 2.7, panel b**). PD patients showed overall comparable performances to elderly controls.

Mixture Model findings

Disentangling the different sources of errors within the Mixture Model framework revealed that healthy ageing is associated with different types of errors: reduced memory Precision, decreased Target detection, increased Misbinding and increased Guessing (**Figure 2.5**). But while filtering out an item was beneficial in increasing the numbers of detected targets, and reducing Misbinding, it was not able to help maintaining a high memory Precision, which declined invariably across all conditions in elderly subjects (**Figure 2.5, panel a**). Filtering out a distractor was also beneficial in reducing Guessing, and that was true specifically for

EHC, with YHC not showing such effect. A possible interpretation is that the task was more taxing for the EHC, who had to deploy a cognitive strategy to free up space in WM.

These results support the idea that memory Precision could be considered as a marker of ageing, as previously reported by other groups [38, 40, 41, 42]. In this experiment, we were also able to show that not only memory Precision declines with ageing, but also that AD patients have a further drop in memory precision compared to age-matched healthy controls, albeit to a lesser extent, while this does not occur in PD patients.

Another memory feature that seems to decline with ageing is the ability to correctly identify a target [38, 241, 242]. Our results are consistent with such a view since a clear decline in Target detection was observed in elderly compared to younger controls across all conditions (**Figure 2.5, panel b**). In healthy ageing, fewer targets were detected when three items had to be remembered compared to if they could be filtered out (**Figure 2.5, panel b**). There was also a synergic detrimental effect in ageing, as reflected by a significant three-way Instruction x Condition x Group interaction in Experiment 1, suggesting that elderly controls were worse at detecting targets when three items had to be remembered, and when this had to be accomplished during the Encoding phase. AD patients were even less likely to correctly identify a target compared to age-matched controls (**Figure 2.9, panel b**). This supports previous findings of reduced correct object identity recognition in familial AD [77] as well as sporadic, late onset AD [78, 243].

In line with previous data [38], Misbinding increased even in healthy ageing and, as previously suggested [40], this effect is specifically marked when more items have to be remembered (**Figure 2.5, panel c**). Moreover, Misbinding was further increased in AD patients compared

to healthy controls (**Figure 2.9, panel c**), unlike in PD patients (**Figure 2.11, panel c**). Previous data show that both familial [77] and sporadic late onset AD patients [76, 78] exhibit increased rates of Misbinding. Whilst confirming that object binding is a feature that is frail in healthy ageing, and more so, in AD, we also found that filtering out an item decreases Misbinding rates in EHC, AD and PD (**Figures 2.5, 2.9, 2.11, panel c**).

Lastly, Guessing rates were higher in healthy ageing, AD and PD patients (**Figures 2.5, 2.9, 2.11, panel d**), and this was the only metric which was able to differentiate between PD patients and elderly controls. This is in line with previous evidence showing that Guessing is primarily affected in PD but can also be increased in AD patients [78]. Filtering out an item was able to help Guessing less and free up WM resources in EHC but not YHC, in Experiment 1, was ineffective in AD patients in Experiment 2, and was shown to be beneficial in EHC and PD in Experiment 3. Therefore, the effect of filtering on random guessing might be highly dependent on the population examined.

Role of the hippocampus

Previous studies have specifically linked increased misbinding in both sporadic, late-onset AD and familial AD to hippocampal dysfunction [132, 143, 244]. This evidence is consistent with the results of several different reports in patients with different pathologies beyond AD, which all point to a potential role of the hippocampus in binding information in WM [138, 244, 245]. However, to the best of our knowledge, there are no previous data on whether hippocampal integrity might be related to filtering ability, either at encoding or during maintenance. The data here demonstrate that filtering ability during maintenance – but not encoding – is related to hippocampal volume (**Figures 2.12, panel a, 2.13**). This suggests that the hippocampus plays a role in protecting the contents of WM, once encoded. This

result is further supported by the lack of correlation between Filtering at Maintenance and whole brain volume, that could suggest a specific contribution of the hippocampus in WM filtering during the maintenance period.

The results presented here also support the idea that misbinding rates are indeed correlated with hippocampal atrophy (**Figure 2.12, panel b**). Previous studies have argued for a role of the hippocampus in relational binding in long-term memory. Our results support findings from several different laboratories that the hippocampus also plays a role in binding information in short-term or WM [132, 138, 144, 244]. The lack of correlation with whole brain volume, unlike other metrics such as MAE or Target detection, also reinforces the idea that this metric reflects hippocampal integrity. In our dataset, Guessing was also a predictor of hippocampal volume and was the best predictor across the WM examined (**Figure 2.12, panel b**). To note though, our sample consisted of a heterogeneous population, comprising also PD patients that are selectively more prone to guessing, as well as patients with well-established late-onset AD, who were much more cognitively impaired compared to similar studies [77, 143]. For misbinding errors to occur, at least a partial memory of the object (its identity) must be retained, while the memory of its location might be lost. When no mnemonic trace is left, Guessing remains the only viable option, and therefore its rates might be higher in more cognitively impaired populations. While higher Target detection was associated with higher hippocampal volumes, this was not the case for memory Precision. Independent findings suggest that other brain structures (including the parietal cortex) rather than the hippocampus could be primarily involved in maintaining information with high detail of precision over time [131, 246], and therefore we cannot exclude that this could be the reason underlying the lack of structural associations with Precision in our sample.

2.7 Conclusion

Taken together, the findings presented here show that healthy ageing was mainly characterized by a reduced capacity at encoding (increased Set Size effect at Encoding) and by a uniform decline in memory Precision. Decreased Target detection, increased Guessing, and increased Misbinding also occur with ageing. AD patients seemed particularly vulnerable during the maintenance period and were particularly impaired at filtering during maintenance. Furthermore, similarly to healthy ageing, they also showed overall increased Misbinding and decreased memory Precision but had a more prominent increase in Guessing and decrease in Target detection. PD patients exhibited preserved filtering abilities and WM capacity both at encoding and maintenance, but they had an overall increased rate of Guessing compared to elderly controls. Finally, hippocampal atrophy was associated with a significant decline in ability to filter distractors during maintenance, but not at encoding, as well as other metrics of WM performance, providing further evidence for the role of the hippocampus in WM.

Chapter 3. Cross-Sectional and Longitudinal assessment of Visual Short-Term Memory in patients at risk of developing Alzheimer's Disease with the "What was where?" Oxford Memory Task §‡¥#

Abstract

Our aim was to investigate whether digital cognitive metrics derived from a tablet-based task, the "What was where?" Oxford Memory Task, were able to clinically stratify patients at risk of developing AD. We also compared the performance of these new digital metrics to traditional neuropsychological tests of cognition. Finally, we tested whether these new digital metrics relate to biological hallmarks of AD such as hippocampal atrophy. 325 subjects participated in this experiment: 49 patients with subjective cognitive impairment (SCI), 57 patients with mild cognitive impairment (MCI), 63 patients with AD dementia and 156 elderly healthy controls (EHC). Across groups, increasing set size (i.e., remembering one compared to remembering three items) had a major impact on memory recall, showing an effect on all metrics. Most metrics, except for Misbinding, showed also an effect of delay (i.e., being probed after 1 or 4 seconds). With the exception of reaction time measures, i.e., Identification and Localization time, all our digital metrics were able to discriminate between healthy controls and patients with MCI. Metrics derived from the mixture model of working memory (including Target detection, Misbinding and Guessing), and Absolute Error (an index of spatial localization error) also differed between patients with SCI and MCI. Apart from Absolute Error, all digital metrics were also able to discriminate between

patients with MCI and AD. Identification accuracy was the best predictor of hippocampal atrophy, performing as well as standard neuropsychological tests. Memory Imprecision was able to predict cognitive decline on standard tests of cognition over one year.

3.1 Cross-sectional Visual Short-Term memory (VSTM) assessment in patients with SCI, MCI and AD §‡¥#

3.1.1 Delay, Set Size effects and Group differences across eight key VSTM metrics (Identification Time, Localization Time, Identification Accuracy, Absolute Error, Target Detection, Misbinding, Guessing and memory Imprecision)

3.1.1.1 Introduction

Since William James's first description of 'primary' and 'secondary' memory, i.e., events that are in the '*rearward portion of the present space of time'* or that '*already dropped out of consciousness'* [247], memory has been perceived as composed of two different systems, commonly referred to as STM and LTM. Traditionally, hippocampal dysfunction has been more commonly associated with episodic LTM impairment, in contrast with relatively spared STM performances [65]. This idea was initially supported by seminal studies where STM abilities such as digit span were preserved in patients with hippocampal lesions due to conditions such as viral encephalitis or surgical resection for epilepsy, despite LTM being severely affected [65, 66, 248]. These data corroborated the conceptualization of memory being divided into different systems, a STM buffer and a LTM storage, which could function

independently from each other, and anatomically linked exclusively LTM to the hippocampus.

While the distinction between STM and LTM is still valid these days, the role of the hippocampus in STM has gained more recognition throughout the years, thanks to the development of more fine grain measures of STM. Hippocampal processing is now regarded as crucial in combining the representations of elementary stimuli to construct unique representations, whether this occurs during perception, STM or LTM [132, 249], breaking the assumption that hippocampus is involved only in LTM. It has been hypothesized that hippocampal damage leads to impairments on tasks that require complex high-resolution bindings, while is much less likely to lead to poor performance on tasks that require only low-resolution bindings or simple associations [132]. STM performance for single items such as digits, words or visual locations could be well-preserved in patients with hippocampal lesions, while a high-resolution, precise representation that relies on feature binding might be impaired [135, 144, 250, 251, 252]. These binding deficits can be observed both in change-detection tasks and in delayed-reproduction tasks [44, 77]. One important piece of evidence on the hippocampal contribution to STM binding using a delayed reproduction task, the "What was where?" task, came from studying patients with voltage-gated potassium channel antibody mediated limbic encephalitis [143], which is an autoimmune disorder affecting primarily the medial temporal lobe. In that study, Pertzov et al showed that patients with limbic encephalitis have higher chances of incorrectly binding the identity and the location of an object (i.e., misbinding rates) compared to healthy controls, while performed as well as controls in recalling single item features. While limbic encephalitis is a good model of a neurological condition affecting selectively the hippocampus, it is a relatively rare condition.

Hippocampal atrophy is the hallmark of AD [49, 135 253] which is the most prevalent cause of dementia worldwide, far more common than limbic encephalitis [254]. Previous data on the same delayed reproduction task, demonstrated that people with a genetic risk factor for familial Alzheimer's disease (FAD), such as carriers of the Presenilin-1 (PSEN1) mutation, exhibit higher misbinding rates when high-resolution binding needs to be accomplished, and that their performance correlated with the degree of hippocampal atrophy [77]. Recent evidence using the same paradigm in patients with sporadic, late onset AD, reported that misbinding rates are increased also in these patients [76, 78], but how this relates to their hippocampal integrity is still unknown. Moreover, it has been found that different metrics extracted from the mixture model of working memory [17] might be impaired in specific patient populations, (e.g., misbinding in patients with limbic encephalitis and random guessing in patients with PD), while others are spared [76, 255], as also seen in Chapter 2.

While we know that hippocampal atrophy relates to standard screening tests of cognitive function [256], and that delayed-reproduction VSTM metrics also show a good concordance with measures of hippocampal integrity [77], there is surprisingly little data on head-to-head comparisons between these STM experimental behavioral metrics and commonly used clinical cognitive scales, and is it unknown which metrics best predict hippocampal atrophy. Moreover, most studies have focused on misbinding rates, which might or might not be the best marker of hippocampal integrity amongst a wide array of metrics that can be computed using digital delayed reproduction tasks. In standard experimental psychology paradigms, reaction times are used as a proxy of performance; however, it has also not been tested whether these are better or worse than metrics derived from the mixture model, such as misbinding, in reflecting hippocampal atrophy but also in mapping disease progression in AD.

Another important gap in the existing knowledge is also not only *if* but *when* these delayed reproduction metrics can outperform standard measures of cognitive function during the disease course in the AD continuum. This applies to whether they perform better in the early stages of AD, such as subjective and mild cognitive impairment, or during disease progression, as in patients with an established diagnosis of AD dementia.

In this chapter we will investigate several different metrics derived from the mixture model, including but not limited to misbinding, and measures of reaction times, in patients at risk of developing AD, such as SCI and MCI, as well as AD dementia patients. We were also interested to test whether these metrics relate to standard screening tests of cognition and to hippocampal integrity.

3.1.1.2 Methods

Participants

325 participants were enrolled in the study: 49 patients with SCI, 57 patients with MCI, 63 patients with AD dementia and 156 elderly healthy controls. Patients were recruited from cognitive disorders clinics at the John Radcliffe Hospital in Oxford, United Kingdom and Friedrich-Schiller-Universität Klinik, Jena, Germany. Subjective cognitive impairment was defined according to the 2020 criteria from Jessen et al for subjective cognitive decline [257]. Mild Cognitive Impairment patients were classified according to Petersen's criteria of 2014 [258]. Alzheimer's disease dementia patients were defined as Alzheimer's disease clinical syndrome according to the 2018 criteria by Jack et al [61] and will be subsequently referred to as AD. From the Oxford cohort, 26/63 patients had plasma biomarkers analysis as part of

the study described in Chapter 6. All subjects underwent either brain CT or MRI imaging and were excluded from the study if there was evidence of structural abnormalities not compatible with their clinical diagnosis. Inclusion criteria for elderly healthy subjects followed the same principles outlined in previous Chapters. A summary of participants' demographics is presented in **Table 3.1**. All gave written informed consent prior to the start of the experiment. Ethical approval was granted by the local ethics committee in Oxford and Jena.

Neuropsychological and behavioral test assessment

The study protocol included a brief neuropsychological assessment and the "What was where?" Oxford Memory Task [76, 77, 143].

The neuropsychological assessment included measures of global cognition, verbal short-term memory, depression, and sleep quality to rule out a potential neurodegenerative disorder or major depression in the elderly controls and for being able to compare performances at standard cognitive tests to our experimental paradigm. These included Addenbrooke's cognitive examination (ACE-III) [27], subsequently termed ACE, Digit Span (DS) [220], Hospital Anxiety and Depression Scale (HADS) [223], 15-items Geriatric Depression Scale (GDS) and Pittsburgh Sleep Quality Index (PSQI) [228]. Participants' test scores are presented in **Table 3.1**.

A schematic of the "What was where?" Oxford Memory Task is shown in **Figure 3.1**.

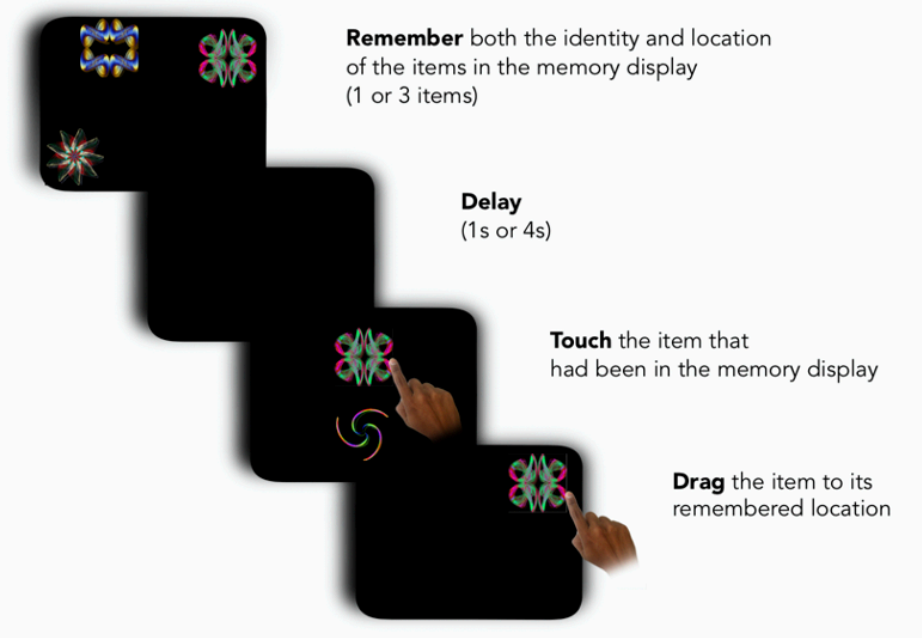


Figure 3.1 | "What was where?" Oxford Memory task (OMT)

Participants were presented with either 1 or 3 fractals randomly distributed on the screen. After a 1 or 4 second delay two fractals appeared at the centre of the screen, one of which had appeared in the memory array whereas the other one was a distractor. Firstly, they needed to identify the object they had seen previously ('what'), and then drag it back to its original location ('where').

Stimuli were presented on a black background and were chosen from a library of 196 fractals (http://sprott.physics.wisc.edu/fractals.htm), including 49 different shapes and each shape containing 4 colour variations. Participants sat approximately 30 cm in front of a tablet (either iPad or Android), yielding 2.3° of visual angle. Stimuli were calibrated using the dimension on the screen to ensure matching of stimuli properties across different tablet models. A fully remote online version of this task, available for computers, laptops, iPADs and phones, has been subsequently developed by Sijia Zhao (https://oxfordcognition.org/) and was used in Chapter 6.

In different trials, participants were presented with either 1 or 3 fractals located randomly on the screen. They were asked to remember the identity of the fractals ('what'), and their

locations ('where'). After a delay of either 1 or 4 seconds, two fractals appeared at the centre of the screen along the vertical axis. One of these had appeared in the memory array (target) whereas the other one was a foil, which had not been shown in the current trial. Participants were required to touch the target and drag it to its original location. They were instructed to be as precise as possible, thus allowed to move the item and adjust its location as many times as they needed. This design yielded four different conditions (1 and 3 items, 1 and 4 seconds). Each participant performed a practice block of 8 trials followed by three test blocks of 40 trials, with a total of 120 test trials (10 trials x 4 conditions x 3 blocks). The order of the trials in a block was randomly chosen for each block. The locations of the fractals were determined by a MATLAB script (MathWorks, Inc) in a pseudorandom manner.

Overall, this task allowed us to extract different working memory metrics (**Figure 3.2**), including (**basic metrics**):

- **Identification time**: the time in seconds taken to identify the correct object.
- **Localization time**: the time in seconds to drag the chosen object to its remembered location.
- **Identification accuracy**: the proportion of trials in which participants correctly identified the target.
- **Absolute Error**: the Euclidian distance from the center of original item location to the center of participant's response location.

The Mixture Model of WM by Bays et al [17] was then fitted to the data to unravel the differential contribution of memory errors to our dataset. The following metrics (**mixture model metrics**) were therefore calculated:

- **Target detection:** the probability of correctly identifying the target.
- **Misbinding**: the probability of mislocalizing a correctly identified item to the remembered location of another item in the memory array.
- **Guessing**: the probability of random guessing responses.
- **Imprecision**: the width of the distribution of the responses around the target.

Model fitting was achieved using a permutation approach, where per each single trial, we calculated the distances between the response location and the location of:

1) Target.
2) Distractor (non-target in that trial).
3) A distractor taken from a randomly chosen trial.

Depending on which of these distances was the shortest, the response was either counted as target (1), distractor (2, i.e., misbinding) or random response (3, i.e., uniform guessing). We repeated this procedure 5000 times per trial, introducing a distractor from a randomly chosen trial each time. This procedure allowed us to calculate proportions for these three sources of response per trial (absolute amount of response type/5000). The introduction of a distractor that was randomly chosen from another trial allowed us to differentiate whether an error was systematically linked to the very specific trial's distractor or whether it could be accounted for even by a randomly chosen distractor that was not present at trial. Importantly, these metrics were calculated uniquely on trials where an object was correctly identified.

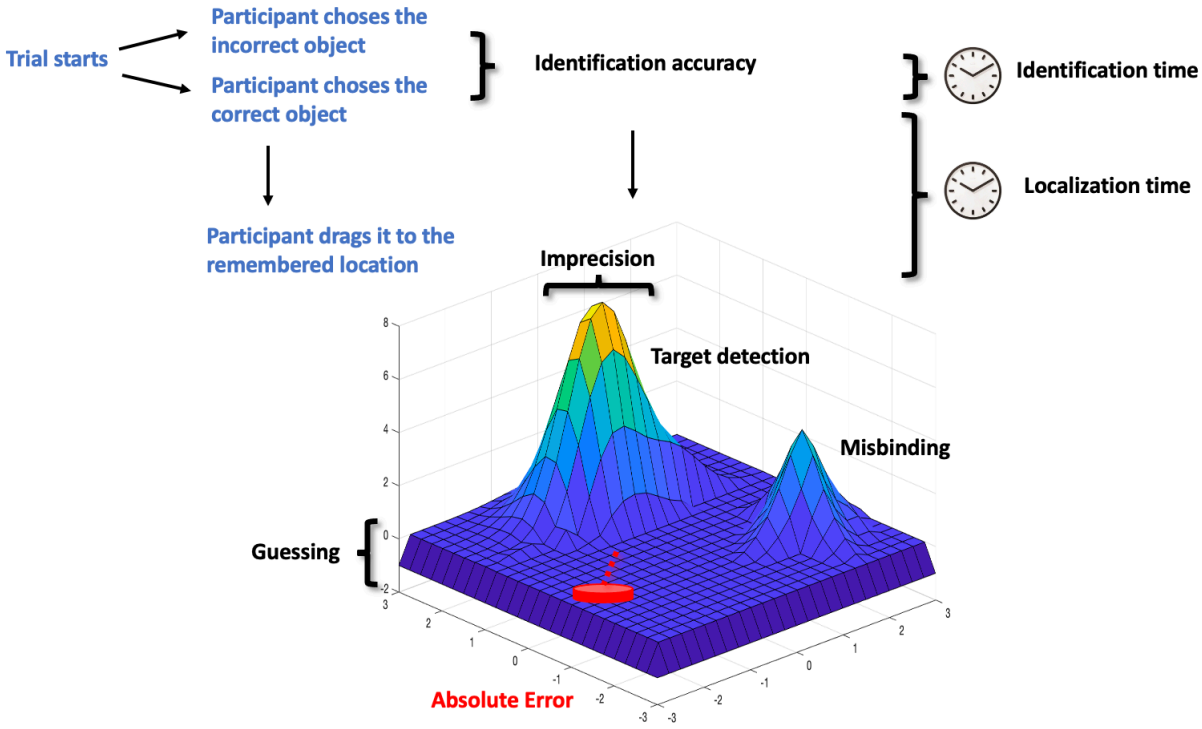


Figure 3.2 | Summary metrics from "What was where?" Oxford Memory task (OMT)

Statistical analysis

All analyses were conducted using MATLAB 2019a, R (version 3.5.2) and JASP (JASP team, 2022). Data visualization was conducted using software Grammar of graphics plotting in MATLAB R2018b (Gramm library) [259] and RStudio (Version 1.1.463). Effect size was quantified using Eta Squared (η2), defined as large (η2 > 0.14), medium (η2 > 0.06) or small (η2 > 0.01) [260].

1) **Demographics and standard neuropsychological tests:** A one-way ANOVA was used to compare subjects' continuous variables in demographics and test scores, with Holm post-hoc correction amongst the four groups. χ^2 test was used to compare categorical variables across groups.

2) **Transdiagnostic** analysis: Firstly, we looked at the effects of Set Size and Delay across all groups (transdiagnostically), calculating a 2 (Set Size: 1 item, 3 items) x 2 (Delay: 1 second, 4 seconds) ANOVA for each of the eight digital working memory metrics.
3) **Between-group** analysis: Subsequently, we derived a cumulative measure for each metric by calculating the mean across the 4 conditions (1 item 1 second, 1 item 4 seconds, 3 items 1 second, 3 items 4 seconds). For the between-group analysis, an ANCOVA, with age, gender, and education as covariates, with subsequent Holm post-hoc correction, was used to test differences across the groups on each cumulative metric while controlling for factors that differed across groups and could impact performance.

3.1.1.3 Results

Demographics and standard neuropsychological tests

Age, gender, and education were not matched across groups, and therefore were included as covariates when comparing group performances. As expected, ACE scores were statistically significantly different across groups, with AD showing the lowest scores, followed by MCI, then SCI and finally EHC. DS was not different between EHC and SCI but declined significantly in MCI and AD patients. Patients with SCI scored higher at tests of depression and quality of sleep compared to the other groups.

	EHC (n = 156)	SCI (n = 49)	MCI (n = 57)	AD (n = 63)	p-value all	p-value EHC/SCI	p-value SCI/MCI	p-value MCI/AD
Age	67.3 (8.6)	57.6 (7.8)	69.7 (8.9)	70.6 (9.0)	*<0.001	*<0.001	*<0.001	0.582
Gender (M/F)	58/98	23/26	39/18	34/29	*<0.001	0.222	*0.025	0.105
Education	14.9 (4.1)	15.3 (4.7)	11.8 (3.2)	13.2 (3.7)	*<0.001	0.480	*<0.001	0.136
Handedness (R/L/A)	128/26/2	45/4/0	49/6/2	57/3/3	0.150	0.305	0.799	0.470
ACE	95.9 (3.1)	92.7 (6.3)	87.9 (3.8)	73.9 (15.7)	*<0.001	*0.013	*0.004	*<0.001
DS	17.6 (4.4)	17.4 (5.2)	14.8 (3.3)	14.2 (4.5)	*<0.001	1.0	*0.008	1.0
HADS	8.3 (5.6)	23.0 (3.7)	11.5 (6.8)	9.2 (6.5)	*<0.001	*<0.001	*<0.001	0.129
GDS	2.4 (2.7)	6.2 (3.8)	5.1 (3.8)	3.8 (3.6)	*<0.001	*<0.001	0.097	0.066
PSQI	5.6 (3.1)	9.3 (4.4)	6.3 (3.6)	5.1 (3.6)	*<0.001	*<0.001	*<0.001	0.163

Table 3.1 | Demographics and tests

TRANSDIAGNOSTIC ANALYSIS - EFFECTS OF SET SIZE AND DELAY

There was a main effect of Set Size, of large magnitude ($\eta 2 > 0.14$) in all metrics, except Imprecision, where it was of medium magnitude ($\eta 2 > 0.06$), (**Table 3.2**, **Figures 3.3 and 3.4**). Set Size for Misbinding was not included because it cannot be computed for 1 item, as at least two objects are required for misbinding to occur. All metrics, except for Misbinding, showed also an effect of Delay, which however was of small ($\eta 2 > 0.01$) magnitude. Identification time, Identification accuracy, Absolute Error and memory Imprecision also showed a significant Set Size by Delay interaction, which was however of small magnitude ($\eta 2 > 0.01$), with a synergistic impact of higher number of items and longer delays determining worse memory performance. Set Size and Delay effects calculated per each group separately can be found in Appendix, **Table 3.1S**.

	Set Size	Delay	Set Size x Delay
Identification time	F = 493.6, ***p < 0.001** η2 = 0.466	F = 122.0, ***p < 0.001** η2 = 0.037	F = 4.1, ***p = 0.043** η2 = 0.001
Localization time	F = 423.2, ***p < 0.001** η2 = 0.416	F = 161.8, ***p < 0.001** η2 = 0.046	F = 1.8, p = 0.186
Identification accuracy	F = 1298.7, ***p < 0.001** η2 = 0.624	F = 168.2, ***p < 0.001** η2 = 0.043	F = 20.7, ***p < 0.001** η2 = 0.006
Absolute Error	F = 1625.5, ***p < 0.001** η2 = 0.726	F = 109.5, ***p < 0.001** η2 = 0.019	F = 35.0, ***p < 0.001** η2 = 0.005
Target detection	F = 3088.8, ***p < 0.001** η2 = 0.845	F = 34.1, ***p < 0.001** η2 = 0.004	F = 0.008, p = 0.927
Misbinding	N/A	F = 0.3, p = 0.577	N/A
Guessing	F = 145.0, ***p < 0.001** η2 = 0.211	F = 96.6, ***p < 0.001** η2 = 0.045	F = 1.3, p = 0.249
Imprecision	F = 49.6, ***p < 0.001** η2 = 0.085	F = 78.8, ***p < 0.001** η2 = 0.043	F = 5.0, ***p = 0.001** η2 = 0.002

Table 3.2 | Summary results from the Transdiagnostic analysis

Set Size (1 vs 3 items), Delay (1 vs 4 seconds), N/A = not applicable.

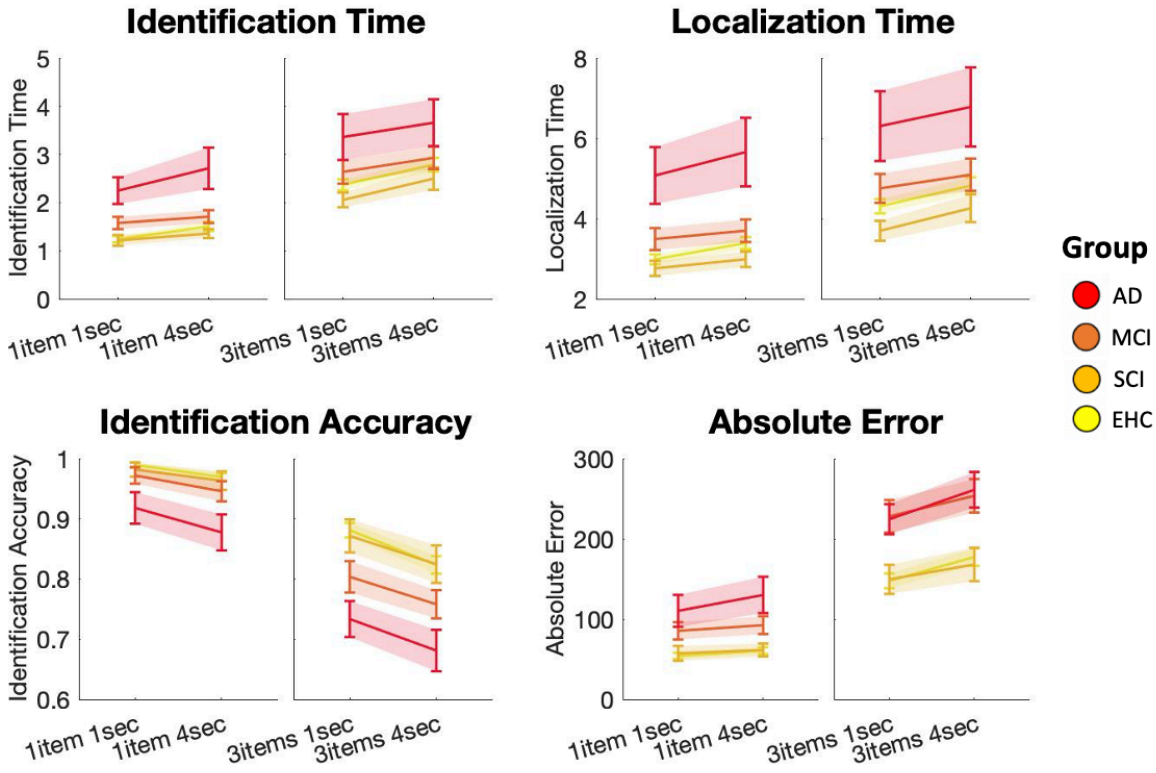


Figure 3.3 | Transdiagnostic analysis: Effects of Set Size and Delay – basic metrics

EHC = yellow, SCI = light orange, MCI = dark orange, AD = red.

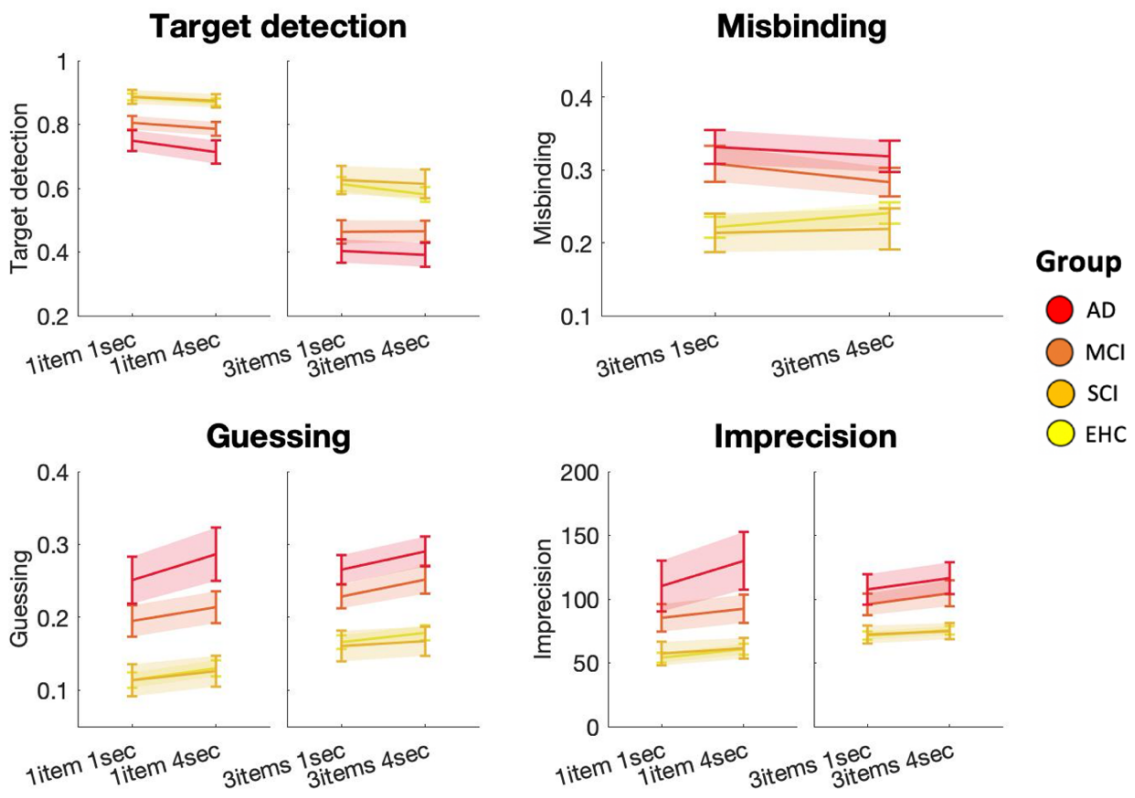


Figure 3.4 | Transdiagnostic analysis: Effects of Set Size and Delay – mixture model metrics

EHC = yellow, SCI = light orange, MCI = dark orange, AD = red.

BETWEEN-GROUP ANALYSIS - GROUP DIFFERENCES

All metrics were able to discriminate between the groups, with high effect size for a between-group difference (**Table 3.3**, column ALL, **Figures 3.5 and 3.6**), and were able to discriminate between EHC and AD and SCI and AD (**Table 3.3**, columns EHC/AD and SCI/AD). On post-hoc comparisons, no metric was able to distinguish between healthy controls and patients with SCI, which highlights the fact that despite these patients' complaints, they do perform normally on the VSTM test used. On the other hand, compared to healthy controls patients with MCI showed lower Identification accuracy, had higher Absolute Error rates, lower rates of Targets detected, higher amount of Guessing, higher rates of Misbinding, and higher memory Imprecision.

Finally, Absolute Error, Target detection, Misbinding and Guessing were also able to discriminate between subjects with SCI and MCI (**Figures 3.5 and 3.6, Table 3.3**, column SCI/MCI). However, the precision of their memory recall (Imprecision), as well as the proportion of correctly identified items (Identification accuracy) were not different between those groups. Absolute Error was the only metric that was not able to track disease progression between patients with MCI and AD dementia, as likely already at ceiling in patients with MCI (**Figure 3.5, Table 3.3**, column MCI/AD). In the transition between MCI and AD dementia, Misbinding was significantly higher in patients with AD dementia, but was the metric with the lowest effect size (**Figure 3.6, Table 3.3**, column MCI/AD).

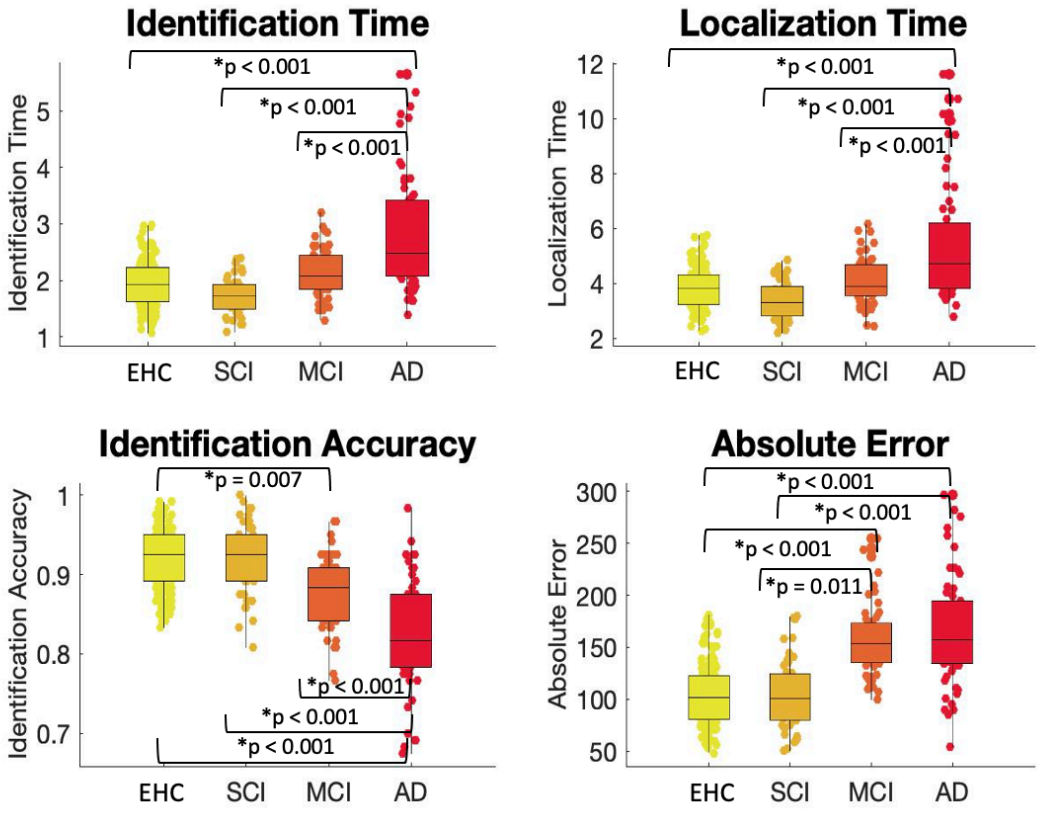


Figure 3.5 | Between-group analysis – basic metrics

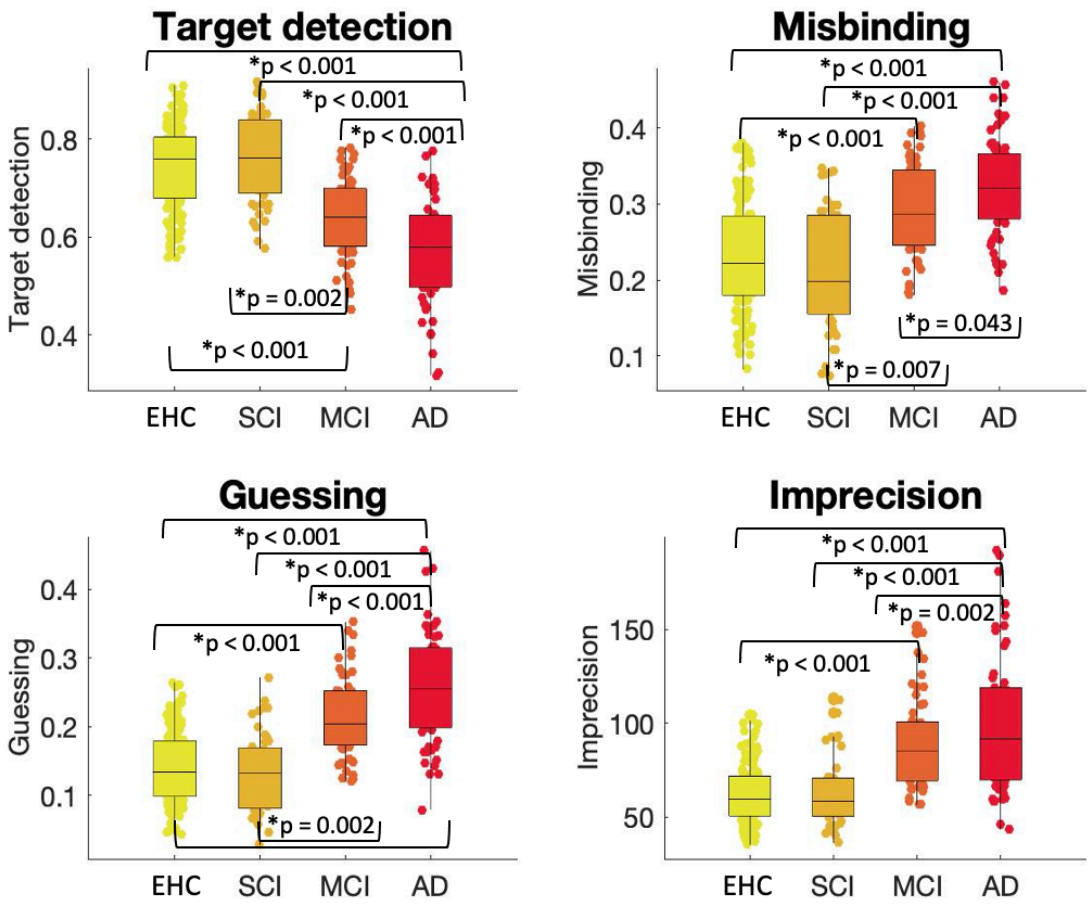


Figure 3.6 | Between-group analysis – mixture model metrics

	ALL	EHC/SCI	EHC/MCI	SCI/MCI	EHC/AD	SCI/AD	MCI/AD
Identification time	F = 27.9, ***p < 0.001** η2 = 0.205	t = 0.6, p = 0.506, d = 0.118	t = -2.3, p = 0.073, d = -0.374	t = -2.3, p = 0.073, d = -0.492	t = -8.8, ***p < 0.001**, d = -1.358	t = -5.3, ***p < 0.001**, d = -1.476	t = -5.3, ***p < 0.001**, d = -0.984
Localization time	F = 27.4, ***p < 0.001** η2 = 0.199	t = 0.8, p = 0.423, d = 0.142	t = -2.2, p = 0.059, d = -0.369	t = -2.8, p = 0.067, d = -0.502	t = -8.7, ***p < 0.001**, d = -1.342	t = -7.0, ***p < 0.001**, d = -1.484	t = -5.3, ***p < 0.001**, d = -0.972
Identification accuracy	F = 38.2, ***p < 0.001** η2 = 0.251	t = 1.8, p = 0.139, d = 0.322	t = 3.0, ***p = 0.007,** d = 0.504	t = 0.8, p = 0.403, d = 0.182	t = 10.7, ***p < 0.001**, d = 1.646	t = 6.3, ***p < 0.001**, d = 1.324	t = 6.2, ***p < 0.001**, d = 1.141

Absolute Error	F = 27.0, ***p < 0.001** η2 = 0.184	t = -1.5, p = 0.124, d = -0.273	t = -5.4, ***p < 0.001**, d = -0.909	t = -2.9, ***p = 0.011**, d = -0.6	t = -8.5, ***p < 0.001**, d = -1.314	t = -4.9, ***p < 0.001**, d = -1.041	t = -2.2, p = 0.058, d = -0.405
Target detection	F = 37.5, ***p < 0.001** η2 = 0.246	t = 0.9, p = 0.361, d = 0.162	t = 5.4, ***p < 0.001**, d = 0.900	t = 3.4, ***p = 0.002**, d = 0.738	t = 10.3, ***p < 0.001**, d = 1.587	t = 6.7, ***p < 0.001**, d = 1.426	t = 3.7, ***p < 0.001**, d = 0.687
Misbinding	F = 20.5, ***p < 0.001** η2 = 0.154	t = -0.3, p = 0.764, d = -0.053	t = -4.3, ***p < 0.001**, d = -0.720	t = -3.1, ***p = 0.007**, d = -0.667	t = -7.4, ***p < 0.001**, d = -1.147	t = -5.2, ***p < 0.001**, d = -1.094	t = -2.3, ***p = 0.043**, d = -0.427
Guessing	F = 42.0, ***p < 0.001** η2 = 0.268	t = -1.2, p = 0.245, d = -0.206	t = -5.6, ***p < 0.001**, d = -0.920	t = -3.3, ***p = 0.002**, d = -0.714	t = -11.0, ***p < 0.001**, d = -1.688	t = -7.0, ***p < 0.001**, d = -0.768	t = -4.2, ***p < 0.001**, d = -0.768
Imprecision	F = 24.7, ***p < 0.001** η2 = 0.179	t = -1.4, p = 0.141, d = -0.261	t = -4.1, ***p < 0.001**, d = -0.672	t = -1.9, p = 0.120, d = -0.411	t = -8.5, ***p < 0.001**, d = -1.311	t = -5.0, ***p < 0.001**, d = -1.050	t = -3.4, ***p = 0.002**, d = -0.639

Table 3.3 | Summary results from the Between-group analysis

3.1.1.4 Discussion

Across groups, Set Size effects had a major impact on memory recall, showing an effect of large magnitude in all metrics, except for memory Imprecision, where it was of medium magnitude. This highlights how consuming WM capacity resources by adding items to remember results in lower performance, not only by reducing the precision with which the items are stored, but also causing slower responses, higher spatial localization error, lower percentages of correctly identified items and targets detected, and higher rates of random guessing and misbinding. Apart from Misbinding, all other metrics also showed an effect of Delay, which however was of small magnitude. This points towards the fact that

remembering an increasing number of items has a much more detrimental impact on memory recall compared to time-dependent memory degradation.

A significant effect of Set Size on memory performance is in line with what previously reported with the same paradigm in healthy controls and patients with sporadic AD and PD [76, 78]. However, the latter studies either did not have sufficient data to examine the effect of delay [76,] or investigated only a limited subset of these metrics and included a large number of healthy controls but a very small number of patients with either AD or PD (n = 19 each) [78]. Interestingly, in the latter study, healthy controls showed higher misbinding rates at longer delays, which was not present when looking at healthy controls and patient groups together. In Appendix, **Table 3.1S**, it is shown that also in our study there is a similar effect of delay on misbinding in healthy controls.

Regarding the between-group analysis, all metrics were able to discriminate between healthy controls and AD dementia patients, as well as between patients with SCI and AD patients. However, no metric was able to distinguish between healthy controls and patients with SCI, possibly suggesting that these metrics are not dependent on psychological determinants of self-reported impaired cognition. This finding is consistent with data that show that the majority of patients with SCI will not subsequently develop AD dementia [257]. Another important finding is that Absolute Error, Target Detection, Misbinding and Guessing were able to discriminate between patients with subjective and mild cognitive impairment. This might have potentially useful clinical implications, as in clinical practice it can be hard to distinguish between patients with SCI from those with MCI on standard tests of cognition.

Both Identification accuracy and memory Imprecision did not differ significantly between healthy controls and patients with SCI, nor between patients with SCI and MCI, while there was a significant difference between healthy controls and patients with MCI. This might suggest that there could be a more graded difference between the three groups, which is too small to be detected with the current sample sizes.

However, in the comparison between MCI and AD, all metrics, except Absolute Error, were significantly different. Absolute Error seems to be sensitive to early stages of disease, as in the transition between healthy controls and patients with MCI, as well as between patients with subjective and mild cognitive impairment, but beyond the MCI stage it does not deteriorate significantly further in patients with AD. However, is worth noticing that the difference was just above the threshold for significance ($p = 0.058$), and therefore we cannot rule out that this measure might be able to detect such difference in another dataset.

Overall, Absolute Error can be seen as an '*early*' digital marker, as changes more prominently in the early phases of cognitive impairment. Interestingly, it was the only marker which was able to highlight disease progression in asymptomatic mutation carriers at risk of developing familial AD in a study by Pavisic et al [261], while in the same report Identification accuracy only declined in symptomatic patients. These findings might suggest that Absolute Error can indeed be seen as an early marker of cognitive impairment.

On the other hand, Identification time and Localization time can be viewed as '*late*' markers, as in our sample they changed only when patients reached the AD dementia stage, in line with known effects of slowing of responses, e.g., slower reaction times (RT), in patients with AD dementia compared to age-matched elderly controls [262] and patients with MCI [263,

[264]. Evidence of slower reaction times in patients with MCI compared to age-matched controls has been more mixed, and whilst a recent metanalysis reported reduced RT in this patient population [265], our current study does not support this finding. However, the investigations included in the metanalysis included only two diagnostic groups, whilst in our analysis four groups were compared and the analyses were corrected for multiple comparisons, with a lower risk of publication bias but also possibly a reduced statistical significance in head-to-head differences.

Mixture model metrics including Target detection, Guessing and Misbinding, were able not only to distinguish between a healthy status and MCI, but also to discriminate between patients with SCI and MCI and MCI and AD, so were useful metrics to provide an accurate stratification of patients and can be considered as both '*early*' and '*late*' markers.

3.1.2 Linear Support Vector Machine classifier using VSTM metrics for patients' stratification

3.1.2.1 Introduction

Whilst is important to understand the validity of each single digital metric in discriminating between groups, it is unlikely that any of these metrics will be used in isolation to aid diagnostic classification. Machine learning approaches, whether completely unsupervised or supervised, can help us estimate how the combination of different parameters impact overall accuracy in patients' stratification. For this study, as we were not blinded to the diagnostic groups, we used a supervised method, the linear support vector machine [266], to

test the performance of our digital metrics in patients' classification. We then compared the resulting model to one using standard neuropsychological tests.

3.1.2.2 Methods

A linear support vector machine classifier was used to test the performance of a model including age, gender, education and the eight digital metrics derived from the "What was where?" Oxford memory Task (OMT) described in 3.1.1, in discriminating between the four different diagnostic groups (EHC, SCI, MCI and AD), and a competing model using age, gender, education, and ACE scores, also derived from the same dataset.

3.1.2.3.1 Results - overall classifier

For the model using the combination of our eight metrics (here labeled as OMT), overall accuracy was 61.8%, and the area under the curve (AUC) for predicting group classification was respectively 0.82 for healthy controls, 0.85 for patients with SCI, 0.80 for MCI, and 0.87 for AD dementia. In comparison, overall accuracy for the model including ACE was 71.4%, and the AUCs were respectively 0.92 for healthy controls, 0.83 for patients with SCI, 0.90 for MCI and 0.96 for AD dementia (**Figures 3.7 and 3.8**).

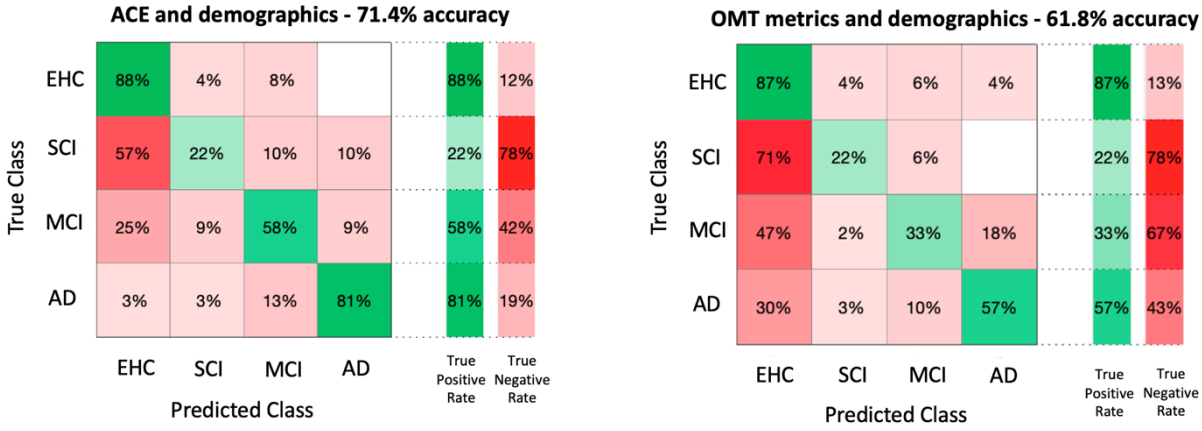


Figure 3.7 | Linear support vector machine classifier – all groups

We computed pairwise comparisons using the DeLong method for comparing receiver operating characteristic (ROC) curves [267] between the model including OMT and the one including ACE for each group, which were respectively: EHC: Z = 4.83, *p < 0.001, SCI: Z = - 1.34, p = 0.179, MCI: Z = 3.81, *p < 0.001, AD: Z = 3.89, *p < 0.001. Therefore, the model using our digital metrics did not perform better than ACE in classifying healthy controls, patients with MCI and AD dementia, but was as good as the ACE in patients with SCI (**Figure 3.8**).

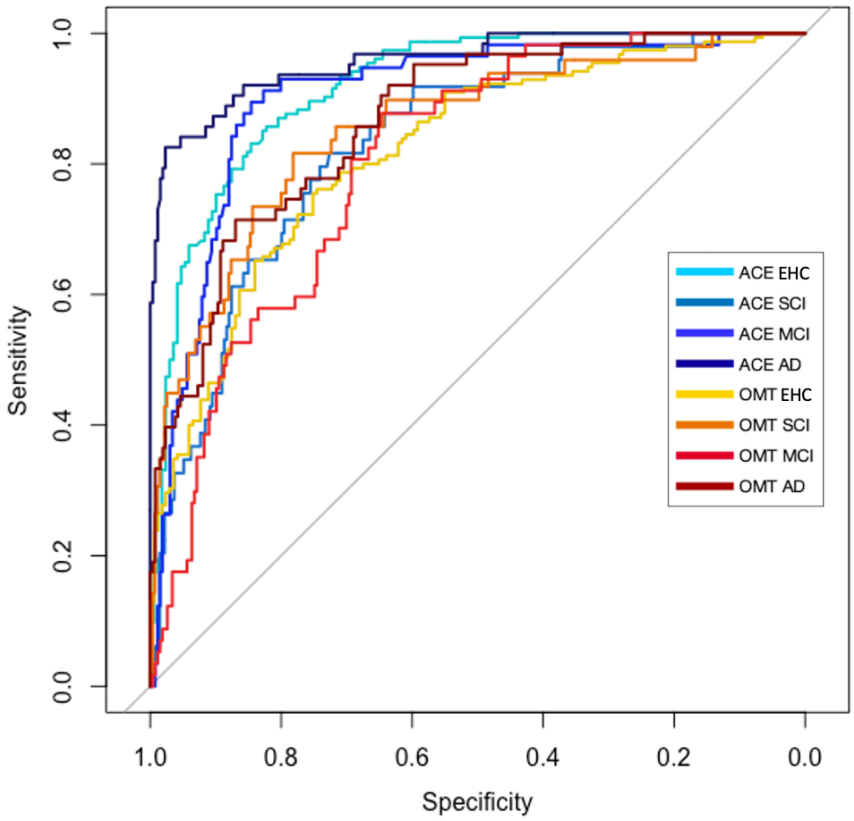


Figure 3.8 | Head-to-head comparison of ROC curves for ACE and OMT models

The first model included ACE, age, gender and education, and the second all the eight OMT metrics, age, gender and education. Both models were tested for predicting group classification separately for EHC, SCI, MCI and AD. Graded blue colours belong to ACE model: light blue = EHC, turquoise = SCI, blue = MCI, dark blue = AD. Graded red colour belong to the OMT model: yellow = EHC, orange = SCI, red = MCI, dark red = AD.

3.1.2.3.2 Results - pairwise classifier

Despite these multigroup classifiers being useful to blinded automatic diagnostic classifications, in a clinical setting we are less likely presented with a scenario when we have four different diagnoses that could fit the subject's clinical profile. We therefore performed specific pairwise group comparisons where our metrics could potentially be clinically useful.

As shown in **Figure 3.9**, the model including ACE and the one using our OMT metrics perform similarly in the differential diagnosis between healthy elderly controls and patients with SCI (AUC ACE = 0.86, AUC OMT = 0.82, Z = 1.33, p = 0.182) and between SCI and MCI (AUC ACE = 0.91, AUC OMT = 0.92, Z = -0.757, p = 0.449). On the contrary, when looking at the differential diagnosis between patients with MCI and AD dementia, whilst the ACE performs extremely well, our OMT metrics show a steep drop in accuracy (AUC ACE = 0.91, AUC OMT = 0.75, Z = 2.84, *p = 0.004).

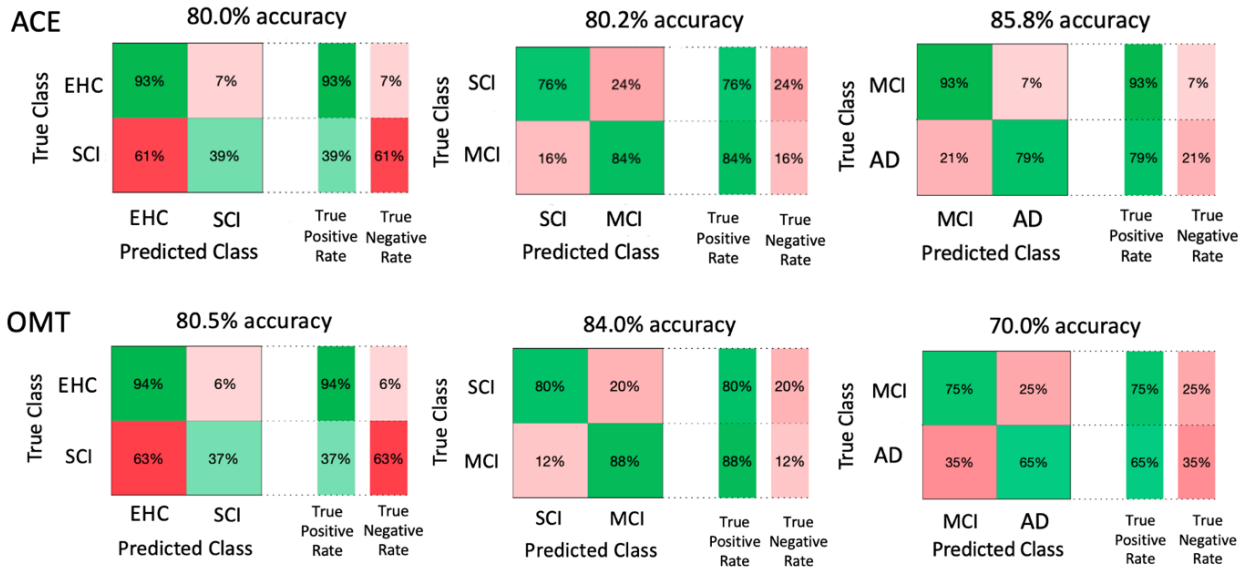


Figure 3.9 | Linear support vector machine classifier – pairwise

The first model (ACE) includes ACE, age, gender and education, while the second (OMT) includes all eight OMT metrics, age, gender and education.

3.1.2.4 Discussion

To conclude, while ACE is overall superior to our OMT metrics and is better in the discrimination between patients with MCI and patients with AD dementia, our digital metrics seem to perform similarly to traditional pen-and-paper tests of cognition when comparing healthy controls to patients with SCI or patients with SCI to MCI, without requiring dedicated, time-consuming face-to-face appointments since they can be performed remotely.

3.1.3 Relationship between VSTM metrics and baseline hippocampal volumes

3.1.3.1 Introduction

Previous data using a similar behavioral paradigm to the one used in this study suggest that specific metrics, such as misbinding, correlate with measures of hippocampal atrophy, at least in selected patients' populations such as patients with familial AD [77]. However, it is unclear whether this finding can be generalizable also to patients with sporadic, late-onset AD. Moreover, we don't know if that also applies to the other digital metrics that can be extracted by this paradigm, or if only misbinding is a reliable marker of hippocampal integrity. Indeed, while progressive hippocampal atrophy correlates with cognitive decline at standard neuropsychological testing and is commonly associated with conversion to AD dementia in sporadic patients and familial AD [256, 268, 269], little is known about how well our metrics perform in detecting loss of hippocampal volume. Additionally, while the left hippocampus is commonly a stronger marker of ageing and AD pathophysiological changes

[270], visual working memory has traditionally been linked more strongly with functions regulated by the right hemisphere [271].

The main questions I sought to address were:

1) Are these VSTM metrics correlated with hippocampal volume (HV), and if so, which metric is the best predictor?
2) Are these digital VSTM metrics better compared to standard neuropsychological tests such as the ACE in predicting hippocampal atrophy?
3) Are these results generalizable both to the right and left hemisphere, i.e., right and left hippocampus, compared to whole hippocampus?
4) Is this correlation mirroring an AD pathological process or represents a physiological construct, and therefore can be detected in a population of healthy individuals?

3.1.3.2 Methods

Participants

138 participants from the cohort described in 3.1.1 (EHC: n = 61, SCI, n = 31, MCI, n = 9, AD n = 37) agreed to a 3T structural MRI scan. Demographics and standard tests of cognition for this subsample are presented in **Table 3.4**. Group differences in demographics and test scores were calculated using the same principles outlined in 3.1.1.2.

	EHC (n = 61)	SCI (n = 31)	MCI (n = 9)	AD (n = 37)	p-value all	p-value EHC/SCI	p-value SCI/MCI	p-value MCI/AD
Age	66.6 (7.5)	57.7 (7.2)	65.8 (11.3)	69.2 (9.7)	*<0.001	*<0.001	*0.048	0.546
Gender (M/F)	24/37	14/17	4/5	22/15	0.285	0.592	0.969	0.415
Education	16.2 (3.8)	16.0 (4.5)	16.6 (4.4)	14.1 (3.6)	*0.046	1	1	0.411
ACE	97.2 (2.5)	93.9 (5.7)	89.6 (3.9)	70.5 (19.3)	*<0.001	0.335	0.335	*<0.001

Table 3.4 | Demographics and tests – neuroimaging sample

Statistical analysis and brain imaging pipeline

We computed left, right and whole hippocampal volumes, head-size corrected, using the same procedure as in paragraph 2.3.1. A generalized linear model was used to study correlations between neuropsychological measures and hippocampal volumes, while correcting for age, gender and education. Comparisons between regression coefficients were compared using the Cocor package in RStudio [272].

3.1.3.3 Results

All metrics were able to independently predict hippocampal volume in the whole dataset (**Figure 3.10**). As comparison, age had a smaller effect size compared to Identification accuracy in predicting hippocampal volume in the regression analysis (age: t = - 2.823, *p = 0.005, Identification accuracy, t = 3.251, *p = 0.001, h-test, z = - 4.3602, *p < 0.001). Moreover, Identification accuracy was the metric that was more tightly correlated with

hippocampal volumes compared to the other metrics (h-test, z = 3.721, *p < 0.001), and performed as well as ACE (t = 3.761, *p < 0.001) in predicting hippocampal volume (h-test, z = - 0.190, p = 0.8494).

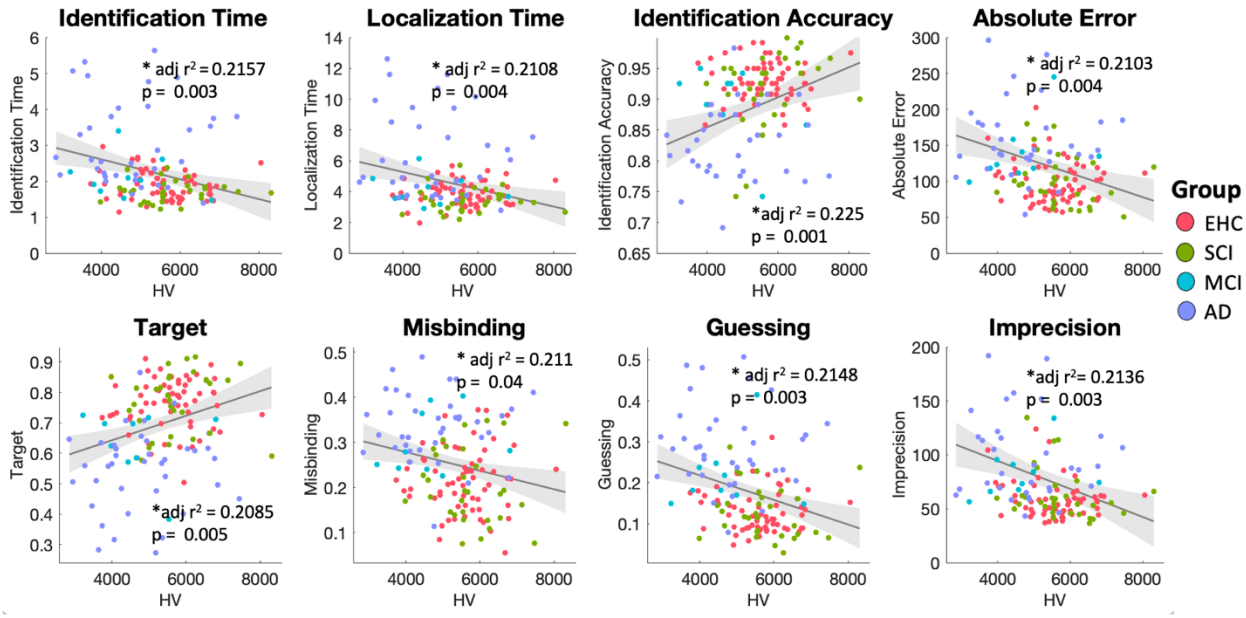


Figure 3.10 | Correlations between hippocampal volumes and our eight metrics

r^2 represents the overall model fit, the p-value refers to the contribution of the metric to the regression. EHC in coral red, SCI in green, MCI in light blue, AD in violet. The regression line has been plotted for the whole dataset, in grey. HV = head-size corrected hippocampal volume, in mm^3.

Additionally, we found that all our metrics, as well as the ACE, were able to predict separately both left and right hippocampal volumes (Appendix, **Table 3.2S**). Finally, neither ACE nor any of our metrics were able to predict HV volume in our healthy control sample.

3.1.3.4 Discussion

Previous evidence before this study pointed towards Misbinding as being the most important marker of hippocampal integrity. Whilst we could replicate in our dataset of patients in the sporadic late-onset AD continuum that misbinding rates are indeed associated with

hippocampal atrophy, this was not the best predictor amongst the eight different digital working memory metrics used for this experiment. Conversely, Identification accuracy stood out as the metric that most strongly correlated with hippocampal atrophy, performing equally as well validated standard neuropsychological test such as the ACE. As Identification accuracy is the metric that could be seen as a proxy of overall performance in our VSTM task, this result is very reassuring, as seems that even if what we deploy is a test of VSTM, it performs as well as a standard test of general cognition, which scales with hippocampal atrophy. It is also worth underlying that these results were corrected for common confounding factors that could impact performance as well as hippocampal integrity, such as age, gender and levels of education. More importantly, Identification accuracy explained more variance compared to ageing in predicting hippocampal volume in this dataset, which is also encouraging. However, these results did not generalize to our cohort of healthy individuals alone, where none of our metrics, nor the ACE, could be used to predict hippocampal volumes. These results were still valid whether analysing either right and left hippocampal volumes separately, and therefore lateralization might not be playing an important part, at least in this analysis.

3.2 Longitudinal evaluation of Visual Short-Term Memory decline after 1 year in patients with SCI, MCI and AD §‡¥#

3.2.1 Introduction

In Chapter 3.1 we have shown how these fine-grain working memory metrics have each one a different profile, as being more or less impacted by trial design variables such as Set Size or Delay effects, and how they can be used cross-sectionally to discriminate between different patient populations, including patients with SCI, MCI and AD. However, longitudinal data on these metrics are surprisingly scarce. At the present time, there is only one published study that tried to assess this question using a similar paradigm, but enrolling patients with familial AD [261]. Thus, longitudinal data on sporadic late onset AD are lacking, as well as data in patients at risk of developing AD such as SCI and MCI.

In the abovementioned paper, Pavisic et al studied a cohort of symptomatic (n = 6) and presymptomatic (n = 23) mutation carriers (PMC) of a gene associated with familial AD (either amyloid precursor protein (APP) or PSEN1 gene) for up to 5 years. Symptomatic carriers were defined as mutation carriers who had cognitive symptoms compatible with AD, while presymptomatic carriers carried the mutation but had no symptoms at the time of enrolment. The authors also estimated the expected year of onset of PMC based on individuals' age at the time of assessment subtracted from the age at which their affected parent developed symptoms and subdivided them in early (> 8.5 years from expected onset), and late (< 8.5 years) presymptomatic carriers. They found that identification accuracy did not change in controls and in PMC carriers but declined over time in symptomatic carriers.

Absolute error was stable over time in controls and early PMC, while declined over time in late PMC and symptomatic carriers, with late PMC being particularly affected at longer delays (4 seconds), and in the 3 items condition. Absolute Error was also greater the closer the PMC were to their estimated year of onset of dementia. Misbinding rates did not change over time in any of the groups. Crucially, a significant difference between late PMC and controls on the best performing standard neuropsychological tests in that study, a delayed memory task, was able to detect a difference between late PMC and controls one year later compared to Absolute Error performance. This has important clinical implications as it means that using these kinds of sensitive tests can give useful predictive information in this population, and potentially speed up diagnostic referrals by 1 year compared to traditional tests, and up to 6 years prior to cognitive symptoms.

However, the study deployed a small but well characterized genetic cohort, where people carrying these pathological mutations will undoubtedly eventually develop the disease. This is very different compared to conditions such as subjective cognitive impairment and mild cognitive impairment, where patients are simply 'at risk' of developing AD dementia, but there is no guarantee of whether and when they will develop an AD dementia phenotype.

The questions I aimed to assess in this chapter were:

1) How stable are these metrics over time during a 1-year period in healthy controls, but also in clinical populations along the AD disease spectrum?
2) How do these changes, or absence of changes over time, compare to decline in standard neuropsychological tests in these patient populations?

3) Which metrics at baseline are the best predictors of cognitive decline at follow-up indexed by a reduction of overall cognition measures by standard neuropsychological changes (as ACE)?

3.2.2 Methods

Participants

A subset of 60 people from the cohort analyzed in Chapter 3.1, took part in the longitudinal part of the study and completed the repeated assessment at 1 year. Of these 60 participants, 21 were healthy elderly controls, 15 were patients with SCI, 12 were patients with MCI, and 12 were patients with AD dementia. It is worth mentioning that these very small numbers are due to the study being prematurely interrupted due to COVID-19 restrictions. Whilst we have subsequently developed a fully remote, online version of this task, the original task we used in Chapter 3.1 was a tablet task, which required face-to-face administration. Enrolment and classification criteria followed the same principles outlined in Chapter 3.1. At both visits all subjects completed the "What was where?" tablet-based task, the ACE-III [59] and Digit Span [220]. The summary of demographics and test scores can be found in **Table 3.5**. Group differences in demographics and test scores were calculated using the same principles outlined in 3.1.1.2.

	EHC (n = 21)	SCI (n = 15)	MCI (n = 12)	AD (n = 12)	p-value all	p-value EHC/SCI	p-value SCI/MCI	p-value MCI/AD
Age	66.6 (7.5)	59.9 (8.6)	69.3 (11.5)	69.8 (11.2)	*0.040	0.165	0.065	1
Gender (M/F)	6/15	8/7	8/4	5/7	0.175	0.714	1	0.877
Education	15.8 (3.6)	15.9 (5.2)	14.0 (4.7)	15.5 (6.2)	0.448	1	1	1
ACE	97.4 (2.0)	94.3 (3.5)	86.7 (7.7)	71.7 (5.2)	*<0.001	0.052	*<0.001	*<0.001
DS	17.7 (2.9)	18.2 (3.8)	17.2 (4.8)	13.8 (2.2)	*0.007	1	1	0.079

Table 3.5 | Demographics and tests - longitudinal subset

Statistical analysis

A 4 (Group) x 2 (Session) ANCOVA, with age, gender and education as covariates, with subsequent Holm post-hoc correction, was used to test differences across the groups and sessions on each cumulative metric while controlling for confounding factors that might differ across groups.

Additionally, we tested whether the baseline values extracted from our tests could be used to predict cognitive decline over one year at standard neuropsychological tests. To this end, we calculated the change between ACE scores from the second visit and the first visit, and looked at whether any of the metrics was able to predict cognitive decline after 1 year, and we also tested which metric performed the best. To explore this question, we run linear regression for each metric separately and compared the results with Cocor package in RStudio.

3.2.3 Results

Group and Session effects

All metrics showed a significant effect of Group, with high effect size ($\eta 2 > 0.14$), even if the sample size was much smaller compared to the bigger cross-sectional dataset in 3.1 (**Table 3.6, Figures 3.1** and **3.12**). As the effect of Group was investigated in Chapter 3.1 and was not the main aim of Chapter 3.2, post-hoc comparisons across groups are presented in Appendix, **Table 3.3S**, and **Figures 3.1S** and **3.2S**.

Only Localization time showed a main effect of Session ($F = 4.3$, $*p = 0.039$, $\eta 2 = 0.024$) (**Table 3.6, Figure 3.11)**. Absolute Error and memory Imprecision showed a significant Group by Session interaction (**Table 3.6, Figures 3.11** and **3.12**). Post-hoc analysis for Absolute error and Imprecision showed a significant difference between sessions only in the AD group (Absolute error: $t = -3.120$, $*p = 0.031$, $d = - 1.274$, Imprecision: $t = - 4.064$, $*p = 0.002$, $d = -1.660$).

	Group	Session	Group x Session
Identification time	F = 35.0, ***p < 0.001**	n.s	n.s
	η2 = 0.472		
Localization time	F = 20.3, ***p < 0.001**	F = 4.3, ***p = 0.039**	n.s
	η2 = 0.333	η2 = 0.024	
Identification accuracy	F = 46.6, ***p < 0.001**	n.s	n.s
	η2 = 0.551		
Absolute Error	F = 64.9, ***p < 0.001**	n.s	F = 3.2, ***p = 0.025**
	η2 = 0.614		η2 = 0.031
Target detection	F = 72.7, ***p < 0.001**	n.s	n.s
	η2 = 0.667		
Guessing	F = 67.6, ***p < 0.001**	n.s	n.s
	η2 = 0.640		
Misbinding	F = 28.9, ***p < 0.001**	n.s	n.s
	η2 = 0.447		
Imprecision	F = 45.9, ***p < 0.001**	n.s	F = 4.8, ***p = 0.004**
	η2 = 0.507		η2 = 0.053

Table 3.6 | Overview of Group and Session effects for our eight digital working memory metrics

Group reflects four diagnostic categories: healthy elderly controls, subjective cognitive impairment, mild cognitive impairment, Alzheimer's Disease dementia. Session represents baseline session (Time 0) and follow-up session after 1 year. n.s. = not significant.

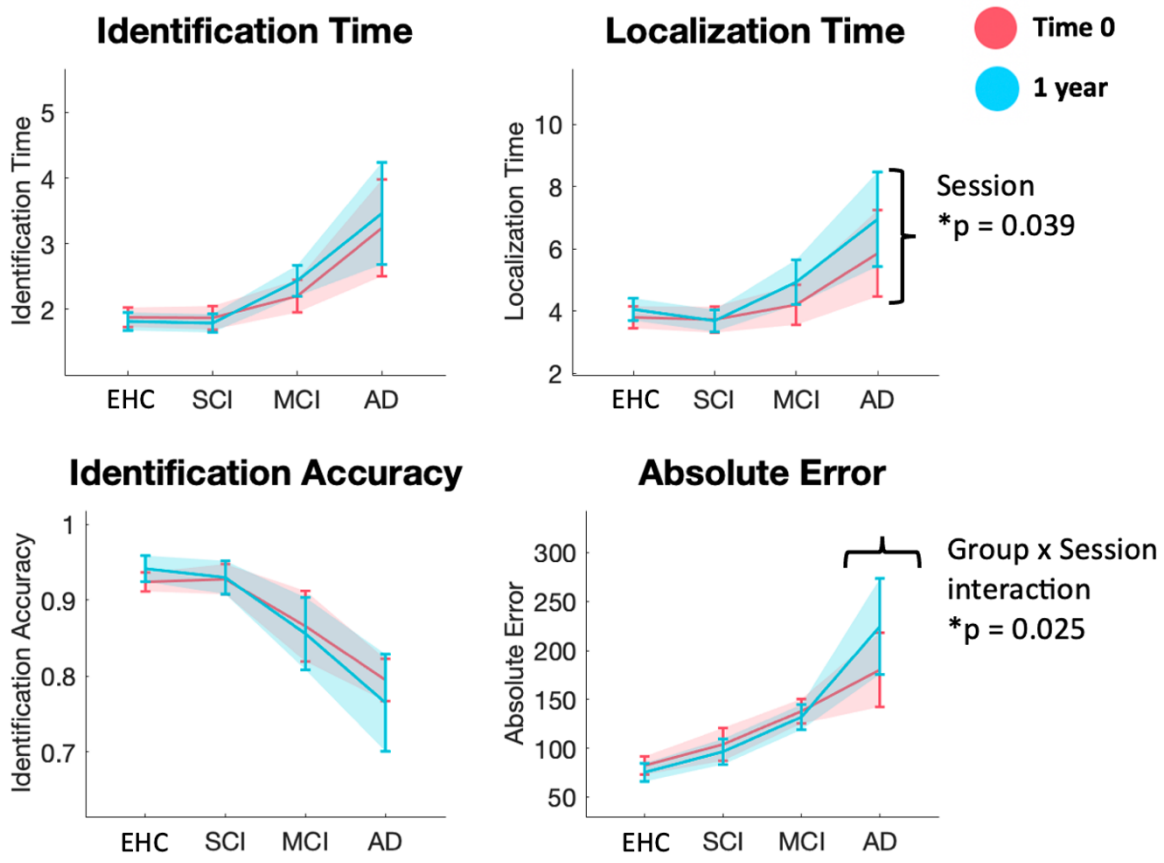


Figure 3.11 | Longitudinal analysis - basic metrics

Baseline session (Time 0) in light blue, follow-up session after 1 year (1 year) in coral red.

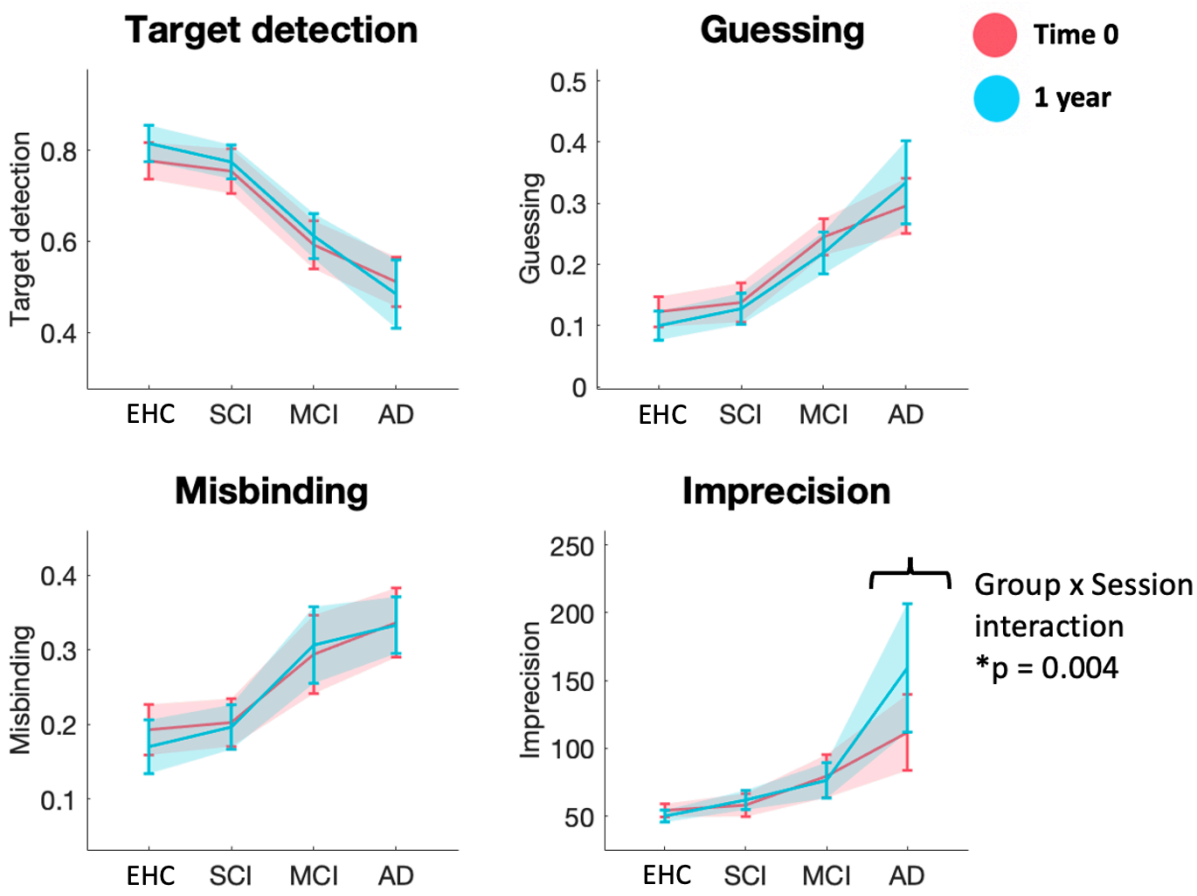


Figure 3.12 | Longitudinal analysis - mixture model metrics

Baseline session (Time 0) in light blue, follow-up session after 1 year (1 year) in coral red.

As a comparison, we also computed the effect of Group and Session for the two neuropsychological tests used, the ACE and DS, (**Figure 3.13**). For ACE there was an effect of Group ($F = 206.04$, $*p < 0.001$, $\eta 2 = 0.812$), an effect of Session ($F = 10.16$, $*p = 0.002$, $\eta 2 = 0.013$), and a Group x Session interaction ($F = 9.41$, $*p < 0.001$, $\eta 2 = 0.037$). For DS there was only an effect of Group ($F = 6.67$, $*p < 0.001$, $\eta 2 = 0.160$), but no effect of Session or a Group x Session interaction. Post-hoc analysis was significant for ACE with respect to session for the AD group ($F = 5.82$, $*p < 0.001$), and for all groups comparisons ($*p < 0.001$ except EHC vs SCI $*p = 0.009$), and for DS only for group comparisons between the AD

group and the other groups (EHC vs AD *p <0.001, SCI vs AD *p = 0.003, MCI vs AD *p = 0.009).

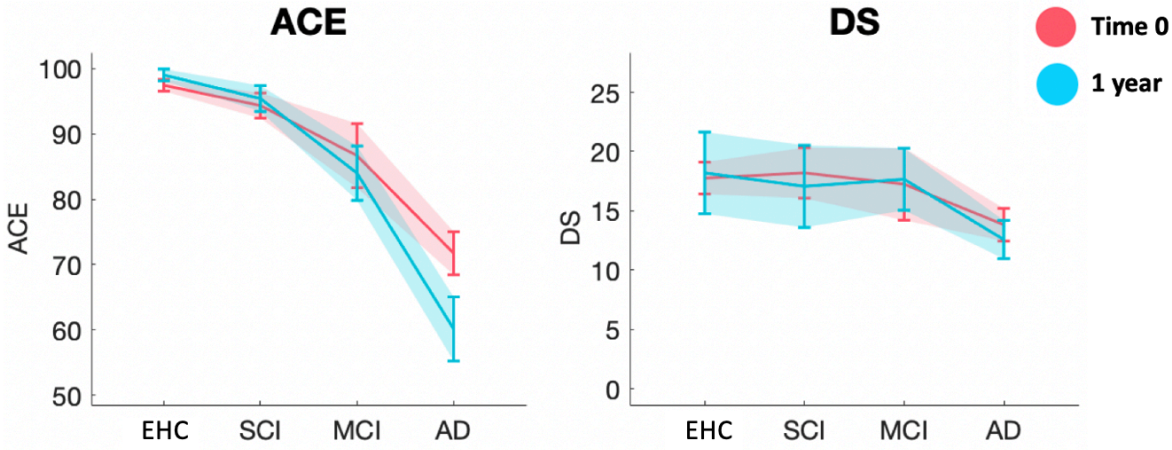


Figure 3.13 | Longitudinal analysis – standard neuropsychological tests

Digital metrics in the prediction of cognitive decline

All metrics at baseline were able to independently predict cognitive decline after 1 year (Appendix, **Table 3.4S**). The metric which performed the best in predicting a decline in ACE scores after 1 year was memory Imprecision, t = 5.5, *p < 0.001, R^2 = 0.411. As a comparison, both ACE itself at baseline and hippocampal volume performed worse than memory Imprecision (ACE: h-test ACE, z = 6.34, *p < 0.001, Hippocampal volume: h-test, z = 5.74, *p < 0.001).

3.2.4 Discussion

Measuring changes in our cognitive metrics after one year revealed that these metrics show different profiles: some tend to remain very stable over time, while others follow a similar trajectory as the decline in performance on the ACE. While Identification accuracy in Chapter 3.1 seemed the metric which was able to reflect most faithfully hippocampal

atrophy, it was not the metric that best tracked individual changes in cognition over time. These were respectively Absolute Error, but more so, memory Imprecision. Indeed, memory Imprecision performed very well in the longitudinal tracking of cognitive decline, so well that its baseline values outperformed baseline levels of ACE and also hippocampal volumes.

One of the most interesting results was that despite a decline in cognition, indexed by a reduction in ACE scores, mixture model metrics (Target detection, Misbinding and Guessing) did not change over time. Why is that so? To answer this question, we need to bear in mind how these metrics are calculated. They do not reflect absolute values of errors, but they calculate the probability of a response belonging to correct target identification, misbinding and guessing, which all add up to 1, as they are relative to each other. Therefore, the most likely explanation is that the proportion of mistakes made, relative to each other, remains the same across time.

Regarding the post-hoc analysis of Group comparisons in Appendix (**Table 3.3S**, **Figures 3.1S** and **3.2S**), is important to emphasize that the fact that most results seen in the bigger dataset were replicated in this much smaller dataset, with less than 22 participants per group, makes the use of these metrics very encouraging as it suggests that only very small numbers are needed to detect groups differences, and therefore these metrics might indeed be very sensitive to subtle changes in cognition. Findings that were not replicated in the bigger cross-sectional dataset presented in Chapter 3.1 can be found in Appendix (**Table 3.3S, Figures 3.1S** and **3.2S**) included that Absolute Error was able to discriminate between healthy controls and patients with SCI, and also between patients with MCI and AD dementia. However, in the bigger dataset there was a trend ($p = 0.058$) towards significance for the comparison between MCI and AD, so we cannot exclude that such an effect could be

observed with different sample sizes. On the same line, Identification accuracy in this smaller dataset was able to distinguish between patients with subjective and mild cognitive impairment. The mixture model metrics (Target detection, Misbinding and Guessing) were not able to distinguish between patients with MCI and AD in this smaller sample. This is not surprising as in the bigger dataset these metrics had much smaller effect sizes compared to late metrics such as Identification, Localization time, or Identification accuracy.

3.2.4 Conclusion

In Chapter 3, we have shown that digital visual working memory metrics extracted by the "What was Where?" Oxford memory task are useful to detect early signs of cognitive impairment in a large cohort of subjects in the AD spectrum. They are also able to detect disease progression cross-sectionally, as in the transition between MCI and AD, but also longitudinally after 1 year. Using a linear support vector machine classifier we found that these metrics seem to perform equally as good as standard tests of cognition such as the ACE in discriminating between healthy controls, patients with SCI and MCI, while they are not as accurate as ACE in later stages of the disease. Identification accuracy seems the metric that best reflects the degree of underlying hippocampal atrophy, while memory Imprecision is useful to predict cognitive decline at standard tests of cognition over one year.

Chapter 4. Structural correlates of Visual Short-Term Memory using Voxel-Based Morphometry (VBM) in AD and EHC §#

Abstract

In this Chapter we applied whole brain voxel-based morphometry (VBM) to investigate brain-behavioural correlations in a cohort of elderly healthy controls and patients with AD. While several brain areas were associated with poor performance at VSTM tasks, four different spatial patterns emerged across the eight VSTM metrics used: a) Target Detection and Guessing, b) Identification and Localization time, c) Identification accuracy, and d) Absolute Error and memory Imprecision. The Precuneus grey matter (GM) volume was associated to all metrics that carried spatial localization, but not object identity information (Identification accuracy and Identification time). Identification time, Localization time and Identification accuracy showed a prominent involvement of the frontal lobes, including the Anterior Cingulate Cortex and Frontal Medial Cortex. In EHC, increased Identification Time was associated with reduced GM volume in the right Frontal Pole, bilateral Orbitofrontal cortex, left Middle Frontal Gyrus, and in the right Middle Temporal Gyrus while memory Imprecision was associated with greater GM atrophy in the left Supramarginal Gyrus.

4.1 Introduction

Several distinct brain areas have been implicated in different aspects of WM function, depending on the nature of the material (verbal, spatial and objects), task requirements (including storage and manipulation) and task complexity [273, 274, 275, 276]. Bilateral fronto-

parietal networks are necessary for core functioning of WM and their contribution has been consistently reported across studies [97]. However, latest trends, aided by the development of multivoxel pattern analysis (MVPA), which enables stimuli decoding from multiple brain areas [277], support the idea that WM processes share the same areas involved in representing information in sensory perception, situated in posterior cortices [275, 278, 279, 280]. It is thought that the executive component of WM is what sustains and transforms posterior STM representations into the 'working' component of WM [281]. Prefrontal cortex (PFC), whether ventral or dorsal, has received great attention in terms of contribution to executive processes of WM, but parietal regions have also been increasingly recognized as playing a role [282], as well as PFC interactions with basal ganglia and thalamus [274], especially during learning [283].

Another way to consider anatomical specialization is that, while a widely distributed network underlies general WM processes, specific localization might reflect different levels of abstraction, with early sensory regions encoding simple features of stimuli in low-detail, sensory-like representations and prefrontal regions storing memories using more abstract, categorical representations, supporting upcoming behavioural responses [275].

Some studies in both monkeys and humans substantiate a *dorsal-ventral* segregation between spatial ''*where*'' information and identity ''*what*'' information, going from primary visual areas to the frontal lobes [284, 285, 286, 287]. According to these reports, the dorsal visual stream projections feed forward from the extrastriate cortex to the parietal lobule and the intraparietal sulcus (IPS) to dorsal regions of the PFC such as the superior frontal sulcus (SFS), while the ventral stream extends from temporal regions to ventral areas of the PFC, up to the inferior surface of the frontal pole. However, a large metanalysis of 60 neuroimaging (PET and fMRI) studies by Wager et al [273] showed the expected dorso-ventral

dissociation between spatial and nonspatial storage in the posterior cortex, but not in the PFC. Specifically, the Fusiform Gyrus (ventral) showed preference for object storage, while the Superior Parietal Lobule (dorsal) for spatial information. The Precuneus, coded verbal and spatial, but no object information.

Another large metanalysis of 189 neuroimaging (PET and fMRI) experiments by Rottschy et al [274] showed that memory for (non-verbal) object identity as compared to object location was significantly more likely to recruit bilateral Inferior Frontal Gyrus, left cerebellar lobule VI and left ventral visual cortex [274]. In the same analysis, object location as compared to identity localized bilaterally in the Superior Frontal Gyrus, the Superior Parietal Lobule and the Precuneus as well as the right Inferior Parietal Lobe [274]. Taken together, these studies suggest that in humans some dorsal parietal areas such as the Superior Parietal Lobule and the Precuneus might be important for object location, while more ventral areas might be implicated in storing object identity.

Another theory postulates that a dorsal-ventral specialization is linked not to the content of the memoranda per se but is rather secondary to the type of processing required: dorsal areas are more active during tasks with higher level of task demands and manipulation of information, while ventral areas are more involved in rehearsal during simple storage [276, 288]. A metanalysis on 36 experiments examined specifically areas involved in executive control in WM [276]. The authors found that two key regions were involved in several WM processes such as distractor resistance, attention shifting and updating: a dorsal region in the SFS, selective to spatial content, and a second, ventral region in the midlateral PFC, which showed sensitivity to nonspatial content. Therefore, contrary to metanalysis of Wager et al [273] but in line with the one by Rottschy et al [274], they concluded that the dorso-ventral

segregation theory also extends to the frontal lobe. Most importantly, executive components of WM follow the same anatomical specialization for spatial versus nonspatial content, rather than showing a clear-cut difference between different executive functions within WM.

Most functional imaging studies also found a preferential *lateralization*, left for verbal and object WM, right for spatial WM [273, 274, 275, 276]. This function segregation is however not consistently found across reports, and while there is more convincing evidence that verbal tasks activate preferentially areas within the left lateral frontal cortex (particularly Brodmann Areas (BAs) 44, 45, 46, and 6) [274, 276], some evidence showed no effect of lateralization for spatial working memory in posterior regions, but a preference for the right frontal regions for spatial tasks if the task demands are high [276].

Feature *binding* in WM has been frequently associated with hippocampal activity in healthy individuals as well as with disruption of hippocampal integrity in patients with hippocampal lesions [77, 143, 144, 210, 279, 289, 290, 291]. It has been suggested that the functional role of sustained activity observed in the hippocampus in WM tasks could be attributable to the recently proposed additional WM construct, the episodic buffer [2, 292, 293].

Sustained hippocampal activity during WM binding might be crucial for both WM maintenance of associative information and the encoding of new declarative memories [294, 295]. However, not all tasks that require binding rely necessarily on the hippocampus. Indeed, *conjunctive binding*, i.e., the integration of features within an object (such as color and shape) does not seem to be dependent on the hippocampus [290, 291, 252], while other posterior parieto-temporo-occipital areas, particularly the posterior parietal cortex (e.g., the Superior Parietal Lobule) might play a role [296, 297, 298, 299]. Conversely, *relational binding*, i.e., the

association of an object identity to other distinct features such as its spatial location or background scene, seems to be highly dependent on the hippocampus [77, 145, 252, 290, 291]. This has been found not only in LTM but also in WM, and particularly VSTM [77, 145, 252, 290, 291], and seems beneficial for subsequent successful retrieval [295]. Some evidence supports a selective role solely of the right hippocampus in relational binding [252, 291], but most studies found a bilateral involvement.

Beyond binding, another key feature of VSTM is the *precision* of memory representations. Multiple lines of evidence converge onto posterior parietal areas as being associated rather than with pure retrieval success, with the precision by which items are stored in memory. An fMRI study by Xu et al concluded that while the inferior IPS is crucial in maintaining spatial attention over a fixed number of objects at different spatial locations, the superior IPS and the lateral occipital complex (LOC) encoded a variable subset of the attended objects, depending on their complexity [298]. Deploying the mixture model of WM [17], Zhao et al found that the LOC was not affected by memory load but rather coded the precision by which participants remembered the stimuli [300]. Similarly, Richter et al showed that the Angular Gyrus might be coding the precision of memory representations, while the hippocampus could contribute to overall retrieval success [131].

From a behavioral standpoint, the study of reaction times (RT) is usually adopted to infer differences across conditions (e.g., low and high memory load). However, most brain-behavioral studies are designed to look at specific contrasts between experimental conditions, which have not always used RTs as outcome variables. RT variance may be due to latency variations in different regions of the brain, e.g., areas involved in sensory perception, sensory-motor transformation, or motor execution, and are thought to be

generated rather than from a single defined brain area, from the interplay between different interconnected regions [301], even if some evidence from a preferential role of the frontal cortex is emerging [302, 303, 304, 305]. However, data looking separately at reaction times for object identity (i.e., Identification time) and object spatial localization (i.e., Localization time), and comparing these to measures that are independent of RT, are currently scarce.

It is important to bear in mind that most of the abovementioned studies focus on healthy young individuals, and data on patients with AD are relatively rare [306], and no investigation so far deployed the mixture model of WM [17]. The paucity of studies is partially due to the widespread use of fMRI, where participants are asked to perform an experimental task inside the scanner, which can prove to be difficult for patients with AD.

In this Chapter, we explored whether out-of-scanner performance on a VSTM task, the "What was Where?" Oxford memory Task described in Chapter 3, was correlated with GM integrity in EHC and patients with AD, as AD primarily affects GM. We used voxel-based morphometry (VBM) to investigate such association. VBM has extensively been used in patients with AD for its ability to provide spatial information on GM loss that can help differentiate it from other types of dementia [307, 308]. GM atrophy measured by VBM is also associated with cognitive decline, as measured by standard tests of cognition such as the MMSE or Rey Auditory Verbal Learning Test (RAVLT) in older adults and AD patients [307, 309, 310]. In this Chapter, the Addenbrooke's Cognitive Examination-III (ACE) was used.

The key outstanding questions we sought to investigate in this Chapter were:

1) Are successful retrieval, memory precision, binding abilities and reaction times associated with different areas of the brain in a cohort of EHC and AD patients?

2) Are there any areas that contribute selectively to object identity or to spatial localization in this cohort?
3) Do these areas differ between the two groups?
4) Are these constructs independent of global cognitive impairment?

4.2 Methods

Participants

52 AD and 60 EHC were recruited from the Cognitive Disorders Clinic at the John Radcliffe Hospital in Oxford. Inclusion criteria for EHC and AD patients' classification criteria followed the same principles of Chapter 2 and 3. 30/52 AD patients had plasma biomarkers analysis as part of the study described in Chapter 6. Patients' basic demographics and tests are presented in **Table 4.1**. The two groups did not differ significantly in terms of age, but there was a higher proportion of males and lower education levels in patients with AD compared to EHC. As expected, performance on the ACE and the "What was where" Oxford Memory Task was lower in AD compared to EHC, across all the eight VSTM metrics described in Chapter 3. Statistical significance was derived from either independent sample t-test or Mann–Whitney U test according to the normality of the data. χ^2 test was used to compare gender distribution across groups. Rank-Biserial correlation was used as a measure of effect size for Mann–Whitney U derived statistics.

	EHC (n = 60)	**AD (n = 52)**	**p-value**	**Rank-Biserial Correlation**
Age	66.45 (7.9)	68.3 (10.1)	0.465	
Gender (M/F)	22/38	30/22	*0.026	
Education	17.3 (6.5)	14.6 (4.2)	*0.005	

ACE	97.3 (2.5)	74.6 (18.1)	*<0.001	
Identification time	1.89 (0.4)	3.14 (1.6)	*<0.001	-0.590
Localization time	3.8 (0.8)	6.3 (3.6)	*<0.001	-0.517
Identification accuracy	0.93 (0.0)	0.81 (0.1)	*<0.001	0.765
Absolute Error	96.4 (28.3)	175.5 (83.3)	*<0.001	-0.740
Target detection	0.76 (0.1)	0.56 (0.1)	*<0.001	0.793
Guessing	0.13 (0.05)	0.28 (0.1)	*<0.001	-0.789
Misbinding	0.21 (0.7)	0.32 (0.1)	*<0.001	-0.703
Imprecision	58.2 (16.8)	115.5 (73.4)	*<0.001	-0.719

Table 4.1 | Demographics and tests

MRI protocol

MRI was performed at 3T (Siemens Magnetom, Verio syngo MR B17) with the following sequences, harmonized to the UK Biobank protocol [311, 312]:

- T1-weighted sequence (3D MPRAGE, sagittal, TR = 2000 ms, TE = 1.94 ms, TI = 880 ms, in-plane acceleration factor (R) = 2, flip angle 8°, voxel-size = 1.0 x 1.0 x 1.0 mm, 208 x 256 x 256 matrix).
- T2-FLAIR sequence (FLAIR, 3D SPACE, sagittal, TR = 5000 ms, TE = 397 ms, TI = 1800 ms, partial fourier (PF) = 7/8, fatsat, voxel-size = 1.05 x 1.0 x 1.0 mm, 192 x 256 x 256 matrix).
- Diffusion-weighted images (dMRI) (Multiband acceleration factor (MB) = 3, TR = 3600 ms, TE = 92 ms, PF = 6/8, fatsat, voxel size 2.0 x 2.0 x 2.0 mm, 104 x 104 x 72 matrix) with opposite phase encoding directions (AP and PA) to allow for more robust distortion correction (AP: 104 directions, diffusion weighting: b1=0 s/mm^2, b2=2000 s/mm^2 and PA: 7 directions, diffusion weighting: b1 = 0 s/mm^2, b2 = 1000 s/mm^2).

- Resting-state fMRI (rsfMRI) images (MB = 8, flip angle 52°, fat sat, TR = 735 ms, TE = 39 ms, voxel size = 2.4 × 2.4 × 2.4 mm, 88 x 88 x 64 matrix).

Voxel-based morphometry

According to the inclusion criteria, T1 and T2-FLAIR scans were reviewed to exclude the presence of remarkable macroscopic brain abnormalities. Structural data was analysed with FSL-VBM [313] an optimised VBM protocol [314] carried out within the FSL toolbox [315]. First, T1 structural images were brain-extracted using Brain Extraction Tool (BET) [316]. Subsequently, images were segmented into GM before being registered to the Montreal Neurological Institute (MNI) 152 standard space using non-linear registration [317]. The resulting images were averaged and flipped along the x-axis to create a left-right symmetric, study-specific, GM template. Subsequently, all native GM images were non-linearly registered to this study-specific template and Jacobian-modulated to correct for local expansion or contraction due to the non-linear component of the spatial transformation. The modulated GM images were then smoothed with an isotropic Gaussian kernel with a sigma of 4 mm.

Finally, a voxelwise general linear model (GLM) was applied using permutation-based non-parametric testing (n permutations = 5000), correcting for multiple comparisons across space. Multiple comparisons correction between contrasts is embedded within the FSL toolbox. Age, gender and education were included in the GLM as covariates.

Firstly, a main effect of group was computed, to see whether our sample showed the expected pattern of GM atrophy seen in patients with AD compared to EHC. Secondly, we performed a whole brain voxelwise correlation analysis between the eight VSTM metrics

from the "What was Where?" Oxford memory Task described in Chapter 3 across the whole sample (AD and EHC). For each metric we calculated both contrasts (areas where a metric was associated with areas of higher and lower GM volumes), to check the robustness of our findings. We subsequently repeated the same analysis including also total ACE scores in the covariates, to test whether the results were independent of global cognitive impairment. We also separately calculated these brain-behaviour associations within each group (AD and EHC), to infer generalizability to physiological/pathological processes.

Second-level inference was thresholded at a family-wise error (FWE) corrected probability of $p < 0.05$ with threshold-free cluster-enhancement (TFCE) which provides both strong Type I and Type II error control [318].

4.3 Results

Group effect

As expected, there were widespread GM areas where AD patients showed higher atrophy compared to EHC (AD < EHC), with the peak of atrophy being in the left hippocampus (**Figure 4.1**, **Table 4.2**). No areas were found in the reverse contrast.

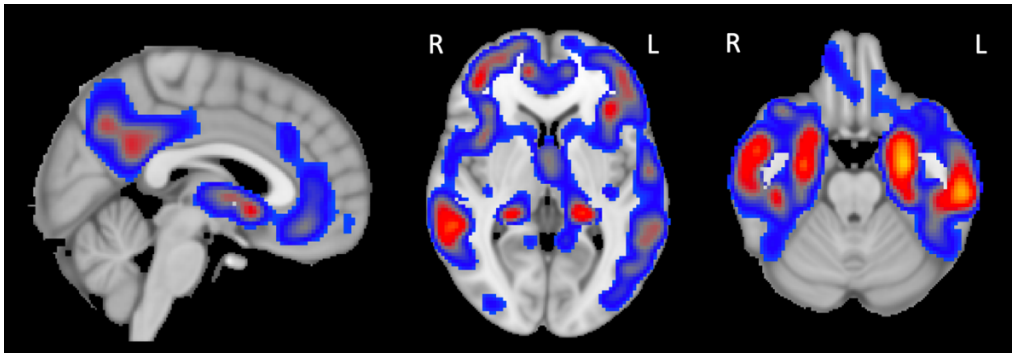


Figure 4 .1 | Group effect (AD < EHC)

Graded color-coded map within the significant cluster: yellow = highest Z-score (corresponding to higher atrophy), red = high, grey = intermediate, blue = lower atrophy. R= right, L = left.

	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Group	1	Hippocampus (L), (R) Amygdala (L), (R) Parahippocampal Gyrus (L), (R) Temporal Pole (L), (R) Inferior Temporal Gyrus (L), (R) Middle Temporal Gyrus (L), (R) Cingulate Gyrus (L), (R) Putamen (L), (R) Angular Gyrus (L), (R) Orbito-Frontal Cortex (L), (R) Frontal Pole (L), (R) Paracingulate Gyrus (L), (R) Lateral Occipital Cortex (L) Middle Frontal Gyrus (L), (R) Inferior Frontal Gyrus (L), (R) Fusiform Gyrus (L), (R)	85893	-24	-34	-4	8.05

Table 4.2 | Group effect

Regions are reported in order of decreasing Z-score (a measure of the magnitude of the effect). R= right, L = left. x,y,z coordinates are reported in the MNI152 space. Only one large cluster was found, encompassing several different confluent brain regions.

VSTM metrics brain-behaviour correlation across the whole cohort

Our whole brain voxel-wise analysis showed that all VSTM metrics, except Misbinding, mapped onto several, largely overlapping GM areas (**Figure 4.2**, **Tables 4.3 - 4.6**). All metrics displayed the expected effect direction, with lower Target detection and Identification accuracy and higher Guessing, Identification Time, Localization Time, Absolute Error and memory Imprecision being associated with reduced GM volume. No areas were found in the reverse contrast for any metric.

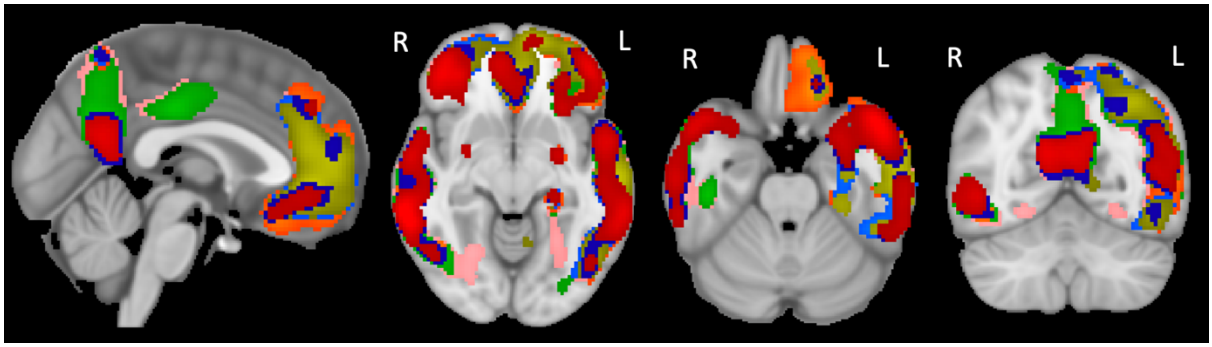


Figure 4.2 | VSTM metrics correlations with GM atrophy

Statistically significant maps are overlaid, with the metrics with lower cluster size at the top: red = Target detection, blue = Guessing, yellow = Localization time, light blue = Identification time, orange = Identification accuracy, green = Absolute Error, pink = memory Imprecision. R= right, L = left.

According to their spatial distribution and the type of information carried by each metric, VSTM metrics were grouped as follows:

1) *Spatial localization - Identity metrics:* The mixture model metrics Target detection and Guessing.
2) *Reaction Times metrics*: Identification (Object) and Localization (Spatial localization) times.
3) *Object identity metric*: Identification accuracy.
4) *Spatial localization - Precision metrics*: Absolute Error and memory Imprecision.

Spatial localization - Identity Metrics

Target detection and Guessing showed a very similar pattern of association with GM atrophy, with involvement of several brain areas including bilateral Middle Temporal Gyrus, Temporal Pole, Precuneus, Paracingulate Cortex, Frontal Pole, Amygdala, left Middle Frontal Gyrus, left Hippocampus and left Supramarginal Gyrus (**Figure 4.3**, **Table 4.3**). Higher Guessing was also associated with lower GM volume in an additional independent cluster in the left Superior Parietal Lobule (**Figure 4.3**, in blue, and **Table 4.3**, Guessing -

Cluster 5). The involvement of the right Amygdala (**Table 4.3**, Target detection - Cluster 2 and Guessing - Cluster 6) was unique to these metrics.

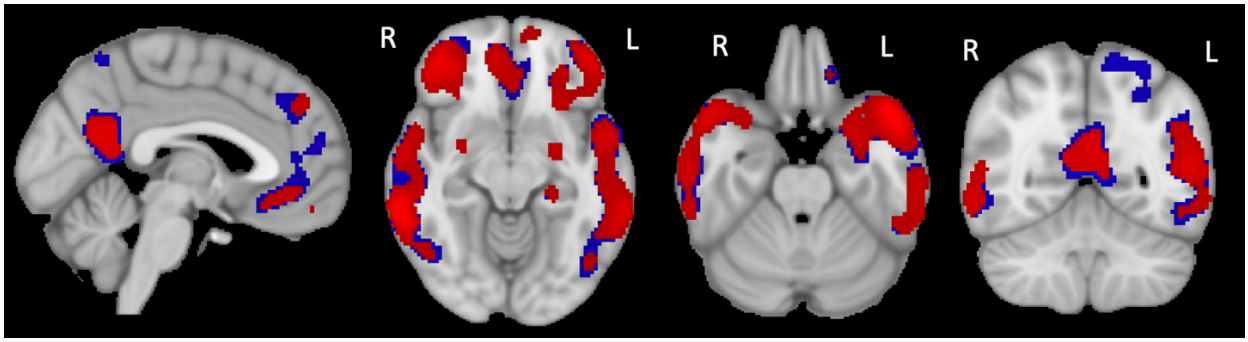


Figure 4.3 | Spatial localization – Identity metrics

Statistically significant maps are overlaid with Target detection (lower cluster size) on top. Red = Target Detection, blue = Guessing. R= right, L = left.

VSTM metric	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Target detection	1	Middle Frontal Gyrus (L) Superior Frontal Gyrus (L) Inferior Frontal Gyrus (L) Orbito-frontal cortex (L) Frontal Pole (L) Temporal Pole (L) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L) Thalamus (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Insula (L) Heschl's Gyrus (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Lateral Occipital Cortex (L)	14056	-44	34	22	5.53
	2	Middle Temporal Gyrus (R) Inferior Temporal Gyrus (R) Angular Gyrus (R) Superior Temporal Gyrus (R) Temporal Pole (R) Amygdala (R) Frontal pole (R)	5411	62	-40	-12	5.28
	3	Precuneus	964	12	-62	22	4.1

	4	Paracingulate Cortex Superior Frontal Gyrus (L), (R)	541	8	48	-8	3.55
	5	Supramarginal Gyrus (L)	123	-54	-44	42	3
Guessing	1	Temporal Pole (L) Middle Frontal Gyrus (L) Superior Frontal Gyrus (L) Inferior Frontal Gyrus (L) Orbito-frontal cortex (L) Frontal Pole (L) Paracingulate Cortex Superior Frontal Gyrus (L), (R) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L) Thalamus (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Insula (L) Heschl's Gyrus (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Supramarginal Gyrus (L) Lateral Occipital Cortex (L)	17766	-48	12	-26	5.92
	2	Middle Temporal Gyrus (R) Inferior Temporal Gyrus (R) Angular Gyrus (R) Superior Temporal Gyrus (R) Temporal Pole (R)	4479	62	-40	-12	5.23
	3	Precuneus	1444	0	-58	20	3.97
	4	Frontal Pole (R)	1168	40	50	-6	4.61
	5	*Superior Parietal Lobule (L)	698	-28	-54	64	3.21
	6	Amygdala (R)	48	28	-2	-12	2.74

Table 4.3 | Correlations between Spatial localization - Identity metrics and GM areas in AD and EHC

Cluster Size = number of significant voxels in that cluster. Coordinates are defined in the MNI152 space. Z-scores represent the magnitude of the effect. R = right, L = left.

Reaction Times metrics

Higher Identification time and Localization time were correlated with higher GM atrophy in several areas withing the frontal cortex, including bilateral Frontal Pole, Anterior Cingulate Cortex, Paracingulate Cortex, left Middle and Inferior Frontal Gyrus, and areas in

the left temporo, parietal and occipital cortices including Middle Temporal Gyrus, Temporal Pole, Hippocampus, Superior Parietal Lobule, and the right Middle Temporal Gyrus (**Figure 4.4, Table 4.4**). Localization Time was additionally associated with atrophy of the left Intracalcarine cortex, extending to the left Lingual Gyrus, the Posterior Cingulate Cortex, and the Precuneus (**Figure 4.4**, yellow overlay, **Table 4.4** Localization time - Cluster 1).

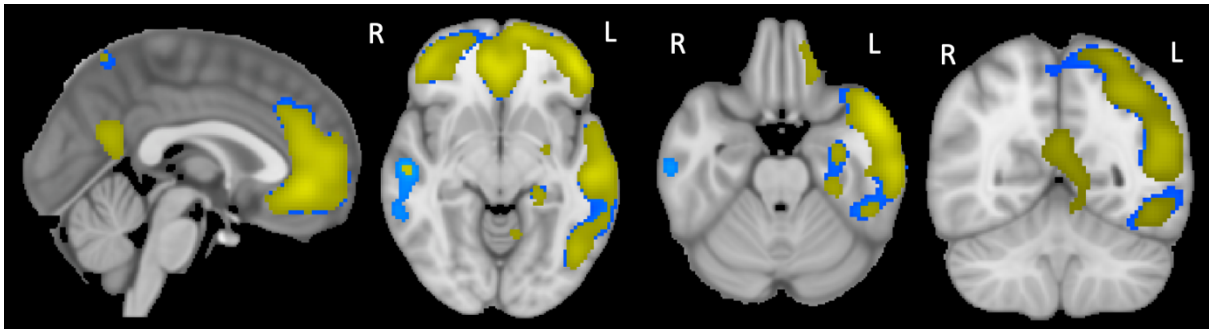


Figure 4.4 | Reaction Times metrics

Statistically significant maps are overlaid with Localization time (lower cluster size) on top. Yellow = Localisation time, light blue = Identification time. R= right, L = left.

VSTM metric	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Localization time	1	Frontal Pole (L), (R) *Anterior Cingulate Cortex Paracingulate Gyrus Orbito-frontal cortex (L) Middle Frontal Gyrus (L) Inferior Frontal Gyrus (L) Superior Frontal Gyrus (L) Frontal Operculum Cortex (L) Insula (L) Temporal Pole (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L) Thalamus (L) Parietal Operculum Cortex (L) Heschl's Gyrus (L) Central Operculum Cortex (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Superior Parietal Lobule (L) Supramarginal Gyrus (L) Lateral Occipital Cortex (L) *Intracalcarine Cortex (L)	28629	-32	54	6	5.7

		* Posterior Cingulate Gyrus (L) *Lingual Gyrus (L) *Precuneus (L)					
	2	Middle Temporal Gyrus posterior division (R)	16	56	-16	-10	3.94
Identification time	1	Frontal Pole (L), (R) Anterior Cingulate Cortex Paracingulate Gyrus Orbito-frontal cortex (L) Middle Frontal Gyrus (L) Inferior Frontal Gyrus (L) Superior Frontal Gyrus (L) Frontal Operculum Cortex (L) Insula (L) Temporal Pole (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L) Thalamus (L) Parietal Operculum Cortex (L) Heschl's Gyrus (L) Central Operculum Cortex (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Supramarginal Gyrus (L) Lateral Occipital Cortex (L)	31254	-28	56	8	6.15
	2	Middle Temporal Gyrus posterior division (R)	750	56	-16	-10	4.22

Table 4.4 | Correlations between Reaction Times metrics and GM areas in AD and EHC

* Denotes areas discussed in the text that differ across metrics within the table or compared to the previous table.

Object identity metric

Lower Identification accuracy was associated with GM atrophy in the same areas as Identification time (**Figure 4.5**, **Table 4.5**), with the additional contribution of the left Frontal Medial Cortex and Subcallosal Cortex (orange overlay in **Figure 4.2**, **Table 4.5** - Cluster 1), and the right Inferior Temporal Gyrus (**Table 4.5** - Cluster 3).

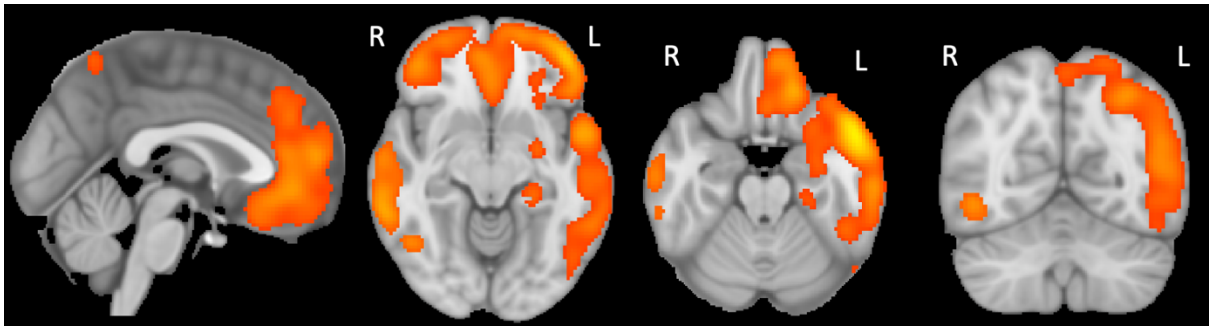


Figure 4.5 | Object Identity metric
Identification accuracy in orange. R= right, L = left.

VSTM metric	N clusters	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Identification Accuracy	1	Temporal Pole (L) *Frontal Medial Cortex (L) *Subcallosal Cortex (L) Frontal Pole (L) and (R) Anterior Cingulate Cortex Paracingulate Gyrus Orbito-frontal cortex (L) Middle Frontal Gyrus (L) Inferior Frontal Gyrus (L) Superior Frontal Gyrus (L) Frontal Operculum Cortex (L) Insula (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L) Thalamus (L) Parietal Operculum Cortex (L) Heschl's Gyrus (L) Central Operculum Cortex (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Supramarginal Gyrus (L) Lateral Occipital Cortex (L)	32614	-52	5	-26	6.33
	2	Middle Temporal Gyrus (R)	1702	62	-42	-8	4.4
	3	*Inferior Temporal Gyrus (R)	144	46	-60	-6	3.78

Table 4.5 | Correlations between Object Identity metric and GM areas in AD and EHC
* Denotes areas discussed in the text that differ compared to the previous table.

Spatial localization – Precision metrics

Increased Absolute Error and memory Imprecision were associated with lower GM volumes in the same areas as Identification accuracy and showed an additional involvement of the Precuneus, the Posterior Cingulate Cortex, right Lateral Occipital Cortex, right Temporal Fusiform Gyrus, bilateral Occipital Fusiform Gyrus, and to a much lesser extent left Caudate and Putamen (**Figure 4.6**, **Table 4.6**). Higher memory Imprecision was also associated with lower GM volume in bilateral Lingual Gyrus (**Figure 4.6**, pink overlay).

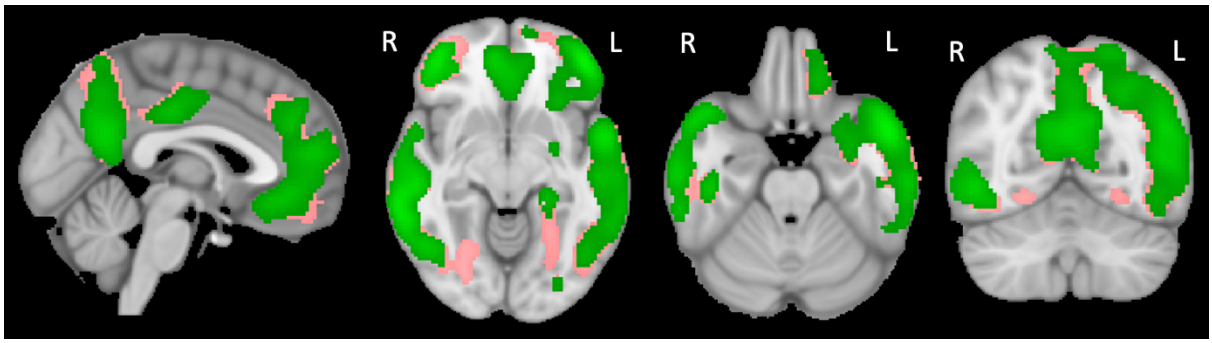


Figure 4.6 | Spatial localization - Precision metrics

Statistically significant maps are overlaid with Absolute Error (lower cluster size) on top. Green = Absolute Error, pink = memory Imprecision. R= right, L = left.

VSTM metric	N clusters	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Absolute Error	1	Middle Frontal Gyrus (L) Frontal Medial Cortex (L) Subcallosal Cortex (L) Frontal Pole (L) Anterior Cingulate Cortex Paracingulate Gyrus Orbito-frontal cortex (L) Middle Frontal Gyrus (L) Inferior Frontal Gyrus (L) Superior Frontal Gyrus (L) Frontal Operculum Cortex (L) Insula (L) *Caudate (L) *Putamen (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L)	33812	-44	34	24	6.01

		Thalamus (L) Parietal Operculum Cortex (L) Heschl's Gyrus (L) Central Operculum Cortex (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Supramarginal Gyrus (L) Lateral Occipital Cortex (L) Precuneus Posterior Cingulate Cortex *Occipital Fusiform Gyrus (L)					
	2	Middle Temporal Gyrus (R) Superior Temporal Gyrus (R) Inferior Temporal Gyrus (R) Heschl's Gyrus (R) Temporal Pole (R) *Temporal Fusiform Cortex (R) *Lateral Occipital Cortex (R) *Occipital Fusiform Gyrus (R)	5166	62	-38	-14	5.35
	3	Frontal Pole (R)	854	40	48	-4	5.02
Imprecision	1	Middle Frontal Gyrus (L) Frontal Medial Cortex (L) Subcallosal Cortex (L) Frontal Pole (L) Anterior Cingulate Cortex Paracingulate Gyrus Orbito-frontal cortex (L) Middle Frontal Gyrus (L) Inferior Frontal Gyrus (L) Superior Frontal Gyrus (L) Frontal Operculum Cortex (L) Insula (L) Caudate (L) Putamen (L) Middle Temporal Gyrus (L) Inferior Temporal Gyrus (L) Hippocampus (L) Amygdala (L) Parahippocampal Gyrus (L) Thalamus (L) Parietal Operculum Cortex (L) Heschl's Gyrus (L) Central Operculum Cortex (L) Angular Gyrus (L) Superior Temporal Gyrus (L) Supramarginal Gyrus (L) Lateral Occipital Cortex (L) Precuneus Posterior Cingulate Cortex Occipital Fusiform Gyrus (L) *Lingual Gyrus (L)	37134	-44	34	-24	5.8
	2	Middle Temporal Gyrus (R) Superior Temporal Gyrus (R) Inferior Temporal Gyrus (R) Heschl's Gyrus (R) Temporal Pole (R)	5440	64	-36	-12	4.82

Temporal Fusiform Cortex (R)
Lateral Occipital Cortex (R)
Occipital Fusiform Gyrus (R)
*Lingual Gyrus (R)

Table 4.6 | Correlations between Spatial localization – Precision metrics and GM areas in AD and EHC

VSTM metrics brain-behaviour correlation in EHC

Increased Identification Time was associated with reduced GM volume in several areas within the Frontal cortex such as right Frontal Pole, bilateral Orbitofrontal cortex and left Middle Frontal Gyrus, and in the right Middle Temporal Gyrus (**Figure 4.7**, top row, yellow overlay, **Table 4.7**). The right Orbitofrontal cortex was the only area that was present in the EHC alone, and not seen in the previous group analysis. Memory Imprecision was associated with greater GM atrophy in the left Supramarginal gyrus (**Figure 4.7**, bottom row, yellow overlay, **Table 4.7**).

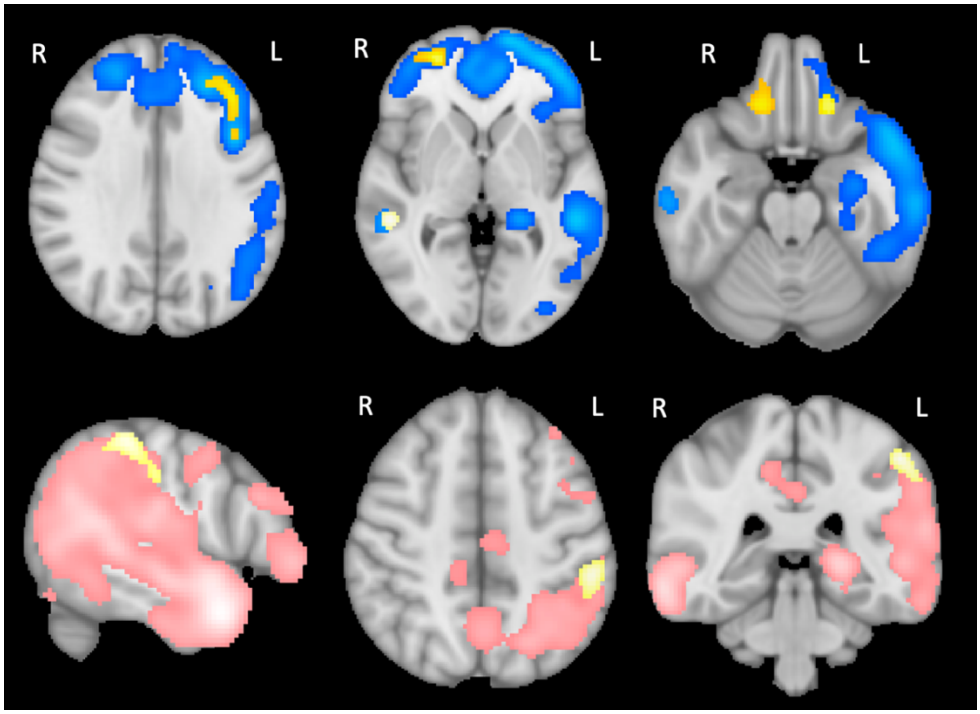


Figure 4.7 | GM atrophy and VSTM in EHC

Statistically significant maps are overlaid with EHC in light yellow for both metrics (Identification time = top row, Imprecision = bottom row). Light blue = Identification Time (top row) and pink = memory Imprecision (bottom row) in the whole sample.

VSTM metric	N clusters	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Identification Time	1	Frontal Pole (R)	673	20	48	-16	3.95
	2	Middle Frontal Gyrus (L)	432	-38	32	28	3.89
	3	Middle Temporal Gyrus posterior division (R)	65	52	-36	0	5.35
	4	Fronto orbital cortex (L)	64	-18	30	-20	3.87
Imprecision	1	Supramarginal Gyrus (L)	325	-56	-30	44	4.29

Table 4.7 | Correlations between VSTM metrics and GM areas in EHC

Brain behavioural analysis within the AD group alone is presented in Appendix. In summary, most of the results seen in the overall sample were driven by AD patients, with few exceptions outlined in **Figures 4.1S - 4.2S, Table 4.1S.**

VSTM contribution independent of general cognition

Patients enrolled in this study had a broad range of disease severity and cognitive impairment, and therefore it's likely that the greater the cognitive impairment, the lower the performance at our VSTM tasks and the greater the GM atrophy. The effect of our VSTM metrics *independent* of general cognitive impairment was explored by controlling all our analyses for ACE scores, including it as a covariate in our GLM. We also investigated these correlations separately for each group. Results are presented in Appendix (**Figures 4.3S - 4.10S, Tables 4.2S - 4.4S**). In summary, VSTM metrics showed a similar segregation pattern as seen in the previous analysis, with four different localization clusters, though with a far less extensive involvement.

4.4 Discussion

In this Chapter, whole-brain VBM was used to investigate whether in a sample of elderly healthy controls and patients with AD there were any areas of the brain particularly informative of a distinctive underlying VSTM process, including memory precision, binding, recognition of object identity (related or irrespective of spatial location) and reaction times.

Shared brain areas across all VSTM metrics

In the whole cohort, several common brain areas were associated with different VSTM metrics (**Figures 4.2 - 4.6**). These included the Middle Frontal Gyrus, Middle and Inferior Temporal Gyrus, Temporal Pole, Hippocampus, Amygdala, Supramarginal Gyrus, Angular Gyrus and Lateral Occipital Cortex in the left hemisphere, bilateral Frontal Pole and Paracingulate Cortex.

The Middle Frontal Gyrus is part of the dorsolateral prefrontal cortex (DLPFC), and has been implicated in executive components of WM, such as distractor resistance, WM manipulation and reorienting of attention [276, 319]. Conversely, the Inferior Temporal Gyrus is considered a key area of the ventral 'what' stream and has been implicated in WM and subsequent LTM retrieval, particularly in maintaining object identity over a delay period [320]. It is one of the core areas of atrophy in AD [321]. The Temporal Pole, while most commonly associated with processing semantics and less commonly correlated with WM performance in healthy individuals, is considered one of signature GM atrophy areas of early AD [321].

The results here show that lower GM volumes in the left Hippocampus were associated with poorer WM performance in AD and EHC, across all VSTM metrics. This extends findings of the association between VSTM and hippocampal integrity in early onset familial AD [77] to sporadic late onset AD. This is particularly relevant as our behavioural paradigm measures relational binding, which is thought to rely on hippocampal integrity [77, 145, 252, 290, 291]. Moreover, this finding supports previous data in Chapter 3 of the association of these metrics with hippocampal integrity. However, it's worth noticing that in the whole brain analysis we did not find any brain area that was associated with Misbinding, either in the hippocampus or elsewhere in the brain. Moreover, once general cognition was factored out, no VSTM metric was specifically associated with hippocampal volumes (**Figure 4.3S**), and we also replicated the absence of correlation of hippocampal integrity with our VSTM metrics in healthy controls from our previous Chapter. It is also worth noticing that even in the study by Piekema et al [291], the whole brain analysis did not reveal a preferential role of the hippocampus in binding, and a specific hippocampal region of interest (ROI) approach had to be deployed to find such an association, while whole brain analysis showed an involvement of the Precuneus.

Importantly, while in Chapter 3 there were no significant lateralization effects of VSTM metrics in predicting hippocampal volumes, in the whole brain analysis presented in this Chapter there was an association with only left hippocampal volumes. This lateralized effect might be potentially explained by higher GM atrophy in the left hippocampus compared to the right hippocampus in AD patients in our sample (**Figure 4.1**), rather than a specific role of the left versus the right hippocampus in WM [252, 291].

Together with the hippocampus, both the Amygdala and the Cingulate Cortex form part of the Papez circuit, which is a neural system that regulates not only emotions but also memory and is highly impaired in AD patients [322]. The contribution of the Amygdala to WM performances is very poorly understood, with some emerging evidence showing that it can be associated with WM performance, regardless of their emotional value, in healthy individuals [323], while Paracingulate and Anterior Cingulate Cortex have been consistently associated with WM abilities [324, 325, 326, 327].

The Frontal Pole, known also as anterior PFC, rostral PFC or frontopolar PFC, corresponding to BA10, constitutes the largest portion of the PFC [328]. It is the prefrontal region that has evolved the most compared to the chimpanzee brain [329]. Lateral portions of the anterior PFC have been consistently associated with working memory, episodic retrieval and perception, while medial areas seem to be more commonly associated with emotion and social cognition [330, 331, 332]. Lateral areas have also been shown to be functionally correlated with Precuneus, Cingulate Cortex, Angular Gyrus, and Inferior Parietal Lobule [331]. Frontal Pole lesions have also recently been associated with a specific impairment of time-based memory tasks that could not be explained by impairments in basic attention, perception, or multiple instruction processing [333], and have been associated with aberrant motor behavior in AD patients [334], which might explain the widespread association with multiple metrics, including and more prominently with our Reaction Times metrics (**Figures 4.3S, 4.5S, 4.8S**). Moreover, this is corroborated by the finding that patients with lesions of the anterior PFC may exhibit difficulties when there are several possible ways to behave if the behaviour is not fully guided by the environment [333], which is consistent with the self-guided, time unconstrained nature of our VSTM task.

Previous studies have reported that the Lateral Occipital Cortex, Supramarginal Gyrus and Angular Gyrus, particularly in the left hemisphere, correlate with measures of memory precision in young healthy individuals, consistent with their involvement in the dorsal, 'where', stream [131, 300]. Across the whole cohort of EHC and AD, our data suggest that these areas are not uniquely linked to memory precision, as they were associated with all our VSTM metrics. However, in EHC controls alone, we were able to replicate the selective association between increased memory Imprecision and lower GM volumes in the Supramarginal Gyrus (**Figure 4.7, Table 4.7**).

Brain areas across VSTM metrics encoding spatial information

Across EHC and AD, the Precuneus was associated with all the WM metrics that encoded spatial information (**Figures 4.3, 4.4** and **4.6**), but not with metrics that only carried object identity information (Identification accuracy and Identification time (**Figures 4.4 - 4.5**)). Remarkably, if general cognition was factored out, the association between the Precuneus and memory Precision metrics (Absolute Error and memory Imprecision) remained significant (**Figure 4.7S**). The Precuneus, whilst traditionally linked to visual imagery [131], can also be involved in tasks that involve abstract, non-imageable stimuli, such as our task [335]. Its main extraparietal projections are directed to the frontal lobes, including the frontal eye fields, and are thought to be essential for visual reaching, i.e., the visual guidance of hand movements [335]. It is also tightly functionally linked to the hippocampus [336], and plays an important role in episodic and WM retrieval [289, 291]. Our data support the view that the Precuneus is essential for maintaining the spatial mapping of an object in WM, while it is not involved in storing the identity of an object, which is in line with dorsal stream functions [273, 274]. While this has been observed in studies involving healthy controls, our data extend

previous findings to patients with AD. We suggest that the progressive degradation of their spatial mapping across disease course is tightly linked to their prominent precuneal atrophy [321, 337].

Similarly, the right Temporal Pole was also associated with metrics that coded for spatial information, either object identity with respect to spatial localization (Target detection and Guessing) or the precision of spatial localization (Absolute Error and memory Imprecision), (**Figures 4.2, 4.3, 4.6)**. However, unlike the Precuneus, when global impairment was regressed out, this was no longer significantly associated with spatial metrics. Therefore, it might represent a general effect of spatial cognitive impairment in patients with AD rather than a more general construct.

The right Middle and Inferior Temporal Gyrus were associated with multiple VSTM metrics in the whole sample, but only with spatial localization metrics if global cognition was factored out (**Figures 4.3S, 4.4S, 4.7S**), and not in EHC alone. We suggest that atrophy in those regions, which are severely affected in AD [321] contributes to the disruption of the spatial and non-spatial properties of an object, but only of spatial properties independently of general cognition.

Another possible explanation for the involvement of right Temporal Pole, Middle and Inferior Temporal Gyrus in metrics coding spatial information lies within the right hemispheric predominance of spatial information compared to the left hemisphere [273, 274, 275, 276], and their contribution might be necessary for spatial tasks in AD patients.

Dissociable brain areas between different VSTM metrics

The eight VSTM metrics examined clustered in the brain into four different spatial patterns, which were biologically plausibly representing four conceptual categories.

The first category was represented by metrics that coded the identity of an object with respect to its spatial localization: the mixture model metrics Target detection and Guessing, as no brain area was specifically associated with Misbinding. Overall, across AD and EHC, these metrics were associated with GM integrity in several brain regions including, beyond the shared common areas and spatial localization areas discussed above, the right Amygdala. Guessing also involved the left Superior Parietal Lobule.

Previous studies have found that the amygdala can be associated with speed of processing in high demand conditions in WM tasks [323]. In our whole cohort, the left Amygdala was correlated with multiple VSTM metrics (**Figures 4.2 - 4.6**), while an association with the right amygdala was found only for these two VSTM metrics (**Figures 4.2 - 4.3**). We propose that the right amygdala encodes information regarding the identity of an object only if that is linked to its spatial location.

The second category is represented by the Reaction Times metrics (Identification time and Localization time). Inefficient response selection and difficulty in maintaining attention to the task can manifest as a greater proportion of trials with prolonged reaction times, which can frequently be observed in patients with AD. Several medial and lateral areas, more prominently in the frontal cortex, were associated with our reaction times across EHC and AD (**Figure 4.4**). After the effect of global cognitive impairment was regressed out, only

areas within the left Frontal Pole and the Paracingulate Gyrus remained significant (**Figure 4.5S**). Intra and inter-individual reaction times variability has been considered an early measure of cognitive decline, has found to be associated with grey matter volume [302], white matter burden [303, 304] and amyloid pathology [305] in elderly healthy individuals and is impaired in neurological conditions affecting the frontal cortex, such as frontotemporal dementia [338, 339]. It has also been linked to the specific involvement of the Anterior Cingulate Cortex in healthy adults [326]. While the left Frontal Pole was correlated with several different VSTM metrics (**Figures 4.2** and **4.3S**), our data suggest that the Anterior Cingulate and Paracingulate Gyrus might be brain areas particularly important to determine reaction times across EHC and AD, either dependent (Anterior Cingulate) or independent (Paracingulate Gyrus) of cognitive impairment. We also propose that a correlation between reaction times with frontal structures represents the late neocortical stage of AD, which is in line with our finding of Chapter 3, showing that reaction times are altered only in later stages of the disease.

Regarding Identification accuracy, across the whole cohort there was an extensive involvement of the frontal cortex, and particularly the Frontal Medial Cortex (**Figures 4.2**, **4.5)**. However, when general cognitive impairment was regressed out, only the lateral part of the rostral PFC (left Frontal Pole) remained significantly associated to memory performance (**Figures 4.3S, 4.6S**). Therefore, both structures might be important to recall specifically object identity, whether dependent (Frontal Medial Cortex), or independent (Frontal Pole) of global cognitive impairment. Whilst we did not find any areas in EHC that were associated with Identification accuracy, it is notable that a similar, mirroring area, in the right Frontal Pole was associated in healthy elderly controls with Identification Time

(**Figure 4.7**), thus reinforcing the idea that these structures are important to maintain object identity.

Finally, our data suggest that memory precision metrics (Absolute Error and memory Imprecision) were associated with shared common areas and brain regions coding spatial information as previously discussed (as the Precuneus and right Temporal Pole), but also to the Posterior Cingulate Cortex, Temporal and Occipital Fusiform Gyrus and Lingual Gyrus (**Figure 4.6**). The involvement of these areas, particularly on the right (**Figures 4.2, 4.6**), is in line with existing evidence showing that encoding shapes activates regions within the ventral visual stream [299]. Unilateral left involvement of the Posterior Cingulate Cortex and Lingual Gyrus was also seen in Localization time, which supports the role of the ventral stream in coding spatial information. However, our study was able to show that bilaterally, these areas are particularly relevant for storing the *precision* by which objects are stored, particularly in patients with AD. Conversely, in EHC it seems that memory precision is more highly dependent on the left Supramarginal Gyrus. These data replicate a previous finding in young adults, using the mixture model metric of WM, which showed that memory precision is associated with left posterior parietal cortex function [131].

4.5 Conclusion

To conclude, the data presented here show that several brain regions are associated with different VSTM functions. However, specific spatial patterns were preferentially associated with four different categories: object identity with respect to their spatial localization, reaction times, object identity and memory precision.

Chapter 5. Diffusion correlates of Visual Short-Term Memory using Tract-based spatial statistics (TBSS) in AD and EHC §#

Abstract

In Chapter 4 we found that the VSTM metrics derived from the "What was where?" Oxford memory task showed an extensive network of GM areas associated with performance in AD patients. Yet, little is known about how a disruption in white matter integrity impacts different VSTM abilities in AD. In Chapter 5, we applied whole brain Tract-based spatial statistics (TBSS) analysis to examine voxelwise associations between white matter integrity and the previously described VSTM metrics, in the same cohort of Chapter 4. Our data show that a key common area, comprising the left Optic Radiation, Forceps Major and Middle Longitudinal Fasciculus, is associated with performance across multiple VSTM metrics in patients with AD. In healthy elderly controls, white matter integrity of a homologous region in the right hemisphere was identified to play a role in maintaining the correct object representation in VSTM. Additionally, specific VSTM functions, such as maintaining the right item in memory and selecting it, might involve additional white matter tracts in patients with AD such as the left Vertical Occipital Fasciculus, Superior Longitudinal Fasciculus, Superior Thalamic Radiation, Arcuate Fasciculus, Inferior Fronto-occipital Fasciculus and Cingulum.

5.1 Introduction

In Chapters 2, 3 and 4 structural MRI was used to detect macroscopic changes in brain tissue, including loss of GM integrity. This revealed how measures of atrophy in AD can map biologically to cognition along different stages of AD. However, macroscopic changes occur relatively late during the AD course, when neurodegeneration is already widespread [340]. As a result, atrophy becomes apparent often too late to detect subtle pathophysiological changes in people at risk of developing AD. Macroscopic changes are usually preceded by microstructural changes in the brain tissue [341], which can be detected before atrophy is overt [342, 343, 344]. Therefore, studying these alterations may be important to fully understand the underlying pathophysiological processes leading to subsequent macroscopic changes in AD.

Diffusion Tensor Imaging (DTI) is a reliable imaging method to examine microscopic changes and quantify the breakdown of tissue's architecture [345]. When the molecules are free to move in all directions, this is termed *isotropic* diffusion, while if they find any anatomical boundary that constrain them to move following a specific direction, the diffusion will be considered *anisotropic* [346]. DTI involves fitting of the diffusion property to a tensor model, i.e., an array of several coefficients that fully describes molecular mobility along each direction. Fractional Anisotropy (FA) is the fraction of the tensor that can be assigned to anisotropic diffusion, according to **Formula 1**.

Formula 1

$$\mathrm{FA} = \sqrt{\frac{1}{2}} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

It is considered a measure of global white matter integrity, with values closer to zero reflecting higher isotropic diffusion, and therefore higher microstructural damage.

Diffusion metrics have been widely studied in AD, for their ability to detect white matter damage even at its earliest stages [347, 348] . In AD, the physiological neat microstructural milieu is disrupted, leading to a loss of anisotropic diffusion due to degradation of microtubules, loss of axonal structure, demyelination, gliosis and microvascular disarray [349]. DTI markers have been found to be altered also in early phases of the disease, such as in mild cognitive impairment, and be also able to predict which of these patients will convert to AD dementia compared to patients whose cognition will remains table over time [350, 351, 352, 353, 354]. In terms of anatomical localization, while most studies report decreased FA in AD patients especially in the Fornix, Cingulum and Corpus Callosum, these alterations are widespread and can be found in several other white matter tracts [355, 356].

Although diffusion imaging can give us useful information regarding disease progression in AD, brain-behavioral correlations between white matter integrity and cognition are relatively scarce. A previous study using the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [357] reported that reduced FA correlated with worse cognitive scores on standard tests of cognition, more prominently with memory and executive functions, in a mixed sample of healthy controls and patients with AD [358]. However, these correlations were not found if the two groups were considered in isolation. One possible explanation could be that these tests are not sensitive enough to detect brain white matter microstructural deficits.

Parra et al investigated associations between white matter integrity and VSTM performance in asymptomatic carriers of a causative mutation for familial AD (FAD), symptomatic patients with FAD and healthy controls, deploying a shape-color binding paradigm [359]. They found that better performance on their VSTM binding task correlated with diffusivity changes in bilateral genu of the Corpus Callosum (corresponding to the Forceps Minor), left Superior Longitudinal Fasciculus, and left hippocampal region of the Cingulum. These

correlations were found only in symptomatic FAD patients. However, that study deployed an *a priori* ROI approach, based on ROIs primarily involved in FAD, and did not explore brain-behavioural correlations in the whole brain.

In this Chapter, we conducted a whole brain analysis to investigate which areas of the brain were associated with our VSTM metrics. Tract-based spatial statistics (TBSS) is an ideal method to answer this question as, similarly to VBM, allows the voxelwise statistics to be calculated between each voxel and behavioural measure directly. Crucially, TBSS is better than VBM when using diffusion imaging data, as overcomes problems in alignment between different subjects' diffusion images and in spatial smoothing (local averaging of noise used in cross-subject registration) [360]. Moreover, it is a robust and validated tool in AD patients [361, 362].

Similarly to Chapter 4, the main questions we aimed to answer in this Chapter were:

1) Are successful retrieval, memory precision, binding abilities and reaction times, associated with different white matter tracts in a cohort of EHC and AD patients?
2) Are these areas specific for each VSTM metric or are they shared across multiple metrics?
3) Do these white matter tracts differ between the two groups, i.e., EHC and AD?
4) Are these brain-behavioural associations independent of global cognitive impairment?

5.2 Methods

Participants

Participants' demographics and tests are described in Chapter 4.

Tract-based spatial statistics (TBSS) pipeline

We applied the following standard processing steps (outputs from our dataset shown below):

1) Nonlinear registration

Firstly, nonlinear registration from all FA images to a target image was performed, followed by an affine alignment to a standard space, the 1x1x1 mm MNI152 space, to create a mean FA mask (**Figure 5.1**).

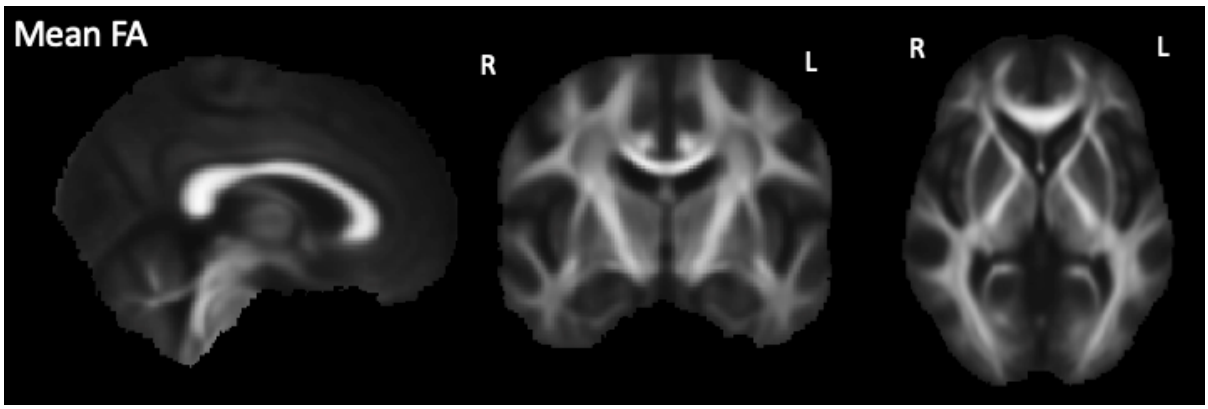


Figure 5.1 | Mean FA

R= right, L = left.

2) Skeletonization

Secondly, the mean FA mask was skeletonized, where the skeleton represents the centre of all tracts (as delineated on an FA map), which are common to all subjects (**Figure 5.2**).

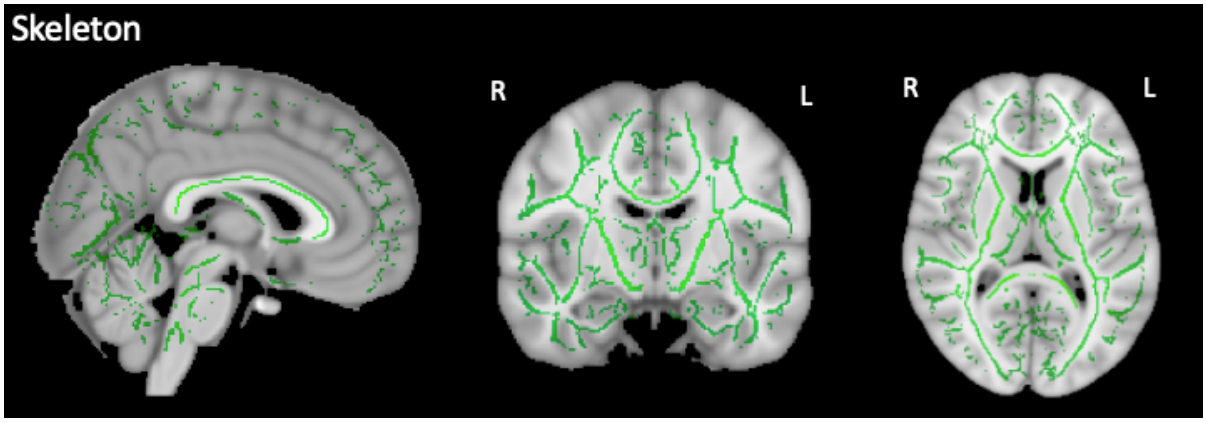


Figure 5.2 | Skeletonized FA

R= right, L = left.

3) Thresholding

The mean FA skeleton was subsequently thresholded to restrict the analysis to voxels that have been successfully aligned across subjects, creating a binarized skeleton mask (**Figure 5.3**).

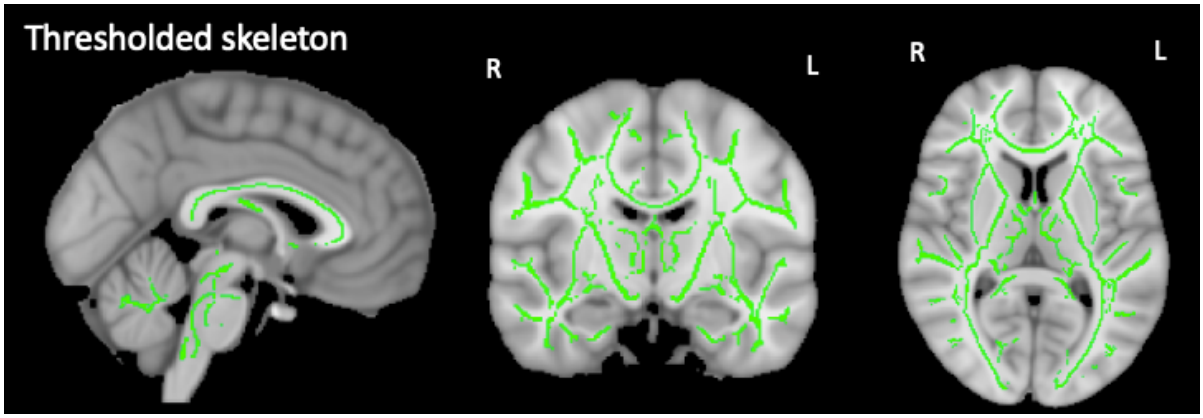


Figure 5.3 | Thresholded skeleton

R= right, L = left.

4) Projection

Each subject's FA was projected to the main skeleton, and a 4D image containing the projected skeletonized FA data is produced.

5) Voxelwise analysis

Voxelwise analysis was carried out within the skeletonized FA data, using the *randomise* function of FSL.

A *group* contrast between EHC and AD was first performed, examining areas in which AD had higher (AD > EHC) or lower (AD < EHC) FA compared to healthy controls, to check the extent of diffusivity changes in our sample. Subsequently, we tested brain-behavioural correlations by applying a voxelwise GLM, using permutation-based non-parametric testing (n permutations = 5000) for correcting for multiple comparisons across space. Correlations between the eight VSTM metrics described in Chapter 4 and FA were calculated in the whole sample (EHC and AD), and separately for each group (EHC, AD), and whether the VSTM metric was associated positively or negatively with FA. Age, gender and education were included in the GLM as covariates. All correlation analyses between VSTM metrics and FA were restricted within the group difference mask derived from the first contrast (AD < HC).

Results reported have as minimum cluster size of n = 10, considered to provide a good balance of Types I and II error rates [363]. The XTRACT HCP Probabilistic Tract Atlas was used as reference labelling system for white matter tracts [364]. This new atlas, developed by Warrington et al in 2020, has 42 probabilistic tracts derived from 1021 subjects from the Human Connectome Project (HCP) [365]. It was our atlas of choice as it has a much bigger testing dataset compared to the other commonly used white matter atlases such as the JHU DTI-based white-matter atlases [366], that have a training sample of respectively 81 subjects (ICBM-DTI-81 white-matter labels atlas) or 28 subjects (JHU white-matter tractography atlas). It must be considered that as it has been released relatively recently, most papers in

the literature up to now would not have necessarily employed the same labelling system. Alternative notations from JHU white matter atlases can also be found in this Chapter, particularly if no match was found within a specific tract and the XTRACT atlas. Finally, for each correlation the probability of the overall cluster of belonging to the tracts within the XTRACT atlas was extracted using FSL *atlasquery* function.

5.3 Results

Group difference

We first assessed whether our sample showed a main effect of group in line with previous literature, where it has been reported that there is a reduced FA in AD patients compared to healthy elderly controls in several white matter tracts [367, 368]. In this cohort, AD patients had lower FA in multiple white matter bundles including the Optic Radiation (Fornix in JHU atlases), the temporal section of the Cingulum, Forceps Major, Forceps Minor, Superior and Inferior Longitudinal Fasciculus, Middle Longitudinal Fasciculus, Superior and Inferior Fronto-occipital Fasciculus, Anterior Thalamic Radiation, Vertical Occipital Fasciculus and Corticospinal Tract (**Figure 5.4**, **Tables 5.1 - 5.2**). The two biggest clusters had maximum peaks respectively in the left Optic Radiation (**Table 5.4**, Cluster 1, **Figure 5.4, top row**) and in the right temporal section of the Cingulum bundle (**Table 5.4,** Cluster 2, **Figure 5.4,** bottom row). No areas had lower FA in healthy elderly controls compared to AD patients.

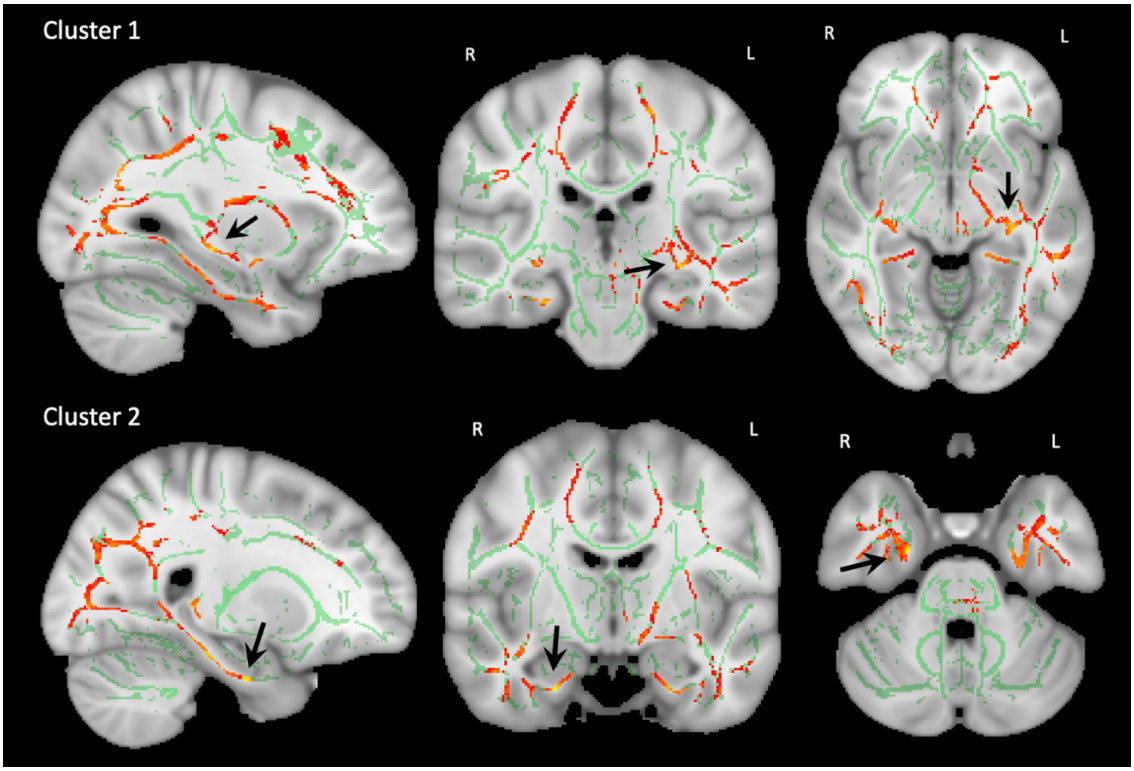


Figure 5.4 | Group differences: Clusters 1 and 2

The arrows represent respectively the peaks of Cluster 1 (top row), in the left Optic Radiation, and of Cluster 2 (bottom row), in the right temporal section of the Cingulum bundle, R= right, L = left.

Cluster	White matter tract	Cluster Size	Coordinates (x,y,z)			Z-score
1	Optic Radiation (L) / Fornix (JHU)	18894	-30	-24	-7	5.61
2	Cingulum - temporal section (R)	11365	24	-9	-31	6.31
3	Forceps minor (L) / Genu of the corpus callosum (JHU)	1206	-5	23	-3	4.05
4	Superior Longitudinal Fasciculus 1 (R)	864	52	-1	19	3.9
5	Superior Longitudinal Fasciculus 2 (L)	750	-31	6	48	4.01
6	Fornix (R)	613	26	-32	1	5.72
7	Forceps Minor (R)	175	17	43	-8	3.49
8	Inferior Fronto-occipital Fasciculus (L)	34	-9	35	47	2.42
9	Corticospinal Tract (L)	19	-3	-33	-44	2.41

Table 5.1 | Group comparisons: Clusters and diffusivity peaks

Cluster Size = number of significant voxels in that cluster. Coordinates are defined in the MNI152 space. Z-scores represent the magnitude of the effect.

	White matter tract	%
Group	Forceps Major	3.5691
	Optic Radiation (L)	3.0312
	Inferior Fronto-occipital Fasciculus (L)	2.9021
	Optic Radiation (R)	2.2746
	Superior Longitudinal Fasciculus 1 (L)	1.9891
	Cingulum – Temporal (L)	1.9024
	Corticospinal Tract (L)	1.7636
	Forceps Minor	1.7201
	Middle Longitudinal Fasciculus (L)	1.6679
	Superior Fronto-occipital Fasciculus 2 (L)	1.6245
	Inferior Fronto-occipital Fasciculus (R)	1.6051
	Inferior Longitudinal Fasciculus (L)	1.5985
	Superior Longitudinal Fasciculus 1 (R)	1.341
	Anterior Thalamic Radiation (L)	1.3328
	Middle Longitudinal Fasciculus (R)	1.2442
	Cingulum – Temporal (R)	1.1725
	Superior Thalamic Radiation (L)	1.0652
	Vertical Occipital Fasciculus (R)	1.0474
	Superior Longitudinal Fasciculus 3 (R)	1.0273

Table 5.2 | Group comparisons: Tracts probability

Voxel-wise correlations between VSTM metrics and white matter tracts in the whole sample (EHC and AD)

Higher Misbinding, higher Guessing and lower Target detection were associated with reduced FA in the same key region, which encompassed three main tracts: the left Optic Radiation, the left Forceps Major and the left Middle Longitudinal Fasciculus (**Figures 5.5** and **5.7**, **Tables 5.3 - 5.4**). Misbinding was associated with the smallest cluster size and had

higher involvement of the Optic radiation compared to Guessing and Target detection (**Figures 5.5** and **5.7**), reflected also by the peak being within the left Optic Radiation (**Table 5.3**) and by a higher probability of the significant cluster to belong to that white matter tract (**Table 5.4**).

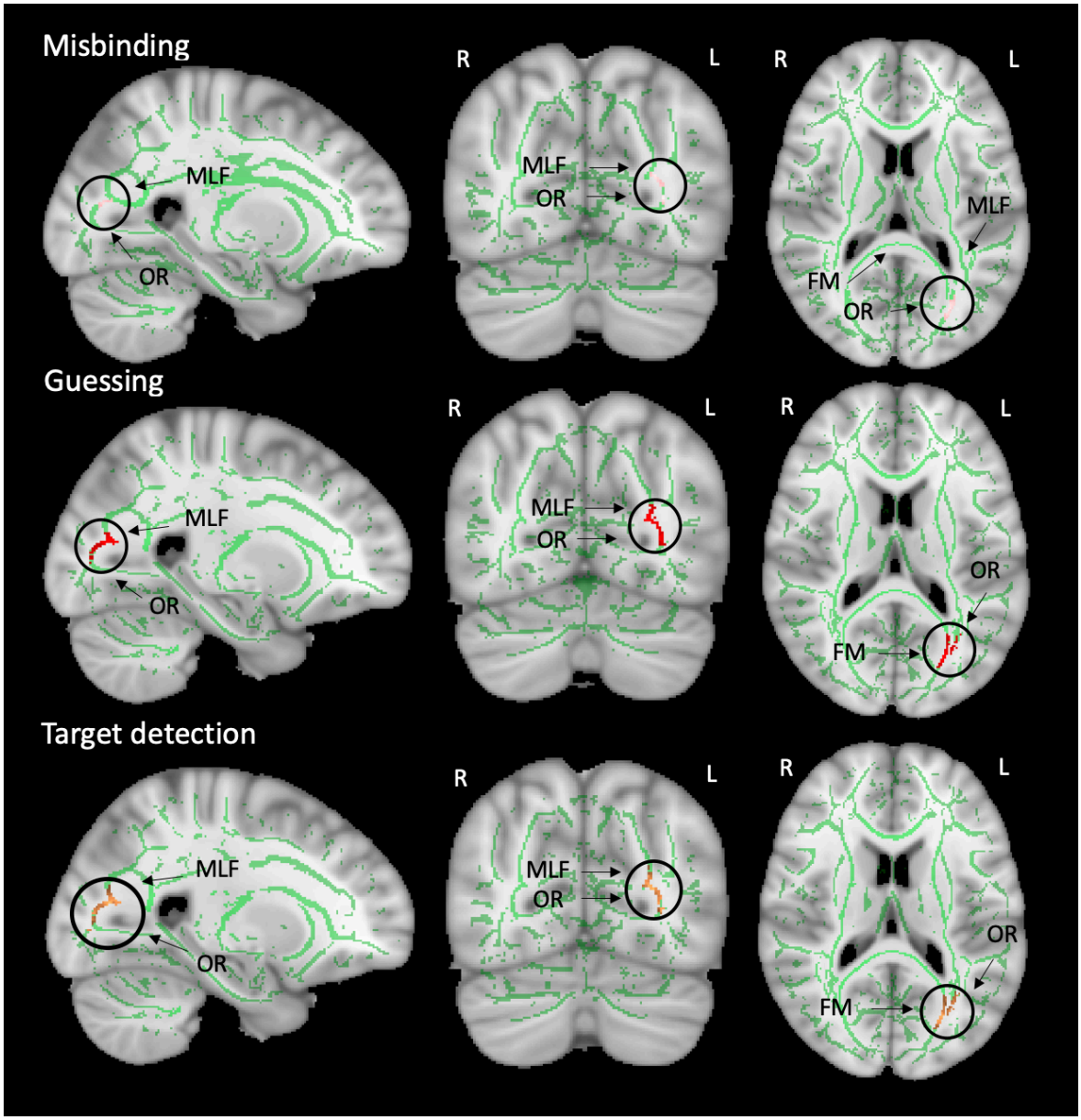


Figure 5.5 | Correlations between VSTM metrics (mixture model metrics) and FA in AD and EHC

Misbinding in pink (upper row), Guessing in red (middle), Target Detection in copper (bottom). From the XTRACT atlas: Optic Radiation (OR), Forceps Major (FM), Middle Longitudinal Fasciculus (MLF). R = right, L = left.

	Cluster	White matter tract	Cluster Size	Coordinates (x,y,z)			Z-score
Misbinding	1	Optic radiation (L)	134	-31	-70	6	4.1
Guessing	1	Forceps Major	423	-24	-74	14	3.77
Target detection	1	Forceps Major	614	-26	-73	13	3.79
Identification time	1	Vertical occipital Fasciculus (L)	806	-34	-78	16	4.65
	2	Superior Longitudinal Fasciculus 1 (L)	276	-15	-58	53	3.38
	3	Superior Longitudinal Fasciculus 1 (L)	66	-22	-66	32	3.46
	4	Superior Thalamic Radiation (L)	42	-18	-12	39	38.7
Identification accuracy	1	Superior Longitudinal Fasciculus 2 (L)	1635	-33	-59	26	3.55
	2	Cingulum - posterior (L) (JHU)	1083	-16	-60	35	3.72
	3	Corticospinal Tract (L)	268	-23	-28	42	3.82

Table 5.3 | Correlations between VSTM metrics and FA in AD and EHC: Clusters and diffusivity peaks

White matter tract	MB	GS	TG	IT	IA
Optic Radiation (L)	47.688	33.250	37.460	11.000	10.811
Forceps Major	28.714	37.236	30.849	16.928	7.967
Middle Longitudinal Fasciculus (L)	20.731	23.034	25.043	9.731	6.754
Inferior Fronto-occipital Fasciculus (L)	7.626	4.294	6.389	0.977	3.763
Vertical Occipital Fasciculus (L)	1.999	5.041	7.547	4.332	2.073
Superior Longitudinal Fasciculus 1 (L)				4.480	4.518
Superior Longitudinal Fasciculus 2 (L)				4.032	3.025
Superior Longitudinal Fasciculus 3 (L)				0.842	1.178
Superior Thalamic Radiation (L)				0.748	0.804
Corticospinal Tract (L)					1.165
Arcuate Fasciculus (L)					0.764

Table 5.4 | Correlations between VSTM metrics and FA in AD and EHC: Tracts probability

MB = Misbinding, GS = Guessing, TG = Target detection, IT = Identification time, IA =

Identification accuracy. L = left. Values represent probabilities of the entire significant cluster within each metric of belonging to a tract. At visual inspection, Misbinding, Guessing and Target detection had no significant clusters belonging to the left Vertical Occipital Fasciculus and the left Inferior Fronto-occipital Fasciculus.

Higher Identification time and lower Identification accuracy were correlated with lower FA in the same core area as the three mixture model metrics, and in additional tracts including the left Vertical Occipital Fasciculus, the left Superior Longitudinal Fasciculus (1, 2, 3) and left Superior Thalamic Radiation (**Figures 5.6 - 5.7**, **Tables 5.3 - 5.4**). Higher Identification accuracy was associated with reduced FA also in the left Inferior Fronto-occipital Fasciculus, left Corticospinal Tract, left Arcuate Fasciculus and the left Cingulum (**Figures 5.6 - 5.7**, **Tables 5.3 - 5.4**).

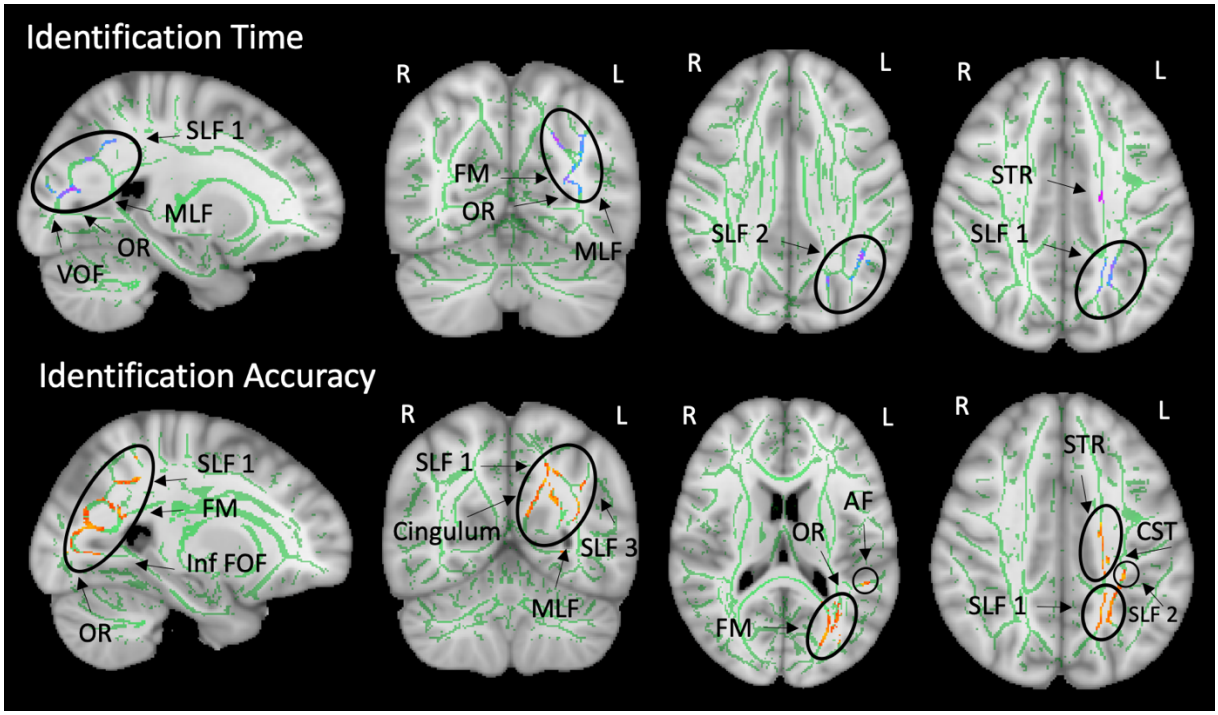


Figure 5.6 | Identification metrics and correlations between all VSTM metrics and FA in AD and EHC

Identification Time in cool colours (upper row), Identification Accuracy in orange (bottom row). Superior Longitudinal Fasciculus 1 (SLF 1), Superior Longitudinal Fasciculus 2 (SLF 2), Superior Thalamic Radiation (STR), Corticospinal Tract (CST), Arcuate Fasciculus (AF), Optic Radiation (OR), Forceps major (FM), Middle Longitudinal Fasciculus (MLF), Inferior Fronto-occipital Fasciculus (Inf FOF), Vertical Occipital Fasciculus (VOF).

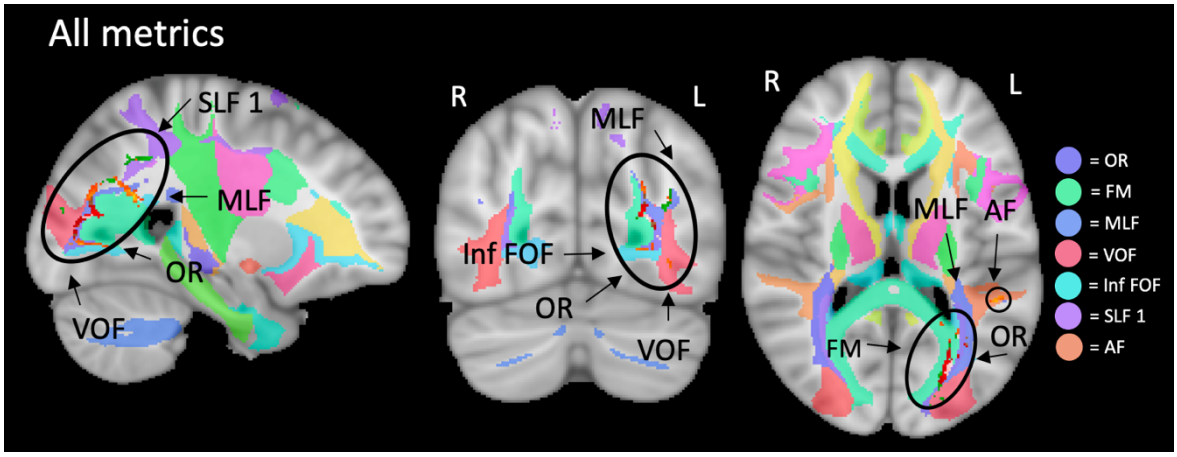


Figure 5.7 | Identification metrics and correlations between all VSTM metrics and FA in AD and EHC

Significant clusters from all metrics are overlaid according to cluster size, with smaller clusters at the top. Misbinding in pink, Guessing in red, Target Detection in copper, Identification Time displayed in green in this figure for visualization purposes, Identification accuracy in orange. From the XTRACT atlas: Optic Radiation (OR) in violet, Forceps Major (FM) in mint green, Middle Longitudinal Fasciculus (MLF) in light violet, Inferior Fronto-occipital Fasciculus (Inf FOF) in light blue, Vertical Occipital Fasciculus (VOF) in coral red, Superior Longitudinal Fasciculus 1 (SLF) in lilac, Arcuate Fasciculus (AF) in orange.

All metrics displayed the expected effect direction, with lower Target detection and Identification accuracy and higher Misbinding, Guessing and Identification time, being associated with reduced FA, and no areas were found in the reverse contrast for any metric. No specific tract was associated with Absolute Error or memory Imprecision.

Correlations between VSTM metrics and white matter tracts in each cohort (AD and EHC)

Next, we investigated whether these results held true in the cohort of AD patients alone. In this group, higher Misbinding and higher Guessing were associated with reduced FA in the same key regions observed in the previous analysis, including the left Optic Radiation, the Forceps Major and the left Middle Longitudinal Fasciculus (**Figure 5.8**, **Tables 5.5** - **5.6**), with the difference that in this case Guessing had the smallest cluster size. However, this analysis confirmed that, similarly to the previous analysis, Misbinding had a much higher

involvement of the left Optic radiation compared to Guessing, while higher Guessing was associated more tightly with lower FA in the Forceps Major (**Figure 5.8**). This was supported also by the localization of the respective peaks and distributions of tracts' probability (**Tables 5.5 - 5.6**). Unlike the whole group analysis, in AD patients alone Target detection not only involved these core areas but also showed a significant involvement of the left Superior Longitudinal Fasciculus 1 and 2, posterior Cingulum, left Vertical Occipital Fasciculus, left Inferior fronto-occipital Fasciculus and also of a small cluster in the right Inferior fronto-occipital Fasciculus (**Figure 5.8**, bottom row, **Tables 5.5 - 5.6**).

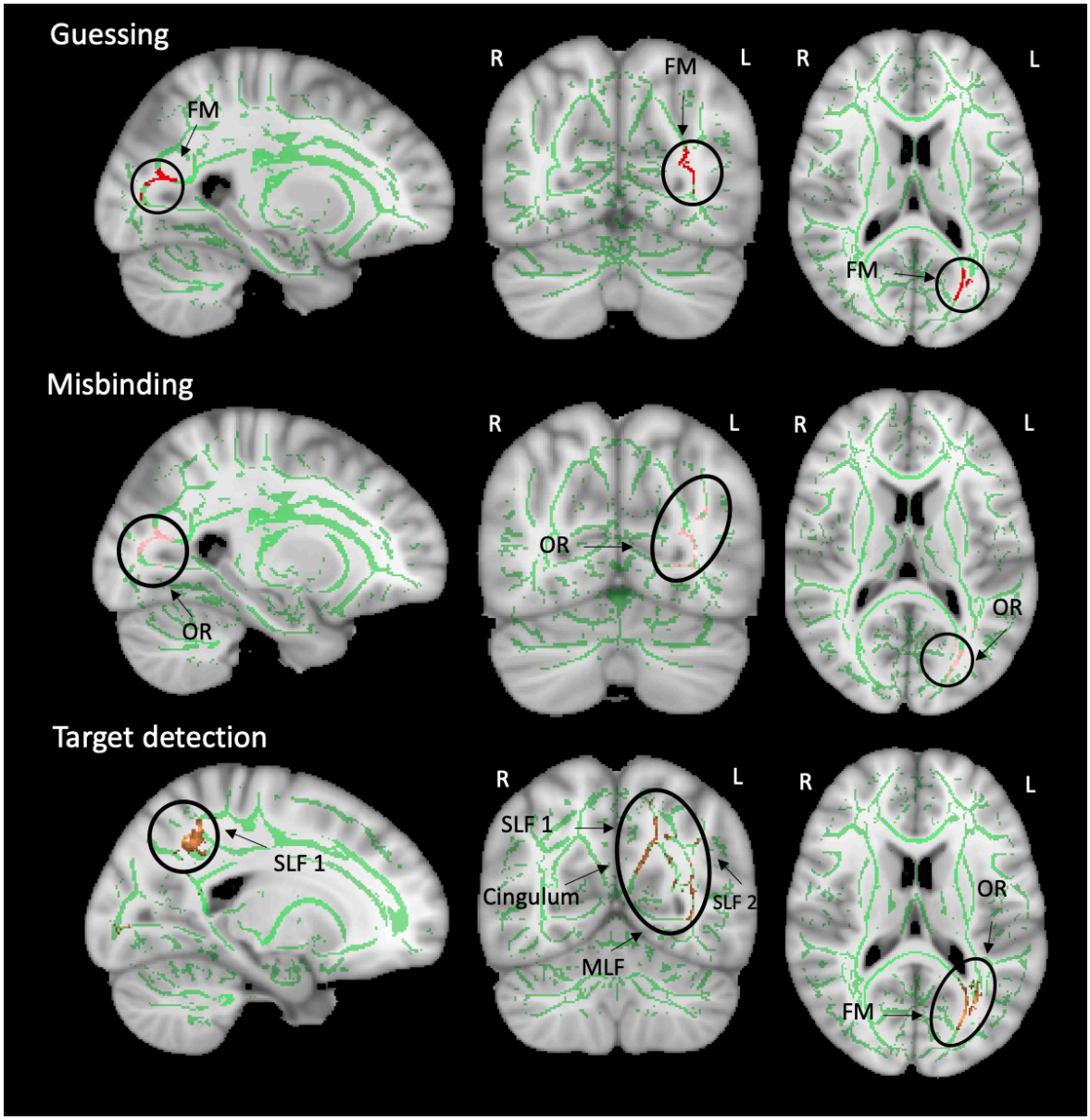


Figure 5.8 | Correlations between mixture model VSTM metrics and FA in AD

Guessing in red (upper row), Misbinding in pink (middle), Target Detection in copper (bottom). From the XTRACT atlas: Optic Radiation (OR), Forceps Major (FM), Middle Longitudinal Fasciculus (MLF), Superior Longitudinal Fasciculus (SLF). R = right, L = left.

	Cluster	White matter tract	Cluster Size	Coordinates (x,y,z)			Z-score
Guessing	1	Forceps Major (82%) Optic Radiation (L) (14%) Middle Longitudinal Fasciculus (L) (9%)	429	-24	-73	14	3.35
Misbinding	1	Optic Radiation (L) (79%) Middle Longitudinal Fasciculus (L) (26%) Inferior fronto-occipital Fasciculus (L) (22%)	817	-32	-67	7	3.68
Target detection	1	Superior Longitudinal Fasciculus 1 (L)	1971	-20	-57	42	3.56
	2	Middle Longitudinal Fasciculus (L)	158	-33	-72	18	3.41
	3	Superior Longitudinal Fasciculus 2 (L)	58	-34	-64	34	2.61
	4	Superior Longitudinal Fasciculus 2 (L)	47	-42	-54	33	2.71
	5	Inferior fronto-occipital Fasciculus (R)	13	35	-57	27	2.73
Identification time	1	Vertical Occipital Fasciculus (L)	2267	-34	-78	16	4.43
	2	Superior Thalamic Radiation (L)	112	-18	-12	39	3.97
	3	Cingulum - temporal (L)	12	-31	-31	-15	3.98
Localization time	1	Optic Radiation (L) (48%) Middle Longitudinal Fasciculus (L) (46%) Forceps major (23%)	235	-25	-75	14	3.75
Identification accuracy	1	Superior Longitudinal Fasciculus 1 (L)	328	-15	-60	35	3.38
	2	Forceps Major	206	-24	-71	15	2.95
	3	Forceps Major (42%) Optic Radiation (L) (13%) Middle Longitudinal Fasciculus (L) (10%)	74	-31	-67	8	2.77
	4	Superior Longitudinal Fasciculus 1 (L)	66	-22	-65	33	2.48
	5	Superior Thalamic Radiation (L)	10	-18	-16	38	3.54

Table 5.5 | Correlations between VSTM metrics and FA in AD: Clusters and diffusivity peaks

White matter tract	GS	MB	TG	IT	LT	IA
Forceps Major	37.498	15.163	13.893	12.233	55.485	17.607
Optic Radiation (L)	35.745	34.581	19.119	10.909	19.970	12.682
Middle Longitudinal Fasciculus (L)	23.295	18.350	10.911	8.104	18.382	8.929
Inferior Fronto-occipital Fasciculus (L)	6.534	11.250	6.349	2.604	0.345	1.961
Vertical Occipital Fasciculus (L)	3.811	7.351	3.766	3.111	3.519	1.542
Superior Longitudinal Fasciculus 1 (L)			5.816	3.889		8.533
Superior Longitudinal Fasciculus 2 (L)			1.482	3.237		0.159
Superior Longitudinal Fasciculus 3 (L)			0.2780	0.537		
Superior Thalamic Radiation (L)				1.405		0.029
Corticospinal Tract (L)			0.012	0.189		
Inferior Longitudinal Fasciculus (L)			0.031			
Cingulum temporal section (L)				0.201		
Inferior Fronto-Occipital Fasciculus (R)			0.010			

Table 5.6 | Correlations between VSTM metrics and FA in AD: Tracts probability

MB = Misbinding, GS = Guessing, TG = Target detection, IT = Identification time, LT = Localization time, IA = Identification accuracy. L = left, R = right. At visual inspection, Misbinding and Guessing had no significant clusters belonging to the left Vertical Occipital Fasciculus and the left Inferior Fronto-occipital Fasciculus.

In the AD cohort, Identification time had the biggest cluster across all VSTM metrics, with longer times correlating with lower FA in the same areas as Target detection, and additional tracts including the left Superior Thalamic Radiation and in a small cluster in the left temporal section of the Cingulum (**Figure 5.9**, **Tables 5.5** - **5.6**). The peak of the main cluster for this metric was in the left Vertical Occipital Fasciculus (**Figure 5.9**, **Table 5.5**). Unlike the previous analysis in the whole cohort, in AD patients there was also a small area that was associated with longer Localization times, corresponding to the key area found in Guessing and Misbinding, with a prominent involvement of the Forceps Major (**Figure 5.9**, **Tables 5.5** - **5.6**).

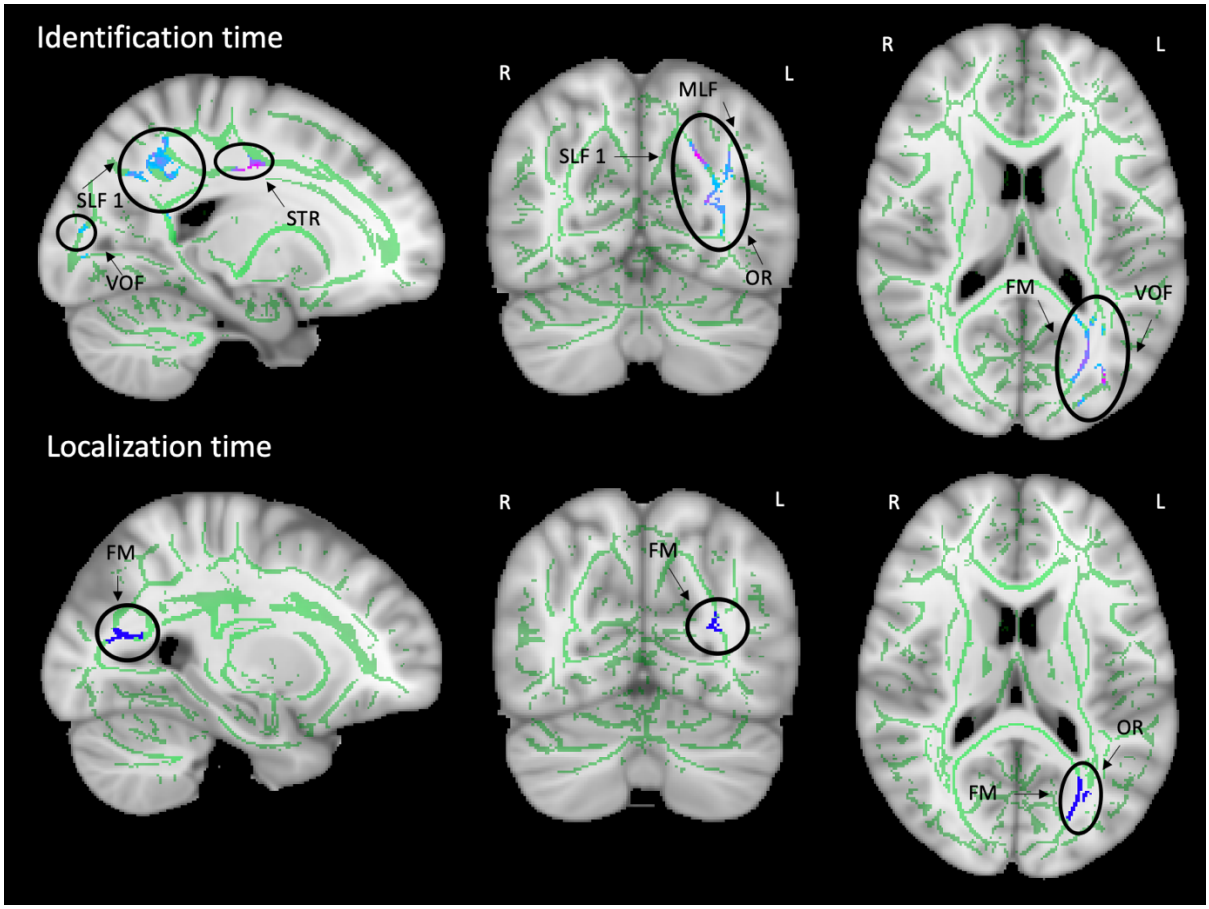


Figure 5.9 | Correlations between reaction times metrics and FA in AD

Identification time in cool colours, the pinker the higher the Z-score (upper row), Localization time in blue (lower row). Superior Longitudinal Fasciculus (SLF), Superior Thalamic Radiation (STR), Vertical Occipital Fasciculus (VOF), Middle Longitudinal Fasciculus (MLF), Forceps Major (FM), Optic Radiation (OR). R = right, L = left.

Lastly, in the AD cohort, lower Identification accuracy was associated with lower FA in the same key areas involved in Misbinding and Guessing, and additionally the left Superior Longitudinal Fasciculus 1 and the left Superior Thalamic Radiation (**Figure 5.10**, **Table 5.5** and **5.6**).

All metrics displayed the expected effect direction, as in the previous analysis, with lower Target detection and Identification accuracy and higher Misbinding, Guessing, Identification and Localization Time, being associated with reduced FA. No areas were

found in the reverse contrast for any metric in AD patients. No specific tract was associated with Absolute Error or memory Imprecision.

We did not find any association between VSTM metrics and white matter diffusivity changes in EHC, except for Identification accuracy. Notably, in EHC, lower Identification accuracy was correlated with lower FA in the right hemisphere, in a homologous region compared to the key common area across VSTM metrics seen on the left in AD patients, including Forceps Major, Optic Radiation and Middle Longitudinal Fasciculus, (**Figure 5.10**, **Tables 5.7 - 5.8**). Additionally, there was an involvement of the right Superior Longitudinal Fasciculus 1, and the left Middle Longitudinal Fasciculus.

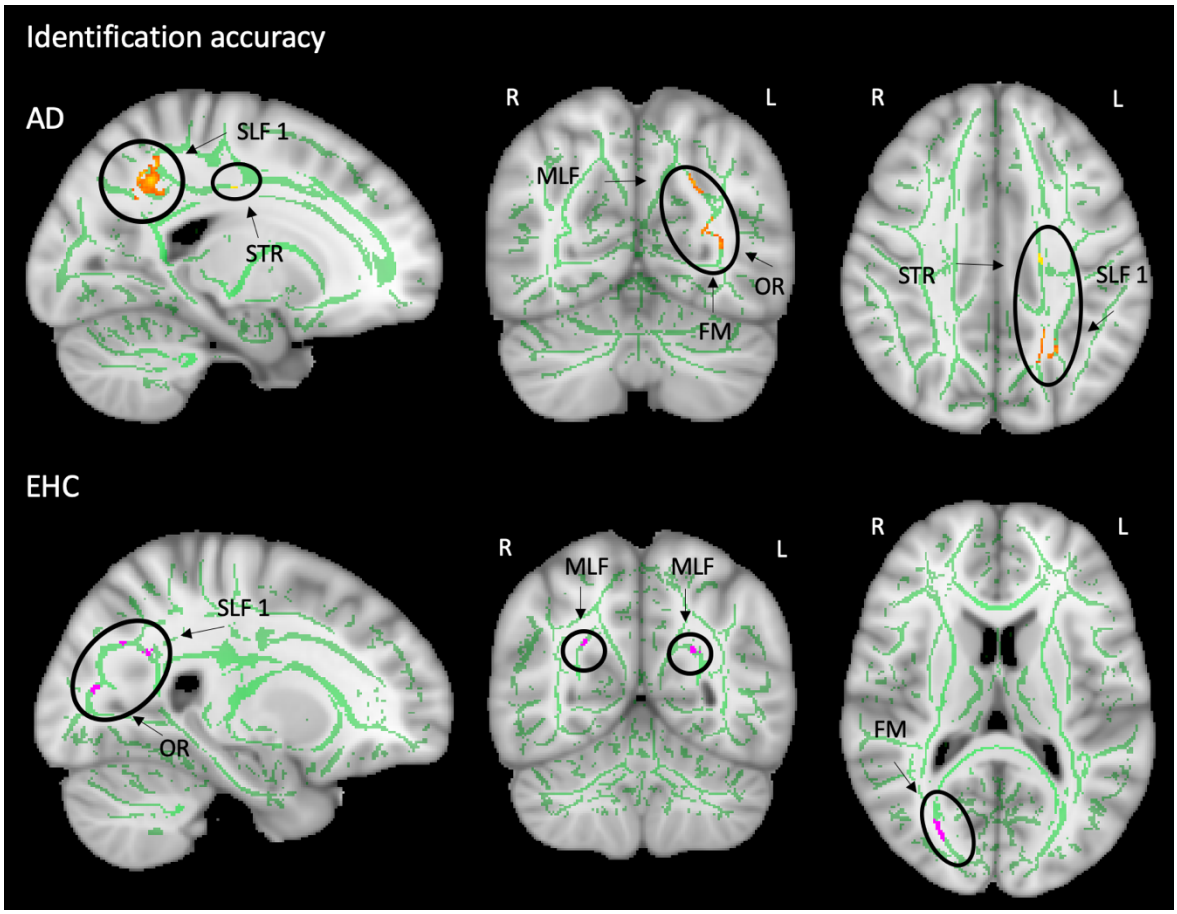


Figure 5.10 | Correlations between Identification accuracy and FA in AD and EHC

Identification accuracy in orange for AD (top row), and in magenta for EHC (lower row). Superior Longitudinal Fasciculus (SLF), Superior Thalamic Radiation (STR), Middle Longitudinal Fasciculus (MLF), Forceps Major (FM), Optic Radiation (OR). R = right, L = left.

	Cluster	White matter tract	Cluster Size	Coordinates (x,y,z)			Z-score
Identification accuracy	1	Forceps Major (83%) Optic Radiation (R) (43%) Middle Longitudinal Fasciculus (R) (24%)	200	32	-67	4	3.49
	2	Superior Longitudinal Fasciculus 1 (R)	42	24	-74	36	4.77
	3	Middle Longitudinal Fasciculus (R)	26	25	-56	28	3.36
	4	Middle Longitudinal Fasciculus (L)	24	-25	-56	24	4.2

Table 5.7 | Correlations between VSTM metrics and FA in EHC: Clusters and diffusivity peaks

White matter tract	IA
Optic Radiation (R)	34.455
Forceps Major	13.410
Middle Longitudinal Fasciculus (R)	16.565
Inferior Fronto-occipital Fasciculus (R)	6.546
Vertical Occipital Fasciculus (R)	10.914
Superior Longitudinal Fasciculus 1 (R)	0.103
Middle Longitudinal Fasciculus (L)	0.223
Optic Radiation (L)	0.024

Table 5.8 | Correlations between VSTM metrics and FA in EHC: Tracts probability

IA = Identification accuracy.

If global cognitive impairment, represented by ACE scores, was regressed out as in Chapter 4, no areas were associated with any of the VSTM metrics in the whole cohort and in AD patient alone, but as expected the correlations remained significant for the EHC only, data presented in Appendix (**Figure 5.1S**, **Tables 5.1S - 5.2S**).

5.4 Discussion

In this Chapter, whole brain TBSS analysis was deployed to investigate voxelwise associations between white matter integrity and the VSTM metrics derived from the "What was where?" Oxford memory task, in a cohort of sporadic late onset AD patients and healthy controls.

Across the whole cohort and in AD patients, the analysis revealed a key white matter area in the left hemisphere to be associated with several VSTM metrics: Guessing, Misbinding, Target Detection, Identification time, Localization time and Identification accuracy. It consisted of left Optic Radiation, Forceps Major and Middle Longitudinal Fasciculus (**Figures 5.5 - 5.10, Tables 5.3 - 5.6**).

The OR (known also as posterior thalamic radiation) is a white matter bundle connecting the primary visual cortex to the lateral geniculate nucleus [369]. The left OR is severely affected in AD, even in the earliest stages of the disease [370]. Indeed, reduced FA in the left OR is considered to be amongst the best predictors for early discrimination between healthy controls and patients with subjective and mild cognitive impairment, but also for clinical progression to AD dementia [351, 370]. In AD it has also been associated with worse performance on multiple standard neuropsychological tests including the MMSE, clinical dementia rating scale sum of boxes (CDR-sob), the Alzheimer's Disease Assessment Scale - Cognitive Subscale (ADAS-cog) and RAVLT [351, 370]. The left OR has also been linked to the precision of color representation in a VSTM task in patients with stroke [371]. However, it is the first time that the disruption of microstructural integrity in the left OR in AD has been associated with worse performance across multiple VSTM metrics.

Similarly, the Forceps Major (FM), a white matter tract connecting the occipital lobes bilaterally, is heavily affected in AD across all disease stages, and shows a very tight correlation with measures of cognitive impairment [351, 370]. DTI changes in the FM have also been found to be amongst the earliest markers of microstructural disruption in patients with FAD, becoming apparent a decade before symptom onset [372].

Our results suggest that the integrity of the OR and the FM is an important predictor of VSTM performance, but this appears to vary across physiological and pathological ageing, with the right hemisphere being more important in healthy elderly controls, as opposed to the left hemisphere in AD patients. Moreover, whilst these two structures were associated with multiple VSTM metrics, Misbinding might have a more prominent involvement of the OR whilst higher Guessing seems more tightly associated with disruption of the FM, which is shown by the peaks and the percentage of tracts involvement across the whole cohort and particularly in AD patients (**Figures 5.5** and **5.8**, **Tables 5.3 - 5.6**). We postulate that an increasingly faulty communication between the two occipital lobes (carried by the FM), causes spatial information of the object-to-be-remembered to be compromised. Conversely, misbinding errors might arise from a disrupted connection from the occipital lobe, through the medial temporal lobe to the thalamus, where some spatial information might be retained but wrongly processed as linked to another object in memory.

The Middle Longitudinal Fasciculus (MLF) is far less studied, as its existence in humans has been unravelled relatively recently by the use of DTI [373]. It connects the Angular Gyrus to the Superior Temporal Gyrus, and whilst its functions have not been completely understood, it has been proposed as being likely associated with a traditional right – left lateralization for visuospatial processing in the right hemisphere, and speech in the left

hemisphere [373, 374]. Our data might shed a light on the differential function of the MLF in health and disease. In healthy subjects the integrity of both the right and the left MLF is important for correctly identifying an item, measured by Identification accuracy of our VSTM task, (**Figure 5.10**, **Tables 5.7 - 5.8**). On the other hand, in AD patients only the left MLF seems to be important to guarantee optimal performance across different VSTM metrics.

Two other white matter tracts that have been involved in VSTM performance at our task, and particularly associated with Target Detection, Identification time and Identification accuracy, were the Inferior Frontal-occipital Fasciculus (Inf FOF or IFOF) and the Vertical Occipital Fasciculus (VOF). The IFOF is a tract that connects directly occipital and frontal lobes, it originates from the lingual, posterior fusiform, cuneus and polar occipital cortex, passes through the internal capsule, and terminates in the Inferior Frontal Gyrus, medial fronto-orbital region and Frontal Pole [375]. The IFOF has been associated with visual working memory performance in healthy individuals, and particularly with visual learning, and is thought to be negatively impacted by ageing [376]. Performance on a VSTM task in a cohort of healthy controls and patients with multiple sclerosis, showed that the IFOF was a strong predictor of performance for object but not for spatial localization accuracy in both groups [377]. Our data support the idea the integrity of the left IFOF supports correct object selection in patients with AD.

The VOF is the only white matter tract connecting dorsolateral and ventrolateral visual cortex [378]. It is thought to be crucial to integrate ventral regions, that are important for the perception of visual stimuli, and dorsal regions, involved in the control of eye movements, attention, and motion perception [378]. Our data support that its disruption contributes to

several VSTM deficits in AD patients, and particularly seems to have the highest association (highest peak) with Identification time (**Figure 5.9**, **Tables 5.5 - 5.6**). This might reflect the complex interaction between visual perception and the execution of the movement needed to select the correct object, which can be degraded in patients with AD, leading to longer Identification times.

The left Superior Longitudinal Fasciculus (SLF) has been shown to be associated with Identification accuracy and Identification time across the whole cohort and additionally with Target detection in patients with AD (**Figures 5.6 - 5.10, Tables 5.3 - 5.6**), while the right SLF 1 was associated with Identification accuracy in our healthy elderly controls (**Figure 5.10, Tables 5.7 - 5.8**). The SLF can be divided into dorsal (SLF 1), middle (SLF 2), and ventral sections (SLF 3), with different patterns of connections to other brain regions and functional specialization [379, 380]. According to these studies, the SLF 1 connects the Superior Parietal Lobule to superior frontal lobe (motor and premotor areas) and has been associated with higher order motor control. Conversely, the SLF 2 conveys information from the inferior parietal lobe (particularly, the intraparietal sulcus) to the prefrontal cortex (as the Superior and Middle Frontal Gyrus) and is particularly involved in visual space perception and re-orienting of spatial attention.

Finally, the SLF 3 connects the rostral part of the Inferior Parietal Lobule with the Inferior Frontal Gyrus and has been associated with monitoring orofacial and hand actions, working memory and ideomotor apraxia in humans [381]. The Arcuate Fasciculus (AF) in humans is thought to follow a similar trajectory to SFL 2 and have similar functions [379]. A large metanalysis [380] found that the SLF 1 carries primarily spatial information, the SLF 3 encodes non-spatial motor information [380], while SLF 2 holds both types of content. VSTM in that

study was represented by spatial and non-spatial clusters, and therefore all three tracts could carry VSTM information. In healthy individuals, right predominance of the SLF has been associated with superior performance on visuospatial tasks [382]. Misbinding errors, as estimated by the mixture model, have been found to be associated with white matter integrity in the SLF (1, 2, 3), IFOF and Inferior Longitudinal Fasciculus in young adults, while in the same study (preprint) a global reduction of axonal diffusivity in the same tracts was associated with lower memory precision [383]. We propose that the left SLF 1 and 2, and to a much lesser extent SLF 3 and AF, play a role in the mechanisms underlying the correct representation of an item in VSTM and are also essential to execute the action of selecting the correct item. In healthy controls this function is carried out uniquely by the right SLF 1, while in AD this is accomplished through the left SLF 1, 2 and 3 and left AF.

The Superior Thalamic Radiation (STR), or thalamo-parietal bundle, connects anterior and intermediate ventro-lateral (VL) nuclei of the thalamus to the prefrontal cortex, and ventral postero-lateral (VPL) nuclei to the parietal cortex [384]. These are involved respectively in coordination and planning of movement (VL) and conveying somatosensory input from different body areas (VPL) [384]. Our data suggests that, similarly to SLF, disruption of the left STR in AD patients is associated with worse ability to hold the representation of an object in VSTM (Identification accuracy) and execute the correct action to select it (Identification time), (**Figures 5.6** and **5.9 - 5.10, Tables 5.3 - 5.6**).

Lastly, we found that the integrity of the Cingulum is also important for VSTM performance in AD patients. In our sample, a disruption of posterior sections of the Cingulum was associated with worse Identification accuracy and Target Detection, while the temporal sections showed a correlation with Identification Time (**Figures 5.6, 5.8, Tables 5.3, 5.5 -**

5.6). The prominent involvement of the Cingulum in AD patients has been shown by multiple studies [348, 385, 386] and also by our own group contrast (**Figure 5.4**, **Tables 5.1 - 5.2**). However, for the first time we were able to link specific lesions in the posterior sections to maintaining the correct information of the identity of an object in VSTM (Identification accuracy, Target Detection), while temporal sections might be more important to determine the time taken to evaluate and carry out the response (Identification time).

Significant brain-behavioural correlations between FA in several white matter tracts and Identification accuracy on this VSTM task, primarily in the right hemisphere, were found in the cohort of healthy elderly adults. These tracts included on the right the FM, OR, SLF 1 and bilaterally the MLF.

VSTM has been associated with the integrity of the Superior and Inferior Longitudinal Fasciculus, Inferior Fronto-occipital Fasciculus, Uncinate Fasciculus and Forceps Major in the right hemisphere, in children and adolescents [387, 388, 389]. According to the "last in, first out" theory of brain aging, the same areas that regulate higher cognitive functions, such as executive functions, mature later but also are the ones that are more vulnerable to physiological and pathological ageing [390]. In another study, VSTM capacity, measured by the Symbol Span subtest of the Wechsler Memory Scale [391], was found to be associated with reduced axonal density as measured by diffusion kurtosis imaging (DKI) in several white matter tracts in healthy adults, including Superior and Inferior Longitudinal Fasciculi, Optic Radiation, Inferior Occipito-frontal Fasciculi, Cingulum bundles, and Uncinate Fasciculi [392].

Our results are encouraging as even by using a relatively crude, but easily accessible measure such as FA, it was possible to detect a preferential involvement of specific white

matter tracts in VSTM abilities in elderly healthy controls. Our interpretation is that by deploying more fine-grain measures of cognition, it is possible to track subtle white matter microstructural changes even in healthy people. The right OR has previously been associated with performance on a VSTM task (Visual Pattern Test, Juelich version) in elderly adults [393]. We were able to replicate this finding in our cohort, where it was associated with the accuracy of object identification.

No specific tract was associated with VSTM spatial localization properties (Absolute Error and memory Imprecision). However, this might be in line with findings in a study using the mixture model of white matter in patients with stroke, which found that several brain areas were associated with memory binding and color precision, but none with the precision of spatial localization [371].

5.5 Conclusion

The analysis of white matter integrity reveals that a core area, comprising the left Optic Radiation, Forceps Major and Middle Longitudinal Fasciculus is important across different VSTM metrics in patients with AD. In healthy elderly controls a homologous region in the right hemisphere might be important in the processes required for correctly identifying an item. Additionally, specific VSTM functions, such as holding the correct item in memory and selecting it, might require the integrity of the left Superior Longitudinal Fasciculus, Superior Thalamic Radiation, Vertical Occipital Fasciculus, Inferior Fronto-occipital Fasciculus, Arcuate Fasciculus and Cingulum in AD.

Chapter 6. Combining Plasma Biomarkers and digital cognitive metrics in AD and EHC §#

6.1 Relationship between Plasma Biomarkers and digital VSTM metrics §#

Abstract

Plasma biomarkers have revolutionized the field of AD research, proving a cost-effective, minimally invasive, and easily scalable screening tool for early diagnosis. In this Chapter the relationship between a range of digital cognitive metrics (measuring VSTM, executive functions, visuospatial abilities, processing speed and LTM), and plasma biomarkers (pTau181, GFAP, NfL and Aβ42/40 ratio) was investigated in a cohort of elderly healthy controls and AD patients. Additionally, we evaluated the test-retest reliability of these digital tests and of plasma biomarkers over a short time frame (0, 3 and 6 months). Our data revealed an association between several digital cognitive metrics and the plasma biomarkers examined, particularly with pTau181, and to a lesser extent with GFAP and NfL, while the Aβ42/40 ratio showed the weakest correlations with cognition. The test-retest analysis showed that pTau181 had a high coefficient of variation (CV) of around 35.8%, and therefore poor test-retest reliability. GFAP (18.4%) and NfL (15.8%) had moderate reliability, while the Aβ42/40 ratio had the lowest CV (9.5%), denoting higher stability over repeated measures. These did not differ between healthy controls and AD patients. Amongst the different digital metrics, Identification accuracy and Target detection on the "What was

Where?" Oxford Memory Task (OMT) and Immediate Object Accuracy and Immediate Semantic Accuracy of the Object-in-Scene Memory Task (OIS) showed high test-retest reliability.

6.1.1 Introduction

The advent of new disease-modifying treatments for AD has highlighted the need for sensitive screening tests to stratify those who might benefit from early treatment, as well as to track the effects of interventions [159]. Traditional cognitive tests are able to detect changes only several years after pathological accumulation of amyloid and tau, a factor that is considered to be crucial for clinical trial failure [394]. Digital metrics can detect subtle signs of cognitive impairment that cannot be captured by standard tests of cognition and are particularly valuable tools in early phases of the disease, when cognitive impairment is at subthreshold levels using commonly used clinical scales [77]. These could potentially bridge the gap between assessing cognition too late - where coarse tests such as the MMSE or MoCA reveal presence of cognitive impairment - and discarding the contribution of cognition altogether [395, 396, 397].

Plasma Aβ42/40 ratio has been found to correlate well with its ratio in the CSF and with amyloid PET, and its reduction is associated with cognitive decline and risk of progression to AD dementia amongst cognitively unimpaired individuals, SCI, and MCI patients [398, 399, 400, 401, 402]. However, the difference in detecting A+ versus A- individuals is much smaller in plasma (around 10%) compared to CSF (around 37.7%), and other biomarkers such as phosphorylated tau (p-tau) 231 (where they are respectively around 166% and 85.6%) have

been proposed as alternative markers of amyloid pathology [403, 404]. However, plasma Aβ42/40 has the potential to reflect the 'A' in the ATN system.

Plasma glial fibrillary acidic protein (GFAP), which reflects astrocytosis, is also considered an early marker of amyloid deposition in healthy controls, SCI, MCI and AD dementia patients [405, 406, 407, 408]. This can be observed even in cognitively normal individuals with a normal amyloid status [405], and some evidence suggests that it is better than direct measures of amyloid such as the Aβ42/40 ratio in predicting positivity of amyloid PET [407, 408]. Recent data showed that GFAP levels are associated with amyloid PET burden and importantly are not associated with higher tau levels measured by tau PET independently of amyloid burden [405]. It has been suggested that GFAP can act as a mediator between amyloid and tau deposition, where reactive astrocytosis following amyloid plaque deposition might promote tau accumulation [405]. Being a marker of neuroinflammation, it is however not specific for AD, and is increased also in Lewy body disease [409], Frontotemporal dementia (particularly symptomatic progranulin mutation carriers) [410, 411] , traumatic brain injury [412], hemorrhagic stroke [413], and many other neurological diseases [414].

Conversely, plasma pTau181 has been emerging as a specific and sensitive biomarker for AD, possibly reflecting the 'T' section of the ATN criteria similarly to CSF pTau181[180, 188, 415]. It shows good correlation with its levels in the CSF [415], and unlike GFAP, it has been associated with both amyloid and tau PET positivity [188, 415], which is in line with the expected biomarker cascade hypothesis [394]. It is estimated that plasma pTau181 can detect pathological accumulation of tau approximately 6 years earlier than tau PET [416]. It has also been validated in neuropathologically confirmed AD cases and has been shown to associate with Braak staging [180, 415]. Increased levels of plasma pTau181 are found in amyloid positive

cognitively unimpaired individuals, and it increases further in amyloid positive MCI and AD patients, while is not increased in several other neurodegenerative diseases and clinical mimics of AD [180, 415].

Similarly to GFAP, another marker that can be altered across several neurological disorders is Neurofilament Light Chain (NfL). High baseline levels of NfL, an index of the rate of axonal injury, are strongly linked to N+ status according to the ATN criteria and is tightly associated with other measures of neurodegeneration such as CSF total tau (t-tau), MRI atrophy and FDG-PET hypometabolism [417, 418, 419]. It is however not specific for AD and can be increased across multiple neurodegenerative diseases.

Amyloid burden has traditionally shown a weaker association with cognition compared to tau [420, 421]. A large metanalysis looking at different methods measuring amyloid positivity in cognitively unimpaired elderly adults, not including blood-based biomarkers, showed that episodic memory might be correlated with amyloid burden, but global cognition and executive functions might not if assessed by amyloid PET [422]. There is also emerging evidence that the Aβ42/40 ratio can be associated with more sensitive neuropsychological tests such as the Face Name Associative Memory Exam cross-sectionally, at least in patients with subjective cognitive decline [423]. This test has been previously shown to correlate with amyloid burden measured by amyloid PET in cognitively unimpaired individuals [424]. Therefore, more fine grain cognitive tests might be needed to detect early signs of amyloid deposition.

GFAP has been linked to cognitive performance on standard tests of cognition in plasma and serum cross-sectionally [407, 409], to a decline in cognition over time [425], and to clinical conversion to AD [406]. Neuropathological evidence suggests that the extent of GFAP activation might be correlated to Braak staging [426], thus possibly representing increased

amyloid burden during disease course. However, using amyloid PET, GFAP has been found to be associated with an initial increase in amyloid burden in asymptomatic or mildly symptomatic individuals, but then its levels seem to follow an inverted U shape trajectory later in the disease course, with increasing GFAP being associated with lower amyloid burden in patients with advanced dementia [427].

Baseline and longitudinal increases of plasma pTau181 have shown associations with a decline in cognition, amyloid and tau accumulation, and brain atrophy [180, 187, 403, 428]. Crucially, longitudinal changes in plasma pTau181 reflect neurodegeneration only in Aβ + participants [416, 429], while plasma NfL is associated with neurodegeneration independent of Aβ status [430]. It has also shown to correlate with longitudinal decline in cognitively healthy participants [431]. The longitudinal increase in plasma pTau181 levels slows as individuals reach the advanced stages of disease, i.e., the association between plasma pTau181 concentrations and tau PET tracer retention is stronger in individuals with MCI than in those in the AD dementia stage, and its ability to discriminate between amyloid positive MCI and AD dementia patients might be suboptimal [184, 188, 432]. Plasma pTau217 has been found to increase longitudinally in cognitively unimpaired and MCI subjects with evidence of amyloid positivity (A+), and in MCI patients who converted to AD, but remained stable in the control groups (A– cognitively unimpaired individuals and MCI who did not convert to AD) over the course of 6 years [433]. In the same study, the longitudinal increase of pTau217 was correlated with worsening cognition on the MMSE and modified Preclinical Alzheimer's Cognitive Composite (mPACC) [434], in cognitively unimpaired and MCI participants. A recent report by the same group, compared head-to-head performance of different plasma biomarkers in cognitively unimpaired A+ individuals [435]. They found that in those subjects, pTau217, pTau181 and GFAP were associated with cognitive decline (on

the mPACC and MMSE), while pTau 231 and NfL did now show such association [435]. The authors found that pTau217 alone was the strongest individual predictor of cognitive impairment. This was observed also in the whole sample of cognitively unimpaired individuals (whether A+ or A-).

High baseline levels of NfL are correlated with greater cognitive impairment cross-diagnostically [436], and a longitudinal increase in NfL seems to be associated with worse cognitive scores in in cognitively unimpaired amyloid positive individuals [419] and patients with already evidence of cognitive deficits (such as MCI and AD) [417, 437]. These changes however do not seem to be specific for AD [430].

Beyond single biomarkers, a decline in cognition is likely to be multifactorial, rising from a combination of initial amyloid accumulation, followed by tau deposition, and finally neurodegenerative changes [438].

In this Chapter, my aim was to investigate whether plasma biomarkers levels are associated with several digital cognitive metrics, measuring VSTM, executive functions, processing speed, visuospatial abilities and LTM recall, in a cohort of elderly healthy individuals and AD patients. I also explored whether these association are biomarker-specific and whether there is a cognitive test or a combination of tests which could best predict group classification, amyloid burden and tau accumulation.

6.1.2 Methods

53 EHC and 46 AD patients were recruited from the Cognitive Disorders Clinic at the John Radcliffe Hospital in Oxford. Inclusion criteria for EHC and AD patients' classification

criteria followed the same principles outlined in previous Chapters. Participants' basic demographics and tests are presented in **Table 6.1**. All participants gave written informed consent prior to the start of the study. Ethical approval was granted by the University of Oxford's MS IDREC.

Outline of the metrics collected

1) Plasma biomarkers

- *Amyloid pathology ('A')*: Aβ42, 40 and the Aβ42/40 ratio, which is a better measure of amyloid pathology than Aβ42 alone [402, 439].
- *Tau pathology ('T')*: pTau181, which is a specific and sensitive marker of tau pathology in the blood and is highly predictive of tau PET positivity [180].
- *Neurodegeneration ('N'):* NfL, the most commonly used blood-based biomarker reflecting the rate of neurodegeneration occurring in the brain [440].
- *Astrocytosis:* GFAP, an established marker of astrocytosis and synaptic plasticity [405].

2) Digital cognitive metrics

Six tests from a digital cognitive portal developed within our team by Sijia Zhao, available at https://oxfordcognition.org/ were included in the project: Oxford Memory Test, Object-in-Scene Memory Task, Rey-Osterrieth Complex Figure, Trail Making Test, Digit Symbol Substitution Test, Freestyle Corsi Block Task.

3) Standard test of cognition

The Addenbrooke's Cognitive Examination-III (ACE) was collected in person at the time of the visit as standard test of cognition.

6.1.2.1 Plasma biomarkers protocol

Blood was collected in 6 ethylenediaminetetraacetic acid (EDTA) tubes (10 ml each), and centrifuged (1800 g, room temperature, 10 minutes). The EDTA tubes were filled completely and gently inverted after collection to avoid coagulation. After centrifugation, plasma from all 6 tubes was transferred into one 50-mL polypropylene tube, mixed, aliquoted into 0.5 mL polypropylene tubes (Fluid X, Tri-coded Tube, Azenta Life Sciences) and stored at 4°C, until (< 8 hours) it was transferred into a -80°C freezer. Time between blood collection and centrifugation was < 30 minutes. Transfer time between 4°C and -80°C storage was < 20 minutes and the samples were kept refrigerated during transport. All cryovials were anonymized and the unique cryovial code was logged into a secure database, linked to the participant's anonymized code and visit number. A separate paper-based document (i.e., sample log) was filled in at collection, with the unique participant's anonymized code, gender, date of birth, hospital number, date and time of different pre-processing steps (collection, start and end time of centrifugation, time into the 4°C fridge and -80°C storage). An outline of the process in presented in **Figure 6.1**.

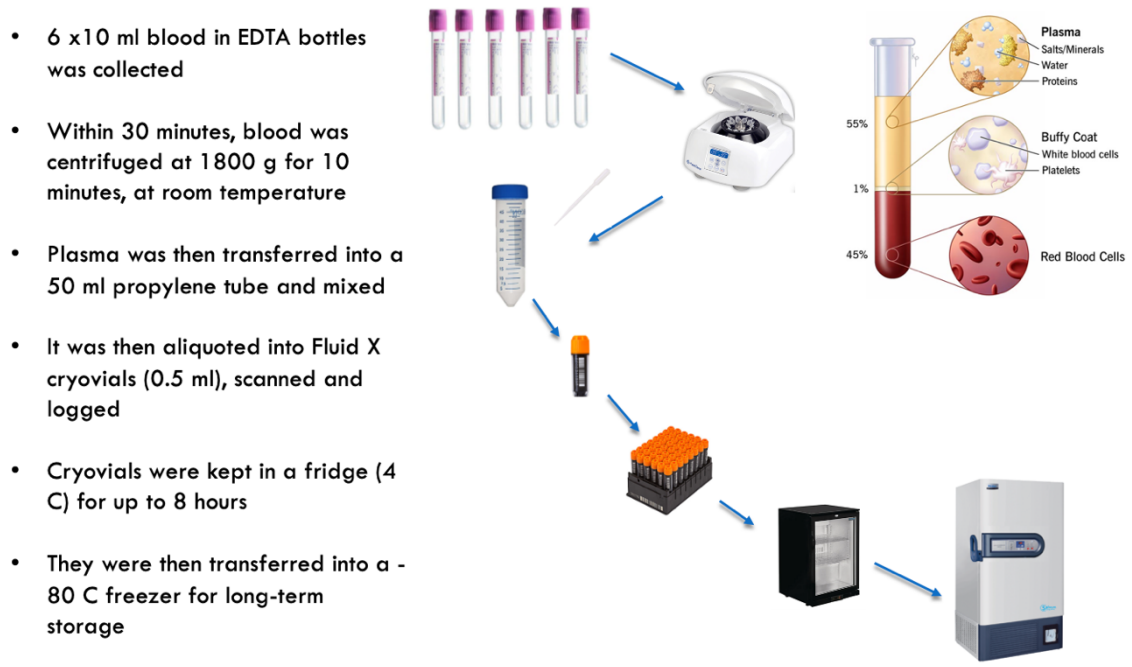


Figure 6.1 | Preprocessing pipeline for plasma biomarkers adopted for this study

Samples were sent to the Biomarker Factory / Fluid Biomarker Laboratory, UK Dementia Research Institute at UCL, London. Analyses were carried out by Dementia Research Institute (DRI) laboratory staff. Plasma Aβ40, Aβ42, GFAP, and NfL part of Neurology 4-plexE assay were measured by single molecule array (SIMOA) on an HD-x analyser (Quanterix), according to manufacturer's instructions. Plasma pTau181 was also measured by single molecule array (SIMOA) on an HD-x analyser (Quanterix). Briefly, samples were thawed at 21°C, and centrifuged at 10,000 RCF for five minutes at 21°C. On-board the instrument, samples were diluted 1:4 with sample diluent and bound to paramagnetic beads coated with a capture antibody specific for human Aβ40, Aβ42, GFAP, NfL and pTau-181. Aβ40, Aβ42, GFAP, NfL and pTau181bound beads were then incubated with a biotinylated Aβ40, Aβ42, GFAP, NfL and pTau181 detection antibody in turn conjugated to streptavidin-β-galactosidase complex that acts as a fluorescent tag. Subsequent hydrolysis reaction with a resorufin β-D-galactopyranoside substrate produces a fluorescent signal proportional to the concentration of Aβ40, Aβ42, GFAP, NfL and pTau-181 present. Singlicate measurements were taken of each sample. Sample concentrations were extrapolated from a standard curve, fitted using a 4-parameter logistic algorithm. Intra-assay and inter-assay CVs were less than 10% and 15% respectively, as determined by 8 quality controls following the same principles.

6.1.2.2 Digital cognitive metrics

A link with a unique patient and visit identifier was sent to the participants' email address the same day as the in-person visit when blood was collected. If on-site testing was previously requested, it was performed at the same day of the blood sample. Participants were required to use Chrome or Firefox browsers on a desktop, laptop or tablet, the operating system was not restricted. They were encouraged to complete the online tests within a week

maximum. After two weeks, completion of the online tests was reviewed and participants who did not complete the tasks within that timeframe were prompted via email to do so. Tests completed more than 3 weeks from the blood sample were discarded.

Oxford Memory Task (OMT)

A modified online version of the "What was Where?" Oxford Memory Task (OMT) described in the previous chapters which can be done remotely on a digital platform was used. In this version, only two experimental conditions (1 and 3 items), using a fixed delay of 4 seconds were included. Each participant performed a practice block of 6 trials including 3 trials with 1 item followed by 3 trials with 3 items. This was followed by a test block of 40 trials, including 20 trials of 1 item and 20 trials of 3 items. The order of trials was randomised.

As participants did the task remotely with their own device, to ensure that the size of stimuli was comparable a card calibration procedure was used, previously described and validated by Li et al. (2020) [441]. Participants were instructed to place a bank card or card of comparable size on the screen and adjust the slider until the size of the image of the card on the screen matched the size of the physical card. This allowed us to estimate screen distance by calculating the display's logical pixel density in pixels per centimetre. A MATLAB script (MathWorks, Inc.) was used to determine the fractals' locations in a pseudorandom manner with a few constraints. To avoid crowding and create a clear zone around the items' original locations, which is essential for the analysis of localization errors, fractals were never placed closer than 3 cm to one another. Additionally, they were placed at least 10% of the screen's total height/width away from the screen's edges. The experiment was implemented using PsychoPy v2022.2.4. along with some custom written codes developed in our lab by Sijia Zhao. The metrics extracted from this task are the same eight VSTM metrics described in

the previous Chapters (i.e., Identification time, Localization time, Identification accuracy, Absolute Error, Target detection, Misbinding, Guessing and memory Imprecision).

Object-in-Scene Memory Task (OIS)

This test provides measures of identification accuracy, precision of spatial localization and semantic accuracy in VSTM and LTM (**Figure 6.2**). Participants were presented with a photo of an everyday scene and instructed to remember a particular object placed in the picture. To aid effective encoding, the participant was also asked an unrelated question about the scene and to click on the displayed object. Subsequently, 20 different objects were presented, and participants were asked to choose the correct object and place it in the remembered location in the scene. To ensure that they were not simply remembering the semantic information or name of the object, the object pool contained a foil that matched the target's category (e.g., two guitars of different color and shape). After 5 different object and scene pairs, participants were asked to reproduce the object-scene associations probed (delayed recall, after 3-4 minutes). There was a total of 20 trials divided into four blocks, and the order of the pairs was randomized.

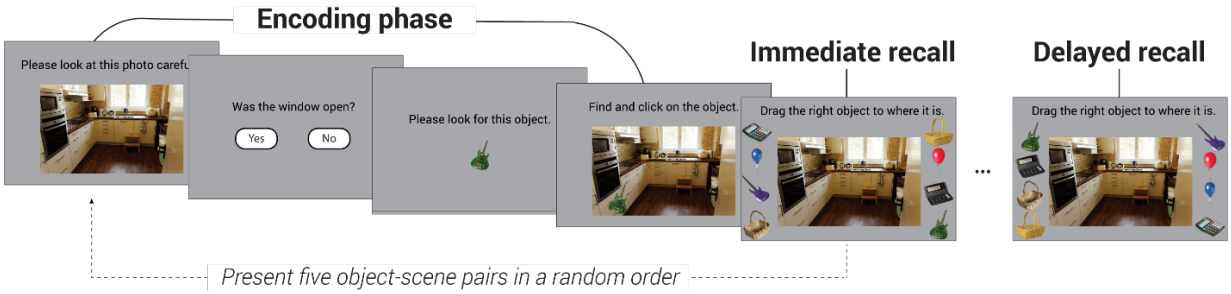


Figure 6.2 | Object-in-Scene Memory Task (OIS)

The following metrics were extracted from the task:

- **Object Accuracy - immediate and delayed**: the proportion of trials in which participants correctly identified the original object at immediate and delayed recall.
- **Location Error - immediate and delayed**: the distance from the original target item location to the center of participant's response location.
- **Semantic Accuracy - immediate and delayed**: proportion of items correctly recalled as belonging to the same semantic category as the target.

Rey-Osterrieth Complex Figure (ROCF)

This task is a digitized version of the traditional pen-and-paper test [442], which is an established measure of visuospatial abilities (**Figure 6.3**). The original ROCF task requires the participant to draw a complex line-drawing freehand, first by replicating an existing figure (copy), and then again from memory (immediate recall, and delayed recall after a 20–30-minute delay). Our digitised version does not require hand drawing. Instead, the figure is split into 13 independent elements, and participants are required to drag each element into an empty canvas to copy the figure. Each test was automatically scored using an offline MATLAB-based algorithm. In contrast to the discrete score used in the pen-and-paper version, the score of our digital version provides a continuous measure of precision. The middle large rectangle is selected as the anchor point as a reference element. If the element is not placed (not present on the canvas), there will be no score; otherwise, the distance from the large rectangle is computed. As a measure of imprecision for each element, the absolute difference between the ideal distance and the actual distance is then calculated. The absolute error is then scaled using a logarithmic function: if the element is placed relatively correctly, the difference in the distance from the big rectangle is computed; if the element is placed too far, the score is zero. The normalised absolute difference is then subtracted from 1 to

calculate the score for this element. The sum of all element scores is 13, but the results are scaled to a percentage to match the original 36 element picture. Our version's scoring is consistent with the pen-and-paper scoring guide, as the participant receives one point for correctly positioning the element and no score if the element is placed incorrectly or not at all on the canvas. This task and scoring have been validated with the in-person traditional 36 items pen-and-paper test and manual scoring with standard scoring guide [443] in healthy participants before the start of the study. The percentages obtained at the copy and immediate recall – ROCF copy and recall scores – were used as metrics of interest.

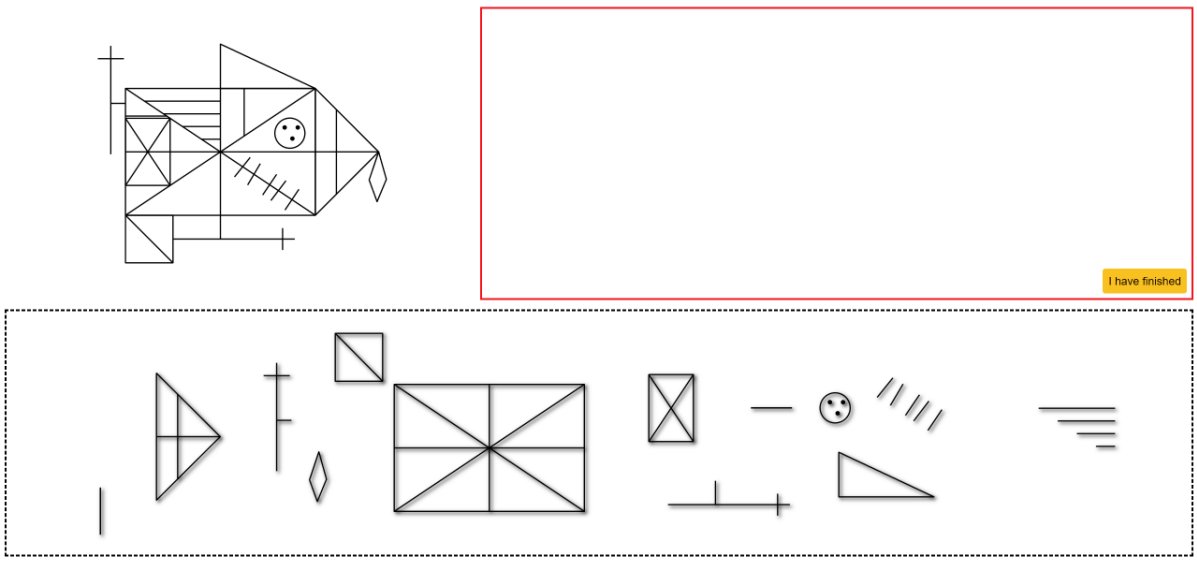


Figure 6.3 | Rey–Osterrieth Complex Figure's copy task

The participant is instructed to drag elements from the bottom black rectangle to the red rectangle on the top right.

Digit Symbol Substitution Test (DSST)

The Digit Symbol Substitution Test is a pen-and-paper measure of processing speed. In this digitized version, participants were required to match symbols to digits according to a key located at the top of the page (**Figure 6.4**). The key consisted of 9 symbols next to the digits 1–9. At the bottom of the screen, there was a row of 9 randomised symbols. Participants

reported the digit that corresponded to each symbol by clicking on the correct digit. The row was refreshed once all 9 were answered. Participants were allowed two minutes to answer as many as they could and the number of correct matches within the allowed time (DSST - correct responses) and their average reaction time were calculated (DSST - RT).

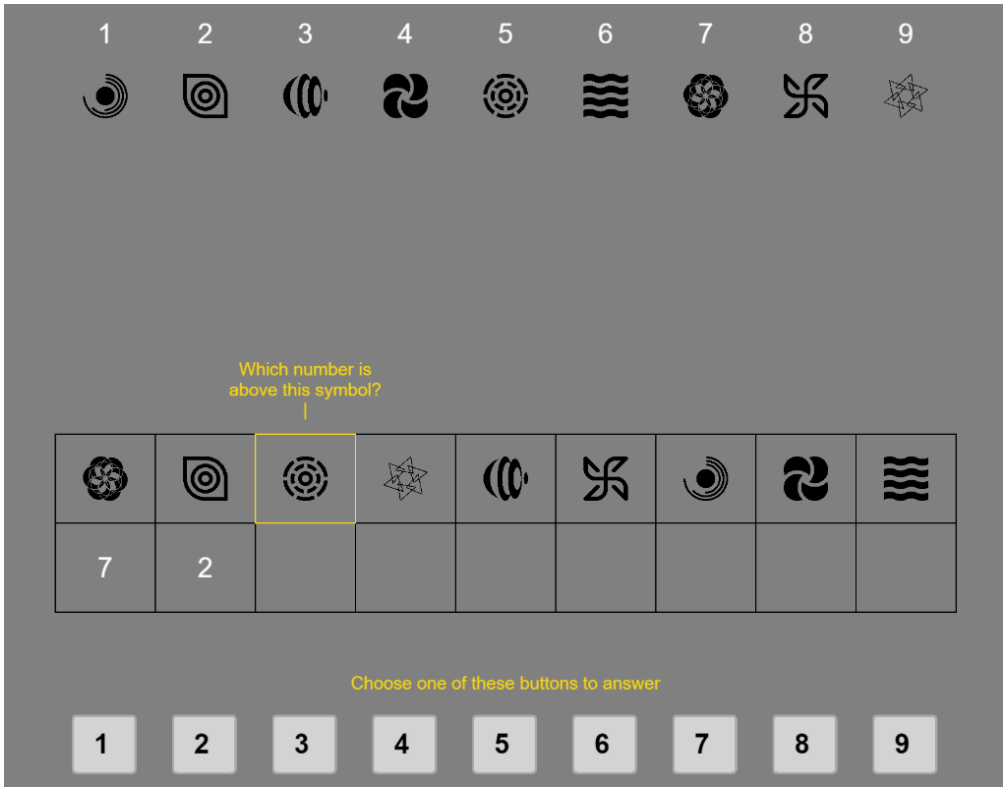


Figure 6.4 | Digit Symbol Substitution Test

Trail Making Test

The Trail Making Test (TMT) is a standard test of processing speed and executive functions [444]. In this online version, 25 circled numbers are presented on-screen, and participants are instructed to connect them by clicking the circles in order as fast as possible. It contains three trials of Task A (where the order is 1-2-3-4-5-6-…) and three trials of Task B (order 1-A-2-B-3-C-…). Each participant sees six different trail maps randomly chosen from a pool of 100 pre-made maps, generated using a "divide-and-combine" approach [445]. The task also

included a control condition of four trials to assess basic motor speed, where participants are presented with two circles located at two opposite corners of the screen. One is labelled with 1 and the other with 2, and participants were instructed to connect 1 with 2 as quickly as possible. The average of the reaction times of the TMT was used as variable of interest.

Freestyle Corsi Block Task

This task is a modification of the Corsi Block Tapping Task [446], which is a standard measure of visual STM. In the original version, participants were presented with a set of 9 identical wooden blocks positioned on a board. Subjects were required to point at the blocks in the order they were presented. They started with sequences of smaller blocks, and the number of blocks increased during the test. In most computerized version of the task the participant is shown several identical blocks that are in fixed locations spread across the screen [447]. Blocks then light up in sequence and the participant must remember which blocks lit up and in what order. In this digital version, the blocks' locations were not fixed ('Freestyle Corsi Block Task'). In an n-location trial, a 1-cm-wide red dot appeared at a random location on the screen, disappeared after 1 second, and reappeared at another random location on the screen (for n >1), and this process is repeated n times up to a sequence of 5 items. Once the sequence has finished, after a 1 second break the participant could freely click anywhere on-screen to indicate where each block appeared in sequence. The location error was calculated as the average distance between the response and the target location. The task was divided into five blocks, each block having five trials of a n-location sequence (i.e., 5 blocks of 1-item, 5 blocks of 2 item, up to 5 blocks of 5 items). The average of the reaction times of the 5 conditions was chosen as variable of interest.

6.1.2.3 Statistical analysis

For analysis and data visualization purposes MATLAB (version R2023a), R studio (version 12.0), JASP (version 0.16.4) and SPSS (version 29.0) were used. Demographics, cognitive tests and plasma biomarkers levels were compared using a t-test or Mann-Whitney U test, according to the normality of the data, for continuous variables, while χ^2 test was used for categorical variables such as gender. Rank-Biserial Correlation (rrb) was used to assess effect size. Plasma biomarkers' log10 transformed values and Z-scores of digital cognitive measures were used for correlation and linear regression analyses. Correlations between digital cognitive metrics and plasma biomarkers were assessed with Spearman's rank test, using age, sex, and education as covariates. The Benjamini–Hochberg method, which controls the False Discovery Rate (FDR), was used to correct for multiple comparisons across tests.

Additionally, machine learning was applied to predict group classification and plasma biomarkers levels, firstly using MATLAB-based algorithms for feature ranking, to estimate the absolute contribution of each variable. The fscchi2 function in MATLAB was used to predict group classification, while the fsrftest function was used to predict continuous variables, i.e., plasma biomarkers' levels. Rank's importance scores where then transformed into p-values by calculating the negative value of the exponential score. Secondly, we applied the R-based MuMIn package to test which combinations of biomarkers would best predict group classification and plasma biomarkers levels. For predicting group, we used logistic regression, while for predicting pTau181 levels and Aβ42/40 ratio linear regression. The MuMIn package then uses the dredge function to achieve model selection, with the best performing model having the lowest corrected Akaike information criterion (AIC). ROC curves and AUCs were then computed for the model of interest. The pROC package in R

with De Long's test was used to compare model performance in direct comparisons between two ROC curves.

6.1.3 Results

Participants and tests

AD patients and EHC were not significantly different in terms of age, gender, and education (**Table 6.1**). As expected, patients with AD had higher mean levels of pTau181, GFAP and NfL and lower Aβ42/40 ratio compared to controls. All the digital cognitive metrics used were also statistically significantly different across groups, with high effect sizes (|rrb| > 0.5). ACE was significantly different between AD and EHC and had a high effect size in group comparisons, which was expected as it was the only test used for diagnosis. Averaged values per participant across the three visits were used if participants took part in the longitudinal study described in 6.2.

	EHC (n= 53)	**AD (n= 46)**	**p-value**	**Rank-Biserial Correlation**
Age	68.6 (7.0)	68.3 (10.2)	0.861	
Gender (M/F)	24/29	23/23	0.639	
Education	15.8 (3.1)	14.5 (3.5)	0.088	
pTau181 (pg/ml)	2.6 (1.3)	5.4 (3.4)	* < 0.001	-0.702
GFAP (pg/ml)	111.8 (57.9)	224.7 (118)	* < 0.001	-0.686
NfL (pg/ml)	17.1 (11)	29.7 (19)	* < 0.001	-0.582
Aβ42/40 ratio	0.067 (0.01)	0.059 (0.01)	* < 0.001	0.529
Aβ42 (pg/ml)	7.1 (1.5)	6.6 (1.5)	0.059	0.221
Aβ40 (pg/ml)	108.4 (19.9)	115.2 (25.7)	0.278	-0.127

ACE	97.4 (2.0)	63.5 (20.0)	* < 0.001	0.934
OMT - Identification time	2.61 (1.3)	4.25 (2.0)	* < 0.001	-0.578
OMT - Localization time	5.5 (1.9)	8.5 (3.2)	* < 0.001	-0.600
OMT - Identification accuracy	0.92 (0.0)	0.81 (0.1)	* < 0.001	0.686
OMT - Absolute Error	137.5 (56.8)	242.2 (85.6)	* < 0.001	-0.658
OMT - Target detection	0.73 (0.1)	0.52 (0.2)	* < 0.001	0.680
OMT - Guessing	0.15 (0.1)	0.29 (0.1)	* < 0.001	-0.651
OMT - Misbinding	0.24 (0.1)	0.37 (0.1)	* < 0.001	-0.686
OMT - Imprecision	84.4 (45.7)	144.1 (71.7)	* < 0.001	-0.525
OIS - Immediate Object Accuracy	95.7 (4.9)	75.2 (16.5)	* < 0.001	0.813
OIS - Delayed Object Accuracy	68.6 (18.0)	31.4 (23.0)	* < 0.001	0.763
OIS - Immediate Location Error	0.95 (0.3)	3.0 (1.8)	* < 0.001	-0.766
OIS - Delayed Location Error	1.9 (1.0)	5.1 (2.8)	* < 0.001	-0.653
OIS - Immediate Semantic Accuracy	99.7 (0.9)	91.2 (10.7)	* < 0.001	0.621
OIS - Delayed Semantic Accuracy	79.8 (15.0)	41.0 (27.2)	* < 0.001	0.727
ROCF - copy	90.0 (5.6)	73.8 (24.9)	* < 0.001	0.644
ROCF - recall	84.8 (8.7)	49.4 (21.2)	* < 0.001	0.897
DSST - correct responses	34.3 (9.9)	14.7 (7.2)	* < 0.001	0.857
DSST - reaction times	3.6 (2.1)	11.6 (12.7)	* < 0.001	-0.828
TMT	34.5 (9.9)	99.4 (87.1)	* < 0.001	-0.839
CORSI	10.9 (3.7)	19.6 (6.6)	* < 0.001	-0.759

Table 6.1 | Demographics, plasma biomarkers and digital cognitive metrics

After correcting for multiple comparisons across all digital tests, in the OMT task, Identification time and Localization time were correlated with all four biomarkers (**Figure 6.5**). Identification accuracy was negatively correlated with pTau181, GFAP and NfL levels. Absolute Error, Target detection and Guessing were associated with pTau181 levels, while

both Guessing and Misbinding were correlated with GFAP levels (**Figure 6.5**). Imprecision showed an association with NfL levels.

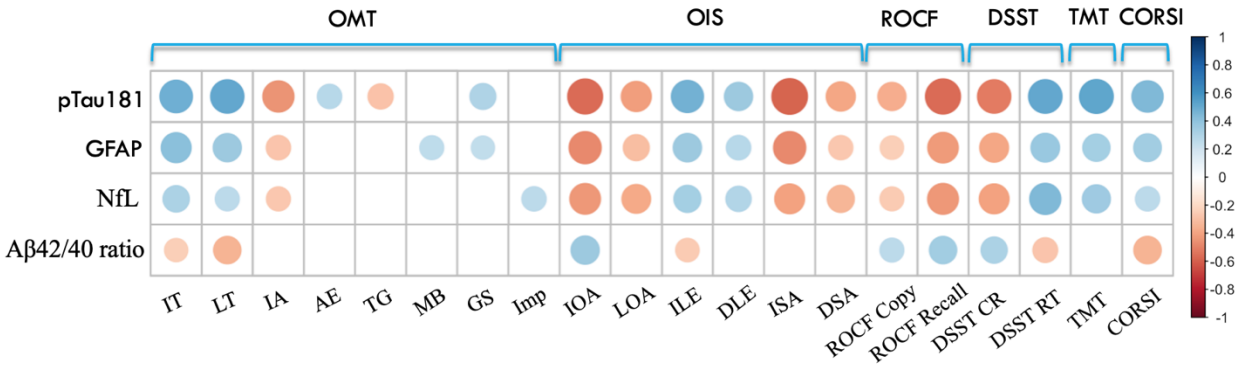


Figure 6.5 | Correlations between digital cognitive tests and plasma biomarkers

Strength of the correlation is given by diameter of the circle, with positive correlations in blue and negative in red. All displayed correlations are significant after Benjamini–Hochberg correction for multiple comparisons. OMT: IT = Identification time, LT = Localization time, IA = Identification accuracy, AE = Absolute Error, TG = Target Detection, MB = Misbinding, GS = Guessing, Imp = Imprecision. OIS: IOA = Immediate Object Accuracy, DOA = Delayed Object Accuracy, ILE = Immediate Location Error, DLE = Delayed Location Error, ISA = Immediate Semantic Accuracy, DSA = Delayed Semantic Accuracy. ROCF = Rey–Osterrieth Complex Figure, DSST = Digit Symbol Substitution Task, CR = correct response, RT = reaction times, TMT = Trail Making Test, CORSI = Freestyle Corsi Block Task.

In the OIS task, all metrics were correlated with pTau181, GFAP and NfL (**Figure 6.5**). Immediate Object Accuracy and Immediate Location Error were also significantly associated with the Aβ42/40 ratio.

Both the copy and the recall of the ROCF and the number of correct responses and the reaction times of the DSST were correlated with levels of all four biomarkers (**Figure 6.5**). TMT scores were correlated with pTau181, GFAP, NfL but there was no association with the Aβ42/40 ratio. Performance at the CORSI showed a significant association with all four biomarkers.

Subsequently, the relationship between all plasma biomarker and digital cognitive metrics was investigated (**Figure 6.6**)**.** Only one metric was selected per each test, according to the

highest effect size in discriminating between groups (**Table 6.1**): Identification accuracy for OMT, Immediate Object Accuracy for OIS, the recall of the ROCF, the number of correct responses of the DSST, and the average times for the TMT and the CORSI. In this network plot it can be appreciated that pTau181 was the plasma biomarker that most strongly correlated with our digital cognitive metrics, which all clustered on the right of the plot. Amongst the digital cognitive metrics, OIS and the ROCF were the closest to the plasma biomarkers, particularly pTau181. OMT and TMT were the tests that showed the weakest correlations with plasma biomarkers (farthest on the right).

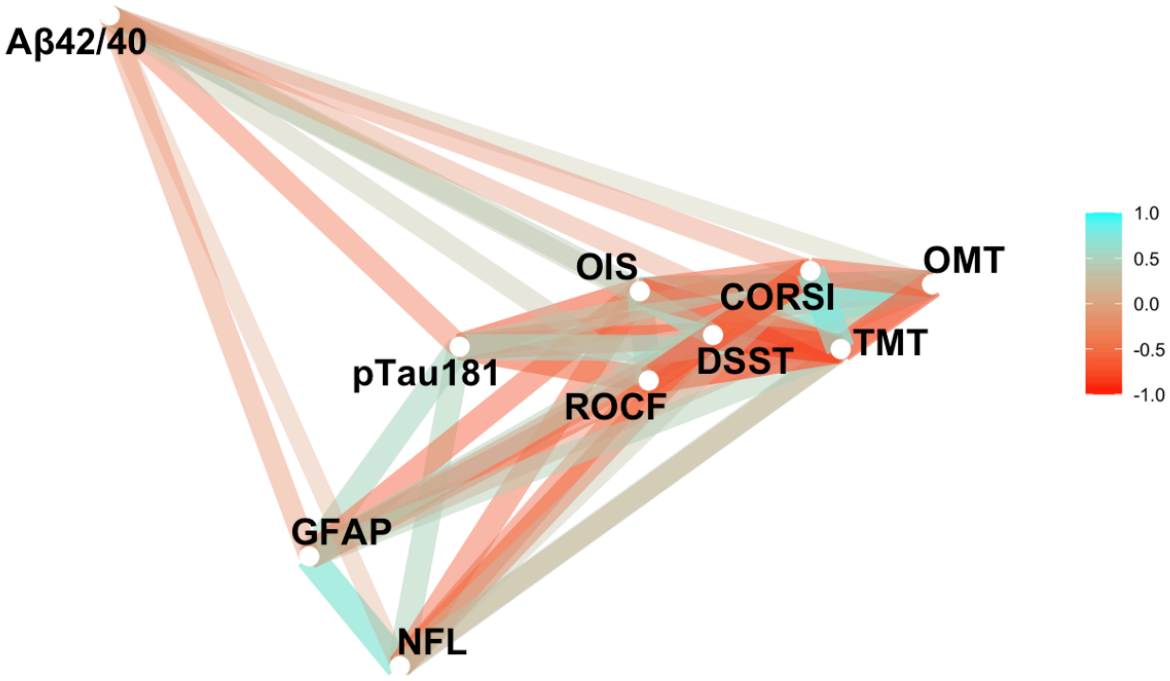


Figure 6.6 | Correlations between biomarkers and digital cognitive metrics

OMT = Identification accuracy of the Oxford Memory Task, OIS = Immediate Object Accuracy of the Object-in-Scene Memory Task, ROCF = recall of the Rey–Osterrieth Complex Figure, DSST = number of correct responses of the Digit Symbol Substitution Task, TMT = average reaction time of the Trail Making Test, CORSI = average reaction time of the Freestyle Corsi Block Task. Strength of the correlation is represented by the distance between the items and by graded color coding, where red represents a negative correlation and light blue a positive correlation.

The selected variables were then ranked according to their importance in predicting group classification, i.e., AD or EHC (**Figure 6.7**). The rank represents the negative log of the p-

values. In this sample, the recall of the ROCF and the Immediate Object Accuracy of the OIS had a higher rank compared to pTau181, and all cognitive metrics ranked higher than all the other plasma biomarkers. All tests and biomarkers were significant predictors of group (all *p < 0.001 except Aβ42/40 ratio *p= 0.002).

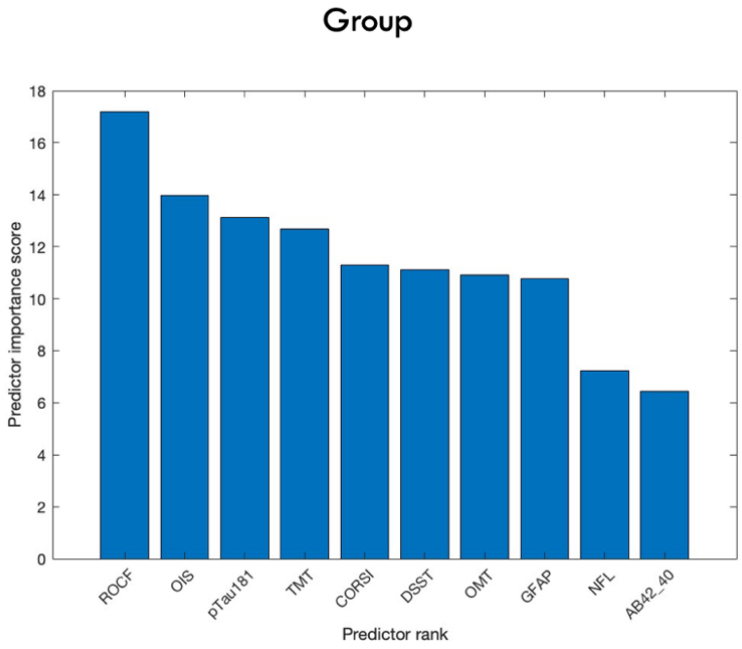


Figure 6.7 | Ranked biomarkers and digital cognitive metrics in predicting group

OMT = Identification accuracy of the Oxford Memory Task, OIS = Immediate Object Accuracy of the Object-in-Scene Memory Task, ROCF = recall of the Rey–Osterrieth Complex Figure, DSST = number of correct responses of the Digit Symbol Substitution Task, TMT = average reaction time of the Trail Making Test, CORSI = average reaction time of the Freestyle Corsi Block Task.

We then explored which were the best predictors of amyloid and tau burden. Most of our tests, except the CORSI, were better predictors of pTau181 levels compared to the other plasma biomarkers (**Figure 6.8**). The best performing tests were the ROCF and TMT (*p < 0.001), followed by OIS (*p = 0.003), OMT (*p = 0.004), DSST (*p = 0.004), NfL (*p = 0.005), Aβ42/40 (*p = 0.019), while GFAP (p = 0.056) and CORSI (p = 0.125) were not statistically significant predictors of tau burden.

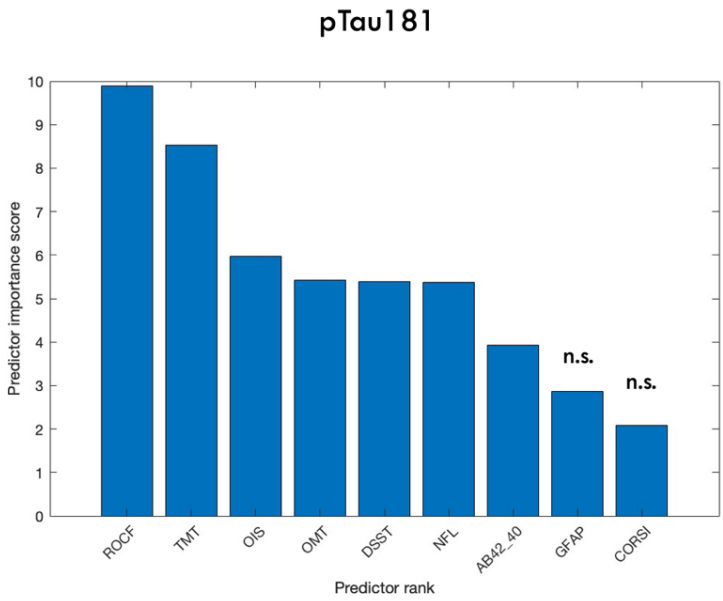


Figure 6.8 | Ranked biomarkers and digital cognitive metrics in predicting pTau181 levels

OMT = Identification accuracy of the Oxford Memory Task, OIS = Immediate Object Accuracy of the Object-in-Scene Memory Task, ROCF = recall of the Rey–Osterrieth Complex Figure, DSST = number of correct responses of the Digit Symbol Substitution Task, TMT = average reaction time of the Trail Making Test, CORSI = average reaction time of the Freestyle Corsi Block Task. n.s. = not significant.

Conversely, pTau181 was the only statistically significant predictor of amyloid burden, as measured by the Aβ42/40 (*p = 0.022), (**Figure 6.9**). The best performing digital cognitive test to predict amyloid burden was the OIS task but was not statistically significant.

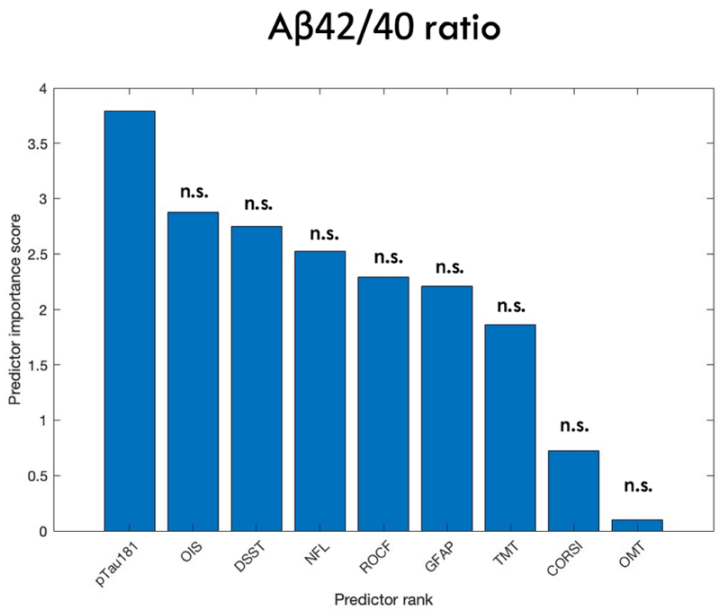


Figure 6.9 | Ranked biomarkers and digital cognitive metrics in predicting Aβ42/40 ratio

OMT = Identification accuracy of the Oxford Memory Task, OIS = Immediate Object Accuracy of the Object-in-Scene Memory Task, ROCF = recall of the Rey–Osterrieth Complex Figure, DSST = number of correct responses of the Digit Symbol Substitution Task, TMT = average reaction time of the Trail Making Test, CORSI = average reaction time of the Freestyle Corsi Block Task. n.s. = not significant.

Model comparisons using the MuMIn R package was then used to choose the best combination of plasma biomarkers and cognitive metrics in predicting group classification, pTau181 levels and the Aβ42/40 ratio. The best model for predicting group consisted of pTau181, recall of the ROCF and Immediate Object Accuracy of the OIS (**Figures 6.10** and **6.11**), with an AUC of 1. In comparison, the recall of the ROCF alone had an AUC of 0.946, while pTau181 had an AUC of 0.911, which however were not statistically significantly different from the best model (ROCF: Z = 1.6056, p-value = 0.1084, pTau181: Z = 1.8957, p-value = 0.058), nor between them (Z = 0.60346, p-value = 0.5462) as shown by the largely overlapping confidence intervals (**Figure 6.11**). When Lasso penalization was introduced to avoid perfect separation of the best model, it still performed significantly well (AUC = 0.92).

The best model for predicting pTau181 levels consisted of Aβ42/40, NfL, OIS and ROCF, which had an adjusted R^2 of 0.44 (**Figure 6.10**). The winning model for predicting Aβ42/40 levels consisted of pTau181, CORSI, OIS, OMT and TMT, but had an overall poor model fit with an adjusted R^2 of 0.24 (**Figure 6.10**).

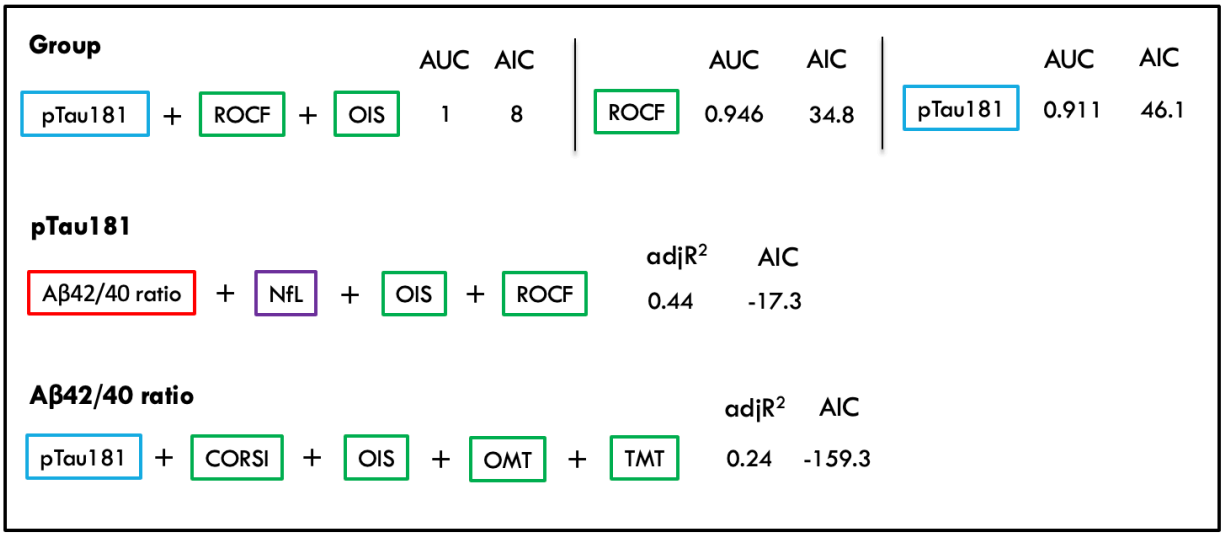


Figure 6.10 | Best combinations of biomarkers and digital cognitive metrics in predicting Aβ42/40 ratio

AUC = area under the curve. AIC = Akaike information criterion.

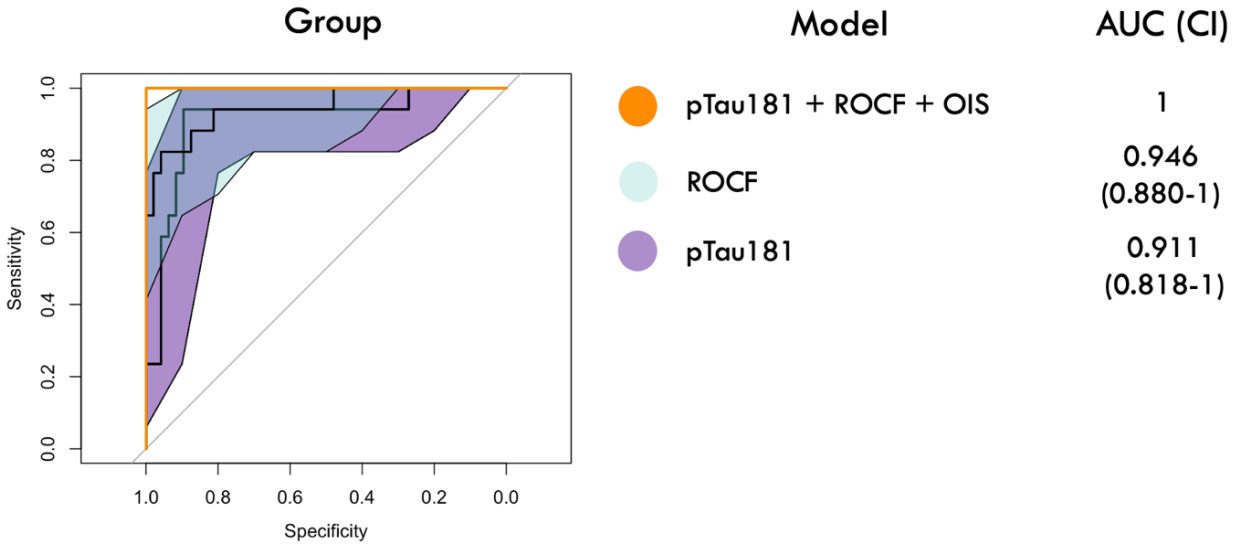


Figure 6.11 | ROC curves for Group classification

AUC = area under the curve, CI = confidence interval in light blue (ROCF alone) and violet (pTau181 alone). Orange: combined model with pTau181, ROCF and OIS.

6.1.4 Discussion

The findings reported here suggest that pTau181, GFAP and NfL were highly correlated with several digital cognitive metrics. pTau181 was the plasma biomarker which showed the closest association with the digital cognitive metrics examined (**Figure 6.6**), while the Aβ42/40 ratio was the biomarker which showed the weakest association with cognition. As single metrics, the recall of the ROCF and Immediate Object Accuracy of the OIS were better predictors of group classification, i.e., AD or EHC, compared to all plasma biomarkers (**Figure 6.7**), and TMT, CORSI, number of correct responses of the DSST and Identification accuracy of the OMT predicted group better than all biomarkers except pTau181. If used in combination, the best performing model to predict group was the one combining pTau181, recall of the ROCF and Immediate Object Accuracy of the OIS, with an AUC of 1, achieving perfect separation (**Figures 6.10 - 6.11**). However, it was not statistically significantly better than the model containing ROCF or pTau181 alone.

The very high AUCs of these metrics in predicting group in this small, highly selected sample might partially explain the lack of positive contribution in adding cognition to plasma biomarkers, accuracy being at ceiling, and further evidence is needed to see whether combining pTau181 to digital metrics might be beneficial in a larger dataset and in different clinical populations such as patients with SCI or MCI. Data presented in Chapter 3 support the positive contribution of digital cognitive metrics in early phases of the disease, when cognitive impairment is below clinical threshold of significance. Therefore, it needs to be tested whether the combination of digital cognitive metrics and plasma biomarkers can be useful to stratify which patients in the preclinical or prodromal phase of AD might be at risk of developing AD dementia.

Independent measures of ATN status (such as amyloid or tau measured by PET or CSF) and longitudinal data of progression rates to dementia in healthy controls are not available for our dataset, so it is unknown at present whether these digital metrics are good predictors of more validated measures of amyloid and tau pathology, or of clinical deterioration in healthy participants. A bigger sample would be needed also to assess the performance of these metrics in patients with preclinical AD versus amyloid and tau negative healthy controls. Nevertheless, it is worth noticing that the recall of the ROCF was as good as pTau181 in predicting group classification in this sample. Additionally, all our digital cognitive metrics, and particularly ROCF, OIS and TMT, were good predictors of pTau181 levels (**Figure 6.8**), while performed poorly in predicting amyloid status (**Figure 6.9**). This is in line with previous evidence that show that tangle deposition is more tightly linked to cognitive impairment compared to amyloid burden [420, 421], and extends these findings to the digital cognitive metrics examined.

The patient population included in this Chapter was already at the AD dementia stage, where cognitive impairment is overt; therefore, these digital metrics could be overperforming compared to patients in the preclinical and prodromal phase of the disease. That is important also in the context not of whether cognitive biomarkers are more important than plasma biomarkers, but also of *which* biomarker in each class might be more informative. In this sample, plasma pTau181 was the biomarker which was more closely associated with cognitive metric and the best predictor of group classification. However, it might be that other biomarkers such as pTau217 could show even higher association with cross-sectional or longitudinal cognitive function in the same population, as some evidence suggest [448, 449]. Moreover, we cannot exclude that early markers of amyloid deposition such as pTau217,

pTau213 and GFAP might be more tightly associated to cognitive changes in early phases of the disease [406, 408 , 448, 449].

In our sample, the recall of the ROCF and the Immediate Object Accuracy of the OIS seemed promising metrics as correlated with pTau181, GFAP, NfL and Aβ42/40 ratio levels, and good predictors of group classification and pTau181 levels (**Figures 6.5-6.8, 6.10-6.11**), but as different plasma biomarkers might be useful either in early or later phases of the disease, we cannot exclude that other digital metrics might be more informative in earlier stages of the diseases. Multiple comparisons correction across multiple digital and plasma biomarkers is also an important factor to be considered, as it is possible that using a restricted panel of digital biomarkers might be more useful to detect specific pathologic changes of interest, targeting specific research questions in different populations.

6.1.5 Conclusion

Our digital cognitive metrics were tightly associated with several plasma biomarkers, particularly pTau181, but also with GFAP and NfL, and to a much lesser extent with the Aβ42/40 ratio. Adding these metrics to pTau181 did not improve group classification in this sample, but the best performing metric, the recall of ROCF, performed at par with pTau181 levels.

6.2 Test-retest variability of Plasma Biomarkers and digital VSTM metrics in AD and EHC §#

6.2.1 Introduction

One of the keys properties of a biomarker is to be able to monitor pharmacologic responses to a therapeutic intervention. However, for effective large-scale studies in the future, it is essential first to determine variability in readouts with repeated testing.

Digital cognitive metrics provide more fine-grain measures of cognition than traditional neuropsychological testing and can be performed remotely and frequently. As the test sets are randomized each time, they are less susceptible to practice effects compared to traditional tests of cognition and are therefore a potential candidate for clinical trials performed at a shorter time scale. Data presented in Chapter 3 show that all digital VSTM from the "What was where?" Oxford Memory Task show good test-retest reliability over a 1-year period, as did not change significantly over time in healthy controls. In the AD cohort, we have also shown that some metrics (Absolute Error and memory Imprecision) might be able to track disease progression over time, while other metrics seem less likely to do so. However, the number of participants being involved in our longitudinal study was very small (< 22 per group), due to the COVID-19 outbreak. Secondly, that dataset was acquired using our tablet-based task, and it is unknown whether the tablet version shares the same properties of the online version of the test in tracking longitudinal progression. Thirdly, while over a one-year course patients with AD showed a clear decline in ACE scores, and therefore

worsening cognitive impairment, is still unclear how much is the test-retest variability of these cognitive metrics, i.e., the amount of noise of the measurements at a shorter time frame.

Plasma biomarkers also share the potential of being administered frequently, as repeated measuring over time might be more feasible compared to repeating CSF sampling or PET imaging. While data on longitudinal evaluations over yearly visits are far more frequent [430, 450], data at shorter time frames are surprisingly scarce. Cullen et al measured the test-retest variability of plasma Aβ42/40 ratio, NfL, GFAP, and pTau217 after 6 - 10 weeks in a cohort of 38 participants, matched by amyloid positive status to a bigger cohort of 399 individuals from the Biomarkers For Identifying Neurodegenerative Disorders Early and Reliably (BioFINDER)-1 study [190]. The focus of the study was to compute an estimate of random noise in the measurements of these plasma biomarkers and quantify how much this random noise impacted the future prediction of abnormal Aβ42/40 ratio in the CSF and clinical conversion to AD dementia in the bigger sample of 399 participants.

The Aβ42/40 ratio had the lowest variability (4.1%), followed by 20% for pTau217, 23.7% for plasma NfL, while GFAP had the highest variability (25%). In the baseline condition, without random noise added, the combination of all biomarkers performed best in group classification (A– vs A+) and AD dementia conversion, followed by pTau217, Aβ42/40 ratio, GFAP and lastly NfL. When the random noise was added to the model, pTau217 was the least affected biomarker in predicting group classification based on Aβ status (difference in AUC, or ΔAUC = -0.98%), followed by the combination of all biomarkers (ΔAUC = -1.2%), NfL (ΔAUC = -1.7%), the Aβ42/40 ratio (ΔAUC = -2.5%), and finally GFAP (ΔAUC = -3.7%).

If prediction of clinical conversion to AD was chosen as outcome, the combination of biomarkers was the least impacted (ΔAUC = −0.88%), followed by pTau217 (ΔAUC = −1.29%), Aβ42/40 (ΔAUC = −2.29%), NfL (ΔAUC = −1.64%) and GFAP (ΔAUC = −3.43%). The percentage of individuals with highly variable predicted outcomes (defined as having > 5% chance of a different outcome) was the lowest (14.3%) if all biomarkers were combined, followed by pTau217 and Aβ42/40 (20.3%), NfL (30.1%) and GFAP (30.6%). In the same study, CSF biomarkers generally had less test-retest variability and performance of CSF biomarkers decreased less when random noise was introduced in the model.

Random error is usually caused by a combination of *biological* variation (i.e., intra-individual variability in the biomarker levels in plasma over time), *pre-analytical* variation (i.e., differences in sample collection and pre-processing steps), and *analytical* variation (i.e., intra- and inter-assay variability).

My aims were to measure the test-retest reliability of plasma biomarkers and digital VSTM cognitive metrics over a short time period (3 monthly over a 6-month period) in a cohort of healthy controls and patients with AD. As practice effects on cognitive tests are usually higher in healthy controls compared to patients with AD, test-retest reliability was measured also in patients, to quantify the noise present specifically in the AD population at repeated cognitive testing, which will be essential for better clinical endpoints in future trials [451]. Our preprocessing protocol was not changed throughout the timeframe of the study, and the same assays were used, to focus our study on estimating true biological variation. At the time of writing, plasma biomarkers levels for the longitudinal study at 6 months are available for 36 EHC and 25 AD patients, whose characteristics did not significantly differ compared to the

cross-sectional analysis (see **Table 6.1S** in Appendix for summary of demographics, plasma biomarkers and digital cognitive metrics).

6.2.2 Methods

The assessments done at each visit are the same described in Chapter 6.1 (digital cognitive tests, plasma biomarkers, ACE). ACE versions A, B and C were used for 1st, 2nd and 3rd visits to minimize learning effects, which is standard practice in studies conducted at shorter timeframes.

Statistical analysis

Only datapoints from participants who completed the blood sample for all three visits were analyzed. The Coefficient of Variation was calculated as follows:

$$\text{Coefficient of Variation (CV\%)} = (\text{Standard Deviation} / \text{Mean}) \times 100$$

We computed the CV% of our plasma biomarkers and digital cognitive metrics per each participant, across the three visits. We then averaged the CV% per metric across the whole sample, and then separately for healthy controls and AD patients (**Table 6.2**, **Figure 6.12**). Group difference in CV% were calculated with t-test or Mann Whitney U test, according to the normality of the data. A Friedman Test with Conover's post hoc comparisons was used to compare differences across tests biomarkers and within each set of metrics belonging to the same digital test.

6.2.3 Results

The test-retest analysis showed that pTau181 had a high CV% of around 35.8% (**Table 6.2**, **Figure 6.12**). GFAP (18.4%) and NfL (15.8%) had good reliability, in line with standard laboratory procedures (< 20%) [452], and the Aβ42/40 ratio had the lowest test-retest variability (9.5%). The CV% were not statistically significantly different between healthy controls and AD patients for any biomarker. CV% were significantly higher for pTau181 (*p < 0.001) and significantly lower for the Aβ42/40 ratio (*p < 0.001) compared to all other biomarkers. There was no difference in CV% between GFAP and NfL (p = 0.364).

Amongst OMT metrics, Identification accuracy had the lowest CV% (3.3%), which however unlike ACE did not vary between EHC and AD patients (**Table 6.2**, **Figure 6.13**). Identification time, Localization time, memory Imprecision and Absolute Error had moderate CV%, with Absolute Error showing high CV% particularly in EHC. Target detection had a low CV% (9.4%), while Misbinding and Guessing showed higher test-retest variability (> 20%). At direct comparison, there was a statistically significant difference across metrics in CV%. Notably, Identification accuracy had lower CV% compared to all the other digital metrics (*p < 0.001 for all, except for Target detection *p = 0.002). Target detection had lower CV% compared to Guessing (*p < 0.001), Misbinding (*p < 0.001), Absolute Error (*p = 0.019), and Imprecision (*p = 0.007).

	CV% EHC	CV% AD	CV% all	p-value
pTau181 (pg/ml)	40.7	28.6	35.8	0.059
GFAP (pg/ml)	18.7	18.0	18.4	0.867
NfL (pg/ml)	15.7	16.0	15.8	0.536
Aβ42/40 ratio	10.1	8.5	9.5	0.150
ACE	1.6	5.9	3.4	* < 0.001
OMT - Identification time	14.1	11.0	13.1	0.653
OMT - Localization time	16.2	13.1	15.2	0.272
OMT - Identification Accuracy	3.3	3.4	3.3	0.348
OMT - Absolute Error	20.0	9.4	16.6	* 0.004
OMT - Target detection	7.7	13.0	9.4	0.560
OMT - Guessing	26.9	25.2	26.4	0.325
OMT - Misbinding	28.4	17.7	25.0	* 0.012
OMT - Imprecision	17.0	11.3	15.2	0.062
OIS - Immediate Object Accuracy	3.6	15.9	7.9	0.101
OIS - Delayed Object Accuracy	19.7	32.9	24.1	0.413
OIS - Immediate Location Error	16.8	30.5	21.5	0.440
OIS - Delayed Location Error	33.2	13.7	26.4	* < 0.001
OIS - Immediate Semantic Accuracy	0.6	7.1	2.8	* < 0.001
OIS - Delayed Semantic Accuracy	12.6	28.0	18.0	0.394
ROCF - copy	3.8	13.8	7.1	0.100
ROCF - recall	8.4	20.8	12.7	0.601
DSST - correct responses	17.8	14.7	17.0	0.416
DSST - reaction times	16.0	13.4	15.3	0.197
TMT	6.8	11.8	8.4	0.591
CORSI	16.7	13.6	15.7	0.487

Table 6.2 | Test-retest reliability of plasma biomarkers and cognitive metrics

For the OIS task, Immediate Object Accuracy had a low CV% (7.9%), which did not differ between EHC and AD patients (**Table 6.2**, **Figure 6.13**). Conversely, Immediate Semantic Accuracy had a remarkable test-retest reliability, which was significantly lower in EHC (0.6%) vs AD patients (7.1%). Immediate Semantic Accuracy was significantly more reliable compared to all other metrics (*p < 0.001, except with Immediate Object Accuracy *p = 0.016). Immediate Object Accuracy had lower CV% compared to all other remaining metrics (*p < 0.001). Delayed Semantic Accuracy was better than Delayed Object Accuracy (*p = 0.036) and Delayed Location Error (*p = 0.017). The other metrics showed moderate to high CV%.

ACE scores varied very little between assessments (3.4%), with a CV% that was lower in healthy controls (1.6%) compared to AD patients (5.9%), (**Table 6.2**, **Figure 6.13**). ROCF, DSST, TMT and CORSI all showed good CV% (< 20%) in healthy controls and AD patients alike. ROCF copy had lower CV% compared to ROCF recall (*p < 0.001).

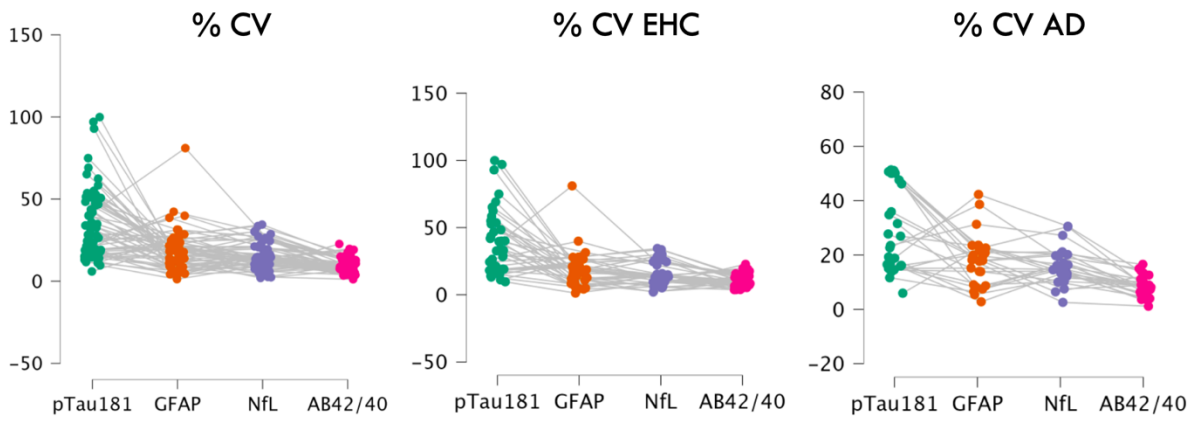


Figure 6.12 | CV% of plasma biomarkers in the whole group, and in EHC and AD alone

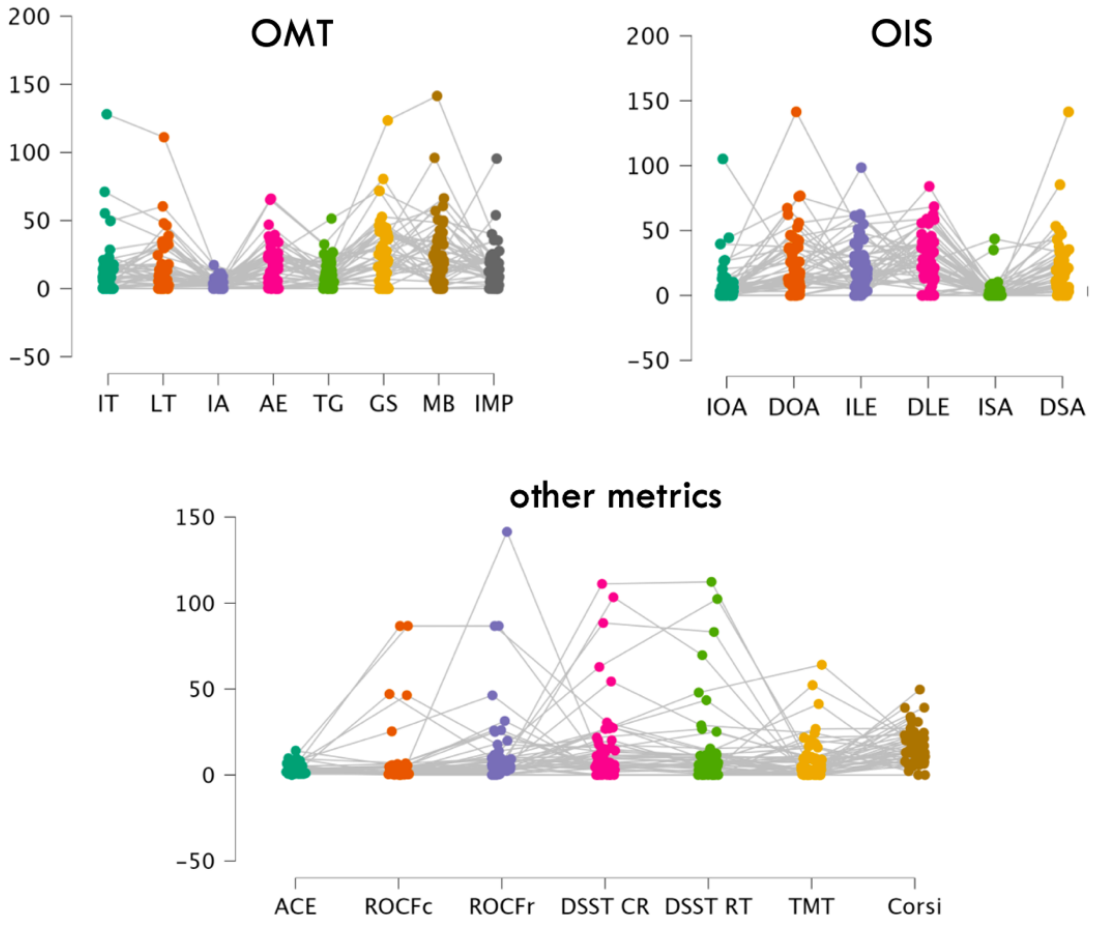


Figure 6.13 | CV% of cognitive metrics in the whole group

OMT: IT = Identification time, LT = Localization time, IA =Identification accuracy, AE = Absolute Error, TG = Target detection, MB = Misbinding, GS = Guessing, IMP = Imprecision. OIS: IOA = Immediate Object Accuracy, DOA = Delayed Object Accuracy, ILE = Immediate Location Error, DLE = Delayed Location Error, ISA = Immediate Semantic Accuracy, DSA = Delayed Semantic Accuracy. ROCFc = Rey–Osterrieth Complex Figure copy, ROCFf = Rey–Osterrieth Complex Figure recall, DSST CR= Digit Symbol Substitution Task correct responses, DSST CR= Digit Symbol Substitution Task reaction times, TMT = Trail Making Test, CORSI = Freestyle Corsi Block Task.

6.2.4 Discussion

The data presented in this chapter suggest that across plasma biomarkers and digital cognitive metrics there is great heterogeneity in test-retest reliability. Amongst plasma biomarkers, pTau181 showed the highest variability (35.8%), which is higher than the recommended threshold for reliability (< 20%) [452]. GFAP and NfL showed intermediate

levels of CV%, while the Aβ42/40 ratio was the most robust plasma biomarker amongst the ones examined (9.5%). In line with the study by Cullen et al [190], the Aβ42/40 ratio in our dataset was the biomarker with the lowest test-retest variability (< 10%). In the abovementioned study, NfL and GFAP showed high CV% (23.7-25%), higher than recommended (< 20%), and worse compared to our dataset (15.8-18.4%).

There are several methodological differences compared to the study by Cullen et al [190]. Firstly, different pre-processing procedures were used, such as centrifugation temperature (+ 4°C in Cullen et al, vs room temperature in our data), aliquots volume (1.5 mL vs 0.5 mL), and time to -80°C storage (30-60 minutes versus two steps + 4°C fridge and then -80°C freezer in our case). However, current standard operating procedures for pre-processing of plasma biomarkers show that amyloid species', NfL, pTau181 and GFAP levels are not affected when the samples are left in the fridge prior to centrifugation for up to 24 h, nor by aliquot volumes, or centrifugation temperature [453]. Therefore, we think these differences are unlikely to explain differences in CV%. Moreover, the timeframe used in that study was 6-10 weeks, whilst we averaged the CV% across three measurements (0-3-6 months). While direct comparison between our timeframe and the one used in the article is not applicable, they are at a relatively short scale in the context of AD pathophysiological changes. Future studies would need to unravel the optimal temporal cut-off for stability of these biomarkers over time in healthy individuals and patient populations.

Measurement variations have been found to differentially affect the predicting accuracy and model fit of combination of plasma biomarkers depending on the assay used and the population studied [191]. Benedet et al investigated whether different variations of CV% impacted the ability of plasma Aβ42/40, NfL, GFAP and pTau181 in predicting amyloid

PET positivity in a cohort of cognitively unimpaired and cognitively impaired individuals [191]. They found that plasma Aβ42/40 as determined by immunoprecipitation mass spectrometry (IP-MS) was the best predictor of Aβ positivity at PET, followed by pTau181 and GFAP, and outperformed other methods such as Simoa Aβ42/40 measurements. However, even small (5%) random simulated variations on the CV% caused the accuracy of IP-MS Aβ42/40 to drop below to that of GFAP and pTau181. In contrast, GFAP and pTau181 performances remained stable even at a 20% CV. When biomarkers were evaluated in combination, IP-MS Aβ42/40 was the most significant contributor in predicting Aβ-positivity in cognitively unimpaired individuals, and adding pTau181, GFAP, or NfL did not significantly improve performance. However, in cognitively impaired individuals a model combing IP-MS Aβ42/40, GFAP, and pTau181 was found to be the best indicator of Aβ-positivity, showing high accuracy (93.5%), and models that included IP-MS Aβ42/40 significantly outperformed models containing Simoa Aβ42/40 as an alternative. With small variations in biomarker measurements, all selected models were not impacted, but for cognitively unimpaired participants, IP-MS measurements were not able to withstand a larger variation (CV > 10%), being subsequently replaced by GFAP. Whilst performance of Simoa Aβ42/40 was inferior to IP-MS Aβ42/40 in this study, Simoa Aβ42/40 was less impacted by differences in CV%.

Test-retest reliability of OMT and OIS have not been tested before. In this dataset, amongst all OMT metrics, Identification accuracy and Target detection showed excellent test-retest reliability (3.3 and 9.4 CV%), similarly for EHC and AD patients (**Figure 6.13**, **Table 6.2**). For OIS, Immediate Object Accuracy and Immediate Semantic Accuracy showed low CV% (7.9 and 2.8 CV%), with Immediate Semantic Accuracy being extremely low (0.6 CV%) in EHC and higher (7.1 CV%) in AD patients compared to EHC.

The ACE-III has been shown to have excellent internal reliability, measured by Cronbach's α coefficient (alpha = 0.88) [59], which is considered very good [454], but its test-retest reliability at short timeframes has not been reported. In our study, ACE had very good (3.4%) CV, particularly in healthy controls (**Figure 6.13**, **Table 6.2**). One report on the original version of the Corsi Block Tapping Task found good reliability (a correlation of r = 0.75) in patients with cognitive impairment [455], but as the Freestyle Corsi Block Task has been newly developed, this is the first time the test-retest reliability is assessed, which was acceptable across groups (15.7%). Previous data in healthy controls on test-retest reliability after 1 day of the paper versions of the TMT and DSST showed high reliability for the DSST (13%), while the TMT obtained intermediate values (22-33%, depending on the test version) [456]. Another study examined test-retest variability after 4-24 months in a cohort of healthy male subjects of several commonly used tests, including TMT, DSST and ROCF, and found that all tests, except the copy of the ROCF showed significant practice effects [457]. Our data replicated the lower CV% in the copy of the ROCF (7.1%) compared to the recall (12.7%) in this digitized version of the ROCF, while in our dataset also the TMT showed low CV (8.4%), probably due to randomized versions of the trail making maps (**Figure 6.13**, **Table 6.2**).

6.2.5 Conclusion

This test-retest study revealed considerable heterogeneity in CV% across plasma biomarkers and cognitive metrics. Amongst the plasma biomarkers examined, the Aβ42/40 ratio was the most robust metric, while pTau181 had the lowest reliability. Identification accuracy and Target detection of OMT and Immediate Object Accuracy and Immediate Semantic accuracy of OIS also showed low CV% and might be the most robust metrics at shorter timeframes.

Chapter 7. GENERAL DISCUSSION

7.1 Overview of findings and contribution to existing literature

Most studies on VSTM apply manipulations of experimental paradigms to young individuals to test differences across conditions of interest. However, cognitive functions change across the lifespan [458], and in healthy ageing different strategies might be adopted compared to younger individuals to achieve the same goal. VSTM performance is further impaired in several neurological conditions such as AD [69, 72] and PD [199]. One of the aims of this DPhil was to examine not only if, but *how* physiological aging and neurodegenerative processes impact VSTM performance.

Chapter 2 focused on how the ability to filter out distractors changes in healthy ageing, patients with AD and PD. By using a continuous delayed reproduction task, it also aimed to dissect different sources of memory errors, i.e., whether there was a lower rate of targets detected, more misbinding errors, higher levels of guessing more or lower memory precision, and on how distractor filtering affected these memory properties. It was also tested if filtering would be differentially impacted depending on when the distractor was presented, i.e., at encoding or during maintenance, and whether distinct memory metrics derived from the mixture model could be affected by varying the presentation phase of the distractor.

Results showed that across all groups participants performed worse when more items had to be remembered compared to when an item could be filtered out. Filtering out an unwanted item improved the rate of targets detected, and reduced misbinding errors, while having no effect on memory precision.

Healthy ageing was associated with lower memory precision across all conditions, suggesting that memory precision is particularly impacted by ageing. Reduced target detection, increased misbinding and increased guessing were also observed. Healthy ageing was also associated with worse performance at remembering multiple items if they were all presented simultaneously, i.e., at encoding.

AD patients were particularly impacted by a distractor appearing during maintenance and performed worse if stimuli were presented sequentially during encoding and maintenance. Similarly to healthy ageing, patients with AD showed a decrease of memory precision and increased misbinding, but in this case a reduced target detection and increased guessing were the main sources of memory errors. Conversely, a diagnosis of PD was associated with generally good filtering abilities, and with only a higher probability of random guessing. Hippocampal volume was negatively correlated with filtering ability during maintenance, but not at encoding, as well as with all mixture model metrics across all groups.

This experimental work extends the emerging knowledge on the decline of memory precision in ageing [41] by evaluating the contribution of the degradation of the precision of memory recall in the context of distractor filtering. This data support the idea that the decline in performance seen in healthy ageing might be countered by distractor filtering, which frees up more VSTM resources, in terms of higher targets detected and lower misbinding rates, but is not able to enhance the precision of memory recall.

If studies of VSTM in elderly subjects are less prevalent compared to young adults, data on patients with AD are even more scarce, particularly using delayed reproduction tasks, and no study so far in this population studied the impact of filtering out a distractor in different presentation stages of the stimuli (i.e., encoding and maintenance). This also helped to

delineate performance metrics across each diagnostic group (healthy ageing, AD and PD), and how these are affected by task demands and the presence of distractors. This is the first time such an extensive paradigm and modelling of responses have been deployed in four different diagnostic groups to measure the effects of distractor filtering. The correlation of filtering at maintenance with hippocampal volumes also reinforces the idea that the decline in hippocampal integrity observed with health ageing, and more so in patients with AD, might contribute to failure in distractor filtering specifically during the STM retention phase.

Whilst Chapter 2 was instrumental in demonstrating how reduced performance in AD patients is multifactorial, as it is the product of the degradation of different distinct memory processes, these patients were already in the AD dementia stage of the disease. In **Chapter 3** we tested a large sample of over 300 individuals using the digital, tablet-based "What was where" Oxford Memory Task (OMT). Performance of patients with AD dementia was compared to that of individuals at risk of developing AD, such as individuals with SCI and MCI, and a comparison between digital and standard neuropsychological tests of cognition was performed.

Across groups, increasing set size and delay had an impact on memory recall, showing an effect on all metrics, except misbinding which showed no effect of delay. Most digital metrics were able to discriminate between healthy controls and patients with MCI, except for the reaction times metrics, which were significantly altered only in AD dementia patients. Mixture model metrics and spatial localization error also differed between patients with SCI and MCI. All digital metrics except for spatial localization error were also able to discriminate between patients with MCI and AD dementia. ACE was overall superior to the digital metrics examined in the discrimination between patients with MCI and those with AD dementia but performed similarly when comparing healthy controls to patients with SCI

or patients with SCI to MCI. Identification accuracy was the best predictor of hippocampal atrophy, performing as well as standard neuropsychological tests such as the ACE, while memory Imprecision was able to predict cognitive decline on standard tests of cognition over one year.

These results suggest that mechanisms that fail in AD are already present on a smaller scale at the MCI stage of the disease, and that those could also be used to differentiate between patients with SCI and MCI. The existing literature is limited by either a comparison of only two diagnostic groups, by relying on simple binary crude measures of performance such as in change detection tasks, or by small sample sizes. Moreover, no study so far explored these changes in different stages of cognitive decline in AD. Additionally, only one investigation in familial AD included biological measures such as hippocampal integrity, and it focused only on a single metric of interest, misbinding [77]. Our data replicated the finding of an association between hippocampal integrity and misbinding in patients with late onset, sporadic AD. It also showed how hippocampal atrophy is associated in this population with deficits in multiple cognitive processes, including but not limited to misbinding, and that overall performance (indexed by Identification accuracy) is a better predictor of hippocampal function in this patient population compared to misbinding, probably due the inclusion in the abovementioned study in familial AD of less cognitively affected participants (asymptomatic mutation carriers and symptomatic carriers).

Beyond the hippocampus, the next step was to understand whether each metric was representative of anatomically distinct brain areas, or if all these different cognitive metrics converged on medial temporal lobe function. In **Chapter 4**, the contribution of whole brain GM integrity underlying distinct VSTM processes measured by the OMT task was investigated in a cohort of elderly healthy individuals and patients with AD. Hippocampal

volume was related to all these metrics across elderly controls and AD, but the hippocampus constituted only one fraction of the multiple brain areas associated with VSTM performance. Evidence was found for clustering of brain behavioural associations into four different spatial patterns: 1) Spatial localization – Identity metrics, i.e., metrics that coded the identity of an object according to its spatial localization (the mixture model metrics Target detection and Guessing, 2) Reaction times metrics (Identification (Object) and Localization (Spatial) times, 3) Object identity metric (Identification accuracy), carrying only identity information and 4) Spatial localization - Precision metrics (Absolute Error and memory Imprecision), which held high-detail granular spatial information.

Metrics that encoded fine-grain spatial information had the most extensive patterns of correlation with brain atrophy (Imprecision and Absolute Error). The GM integrity of the Precuneus was associated with all metrics that carried spatial localization, but not object identity information (Identification accuracy and Identification time). If the contribution of global cognition was factored out, those four spatial patterns were still present, albeit showing less extensive areas involved. Independent of global condition, Reaction times and Object identity metrics had a prominent involvement of the frontal cortex (including the Frontal Pole), Spatial localization - Identity metrics were associated with GM atrophy in the Inferior Temporal Gyrus, and Spatial localization - Precision metrics were correlated with areas including the Precuneus, Middle and Inferior Temporal Gyrus, Lateral Occipital Cortex, Middle Frontal Gyrus and Frontal Pole.

This is the first time the brain anatomical localization of the metrics derived from the mixture model of WM has been investigated in patients with AD. These data are consistent with a large-scale brain network involved in determining memory precision. Precuneal integrity is considered to be important for spatially dependent tasks [335], and this structure is severely

affected in AD patients [459]. However, this is the first time that the degradation of spatial localization has been linked to precuneal integrity in patients with AD using a paradigm including spatial and non-spatially dependent metrics, and deploying the mixture model of WM.

In AD, atrophy is also often accompanied by disruption of white matter integrity, but evidence on the contribution of selective white matter tracts in different VSTM functions is scarce. To own knowledge, at present there are no studies that have applied the mixture model of WM to investigate how white matter disruption impacts performance in patients with AD. Therefore, in **Chapter 5**, whole brain TBSS analysis was applied to examine the contribution of selective white matter tracts in causing specific VSTM errors measured by the OMT task in the same cohort of Chapter 4. Results revealed that a key common area, comprising the left Optic Radiation, Forceps Major and Middle Longitudinal Fasciculus was associated with performance across multiple VSTM metrics in patients with AD. In healthy elderly controls white matter integrity of a homologous region in the right hemisphere was shown to play a role in maintaining the correct object identity in VSTM. Additional white matter tracts were also involved in higher-order functions that are necessary to correctly identifying and localizing the remembered object in patients with AD such as the left Superior Longitudinal Fasciculus, Superior Thalamic Radiation, Inferior Fronto-occipital Fasciculus, Vertical Occipital Fasciculus, Arcuate Fasciculus and Cingulum.

This evidence, supporting a hemispheric dissociation of white matter integrity in determining function in health (right) and disease (left), is important to shed light on fine grain physiological adaptations of white matter function in patients with AD and could be potentially applied to dissociate how a healthy brain transitions to its pathological counterpart.

In **Chapter 6** the relationship between plasma biomarkers of AD and the VSTM metrics derived from the OMT task and a battery of digital tests probing executive functions, visuospatial abilities, and LTM was tested, alongside with their test-retest reliability at a short timeframe (0-3-6 months). Our data revealed an association between several digital cognitive metrics and the plasma biomarkers examined, particularly with pTau181, and to a lesser extent with GFAP and NfL, while the Aβ42/40 ratio showed the weakest correlations with cognition.

Associations differed across tests, and even across different metrics extracted within a test. The OIS task, particularly its metric Immediate Object Accuracy, showed excellent correlations with plasma biomarkers and it was a good predictor of diagnostic classification between AD patients and healthy controls, as well as plasma pTau181 levels. Other digital tests such as the recall of the ROCF also showed promise in group classification and prediction of tau burden. Reaction times and Identification accuracy of the OMT task showed also good correlations with several plasma biomarkers. The test-retest analysis showed that pTau181 had a high coefficient of variation (CV%) of around 35.8%, and therefore poor test-retest reliability. GFAP (18.4%) and NfL (15.8%) had moderate reliability, while the Aβ42/40 ratio had the lowest CV (9.5%), denoting higher stability over repeated measures. These did not differ between healthy controls and AD patients. Amongst the different digital metrics, Identification accuracy and Target detection of OMT and Immediate Object Accuracy and Immediate Semantic Accuracy of the OIS showed high test-retest reliability.

Plasma biomarkers have emerged as a game-changer in AD, and are set to revolutionize the entire field, from aiding large-scale early screening, to inclusion in clinical trials, to

monitoring of disease progression and response to potential disease modifying therapies. However, the roadmap to clinical implementation is still a 'long and winding road', and several unknowns need to be uncovered before these biomarkers can be effectively deployed to replace diagnostic gold-standard such as CSF analysis and PET imaging. Digital metrics also offer similar promise, with their potential of being administered widely and frequently, without the need of attending face-to-face clinical appointments. Data presented in Chapter 6 showed that multiple digital metrics can capture AD-specific pathological cascades. The tighter relationship between pTau181 levels and several cognitive metrics, compared to unspecific markers of neurodegeneration such as NfL, is also promising. Whilst amyloid levels were associated with some of our cognitive metrics (such as the reaction times metrics of OMT, the Immediate Object Accuracy of the OIS task, ROCF, DSST and CORSI), single tests and combination of different tests as well as plasma biomarkers were not good predictors of amyloid burden in this dataset.

7.2 Potential translational value

Existing literature on cognitive changes in AD is polarized between large scale studies that utilize standardized measures of cognitive function (such as MMSE, MoCA or ADAS-cog), and small, often underpowered, experimental investigations that dissect a specific cognitive mechanism in this patient population. Head-to-head comparisons between novel experimental tasks and clinically validated measures are very scarce. The aim of this DPhil was not only to explore fine grain changes in cognition in patients alongside AD progression, but also to test whether these metrics can be used as a proxy for standard clinically available tests, in order to facilitate the roll out of digital cognitive testing on a larger scale, potentially remotely, whilst validating their accuracy compared to established measures of cognition.

Chapter 3 gave us a clearer understanding of the potentials, and limitations, of one of these digital tests, OMT. If all metrics were combined, it showed excellent ability in early phases of the disease to discriminate between healthy subjects and patients with SCI and between SCI and MCI. However, its accuracy dropped in the classification between MCI and AD dementia. Therefore, on a clinical level it might be more useful for screening patients at risk of developing AD. Head-to-head comparison with the ACE, one of the most widespread clinical tests of cognition in the United Kingdom, showed that the latter might still be preferred in later stages of cognitive impairment. Nevertheless, all metrics in isolation, except Absolute Error, were still able to discriminate between patients with MCI and AD, and a more tailored approach, using a subset of metrics, might be suitable to answer different clinical questions in later phases of the disease, even if overall traditional neuropsychological testing seems superior to OMT at this stage. However, the standard neuropsychological test used as a comparator, the ACE, was also the one used to determine diagnostic groups (healthy controls, SCI, MCI, AD), and therefore by comparing the accuracy of these metrics to the ACE in determining group classification we might be overestimating its diagnostic potential. The finding of Chapter 3 that Identification accuracy of OMT performed at par with ACE in predicting hippocampal atrophy is also encouraging.

Face-to-face cognitive assessment is time consuming, and a big limitation of large-scale screening is the bottleneck in access to specialized appointments. Standard cognitive tests are either too short and lack granularity (as MMSE, MoCA, GPCOG [460], often administered at GP surgeries), or too long, and require lengthy neuropsychological testing sessions (as the Free and Cued Selective Reminding Test (FCSRT) or ADAS-cog, often administered in specialized memory clinics). The difficulty in capturing early signs of cognitive impairment in tests that lack granularity is that the changes that occur at the preclinical phase are quite small compared to those in the prodromal phase. These digital tests have the benefit of

exploring fine grain changes across a range of different cognitive domains and offer an array of dissociable indices of cognitive functioning. Moreover, they can be administered at home and they do not require lengthy hospital appointments. Testing patients in their natural environment also carries higher ecological validity and might also mitigate the cognitive 'white coat syndrome' of face-to-face testing [461].

One strength of the metrics derived from the OMT task used in this DPhil is that we have been able to validate them against pathophysiological hallmarks of AD, including hippocampal volumes (**Chapter 3**), whole brain GM atrophy (**Chapter 4**), white matter damage (**Chapter 5**), and plasma biomarkers (**Chapter 6**). Being able to map specific metrics to their anatomical localization in the brain in health and disease, might prove useful to unravel the specific contribution of selected brain areas in determining performance, so that when a process-specific failure is observed behaviourally, that could be reconducted to a selective dysfunction of a defined GM area and white matter bundle. Similarly, it can be mapped onto a biological change (as amyloid or tau accumulation or neurodegeneration). The data from **Chapter 6** aimed to sow the seeds for further development of a combination of plasma biomarkers and digital metrics in the diagnostic process of AD. It also provided essential information to measure the noise in the data, which is needed for large scale predicative modelling of future risks of dementia progression. Whilst our digital metrics were not able to statistically increase diagnostic performance in this current sample, the best performing model to predict group classification, amyloid status and tau burden was always composed of a mixture of plasma biomarkers and digital cognitive tests.

7.3 Future Directions

To overcome the circular argument of a neuropsychological test being used to determine group classification and tested against novel digital metrics, one possible avenue to explore

this is to determine how these digital tests perform against other internationally widely used cognitive scales, such as the ADAS-cog or the PACC. This might give us a better understanding on how generalizable these findings are in a world-wide context.

Data on acceptability of digital tests compared to traditional tests have only recently started to emerge, and whilst they show great promise in terms of scalability and cost-effectiveness, there are still several unknowns [462, 463], including the impact of digital divide, i.e., the lower rates of utilizations of digital devices in older people, which is particularly high in patients with cognitive impairment [464]. We could not estimate in this DPhil the rate of acceptability of remote testing compared to standard neuropsychological testing, as all individuals agreed to undergo both types of testing as part of the study. One limitation of our current protocol is that evaluating several different metrics led to incomplete datasets in some cases, in healthy controls, but more so in patients with cognitive impairment. Therefore, a reduced subset of tests might be implemented in clinical settings to overcome testing fatigue at home. On the positive side, the tests used in **Chapter 6** offer a variety of testing times and difficulties, going from as little as 2 minutes to perform the DSST. It is encouraging to see that even in a relatively small sample, performance on a test that lasts such a short time could mirror levels of pTau181, GFAP, NfL and amyloid.

At the present time, data presented in **Chapter 3** suggest the utility of a digital test such as OMT might be higher in individuals with less prominent cognitive impairment. However, data presented in **Chapter 6** showed that it was not the best performing digital test, neither in the association with plasma biomarkers or with respect to group classification. As AD processes start much before cognitive impairment at standard tests of cognition appears, the whole battery of tests deployed in **Chapter 6** could be tested in the preclinical and prodromal phases of the disease. This is of the utmost importance in order not to miss the window of

opportunity in which disease modifying interventions are more likely to be effective. Testing these digital tests in earlier phases of AD might also give us a better understanding of whether we could find reliable cognitive biomarkers of early amyloid deposition or synaptic dysfunction.

Besides *sensitivity,* assessing *specificity* of these metrics is also crucial. In **Chapter 2** we have been able to show that a different neurological disease, such as PD, carries a different cognitive fingerprint compared to AD. However, patients with PD have lower rates of cognitive impairment compared to patients with a primary dementia such as AD. The higher association with pTau181 rather than NfL is encouraging, but data on how these digital cognitive metrics perform in other clinical mimics of AD such ad Lewy body disease, Frontotemporal dementia and Corticobasal syndrome are much needed. This also applies to their brain-behavioral associations and correlations with plasma biomarkers, which might need to be validated in different clinical populations.

Test-test reliability and effective tracking of cognitive decline are also important factors to be explored. The use of these digital cognitive tests in predicting long-term conversion rates to AD dementia of patients in the preclinical and prodromal phase needs to be evaluated, to identify the temporal evolution of predictive accuracy of such tests.

7.4 Conclusion

In summary, the work presented in this thesis indicates that several different processes lie at the core of STM failures in AD, such as reduced target detection, increased rates of misbinding, increased random guessing, reduced memory precision and lower filtering abilities, particularly if distractors are presented during the maintenance phase. Digital cognitive tests can accurately track these effects already in early phases of the disease (as in

the transition between SCI and MCI, and MCI to AD). They show also good correlation with hippocampal and whole brain GM atrophy, white matter damage, and plasma biomarkers levels. Digital tests need to be validated in diverse clinical populations and at different stages of the disease, to identify tests, and subset of metrics within a test, which best answer specific diagnostic and prognostic questions.

REFERENCES

1. Baddeley AD, Hitch G. Working Memory. *Psychology of Learning and Motivation - Advances in Research and Theory*. 1974;8(C):47-89. doi:10.1016/S0079-7421(08)60452-1
2. Baddeley A. The episodic buffer: a new component of working memory? *Trends Cogn Sci*. 2000;4(11):417-423. doi:10.1016/S1364-6613(00)01538-2
3. Cowan N. What are the differences between long-term, short-term, and working memory? *Prog Brain Res*. 2008;169:323. doi:10.1016/S0079-6123(07)00020-9
4. Peterson L, Peterson MJ. Short-term retention of individual verbal items. *J Exp Psychol*. 1959;58(3):193-198. doi:10.1037/H0049234
5. Posner MI. Components of skilled performance. *Science (1979)*. 1966;152(3730):1712-1718. doi:10.1126/SCIENCE.152.3730.1712/ASSET/B18139BF-4404-46C3-A214-A7D16D45B28B/ASSETS/SCIENCE.152.3730.1712.FP.PNG
6. Schweickert R, Boruff B. Short-term memory capacity: magic number or magic spell? *J Exp Psychol Learn Mem Cogn*. 1986;12(3):419-425. doi:10.1037//0278-7393.12.3.419
7. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev*. 1956;63(2):81-97. doi:10.1037/h0043158
8. Cowan N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci*. 2001;24(1):87-114. doi:10.1017/S0140525X01003922

9. Luck SJ, Vogel EK. Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn Sci*. 2013;17(8):391-400. doi:10.1016/J.TICS.2013.06.006

10. Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature 1997 390:6657*. 1997;390(6657):279-281. doi:10.1038/36846

11. Ma WJ, Husain M, Bays PM. Changing concepts of working memory. *Nat Neurosci*. Published online 2014. doi:10.1038/nn.3655

12. Wheeler ME, Treisman AM. Binding in short-term visual memory. *J Exp Psychol Gen*. 2002;131(1):48-64. doi:10.1037//0096-3445.131.1.48

13. Alvarez GA, Cavanagh P. The Capacity of Visual Short-Term Memory Is Set Both by Visual Information Load and by Number of Objects. Published online 2004.

14. Zokaei N, Ning S, Manohar S, Feredoes E, Husain M. Flexibility of representational states in working memory. *Front Hum Neurosci*. 2014;8(November). doi:10.3389/FNHUM.2014.00853

15. Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. *Nature 2008 453:7192*. 2008;453(7192):233-235. doi:10.1038/nature06860

16. Wilken P, Ma WJ. A detection theory account of change detection. *J Vis*. 2004;4(12):1120-1135. doi:10.1167/4.12.11

17. Bays PM, Catalao RFG, Husain M. The precision of visual working memory is set by allocation of a shared resource. *J Vis*. Published online 2009. doi:10.1167/9.10.7

18. Bays PM. Noise in Neural Populations Accounts for Errors in Working Memory. Published online 2014. doi:10.1523/JNEUROSCI.3204-13.2014

19. Oberauer K. Access to Information in Working Memory: Exploring the Focus of Attention. Published online 2002. doi:10.1037/0278-7393.28.3.411

20. Gorgoraptis N, Catalao RFG, Bays PM, Husain M. Dynamic updating of working memory resources for visual objects. *J Neurosci*. 2011;31(23):8502-8511. doi:10.1523/JNEUROSCI.0208-11.2011

21. Treue S. Visual attention: the where, what, how and why of saliency. *Curr Opin Neurobiol*. 2003;13(4):428-432. doi:10.1016/S0959-4388(03)00105-3

22. Pertzov Y, Bays PM, Joseph S, Husain M. Rapid forgetting prevented by retrospective attention cues. *J Exp Psychol Hum Percept Perform*. Published online 2013. doi:10.1037/a0030947

23. Christiaan Klink P, Jentgens P, Lorteije JAM. Priority Maps Explain the Roles of Value, Attention, and Salience in Goal-Oriented Behavior. *Journal of Neuroscience*. 2014;34(42):13867-13869. doi:10.1523/JNEUROSCI.3249-14.2014

24. Grogan JP, Husain M, Manohar SG, et al. A new toolbox to distinguish the sources of spatial memory error. *J Vis*. 2020;20(13):1-19. doi:10.1167/JOV.20.13.6

25. Brady TF, Konkle T, Alvarez GA. Compression in Visual Working Memory: Using Statistical Regularities to Form More Efficient Memory Representations. Published online 2009. doi:10.1037/a0016797

26. Moorselaar D, Slagter HA. Inhibition in selective attention. *Ann N Y Acad Sci*. Published online 2020. doi:10.1111/nyas.14304

27. Vogel EK, McCollough AW, Machizawa MG. Neural measures reveal individual differences in controlling access to working memory. *Nature*. Published online 2005. doi:10.1038/nature04171

28. Myers NE, Stokes MG, Nobre AC. Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends Cogn Sci*. 2017;21(6):449-461. doi:10.1016/j.tics.2017.03.010

29. Moher J, Egeth HE. The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Atten Percept Psychophys*. Published online 2012. doi:10.3758/s13414-012-0358-0

30. Gaspelin N, Luck SJ. The Role of Inhibition in Avoiding Distraction by Salient Stimuli. *Trends Cogn Sci*. 2018;22(1):79-92. doi:10.1016/J.TICS.2017.11.001

31. Oberauer K, Hein L. Attention to Information in Working Memory: *https://doi.org/101177/0963721412444727*. 2012;21(3):164-169. doi:10.1177/0963721412444727

32. Chatham HC, Badre D. Multiple gates on working memory. *Curr Opin Behav Sci*. 2015;(1). doi:10.1016/j.cobeha.2014.08.001.Multiple

33. McNab F, Dolan RJ. Dissociating distractor-filtering at encoding and during maintenance. *J Exp Psychol Hum Percept Perform*. Published online 2014. doi:10.1037/a0036013

34. Noonan MAP, Adamian N, Pike A, Printzlau F, Crittenden BM, Stokes MG. Distinct Mechanisms for Distractor Suppression and Target Facilitation. *Journal of Neuroscience*. 2016;36(6):1797-1807. doi:10.1523/JNEUROSCI.2133-15.2016

35. Brockmole JR, Logie RH. Age-related change in visual working memory: A study of 55,753 participants aged 8-75. *Front Psychol*. Published online 2013. doi:10.3389/fpsyg.2013.00012

36. Hartshorne JK, Germine LT. When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the lifespan. doi:10.1177/0956797614567339

37. Cowan N, Naveh-Benjamin M, Kilb A, Saults JS. Life-span development of visual working memory: When is feature binding difficult? *Dev Psychol*. Published online 2006. doi:10.1037/0012-1649.42.6.1089

38. Pertzov Y, Heider M, Liang Y, Husain M. Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychol Aging*. Published online 2015. doi:10.1037/a0038396

39. Tas AC, Costello MC, Buss AT. Age-related decline in visual working memory: The effect of non-target objects during a delayed estimation task. *Psychol Aging*. 2020;35(4):565. doi:10.1037/PAG0000450

40. Peich MC, Husain M, Bays PM. Age-related decline of precision and binding in visual working memory. *Psychol Aging*. Published online 2013. doi:10.1037/a0033236

41. Korkki SM, Richter FR, Jeyarathnarajah P, Simons JS. Healthy Ageing Reduces the Precision of Episodic Memory Retrieval. *Psychol Aging*. Published online 2020. doi:10.1037/pag0000432

42. McNab F, Zeidman P, Rutledge RB, et al. Age-related changes in working memory and the ability to ignore distraction. *Proc Natl Acad Sci U S A*. Published online 2015. doi:10.1073/pnas.1504162112

43. Brockmole JR, Parra MA, Sala S Della, Logie RH. Do binding deficits account for age-related decline in visual working memory? *Psychon Bull Rev*. Published online 2008. doi:10.3758/PBR.15.3.543

44. Parra MA, Abrahams S, Fabi K, Logie R, Luzzi S, Sala S Della. Short-term memory binding deficits in Alzheimers disease. *Brain*. Published online 2009. doi:10.1093/brain/awp036

45. Diamond A. Executive Functions. *Annu Rev Psychol*. Published online 2013. doi:10.1146/annurev-psych-113011-143750

46. Ferguson HJ, Brunsdon VEA, Bradford EEF. The developmental trajectories of executive function from adolescence to old age. *Scientific Reports* |. 123AD;11:1382. doi:10.1038/s41598-020-80866-1

47. Rey-Mermet A, Gade M. Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychon Bull Rev*. Published online 2018. doi:10.3758/s13423-017-1384-7

48. 2020 Alzheimer's disease facts and figures. *Alzheimer's and Dementia*. 2020;16(3):391-460. doi:10.1002/alz.12068

49. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol*. 1991;82(4):239-259. doi:10.1007/BF00308809

50. Cummings JL. Toward a molecular neuropsychiatry of neurodegenerative diseases. *Ann Neurol*. 2003;54(2):147-154. doi:10.1002/ANA.10616

51. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. *Cold Spring Harb Perspect Med*. 2011;1(1). doi:10.1101/CSHPERSPECT.A006189

52. Hyman BT, Marzloff K, Arriagada P V. The lack of accumulation of senile plaques or amyloid burden in Alzheimer's disease suggests a dynamic

balance between amyloid deposition and resolution. *J Neuropathol Exp Neurol*. 1993;52(6):594-600. doi:10.1097/00005072-199311000-00006

53. Ingelsson M, Fukumoto H, Newell KL, et al. Early Abeta accumulation and progressive synaptic loss, gliosis, and tangle formation in AD brain. *Neurology*. 2004;62(6):925-931. doi:10.1212/01.WNL.0000115115.98960.37

54. Grothe MJ, Barthel H, Sepulcre J, Dyrba M, Sabri O, Teipel SJ. In vivo staging of regional amyloid deposition. *Neurology*. 2017;89(20):2031. doi:10.1212/WNL.0000000000004643

55. Albert MS, DeKosky ST, Dickson D, et al. 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*. 2011;7(3):270-279. doi:10.1016/j.jalz.2011.03.008

56. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurology*. 2007;6(8):734-746. doi:10.1016/S1474-4422(07)70178-3

57. Carlesimo GA, Perri R, Caltagirone C. Category cued recall following controlled encoding as a neuropsychological tool in the diagnosis of Alzheimer's disease: a review of the evidence. *Neuropsychol Rev*. 2011;21(1):54-65. doi:10.1007/S11065-010-9153-7

58. Belleville S, Fouquet C, Hudon C, Zomahoun HTV, Croteau J. Neuropsychological Measures that Predict Progression from Mild Cognitive Impairment to Alzheimer's type dementia in Older Adults: a Systematic Review and Meta-Analysis. *Neuropsychol Rev*. 2017;27(4):328. doi:10.1007/S11065-017-9361-5

59. Hsieh S, Schubert S, Hoon C, Mioshi E, Hodges JR. Validation of the Addenbrooke's Cognitive Examination III in Frontotemporal Dementia and Alzheimer's Disease. *Dement Geriatr Cogn Disord*. Published online 2013. doi:10.1159/000351671

60. Mathuranath PS, Nestor PJ, Berrios GE, Rakowicz W, Hodges JR. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology*. 2000;55(11):1613-1620. doi:10.1212/01.WNL.0000434309.85312.19

61. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's and Dementia*. Published online 2018. doi:10.1016/j.jalz.2018.02.018

62. Schott JM, Lehmann M, Primativo S, et al. Consensus classification of posterior cortical atrophy. *Alzheimers Dement*. 2017;13(8):870-884. doi:10.1016/J.JALZ.2017.01.014

63. Henry ML, Gorno-Tempini ML. The logopenic variant of primary progressive aphasia. *Curr Opin Neurol*. 2010;23(6):633. doi:10.1097/WCO.0B013E32833FB93E

64. Ossenkoppele R, Pijnenburg YAL, Perry DC, et al. The behavioural/dysexecutive variant of Alzheimer's disease: clinical, neuroimaging and pathological features. *Brain*. 2015;138(9):2732. doi:10.1093/BRAIN/AWV191

65. Drachman DA, Arbit J. Memory and the hippocampal complex. II. Is memory a multiple process? *Arch Neurol*. 1966;15(1):52-61. doi:10.1001/ARCHNEUR.1966.00470130056005

66. Squire LR. The Legacy of Patient H.M. for Neuroscience. *Neuron*. 2009;61(1):6. doi:10.1016/J.NEURON.2008.12.023

67. Wilson RS, Sullivan M, DeToledo-Morrell L, Stebbins GT, Bennett DA, Morrell F. Association of memory and cognition in Alzheimer's disease with volumetric estimates of temporal lobe structures. *Neuropsychology*. 1996;10(4):459-463. doi:10.1037/0894-4105.10.4.459

68. Cave CB, Squire LR. Intact verbal and nonverbal short-term memory following damage to the human hippocampus. *Hippocampus*. 1992;2(2):151-163. doi:10.1002/HIPO.450020207

69. Baddeley A, Logie R, Bressi S, Sala S Della, Spinnler H. Dementia and Working Memory. *The Quarterly Journal of Experimental Psychology Section A*. Published online 1986. doi:10.1080/14640748608401616

70. Collette F, Linden M Van Der, Bechet S, Salmon E. Phonological loop and central executive functioning in Alzheimer's disease. *Neuropsychologia*. 1999;37(8):905-918. doi:10.1016/S0028-3932(98)00148-1

71. Orsini A, Grossi D, Capitani E, Laiacona M, Papagno C, Vallar G. Verbal and spatial immediate memory span: Normative data from 1355 adults and 1112 children. *The Italian Journal of Neurological Sciences*. 1987;8(6):537-548. doi:10.1007/BF02333660

72. Carlesimo GA, Fadda L, Lorusso S, Caltagirone C. Verbal and spatial memory spans in Alzheimer's and multi-infarct dementia. *Acta Neurol Scand*. Published online 1994. doi:10.1111/j.1600-0404.1994.tb01648.x

73. Germano C, Kinsella GJ, Storey E, Ong B, Ames D. Journal of Clinical and Experimental Neuropsychology The episodic buffer and learning in early

Alzheimer's disease. Published online 2008. doi:10.1080/13803390701594894

74. Zokaei N HM. Working Memory in Alzheimer's Disease and Parkinson's Disease. *Curr Top Behav Neurosci*. Published online 2019. doi:10.1007/7854_2019_103

75. Della Sala S, Parra MA, Fabi K, Luzzi S, Abrahams S. Short-term memory binding is impaired in AD but not in non-AD dementias. *Neuropsychologia*. Published online 2012. doi:10.1016/j.neuropsychologia.2012.01.018

76. Zokaei N, Sillence A, Kienast A, et al. Different patterns of short-term memory deficit in Alzheimer's disease, Parkinson's disease and subjective cognitive impairment. *Cortex*. 2020;132:41-50. doi:10.1016/J.CORTEX.2020.06.016

77. Liang Y, Pertzov Y, Nicholas JM, et al. Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex*. Published online 2016. doi:10.1016/j.cortex.2016.01.015

78. Tabi YA, Maio MR, Attaallah B, et al. Vividness of visual imagery questionnaire scores and their relationship to visual short-term memory performance. *Cortex*. 2022;146:186-199. doi:10.1016/J.CORTEX.2021.10.011

79. Gellersen HM, Coughlan G, Hornberger M, Simons JS. Memory precision of object-location binding is unimpaired in APOE ε4-carriers with spatial navigation deficits. *Brain Commun*. 2021;3(2). doi:10.1093/BRAINCOMMS/FCAB087

80. Zokaei N, Grogan J, Fallon SJ, et al. Short-term memory advantage for brief durations in human APOE ε4 carriers. *Scientific Reports 2020 10:1*. 2020;10(1):1-10. doi:10.1038/s41598-020-66114-6

81. Zokaei N, Čepukaitytė G, Board AG, Mackay CE, Husain M, Nobre AC. Dissociable effects of the apolipoprotein-E (APOE) gene on short- and long-term memories. *Neurobiol Aging*. Published online 2019. doi:10.1016/j.neurobiolaging.2018.09.017

82. Sapkota RP, Van Der Linde I, Lamichhane N, Upadhyaya T, Pardhan S. Patients with Mild Cognitive Impairment Show Lower Visual Short-Term Memory Performance in Feature Binding Tasks. *Dement Geriatr Cogn Dis Extra*. 2017;7(1):74-86. doi:10.1159/000455831

83. Fernández-Guinea S, Parra MA, Frank A, Delgado ML, Della Sala S. Short-term memory binding in mild cognitive impairment. Published online September 7, 2011.

84. Baddeley AD, Baddeley HA, Bucks RS, Wilcock GK. Attentional control in Alzheimer's disease. *Brain*. 2001;124(8):1492-1508. doi:10.1093/brain/124.8.1492

85. Simone PM, Baylis GC. Selective attention in a reaching task: Effect of normal aging and alzheimer's disease. *J Exp Psychol Hum Percept Perform*. 1997;23(3):595-608. doi:10.1037//0096-1523.23.3.595

86. Aurtenetxe S, García-Pacios J, del Río D, et al. Interference Impacts Working Memory in Mild Cognitive Impairment. *Front Neurosci*. 2016;10(OCT):443. doi:10.3389/fnins.2016.00443

87. Fernandez-Duque D, Black SE. Attentional networks in normal aging and Alzheimer's disease. *Neuropsychology*. Published online 2006. doi:10.1037/0894-4105.20.2.133

88. Wylie SA, Ridderinkhof KR, Eckerle MK, Manning CA. Inefficient response inhibition in individuals with mild cognitive impairment. *Neuropsychologia*. 2007;45(7):1408-1419. doi:10.1016/j.neuropsychologia.2006.11.003

89. Owen AM. MINI-REVIEW The Functional Organization of Working Memory Processes Within Human Lateral Frontal Cortex: The Contribution of Functional Neuroimaging. *European Journal of Neuroscience*. 1997;9:1329-1339.

90. Veldsman M, Mitchell DJ, Cusack R. The neural basis of precise visual short-term memory for complex recognisable objects. *Neuroimage*. 2017;159:131-145. doi:10.1016/J.NEUROIMAGE.2017.07.033

91. Christophe TB, Hebart MN, Haynes JD. Decoding the Contents of Visual Short-Term Memory from Human Visual and Parietal Cortex. *Journal of Neuroscience*. 2012;32(38):12983-12989. doi:10.1523/JNEUROSCI.0184-12.2012

92. Linke AC, Vicente-Grabovetsky A, Mitchell DJ, Cusack R. Encoding strategy accounts for individual differences in change detection measures of VSTM. *Neuropsychologia*. Published online 2011. doi:10.1016/j.neuropsychologia.2010.11.034

93. Mishkin M, Ungerleider LG. Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*. 1982;6(1):57-77. doi:10.1016/0166-4328(82)90081-X

94. Ungerleider LG, Haxby J V. “What” and “where” in the human brain Introduction.

95. Ungerleider LG, Courtney SM, Haxby J V. Neuroimaging of Human Brain Function. 1998;95:883-890.

96. Goldman-Rakic PS. Cellular basis of working memory. *Neuron*. 1995;14(3):477-485. doi:10.1016/0896-6273(95)90304-6

97. Müller NG, Knight RT. The functional neuroanatomy of working memory: Contributions of human brain lesion studies. *Neuroscience*. 2006;139(1):51-58. doi:10.1016/J.NEUROSCIENCE.2005.09.018

98. Kleist K. Über Form und Ortsblindheit bei Verletzung des Hinterhauptlappens. *Dtsch Z Nervenheilkd*. 1935;138(3-4):206-214. doi:10.1007/BF01736133/METRICS

99. Newcombe F, Ratcliff G, Damasio H. Dissociable visual and spatial impairments following right posterior cerebral lesions: Clinical, neuropsychological and anatomical evidence. *Neuropsychologia*. 1987;25(1 PART 2):149-161. doi:10.1016/0028-3932(87)90127-8

100. Farah MJ, Levine DN, Calvanio R. A case study of mental imagery deficit. *Brain Cogn*. 1988;8(2):147-164. doi:10.1016/0278-2626(88)90046-2

101. Owen AM, Morris RG, Sahakian BJ, Polkey CE, Robbins TW. Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain*. 1996;119 (Pt 5)(5):1597-1615. doi:10.1093/BRAIN/119.5.1597

102. Levine DN, Warach J, Farah M. Two visual systems in mental imagery: dissociation of “what” and “where” in imagery disorders due to bilateral

posterior cerebral lesions. *Neurology*. 1985;35(7):1010-1018. doi:10.1212/WNL.35.7.1010

103. Haxby J V, Horwitz B, Ungerleider LG, Maisog,' JM, Pietrini P, Grady CL. *The Functional Organization of Human Extrastriate Cortex: A PET-RCBF Study of Selective Attention to Faces and Locations*. Vol 74.; 1994.

104. Clark VP, Keil K, Maisog JM, Courtney S, Ungerleider LG, Haxby J V. Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. *Neuroimage*. 1996;4(1):1-15. doi:10.1006/NIMG.1996.0025

105. Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby J V. An area specialized for spatial working memory in human frontal cortex. *Science (1979)*. 1998;279(5355):1347-1351. doi:10.1126/SCIENCE.279.5355.1347

106. Courtney SM, Ungerleider LG, Keil K, Haxby J V. *Object and Spatial Visual Working Memory Activate Separate Neural Systems in Human Cortex*.

107. Petrides M. The role of the mid-dorsolateral prefrontal cortex in working memory. *Exp Brain Res*. 2000;133(1):44-54. doi:10.1007/S002210000399

108. Haxby J V, Ungerleider LG, Horwitz B, Rapoport S 1, Grady CL. Hemispheric Differences in Neural Systems for Face Working Memory: A PET-rCBF Study. *Hum Brain Mapp*. Published online 1995:368-382.

109. Courtney SM, Ungerleider LG, Keil K, Haxby J V. Transient and sustained activity in a distributed neural system for human working memory. *Nature 1997 386:6625*. 1997;386(6625):608-611. doi:10.1038/386608a0

110. Curtis CE, D'Esposito M. Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci*. 2003;7(9):415-423. doi:10.1016/S1364-6613(03)00197-9

111. Leung HC, Gore JC, Goldman-Rakic PS. Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *J Cogn Neurosci*. 2002;14(4):659-671. doi:10.1162/08989290260045882

112. Jha AP, McCarthy G. The influence of memory load upon delay-interval activity in a working-memory task: an event-related functional MRI study. *J Cogn Neurosci*. 2000;12 Suppl 2(SUPPL. 2):90-105. doi:10.1162/089892900564091

113. Rypma B, D'Esposito M. The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. *Proc Natl Acad Sci U S A*. 1999;96(11):6558-6563. doi:10.1073/PNAS.96.11.6558

114. Rowe JB, Toni I, Josephs O, Frackowiak RSJ, Passingham RE. The prefrontal cortex: response selection or maintenance within working memory? *Science*. 2000;288(5471):1656-1660. doi:10.1126/SCIENCE.288.5471.1656

115. Riggall AC, Postle BR. The Relationship between Working Memory Storage and Elevated Activity as Measured with Functional Magnetic Resonance Imaging. Published online 2012. doi:10.1523/JNEUROSCI.1892-12.2012

116. Bettencourt KC, Xu Y. Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience 2016 19:1*. 2015;19(1):150-157. doi:10.1038/nn.4174

117. Sala JB, Rämä P, Courtney SM. Functional topography of a distributed neural system for spatial and nonspatial information maintenance in working memory. *Neuropsychologia*. 2003;41:341-356.

118. Linden DEJ, Bittner RA, Muckli L, et al. Cortical capacity constraints for visual working memory: dissociation of fMRI load effects in a fronto-parietal network. Published online 2003. doi:10.1016/j.neuroimage.2003.07.021

119. Harrison SA, Tong F. Decoding reveals the contents of visual working memory in early visual areas. *Nature*. 2009;458(7238):632-635. doi:10.1038/NATURE07832

120. Serences JT, Ester EF, Vogel EK, Awh E. Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychol Sci*. 2009;20(2):207. doi:10.1111/J.1467-9280.2009.02276.X

121. Murty VP, Sambataro F, Radulescu E, et al. Selective updating of working memory content modulates meso-cortico-striatal activity. *Neuroimage*. 2011;57(3):1264-1272. doi:10.1016/J.NEUROIMAGE.2011.05.006

122. McNab F, Klingberg T. Prefrontal cortex and basal ganglia control access to working memory. *Nat Neurosci*. Published online 2008. doi:10.1038/nn2024

123. Tamber-Rosenau BJ, Esterman M, Chiu YC, Yantis S. Cortical Mechanisms of Cognitive Control for Shifting Attention in Vision and Working Memory. doi:10.1162/jocn.2011.21608

124. Voytek B, Knight RT. Prefrontal cortex and basal ganglia contributions to visual working memory. *Proc Natl Acad Sci U S A*. 2010;107(42):18167-18172. doi:10.1073/PNAS.1007277107/SUPPL_FILE/PNAS.201007277SI.PDF

125. Baier B, Karnath HO, Dieterich M, Birklein F, Heinze C, Müller NG. Keeping memory clear and stable--the contribution of human basal ganglia and prefrontal cortex to working memory. *J Neurosci*. 2010;30(29):9788-9792. doi:10.1523/JNEUROSCI.1513-10.2010

126. Chatham CH, Frank MJ, Badre D. Corticostriatal output gating during selection from working memory. *Neuron*. 2014;81(4):930. doi:10.1016/J.NEURON.2014.01.002

127. Moeller M, Manohar S, Bogacz R. Uncertainty–guided learning with scaled prediction errors in the basal ganglia. *PLoS Comput Biol*. 2022;18(5):e1009816. doi:10.1371/JOURNAL.PCBI.1009816

128. Nakajima M, Schmitt LI, Halassa MM. Prefrontal Cortex Regulates Sensory Filtering through a Basal Ganglia-to-Thalamus Pathway. *Neuron*. 2019;103(3):445-458.e10. doi:10.1016/J.NEURON.2019.05.026

129. Cools R, Miyakawa A, Sheridan M, D'Esposito M. Enhanced frontal function in Parkinson's disease. *Brain*. 2010;133(1):225. doi:10.1093/BRAIN/AWP301

130. Fallon SJ, Mattiesing RM, Muhammed K, Manohar S, Husain M. Fractionating the neurocognitive mechanisms underlying working memory: Independent effects of dopamine and Parkinson's disease. *Cerebral Cortex*. Published online 2017. doi:10.1093/cercor/bhx242

131. Richter FR, Cooper RA, Bays PM, Simons JS. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *Elife*. Published online 2016. doi:10.7554/eLife.18260

132. Yonelinas AP. The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behavioural Brain Research*. Published online 2013. doi:10.1016/j.bbr.2013.05.030

133. Hannula DE, Ranganath C. Behavioral/Systems/Cognitive Medial Temporal Lobe Activity Predicts Successful Relational Memory Binding. Published online 2008. doi:10.1523/JNEUROSCI.3086-07.2008

134. Mitchell KJ, Johnson MK, Raye CL, D'Esposito M. fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. *Brain Res Cogn Brain Res*. 2000;10(1-2):197-206. doi:10.1016/S0926-6410(00)00029-X

135. Olsen RK, Moses SN, Riggs L, Ryan JD. The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front Hum Neurosci*. 2012;0(MAY 2012):1-13. doi:10.3389/FNHUM.2012.00146/XML/NLM

136. Jeneson A, Mauldin KN, Squire LR. Intact Working Memory for Relational Information after Medial Temporal Lobe Damage. *The Journal of Neuroscience*. 2010;30(41):13624. doi:10.1523/JNEUROSCI.2895-10.2010

137. Aly M, Turk-Browne NB. How Hippocampal Memory Shapes, and Is Shaped by, Attention. Published online 2017. doi:10.1007/978-3-319-50406-3_12

138. Koen JD, Borders AA, Petzold MT, Yonelinas AP. Visual short-term memory for high resolution associations is impaired in patients with medial temporal lobe damage. *Hippocampus*. 2017;27(2):184. doi:10.1002/HIPO.22682

139. Emrich SM, Riggall AC, La Rocque JJ, Postle BR. Distributed Patterns of Activity in Sensory Cortex Reflect the Precision of Multiple Items Maintained in Visual Short-Term Memory. *Journal of Neuroscience*. 2013;33(15):6516-6523. doi:10.1523/JNEUROSCI.5732-12.2013

140. Ekstrom AD, Yonelinas AP. Precision, binding, and the hippocampus: Precisely what are we talking about? *Neuropsychologia*. 2020;138:107341. doi:10.1016/J.NEUROPSYCHOLOGIA.2020.107341

141. Graus F, Titulaer MJ, Balu R, et al. A clinical approach to diagnosis of autoimmune encephalitis. *Lancet Neurol*. 2016;15(4):391. doi:10.1016/S1474-4422(15)00401-9

142. Gultekin SH, Rosenfeld MR, Voltz R, Eichen J, Posner JB, Dalmau J. Paraneoplastic limbic encephalitis: neurological symptoms, immunological findings and tumour association in 50 patients. *Brain*. 2000;123(7):1481-1494. doi:10.1093/BRAIN/123.7.1481

143. Pertzov Y, Miller TD, Gorgoraptis N, et al. Binding deficits in memory following medial temporal lobe damage in patients with voltage-gated potassium channel complex antibody-associated limbic encephalitis. *Brain*. Published online 2013. doi:10.1093/brain/awt129

144. Olson IR, Page K, Moore KS, Chatterjee A, Verfaellie M. Working memory for conjunctions relies on the medial temporal lobe. *Journal of Neuroscience*. Published online 2006. doi:10.1523/JNEUROSCI.1923-05.2006

145. Olson IR, Moore KS, Stark M, Chatterjee A. Visual Working Memory Is Impaired when the Medial Temporal Lobe Is Damaged. *J Cogn Neurosci*. 2006;18(7):1087-1097. doi:10.1162/JOCN.2006.18.7.1087

146. Finke C, Braun M, Ostendorf F, et al. The human hippocampal formation mediates short-term memory of colour-location associations. *Neuropsychologia*. 2008;46(2):614-623. doi:10.1016/J.NEUROPSYCHOLOGIA.2007.10.004

147. Hannula DE, Tranel D, Cohen NJ. The Long and the Short of It: Relational Memory Impairments in Amnesia, Even at Short Lags. *The Journal of Neuroscience*. 2006;26(32):8352. doi:10.1523/JNEUROSCI.5222-05.2006

148. Parra MA. Specific deficit of colour-colour short-term memory binding in sporadic and familial Alzheimer's disease. doi:10.1016/j.neuropsychologia.2011.03.022

149. Baddeley A, Allen R, Vargha-Khadem F. Is the hippocampus necessary for visual and verbal binding in working memory? *Neuropsychologia*. 2010;48(4):1089-1095. doi:10.1016/J.NEUROPSYCHOLOGIA.2009.12.009

150. Jeneson A, Wixted JT, Hopkins RO, Squire LR. Visual Working Memory Capacity and the Medial Temporal Lobe. *Journal of Neuroscience*. 2012;32(10):3584-3589. doi:10.1523/JNEUROSCI.6444-11.2012

151. Atkinson AJ, Colburn WA, DeGruttola VG, et al. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95. doi:10.1067/MCP.2001.113989

152. Zetterberg H, Bendlin BB. Biomarkers for Alzheimer's disease—preparing for a new era of disease-modifying therapies. *Molecular Psychiatry 2020 26:1*. 2020;26(1):296-308. doi:10.1038/s41380-020-0721-9

153. Hansson O. Biomarkers for neurodegenerative diseases. *Nat Med*. 2021;27(6):954-963. doi:10.1038/S41591-021-01382-X

154. Sevigny J, Suhy J, Chiao P, et al. Amyloid PET Screening for Enrichment of Early-Stage Alzheimer Disease Clinical Trials: Experience in a Phase 1b Clinical Trial. *Alzheimer Dis Assoc Disord*. 2016;30(1):1-7. doi:10.1097/WAD.0000000000000144

155. Salloway S, Sperling R, Fox NC, et al. Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease A BS TR AC T. *n engl j med*. 2014;370:322-355. doi:10.1056/NEJMoa1304839

156. Monsell SE, Kukull WA, Roher AE, et al. Characterizing Apolipoprotein E ε4 Carriers and Noncarriers With the Clinical Diagnosis of Mild to Moderate Alzheimer Dementia and Minimal β-Amyloid Peptide Plaques. *JAMA Neurol*. 2015;72(10):1124-1131. doi:10.1001/JAMANEUROL.2015.1721

157. Bateman RJ, Klunk WE. Measuring target effect of proposed disease-modifying therapies in Alzheimer's disease. *Neurotherapeutics*. 2008;5(3):381. doi:10.1016/J.NURT.2008.05.009

158. Cummings J, Aisen P, Apostolova LG, Atri A, Salloway S, Weiner M. Aducanumab: Appropriate Use Recommendations. *J Prev Alzheimers Dis*. 2021;8(4):398. doi:10.14283/JPAD.2021.41

159. Van Dyck CH. Lecanemab in Early Alzheimer's Disease. *NEJM*. Published online 2023.

160. McKhann G, Knopman DS, Chertkow H, et al. 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. *Alzheimers Dementia*. 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005.The

161. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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 2. *Neurology*. 1984;34(7):939-944.

162. Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurol*. 2014;13(6):614-629. doi:10.1016/S1474-4422(14)70090-0

163. Dubois B, Villain N, Frisoni GB, et al. Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol*. 2021;20(6):484-496. doi:10.1016/S1474-4422(21)00066-1

164. Jack CR, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol*. 2013;12(2):207-216. doi:10.1016/S1474-4422(12)70291-0

165. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet neurology*. 2012;11(11):1006. doi:10.1016/S1474-4422(12)70191-6

166. Livingston G, Huntley J, Sommerlad A, et al. The Lancet Commissions Dementia prevention, intervention, and care: 2020 report of the Lancet Commission The Lancet Commissions. *thelancet.com*. 2020;396:413-459. doi:10.1016/S0140-6736(20)30367-6

167. Katzman R, Terry R, DeTeresa R, et al. Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques. *Ann Neurol*. 1988;23(2):138-144. doi:10.1002/ANA.410230206

168. Nelson PT, Braak H, Markesbery WR. Neuropathology and Cognitive Impairment in Alzheimer Disease: A Complex but Coherent Relationship. *J Neuropathol Exp Neurol*. 2009;68(1):1. doi:10.1097/NEN.0B013E3181919A48

169. Aizenstein HJ, Nebes RD, Saxton JA, et al. Frequent Amyloid Deposition Without Significant Cognitive Impairment Among the Elderly. *Arch Neurol*. 2008;65(11):1509-1517. doi:10.1001/ARCHNEUR.65.11.1509

170. Brookmeyer R, Abdalla N. Estimation of lifetime risks of Alzheimer's disease dementia using biomarkers for preclinical disease. *Alzheimer's & Dementia*. 2018;14(8):981-988. doi:10.1016/J.JALZ.2018.03.005

171. Palmqvist S, Tideman P, Cullen N, et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nature Medicine 2021 27:6*. 2021;27(6):1034-1042. doi:10.1038/s41591-021-01348-z

172. Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J Psychiatr Res*. 2009;43(4):411-431. doi:10.1016/J.JPSYCHIRES.2008.04.014

173. M H. Blood tests to screen for Alzheimer's disease. *Brain*. 2021;144(2). doi:10.1093/brain/awaa462

174. Janelidze S, Stomrud E, Palmqvist S, et al. Plasma β-amyloid in Alzheimer's disease and vascular disease OPEN. *Nature Publishing Group*. Published online 2016. doi:10.1038/srep26801

175. Ovod V, Ramsey KN, Mawuenyega KG, et al. Amyloid β concentrations and stable isotope labeling kinetics of human plasma specific to central nervous system amyloidosis. *Alzheimers Dement*. 2017;13(8):841-849. doi:10.1016/J.JALZ.2017.06.2266

176. Schindler SE, Bollinger JG, Ovod V, et al. High-precision plasma β-amyloid 42/40 predicts current and future brain amyloidosis. *Neurology*. 2019;93(17):E1647-E1659. doi:10.1212/WNL.0000000000008081

177. Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-β biomarkers for Alzheimer's disease. *Nature*. 2018;554(7691):249-254. doi:10.1038/nature25456

178. Palmqvist S, Janelidze S, Stomrud E, et al. Performance of Fully Automated Plasma Assays as Screening Tests for Alzheimer Disease-Related β-Amyloid Status. *JAMA Neurol*. 2019;76(9):1060-1069. doi:10.1001/JAMANEUROL.2019.1632

179. Palmqvist S, Stomrud E, Cullen N, et al. An accurate fully automated panel of plasma biomarkers for Alzheimer's disease. *Alzheimer's & Dementia*. 2022;14:17. doi:10.1002/ALZ.12751

180. Karikari TK, Pascoal TA, Ashton NJ, et al. Blood phosphorylated tau 181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. *Lancet Neurol*. 2020;19(5):422-433. doi:10.1016/S1474-4422(20)30071-5

181. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. *Alzheimer Dis Assoc Disord*. 2009;23(4):306-314. doi:10.1097/WAD.0B013E3181A6BEBC

182. Thijssen EH, La Joie R, Wolf A, et al. Diagnostic value of plasma phosphorylated tau181 in Alzheimer's disease and frontotemporal lobar degeneration. *Nat Med*. 2020;26(3):387-397. doi:10.1038/s41591-020-0762-2

183. Janelidze S, Mattsson N, Palmqvist S, et al. Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. *Nat Med*. 2020;26(3):379-386. doi:10.1038/s41591-020-0755-1

184. Palmqvist S, Insel PS, Stomrud E, et al. Cerebrospinal fluid and plasma biomarker trajectories with increasing amyloid deposition in Alzheimer's disease. *EMBO Mol Med*. 2019;11(12). doi:10.15252/EMMM.201911170

185. Janelidze S, Palmqvist S, Leuzy A, et al. Detecting amyloid positivity in early Alzheimer's disease using combinations of plasma Aβ42/Aβ40 and p-tau. *Alzheimers Dement*. 2022;18(2):283-293. doi:10.1002/ALZ.12395

186. Tosun D, Veitch D, Aisen P, et al. Detection of β-amyloid positivity in Alzheimer's Disease Neuroimaging Initiative participants with demographics, cognition, MRI and plasma biomarkers. *Brain Commun*. 2021;3(2). doi:10.1093/BRAINCOMMS/FCAB008

187. Mielke MM, Frank RD, Dage JL, et al. Comparison of Plasma Phosphorylated Tau Species With Amyloid and Tau Positron Emission Tomography, Neurodegeneration, Vascular Pathology, and Cognitive Outcomes. *JAMA Neurol*. 2021;78(9):1108-1117. doi:10.1001/JAMANEUROL.2021.2293

188. Mielke MM, Hagen CE, Xu J, et al. Plasma phospho-tau181 increases with Alzheimer's disease clinical severity and is associated with tau- and amyloid-positron emission tomography. *Alzheimers Dement*. 2018;14(8):989-997. doi:10.1016/J.JALZ.2018.02.013

189. Mielke MM, Dage JL, Frank RD, et al. Performance of plasma phosphorylated tau 181 and 217 in the community. *Nature Medicine 2022 28:7*. 2022;28(7):1398-1405. doi:10.1038/s41591-022-01822-2

190. Cullen NC, Janelidze S, Mattsson-Carlgren N, et al. Test-retest variability of plasma biomarkers in Alzheimer's disease and its effects on clinical

prediction models. *Alzheimers Dement*. Published online March 1, 2022. doi:10.1002/ALZ.12706

191. Benedet AL, Brum WS, Hansson O, et al. The accuracy and robustness of plasma biomarker models for amyloid PET positivity. *Alzheimers Res Ther*. 2022;14(1):1-11. doi:10.1186/S13195-021-00942-0/FIGURES/3

192. Nobre AC, Stokes MG. Premembering Experience: A Hierarchy of Time-Scales for Proactive Attention. *Neuron*. Published online 2019. doi:10.1016/j.neuron.2019.08.030

193. Baddeley A. Working memory: Looking back and looking forward. *Nat Rev Neurosci*. Published online 2003. doi:10.1038/nrn1201

194. Bays PM, Gorgoraptis N, Wee N, Marshall L, Husain M. Temporal dynamics of encoding, storage, and reallocation of visual working memory. *J Vis*. Published online 2011. doi:10.1167/11.10.1

195. Nyberg L, Eriksson J. Working memory: Maintenance, updating, and the realization of intentions. *Cold Spring Harb Perspect Biol*. Published online 2016. doi:10.1101/cshperspect.a021816

196. Klencklen G, Banta Lavenex P, Brandner C, Lavenex P. Working memory decline in normal aging: Memory load and representational demands affect performance. *Learn Motiv*. Published online 2017. doi:10.1016/j.lmot.2017.09.002

197. Hasher L, Zacks RT. Working Memory, Comprehension, and Aging: A Review and a New View. *Psychology of Learning and Motivation - Advances in Research and Theory*. Published online 1988. doi:10.1016/S0079-7421(08)60041-9

198. Nyberg L, Lövdén M, Riklund K, Lindenberger U, Bäckman L. Memory aging and brain maintenance. *Trends Cogn Sci*. Published online 2012. doi:10.1016/j.tics.2012.04.005

199. Biundo R, Weis L, Antonini A. Cognitive decline in Parkinson's disease: the complex picture. *NPJ Parkinsons Dis*. Published online 2016. doi:10.1038/npjparkd.2016.18

200. Peters B, Rahm B, Kaiser J, Bledowski C. Differential trajectories of memory quality and guessing across sequential reports from working memory. *J Vis*. Published online 2019. doi:10.1167/19.7.3

201. Jost K, Mayr U. Switching between filter settings reduces the efficient utilization of visual working memory. *Cogn Affect Behav Neurosci*. Published online 2016. doi:10.3758/s13415-015-0380-5

202. Tabi YA, Maio MR, Fallon SJ, et al. Impact of processing demands at encoding, maintenance and retrieval in visual working memory. *Cognition*. 2021;214. doi:10.1016/J.COGNITION.2021.104758

203. Geng JJ, Won BY, Carlisle NB. Distractor Ignoring: Strategies, Learning, and Passive Filtering. *Curr Dir Psychol Sci*. Published online 2019. doi:10.1177/0963721419867099

204. Fallon SJ, Mattiesing RM, Dolfen N, Manohar SG, Husain M. Ignoring versus updating in working memory reveal differential roles of attention and feature binding. *Cortex*. 2018;107:50-63. doi:10.1016/j.cortex.2017.12.016

205. Lin PH, Luck SJ. The influence of similarity on visual working memory representations. *Vis cogn*. 2009;17(3):356-372. doi:10.1080/13506280701766313

206. Oberauer K. Working Memory and Attention – A Conceptual Analysis and Review. *J Cogn*. 2019;2(1):1-23. doi:10.5334/joc.58

207. Cools R, Sheridan M, Jacobs E, D'Esposito M. Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *Journal of Neuroscience*. Published online 2007. doi:10.1523/JNEUROSCI.0601-07.2007

208. Minamoto T, Osaka M, Osaka N. Individual differences in working memory capacity and distractor processing: possible contribution of top-down inhibitory control. *Brain Res*. 2010;1335:63-73. doi:10.1016/J.BRAINRES.2010.03.088

209. Berron D, Frühholz S, Herrmann M. Neural control of enhanced filtering demands in a combined Flanker and Garner conflict task. *PLoS One*. 2015;10(3). doi:10.1371/JOURNAL.PONE.0120582

210. Ranganath C, D'Esposito M. Medial temporal lobe activity associated with active maintenance of novel information. *Neuron*. 2001;31(5):865-873. doi:10.1016/S0896-6273(01)00411-1

211. Stevenson RF, Zheng J, Mnatsakanyan L, et al. Hippocampal CA1 gamma power predicts the precision of spatial memory judgments. *Proc Natl Acad Sci U S A*. 2018;115(40):10148-10153. doi:10.1073/pnas.1805724115

212. Parra MA, Abrahams S, Logie RH, Méndez LG, Lopera F, Della Sala S. Visual short-term memory binding deficits in familial Alzheimer's disease. *Brain*. Published online 2010. doi:10.1093/brain/awq148

213. Fernandez-Duque D, Black SE. Selective attention in early Dementia of Alzheimer Type. Published online 2007. doi:10.1016/j.bandc.2007.08.003

214. Bélanger S, Belleville S, Gauthier S. Inhibition impairments in Alzheimer's disease, mild cognitive impairment and healthy aging: Effect of congruency proportion in a Stroop task. *Neuropsychologia*. 2010;48(2):581-590. doi:10.1016/j.neuropsychologia.2009.10.021

215. Lee EY, Cowan N, Vogel EK, Rolan T, Valle-Inclán F, Hackley SA. Visual working memory deficits in patients with Parkinson's disease are due to both reduced storage capacity and impaired ability to filter out irrelevant information. *Brain*. Published online 2010. doi:10.1093/brain/awq197

216. Zokaei N, Mcneill A, Proukakis C, et al. Visual short-term memory deficits associated with GBA mutation and Parkinson's disease. *Brain*. Published online 2014. doi:10.1093/brain/awu143

217. Chun MM, Turk-Browne NB. Interactions between attention and memory. *Curr Opin Neurobiol*. 2007;17(2):177-184. doi:10.1016/j.conb.2007.03.005

218. Aly M, Turk-Browne NB. Attention promotes episodic encoding by stabilizing hippocampal representations. *Proc Natl Acad Sci U S A*. 2016;113(4):E420-E429. doi:10.1073/pnas.1518931113

219. Jeneson A, Mauldin KN, Hopkins RO, Squire LR. The role of the hippocampus in retaining relational information across short delays: The importance of memory load. *Learning and Memory*. 2011;18(5):301-305. doi:10.1101/lm.2010711

220. Wambach D, Lamar M, Swenson R, Penney DL, Kaplan E, Libon DJ. Digit Span. In: *Encyclopedia of Clinical Neuropsychology*. ; 2011. doi:10.1007/978-0-387-79948-3_1288

221. Ang YS, Lockwood P, Apps MAJ, Muhammed K, Husain M. Distinct subtypes of apathy revealed by the apathy motivation index. *PLoS One*. Published online 2017. doi:10.1371/journal.pone.0169938

222. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry*. Published online 1961.

223. Skapinakis P. Hospital Anxiety and Depression Scale (HADS). In: *Encyclopedia of Quality of Life and Well-Being Research*. ; 2014. doi:10.1007/978-94-007-0753-5_1315

224. Snaith RP, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P. A scale for the assessment of hedonic tone. The Snaith-Hamilton Pleasure Scale. *British Journal of Psychiatry*. Published online 1995. doi:10.1192/bjp.167.1.99

225. Lesher EL, Berryhill JS. Validation of the geriatric depression scale-short form among inpatients. *J Clin Psychol*. Published online 1994. doi:10.1002/1097-4679(199403)50:2<256::AID-JCLP2270500218>3.0.CO;2-E

226. Lerdal A. Fatigue Severity Scale. In: *Encyclopedia of Quality of Life and Well-Being Research*. ; 2014. doi:10.1007/978-94-007-0753-5_1018

227. Lee KA, Hicks G, Nino-Murcia G. Validity and reliability of a scale to assess fatigue. *Psychiatry Res*. Published online 1991. doi:10.1016/0165-1781(91)90027-M

228. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res*. Published online 1989. doi:10.1016/0165-1781(89)90047-4

229. World Health Organisation. Wellbeing Measures in Primary Health Care/ The Depcare Project. *Report on a WHO Meeting*. Published online 1998.

230. Cantril H. Pattern of Human Concerns Data, 1957-1963. *Rutgers University Press*. Published online 2006. doi:10.3886/ICPSR07023.v1

231. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage*. 2011;56(3):907-922. doi:10.1016/J.NEUROIMAGE.2011.02.046

232. Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage*. 2002;17(1):479-489. doi:10.1006/nimg.2002.1040

233. Postuma RB, Berg D, Adler CH, et al. The new definition and diagnostic criteria of Parkinson's disease. *Lancet Neurol*. Published online 2016. doi:10.1016/S1474-4422(16)00116-2

234. Cowan N, Morey CC. Visual working memory depends on attentional filtering. *Trends Cogn Sci*. Published online 2006. doi:10.1016/j.tics.2006.02.001

235. Finke K, Myers N, Bublak P, Sorg C. A biased competition account of attention and memory in alzheimer's disease. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Published online 2013. doi:10.1098/rstb.2013.0062

236. Lenoble Q, Corveleyn X, Szaffarczyk S, Pasquier F, Boucart M. Attentional capture by incongruent object/background scenes in patients with Alzheimer disease. *Cortex*. Published online 2018. doi:10.1016/j.cortex.2018.06.002

237. Peltsch A, Hemraj A, Garcia A, Munoz DP. Saccade deficits in amnestic mild cognitive impairment resemble mild Alzheimer's disease. *European Journal of Neuroscience*. Published online 2014. doi:10.1111/ejn.12617

238. Ben-David BM, Tewari A, Shakuf V, Van Lieshout PHHM. Stroop effects in Alzheimer's disease: Selective attention speed of processing, or color-naming? A meta-analysis. *Journal of Alzheimer's Disease*. Published online 2014. doi:10.3233/JAD-131244

239. Owen AM, Roberts AC, Hodges JR, Robbins TW. Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or parkinson's disease. *Brain*. Published online 1993. doi:10.1093/brain/116.5.1159

240. Fallon SJ, Gowell M, Maio MR, Husain M. Dopamine affects short-term memory corruption over time in Parkinson's disease. *NPJ Parkinsons Dis*. Published online 2019. doi:10.1038/s41531-019-0088-2

241. Feng J, Craik FIM, Levine B, Moreno S, Naglie G, Choi HS. Differential age-related changes in localizing a target among distractors across an extended visual field. *Eur J Ageing*. Published online 2017. doi:10.1007/s10433-016-0399-7

242. Gruber N, Müri RM, Mosimann UP, et al. Effects of age and eccentricity on visual target detection. *Front Aging Neurosci*. Published online 2014. doi:10.3389/fnagi.2013.00101

243. Caterini F, Della Sala S, Spinnler H, Stangalino C, Turnbull OH. Object recognition and object orientation in Alzheimer's disease. *Neuropsychology*. Published online 2002. doi:10.1037/0894-4105.16.2.146

244. Zokaei N, Nour MM, Sillence A, et al. Binding deficits in visual short-term memory in patients with temporal lobe lobectomy. *Hippocampus*. Published online 2019. doi:10.1002/hipo.22998

245. Watson PD, Voss JL, Warren DE, Tranel D, Cohen NJ. Spatial reconstruction by patients with hippocampal damage is dominated by relational memory errors. *Hippocampus*. 2013;23(7):570. doi:10.1002/HIPO.22115

246. Mackey WE, Devinsky O, Doyle WK, Golfinos JG, Curtis CE. Human parietal cortex lesions impact the precision of spatial working memory. *J Neurophysiol*. Published online 2016. doi:10.1152/jn.00380.2016

247. James W. The principles of psychology, Vol I. *The principles of psychology, Vol I*. Published online October 27, 2004. doi:10.1037/10538-000

248. SCOVILLE WB, MILNER B. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry*. Published online 1957. doi:10.1136/jnnp.20.1.11

249. Sutherland RJ, Rudy JW. Configural association theory: The role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology*. Published online 1989. doi:10.3758/BF03337828

250. Shrager Y, Levy DA, Hopkins RO, Squire LR. Working memory and the organization of brain systems. *Journal of Neuroscience*. Published online 2008. doi:10.1523/JNEUROSCI.0710-08.2008

251. Warren DE, Duff MC, Tranel D, Cohen NJ. Observing degradation of visual representations over short intervals when medial temporal lobe is damaged. *J Cogn Neurosci*. Published online 2011. doi:10.1162/jocn_a_00089

252. Parra MA, Fabi K, Luzzi S, Cubelli R, Hernandez Valdez M, Della Sala S. Relational and conjunctive binding functions dissociate in short-term memory. *Neurocase*. Published online 2015. doi:10.1080/13554794.2013.860177

253. Bateman RJ, Xiong C, Benzinger TLS, et al. Clinical and Biomarker Changes in Dominantly Inherited Alzheimer's Disease. *New England Journal of Medicine*. 2012;367(9):795-804. doi:10.1056/NEJMOA1202753/SUPPL_FILE/NEJMOA1202753_DISCLOSURES.PDF

254. Dementia. Accessed September 6, 2022. https://www.who.int/news-room/fact-sheets/detail/dementia

255. Zokaei N, Husain M. Working Memory in Alzheimer's Disease and Parkinson's Disease. In: *Current Topics in Behavioral Neurosciences*. ; 2019. doi:10.1007/7854_2019_103

256. Schuff N, Woerner N, Boreta L, et al. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain*. 2009;132(Pt 4):1067-1077. doi:10.1093/BRAIN/AWP007

257. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. *Lancet Neurol*. 2020;19(3):271. doi:10.1016/S1474-4422(19)30368-0

258. Petersen RC, Caracciolo B, Brayne C, Gauthier S, Jelic V, Fratiglioni L. Mild cognitive impairment: A concept in evolution. *J Intern Med*. 2014;275(3):214-228. doi:10.1111/joim.12190

259. Morel P. Gramm: grammar of graphics plotting in Matlab. *J Open Source Softw*. 2018;3(23):568. doi:10.21105/JOSS.00568

260. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Front Psychol*. 2013;4(NOV). doi:10.3389/FPSYG.2013.00863/ABSTRACT

261. Pavisic IM, Nicholas JM, Pertzov Y, et al. Visual short-term memory impairments in presymptomatic familial Alzheimer's disease: A longitudinal observational study. *Neuropsychologia*. 2021;162. doi:10.1016/J.NEUROPSYCHOLOGIA.2021.108028

262. Gorwn B, Carson K. The Basis for Choice Reaction Time Slowing in Alzheimer's Disease. *Brain Cogn*. 1990;13:148-166.

263. Chen KC, Weng CY, Hsiao S, Tsao WL, Koo M. Cognitive decline and slower reaction time in elderly individuals with mild cognitive impairment. *Psychogeriatrics*. 2017;17(6):364-370. doi:10.1111/PSYG.12247

264. van Deursen JA, Vuurman EFPM, Smits LL, Verhey FRJ, Riedel WJ. Response speed, contingent negative variation and P300 in Alzheimer's disease and MCI. *Brain Cogn*. 2009;69(3):592-599. doi:10.1016/J.BANDC.2008.12.007

265. Andriuta D, Diouf M, Roussel M, Godefroy O. Is Reaction Time Slowing an Early Sign of Alzheimer's Disease? A Meta-Analysis. *Dement Geriatr Cogn Disord*. 2019;47(4-6):281-288. doi:10.1159/000500348

266. Cortes C, Vapnik V, Saitta L. Support-vector networks. *Machine Learning 1995 20:3*. 1995;20(3):273-297. doi:10.1007/BF00994018

267. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*. 1988;44(3):837. doi:10.2307/2531595

268. Fox NC, Warrington EK, Freeborough PA, et al. Presymptomatic hippocampal atrophy in Alzheimer's disease. A longitudinal MRI study. *Brain*. 1996;119 (Pt 6)(6):2001-2007. doi:10.1093/BRAIN/119.6.2001

269. Ridha BH, Barnes J, Bartlett JW, et al. Tracking atrophy progression in familial Alzheimer's disease: a serial MRI study. *Lancet Neurol*. 2006;5(10):828-834. doi:10.1016/S1474-4422(06)70550-6

270. Roe JM, Vidal-Piñeiro D, Sørensen Ø, et al. Asymmetric thinning of the cerebral cortex across the adult lifespan is accelerated in Alzheimer's disease. *Nature Communications 2021 12:1*. 2021;12(1):1-11. doi:10.1038/s41467-021-21057-y

271. Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, Shulman GL. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat Neurosci*. 2000;3(3):292-297. doi:10.1038/73009

272. Diedenhofen B, Musch J. cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS One*. 2015;10(4). doi:10.1371/JOURNAL.PONE.0121945

273. Wager TD, Smith EE. Neuroimaging studies of working memory: a meta-analysis. *Cogn Affect Behav Neurosci*. 2003;3(4):255-274. doi:10.3758/CABN.3.4.255

274. Rottschy C, Langner R, Dogan I, et al. Modelling neural correlates of working memory: a coordinate-based meta-analysis. *Neuroimage*. 2012;60(1):830-846. doi:10.1016/J.NEUROIMAGE.2011.11.050

275. Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes JD. The Distributed Nature of Working Memory. *Trends Cogn Sci*. 2017;21(2):111-124. doi:10.1016/J.TICS.2016.12.007

276. Nee DE, Brown JW, Askren MK, et al. A meta-analysis of executive components of working memory. *Cereb Cortex*. 2013;23(2):264-282. doi:10.1093/CERCOR/BHS007

277. Haynes JD, Rees G. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience 2006 7:7*. 2006;7(7):523-534. doi:10.1038/nrn1931

278. Postle BR. Working Memory as an Emergent Property of the Mind and Brain. *Neuroscience*. 2006;139(1):23. doi:10.1016/J.NEUROSCIENCE.2005.06.005

279. Jonides J, Lewis RL, Nee DE, Lustig CA, Berman MG, Moore KS. The Mind and Brain of Short-Term Memory. Published online 2008. doi:10.1146/annurev.psych.59.103006.093615

280. Gazzaley A, Nobre AC. Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci*. 2012;16(2):129-135. doi:10.1016/J.TICS.2011.11.014

281. Bledowski C, Kaiser J, Rahm B. Basic operations in working memory: Contributions from functional imaging studies. *Behavioural Brain Research*. 2010;214:172-179. doi:10.1016/j.bbr.2010.05.041

282. Collette F, Hogge M, Salmon E, Van der Linden M. Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience*. 2006;139(1):209-221. doi:10.1016/J.NEUROSCIENCE.2005.05.035

283. O'Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput*. 2006;18(2):283-328. doi:10.1162/089976606775093909

284. Wilson FAW, Scalaidhe SPÓ, Goldman-Rakic PS. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*. 1993;260(5116):1955-1958. doi:10.1126/SCIENCE.8316836

285. Smith EE, Jonides J. Storage and executive processes in the frontal lobes. *Science (1979)*. 1999;283(5408):1657-1661. doi:10.1126/SCIENCE.283.5408.1657

286. Levy R, Goldman-Rakic PS. Segregation of working memory functions within the dorsolateral prefrontal cortex. *Exp Brain Res*. 2000;133(1):23-32. doi:10.1007/S002210000397

287. Mishkin M, Ungerleider LG, Macko KA. Object vision and spatial vision:two cortical p hways.

288. Owen AM. The role of the lateral frontal cortex in mnemonic processing: the contribution of functional neuroimaging. *Exp Brain Res*. 2000;133(1):33-43. doi:10.1007/S002210000398

289. Ranganath C, Heller A, Cohen MX, Brozinsky CJ, Rissman J. RAPID COMMUNICATION Functional Connectivity With the Hippocampus During Successful Memory Formation. doi:10.1002/hipo.20141

290. Piekema C, Rijpkema M, Ferná Ndez GN, Kessels RPC. Dissociating the Neural Correlates of Intra-Item and Inter-Item Working-Memory Binding. doi:10.1371/journal.pone.0010214

291. Piekema C, Kessels RPC, Mars RB, Petersson KM, Fernández G. The right hippocampus participates in short-term memory maintenance of object-location associations. *Neuroimage*. 2006;33(1):374-382. doi:10.1016/J.NEUROIMAGE.2006.06.035

292. Berlingeri M, Bottini G, Basilico S, et al. Anatomy of the episodic buffer: a voxel-based morphometry study in patients with dementia. *Behavioural neurology*. 2008;19(1-2):29-34. doi:10.1155/2008/828937

293. Ae MR, Rönnberg J. The role of the episodic buffer in working memory for language processing. doi:10.1007/s10339-007-0183-x

294. Ranganath C, Cohen MX, Brozinsky CJ. Working memory maintenance contributes to long-term memory formation: neural and behavioral evidence. *J Cogn Neurosci*. 2005;17(7):994-1010. doi:10.1162/0898929054475118

295. Schon K, Hasselmo ME, LoPresti ML, Tricarico MD, Stern CE. Persistence of parahippocampal representation in the absence of stimulus input enhances long-term encoding: a functional magnetic resonance imaging study of subsequent memory after a delayed match-to-sample task. *J Neurosci*. 2004;24(49):11088-11097. doi:10.1523/JNEUROSCI.3807-04.2004

296. Baumgartner F, Hanke M, Geringswald F, Zinke W, Speck O, Pollmann S. Evidence for feature binding in the superior parietal lobule. *Neuroimage*. 2013;68:173-180. doi:10.1016/J.NEUROIMAGE.2012.12.002

297. Shafritz KM, Gore JC, Marois R. The role of the parietal cortex in visual feature binding. *Proc Natl Acad Sci U S A*. 2002;99(16):10917-10922. doi:10.1073/PNAS.152694799

298. Xu Y, Chun MM. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*. 2006;440(7080):91-95. doi:10.1038/NATURE04262

299. Parra MA, Della Sala S, Logie RH, Morcom AM. Neural correlates of shape-color binding in visual working memory. *Neuropsychologia*. 2014;52(1):27-36. doi:10.1016/J.NEUROPSYCHOLOGIA.2013.09.036

300. Zhao Y, Kuai S, Zanto TP, Ku Y. Neural Correlates Underlying the Precision of Visual Working Memory. *Neuroscience*. 2020;425:301-311. doi:10.1016/J.NEUROSCIENCE.2019.11.037

301. Paraskevopoulou SE, Coon WG, Brunner P, Miller KJ, Schalk G. Within-subject reaction time variability: Role of cortical networks and underlying neurophysiological mechanisms. *Neuroimage*. 2021;237:118127. doi:10.1016/J.NEUROIMAGE.2021.118127

302. Lövdén M, Schmiedek F, Kennedy KM, Rodrigue KM, Lindenberger U, Raz N. Does variability in cognitive performance correlate with frontal brain volume? *Neuroimage*. 2013;64(1):209-215. doi:10.1016/J.NEUROIMAGE.2012.09.039

303. Bunce D, Anstey KJ, Christensen H, Dear K, Wen W, Sachdev P. White matter hyperintensities and within-person variability in community-dwelling adults aged 60-64 years. *Neuropsychologia*. 2007;45(9):2009-2015. doi:10.1016/J.NEUROPSYCHOLOGIA.2007.02.006

304. Jackson JD, Balota DA, Duchek JM, Head D. White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer disease. *Neuropsychologia*. 2012;50(3):357-366. doi:10.1016/J.NEUROPSYCHOLOGIA.2011.11.024

305. Lu K, Nicholas JM, James SN, et al. Increased variability in reaction time is associated with amyloid beta pathology at age 70. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*. 2020;12(1). doi:10.1002/DAD2.12076

306. Pavisic IM, Suarez-Gonzalez A, Pertzov Y. Translating Visual Short-Term Memory Binding Tasks to Clinical Practice: From Theory to Practice. *Front Neurol*. 2020;11:458. doi:10.3389/FNEUR.2020.00458/BIBTEX

307. Frisoni GB, Testa C, Zorzan A, Sabattoli F, Beltramello A, Soininen H. Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. *J Neurol Neurosurg Psychiatry*. 2002;73:657-664.

308. McCarthy J, Collins DL, Ducharme S. Morphometric MRI as a diagnostic biomarker of frontotemporal dementia: A systematic review to determine clinical applicability. *Neuroimage Clin*. 2018;20:685-696. doi:10.1016/J.NICL.2018.08.028

309. Dinomais M, Celle S, Duval GT, et al. Anatomic Correlation of the Mini-Mental State Examination: A Voxel-Based Morphometric Study in Older Adults. Published online 2016. doi:10.1371/journal.pone.0162889

310. Nho K, Risacher SL, Crane PK, et al. Voxel and Surface-Based Topography of Memory and Executive Deficits in Mild Cognitive Impairment and Alzheimer's Disease. *Brain Imaging Behav*. 2012;6(4):551. doi:10.1007/S11682-012-9203-2

311. Miller KL, Alfaro-Almagro F, Bangerter NK, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci*. 2016;19(11):1523. doi:10.1038/NN.4393

312. Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage*. 2018;166:400. doi:10.1016/J.NEUROIMAGE.2017.10.034

313. Douaud G, Smith S, Jenkinson M, et al. Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain*. 2007;130(Pt 9):2375-2386. doi:10.1093/BRAIN/AWM184

314. Good CD, Johnsrude IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*. 2001;14(1 Pt 1):21-36. doi:10.1006/NIMG.2001.0786

315. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23 Suppl 1(SUPPL. 1). doi:10.1016/J.NEUROIMAGE.2004.07.051

316. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17(3):143-155. doi:10.1002/hbm.10062

317. Andersson JLR, Jenkinson M, Smith S. Non-linear registration aka Spatial normalisation FMRIB Technial Report TR07JA2. Published online 2007.

318. Smith SM, Nichols TE. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*. 2009;44(1):83-98. doi:10.1016/j.neuroimage.2008.03.061

319. Japee S, Holiday K, Satyshur MD, Mukai I, Ungerleider LG. A role of right middle frontal gyrus in reorienting of attention: A case study. *Front Syst Neurosci*. 2015;9(MAR):23. doi:10.3389/FNSYS.2015.00023/ABSTRACT

320. Ranganath C, Cohen MX, Dam C, Wheeler Jr Brain HH. Behavioral/Systems/Cognitive Inferior Temporal, Prefrontal, and Hippocampal Contributions to Visual Working Memory Maintenance and

Associative Memory Retrieval. Published online 2004. doi:10.1523/JNEUROSCI.5053-03.2004

321. Bakkour A, Morris JC, Dickerson BC. The cortical signature of prodromal AD: Regional thinning predicts mild AD dementia. *Neurology*. 2009;72(12):1048. doi:10.1212/01.WNL.0000340981.97664.2F

322. Aggleton JP, Pralus A, Nelson AJD, Hornberger M. Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to Papez circuit. *Brain*. 2016;139(Pt 7):1877-1890. doi:10.1093/BRAIN/AWW083

323. Schaefer A, Braver TS, Reynolds JR, Burgess GC, Yarkoni T, Gray JR. Individual Differences in Amygdala Activity Predict Response Speed during Working Memory. *Journal of Neuroscience*. 2006;26(40):10120-10128. doi:10.1523/JNEUROSCI.2567-06.2006

324. Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci*. 2000;23(10):475-483. doi:10.1016/S0166-2236(00)01633-7

325. Baker SC, Rogers RD, Owen AM, et al. Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia*. 1996;34(6):515-526. doi:10.1016/0028-3932(95)00133-6

326. Johnson BP, Pinar A, Fornito A, Nandam LS, Hester R, Bellgrove MA. Left anterior cingulate activity predicts intra-individual reaction time variability in healthy adults. *Neuropsychologia*. 2015;72:22-26. doi:10.1016/J.NEUROPSYCHOLOGIA.2015.03.015

327. Fornito A, Yücel M, Wood S, et al. Individual Differences in Anterior Cingulate/Paracingulate Morphology Are Related to Executive Functions in

Healthy Males. *Cerebral Cortex*. 2004;14(4):424-431. doi:10.1093/CERCOR/BHH004

328. Ramnani N, Owen AM. Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat Rev Neurosci*. 2004;5(3):184-194. doi:10.1038/NRN1343

329. Semendeferi K, Damasio H, Frank R, Van Hoesen GW. The evolution of the frontal lobes: A volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *J Hum Evol*. 1997;32(4):375-388. doi:10.1006/jhev.1996.0099

330. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci*. 2001;24:167-202. doi:10.1146/ANNUREV.NEURO.24.1.167

331. Bludau S, Eickhoff SB, Mohlberg H, et al. Cytoarchitecture, probability maps and functions of the human frontal pole. *Neuroimage*. 2014;93:260-275. doi:10.1016/J.NEUROIMAGE.2013.05.052

332. Gilbert SJ, Spengler S, Simons JS, et al. Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J Cogn Neurosci*. 2006;18(6):932-948. doi:10.1162/JOCN.2006.18.6.932

333. Volle E, Gonen-Yaacovi G, de Lacy Costello A, Gilbert SJ, Burgess PW. The role of rostral prefrontal cortex in prospective memory: A voxel-based lesion study. *Neuropsychologia*. 2011;49(8):2185-2198. doi:10.1016/J.NEUROPSYCHOLOGIA.2011.02.045

334. Cajanus A, Solje E, Koikkalainen J, et al. The Association Between Distinct Frontal Brain Volumes and Behavioral Symptoms in Mild Cognitive

Impairment, Alzheimer's Disease, and Frontotemporal Dementia. *Front Neurol*. 2019;10:1059. doi:10.3389/FNEUR.2019.01059/FULL

335. Cavanna AE, Trimble MR. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*. 2006;129(3):564-583. doi:10.1093/BRAIN/AWL004

336. Vincent JL, Snyder AZ, Fox MD, et al. Coherent Spontaneous Activity Identifies a Hippocampal-Parietal Memory Network. *Shannon and Buckner*. 2006;96:3517-3531. doi:10.1152/jn.00048.2006

337. Karas G, Scheltens P, Rombouts S, et al. Precuneus atrophy in early-onset Alzheimer's disease: a morphometric structural MRI study. *Neuroradiology*. 2007;49(12):967-976. doi:10.1007/S00234-007-0269-2

338. Murtha S, Cismaru R, Waechter R, Chertkow H. Increased variability accompanies frontal lobe damage in dementia. *J Int Neuropsychol Soc*. 2002;8(3):360-372. doi:10.1017/S1355617702813170

339. Stuss DT, Murphy KJ, Binns MA, Alexander MP. Staying on the job: the frontal lobes control individual performance variability. *Brain*. 2003;126(11):2363-2380. doi:10.1093/BRAIN/AWG237

340. Jack C, Knopman D, Jagust W, et al. O3-03-01: Update on hypothetical model of Alzheimer's disease biomarkers. *Alzheimer's & Dementia*. Published online 2013. doi:10.1016/j.jalz.2013.04.248

341. Le Bihan D. Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci*. 2003;4(6):469-480. doi:10.1038/NRN1119

342. Giulietti G, Torso M, Serra L, et al. Whole brain white matter histogram analysis of diffusion tensor imaging data detects microstructural damage in mild cognitive impairment and alzheimer's disease patients. *Journal of*

Magnetic Resonance Imaging. Published online 2018:1-13. doi:10.1002/jmri.25947

343. Scola E, Bozzali M, Agosta F, et al. A diffusion tensor MRI study of patients with MCI and AD with a 2-year clinical follow-up. *J Neurol Neurosurg Psychiatry*. 2010;81(7):798-805. doi:10.1136/jnnp.2009.189639

344. Gold BT, Johnson NF, Powell DK, Smith CD. White matter integrity and vulnerability to Alzheimer's disease: preliminary findings and future directions. *Biochim Biophys Acta*. 2012;1822(3):416-422. doi:10.1016/j.bbadis.2011.07.009

345. Le Bihan D, Mangin JF, Poupon C, et al. Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*. 2001;13(4):534-546. doi:10.1002/JMRI.1076

346. Beaulieu C. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed*. 2002;15(7-8):435-455. doi:10.1002/NBM.782

347. Bozzali M, Serra L, Cercignani M. Quantitative MRI to understand Alzheimer's disease pathophysiology. *Curr Opin Neurol*. 2016;29(4):437-444. doi:10.1097/WCO.0000000000000345

348. Bozzali M, Giulietti G, Basile B, et al. Damage to the cingulum contributes to alzheimer's disease pathophysiology by deafferentation mechanism. *Hum Brain Mapp*. 2012;33(6):1295. doi:10.1002/HBM.21287

349. Bronge L, Bogdanovic N, Wahlund LO. Postmortem MRI and histopathology of white matter changes in Alzheimer brains: A quantitative, comparative study. *Dement Geriatr Cogn Disord*. 2002;13(4):205-212. doi:10.1159/000057698

350. Lee P, Ryoo H, Park J, Jeong Y. Morphological and microstructural changes of the hippocampus in early MCI: A study utilizing the alzheimer's disease neuroimaging initiative database. *Journal of Clinical Neurology (Korea)*. 2017;13(2):144-154. doi:10.3988/jcn.2017.13.2.144

351. Nir TM, Jahanshad N, Villalon-Reina JE, et al. Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging. *Neuroimage Clin*. Published online 2013. doi:10.1016/j.nicl.2013.07.006

352. Nir TM, Jahanshad N, Villalon-Reina JE, et al. Fractional anisotropy derived from the diffusion tensor distribution function boosts power to detect Alzheimer's disease deficits. *Magn Reson Med*. 2017;78(6):2322-2333. doi:10.1002/mrm.26623

353. Solodkin A, Chen EE, Van Hoesen GW, et al. In vivo parahippocampal white matter pathology as a biomarker of disease progression to Alzheimer's disease. *Journal of Comparative Neurology*. 2013;521(18):4300-4317. doi:10.1002/cne.23418

354. Douaud G, Menke RAL, Gass A, et al. Brain Microstructure Reveals Early Abnormalities more than Two Years prior to Clinical Progression from Mild Cognitive Impairment to Alzheimer's Disease. *Journal of Neuroscience*. Published online 2013. doi:10.1523/JNEUROSCI.4437-12.2013

355. Sexton CE, Kalu UG, Filippini N, Mackay CE, Ebmeier KP. A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging*. 2011;32(12). doi:10.1016/j.neurobiolaging.2010.05.019

356. Acosta-Cabronero J, Nestor PJ. Diffusion tensor imaging in Alzheimer's disease: insights into the limbic-diencephalic network and methodological

considerations. *Front Aging Neurosci*. 2014;6(SEP). doi:10.3389/FNAGI.2014.00266

357. ADNI | Alzheimer's Disease Neuroimaging Initiative. Accessed March 4, 2023. https://adni.loni.usc.edu/

358. Mayo CD, Garcia-Barrera MA, Mazerolle EL, Ritchie LJ, Fisk JD, Gawryluk JR. Relationship Between DTI Metrics and Cognitive Function in Alzheimer's Disease. *Front Aging Neurosci*. 2019;10(JAN). doi:10.3389/FNAGI.2018.00436

359. Parra MA, Saarimäki H, Bastin ME, et al. Memory binding and white matter integrity in familial Alzheimer's disease. *Brain*. 2015;138(Pt 5):1355-1369. doi:10.1093/BRAIN/AWV048

360. Smith SM, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*. 2006;31(4):1487-1505. doi:10.1016/J.NEUROIMAGE.2006.02.024

361. Liu Y, Spulber G, Lehtimäki KK, et al. Diffusion tensor imaging and tract-based spatial statistics in Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging*. 2011;32(9):1558-1571. doi:10.1016/J.NEUROBIOLAGING.2009.10.006

362. Acosta-Cabronero J, Alley S, Williams GB, Pengas G, Nestor PJ. Diffusion tensor metrics as biomarkers in Alzheimer's disease. *PLoS One*. 2012;7(11):e49072. doi:10.1371/journal.pone.0049072

363. Lieberman MD, Cunningham WA. Type I and Type II error concerns in fMRI research: re-balancing the scale. *Soc Cogn Affect Neurosci*. 2009;4(4):423. doi:10.1093/SCAN/NSP052

364. Warrington S, Bryant KL, Khrapitchev AA, et al. XTRACT - Standardised protocols for automated tractography in the human and macaque brain. *Neuroimage*. 2020;217:116923. doi:10.1016/J.NEUROIMAGE.2020.116923

365. Van Essen DC, Ugurbil K, Auerbach E, et al. The Human Connectome Project: A data acquisition perspective. *Neuroimage*. 2012;62(4):2222. doi:10.1016/J.NEUROIMAGE.2012.02.018

366. Mori S. MRI Atlas of Human White Matter. *AJNR Am J Neuroradiol*. 2006;27(6):1384.

367. Bozzali M. White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *J Neurol Neurosurg Psychiatry*. Published online 2002. doi:10.1136/jnnp.72.6.742

368. Amlien IK, Fjell AM. Diffusion tensor imaging of white matter degeneration in Alzheimer's disease and mild cognitive impairment. *Neuroscience*. 2014;276:206-215. doi:10.1016/J.NEUROSCIENCE.2014.02.017

369. Sherbondy AJ, Dougherty RF, Napel S, Wandell BA. Identifying the human optic radiation using diffusion imaging and fiber tractography. *J Vis*. 2008;8(10):12-12. doi:10.1167/8.10.12

370. Wen Q, Mustafi SM, Li J, et al. White matter alterations in early-stage Alzheimer's disease: A tract-specific study. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. 2019;11:576-587. doi:10.1016/J.DADM.2019.06.003

371. Lugtmeijer S, Schneegans S, Lammers NA, et al. Consequence of stroke for feature recall and binding in visual working memory. *Neurobiol Learn Mem*. 2021;179:107387. doi:10.1016/J.NLM.2021.107387

372. Caballero MÁA, Suárez-Calvet M, Duering M, et al. White matter diffusion alterations precede symptom onset in autosomal dominant Alzheimer's disease. *Brain*. 2018;141(10):3065-3080. doi:10.1093/BRAIN/AWY229

373. Makris N, Papadimitriou GM, Kaiser JR, Sorg S, Kennedy DN, Pandya DN. Delineation of the Middle Longitudinal Fascicle in Humans: A Quantitative, In Vivo, DT-MRI Study. *Cerebral Cortex (New York, NY)*. 2009;19(4):777. doi:10.1093/CERCOR/BHN124

374. Menjot De Champfleur N, Lima Maldonado I, Moritz-Gasser S, et al. Middle longitudinal fasciculus delineation within language pathways: A diffusion tensor imaging study in human. *Eur J Radiol*. 2013;82:151-157. doi:10.1016/j.ejrad.2012.05.034

375. Forkel SJ, Thiebaut De Schotten M, Kawadler JM, Dell'acqua F, Danek A, Catani M. The anatomy of fronto-occipital connections from early blunt dissections to contemporary tractography. doi:10.1016/j.cortex.2012.09.005

376. Peters BD, Ikuta T, Derosse P, et al. Age-Related Differences In White Matter Tract Microstructure Are Associated With Cognitive Performance From Childhood to Adulthood. Published online 2013. doi:10.1016/j.biopsych.2013.05.020

377. Walsh M, Montojo CA, Sheu yi S, et al. Object Working Memory Performance Depends on Microstructure of the Frontal-Occipital Fasciculus. *Brain Connect*. 2011;1(4):317. doi:10.1089/BRAIN.2011.0037

378. Yeatman JD, Weiner KS, Pestilli F, Rokem A, Mezer A, Well BA. The vertical occipital fasciculus: A century of controversy resolved by in vivo measurements. *Proc Natl Acad Sci U S A*. 2014;111(48):E5214. doi:10.1073/PNAS.1418503111

379. Makris N, Kennedy DN, Mcinerney S, et al. Segmentation of Subcomponents within the Superior Longitudinal Fascicle in Humans: A Quantitative, In Vivo, DT-MRI Study. doi:10.1093/cercor/bhh186

380. Parlatini V, Radua J, Dell'Acqua F, et al. Functional segregation and integration within fronto-parietal networks. *Neuroimage*. 2017;146:367-375. doi:10.1016/J.NEUROIMAGE.2016.08.031

381. Leiguarda RC, Marsden CD. Limb apraxiasHigher-order disorders of sensorimotor integration. *Brain*. 2000;123(5):860-879. doi:10.1093/BRAIN/123.5.860

382. De Schotten MT, Dell'Acqua F, Forkel SJ, et al. A lateralized brain network for visuospatial attention. *Nat Neurosci*. 2011;14(10):1245-1246. doi:10.1038/NN.2905

383. Li X, Rangelov D, Mattingley JB, Oestreich L, Lévy-3 Bencheton D, O'sullivan MJ. White matter microstructure is associated with the precision of visual working memory 1 2. doi:10.1101/2023.01.23.525278

384. Mai JK, Forutan F. Thalamus. *The Human Nervous System*. Published online January 1, 2012:618-677. doi:10.1016/B978-0-12-374236-0.10019-7

385. Zhang Y, Schuff N, Jahng GH, et al. Diffusion tensor imaging of cingulum fibers in mild cognitive impairment and Alzheimer disease. *Neurology*. 2007;68(1):13-19. doi:10.1212/01.WNL.0000250326.77323.01

386. Xie S, Xiao JX, Wang YH, Wu HK, Gong GL, Jiang XX. Evaluation of bilateral cingulum with tractography in patients with Alzheimer's disease. *Neuroreport*. 2005;16(12):1275-1278. doi:10.1097/01.WNR.0000174061.41897.EE

387. Vestergaard M, Skakmadsen K, Baaré WFC, et al. White matter microstructure in superior longitudinal fasciculus associated with spatial working memory performance in children. *J Cogn Neurosci*. 2011;23(9):2135-2146. doi:10.1162/JOCN.2010.21592

388. Darki F, Klingberg T. The Role of Fronto-Parietal and Fronto-Striatal Networks in the Development of Working Memory: A Longitudinal Study. Published online 2014. doi:10.1093/cercor/bht352

389. Krogsrud SK, Fjell AM, Tamnes CK, et al. Development of white matter microstructure in relation to verbal and visuospatial working memory-A longitudinal study. *PLoS One*. 2018;13(4). doi:10.1371/JOURNAL.PONE.0195540

390. Shaw P, Kabani NJ, Lerch JP, et al. Neurodevelopmental Trajectories of the Human Cerebral Cortex. *Journal of Neuroscience*. 2008;28(14):3586-3594. doi:10.1523/JNEUROSCI.5309-07.2008

391. APA Dictionary of Psychology. Accessed March 5, 2023. https://dictionary.apa.org/wechsler-memory-scale

392. Golestani AM, Miles L, Babb J, Castellanos XFX, Malaspina D, Lazar M. Brief Communications Constrained by Our Connections: White Matter's Key Role in Interindividual Variability in Visual Working Memory Capacity. Published online 2014. doi:10.1523/JNEUROSCI.2317-14.2014

393. Li C, Schreiber J, Bittner N, et al. White Matter Microstructure Underlies the Effects of Sleep Quality and Life Stress on Depression Symptomatology in Older Adults. *Front Aging Neurosci*. 2020;12:410. doi:10.3389/FNAGI.2020.578037/BIBTEX

394. Jack CR, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol*. 2010;9(1):119-128. doi:10.1016/S1474-4422(09)70299-6

395. Thomas KR, Bangen KJ, Weigand AJ, et al. Objective subtle cognitive difficulties predict future amyloid accumulation and neurodegeneration. *Neurology*. 2020;94(4):e397. doi:10.1212/WNL.0000000000008838

396. Bangen KJ, Thomas KR, Weigand AJ, et al. Elevated plasma neurofilament light predicts a faster rate of cognitive decline over 5 years in participants with objectively-defined subtle cognitive decline and MCI. *Alzheimers Dement*. 2021;17(10):1756-1762. doi:10.1002/ALZ.12324

397. Husain M. Screening for Alzheimer's disease: plasma biomarkers and what else? *Brain*. 2023;146(2):407-408. doi:10.1093/BRAIN/AWAC490

398. Verberk IMW, Slot RE, Verfaillie SCJ, et al. Plasma Amyloid as Prescreener for the Earliest Alzheimer Pathological Changes. *Ann Neurol*. 2018;84(5):648-658. doi:10.1002/ANA.25334

399. Giudici KV, De Souto Barreto P, Guyonnet S, Li Y, Bateman RJ, Vellas B. Assessment of Plasma Amyloid-β42/40 and Cognitive Decline Among Community-Dwelling Older Adults. *JAMA Netw Open*. 2020;3(12):e2028634-e2028634. doi:10.1001/JAMANETWORKOPEN.2020.28634

400. Verberk IMW, Hendriksen HMA, van Harten AC, et al. Plasma amyloid is associated with the rate of cognitive decline in cognitively normal elderly: the SCIENCe project. *Neurobiol Aging*. 2020;89:99-107. doi:10.1016/J.NEUROBIOLAGING.2020.01.007

401. Stockmann J, Verberk IMW, Timmesfeld N, et al. Amyloid-β misfolding as a plasma biomarker indicates risk for future clinical Alzheimer's disease in individuals with subjective cognitive decline. *Alzheimers Res Ther*. 2020;12(1):1-13. doi:10.1186/S13195-020-00738-8/FIGURES/4

402. Doecke JD, Pérez-Grijalba V, Fandos N, et al. Total Aβ42/Aβ40 ratio in plasma predicts amyloid-PET status, independent of clinical AD diagnosis. *Neurology*. 2020;94(15):e1580-e1591. doi:10.1212/WNL.0000000000009240

403. Karikari TK, Ashton NJ, Brinkmalm G, et al. Blood phospho-tau in Alzheimer disease: analysis, interpretation, and clinical utility. *Nat Rev Neurol*. 2022;18(7):400-418. doi:10.1038/S41582-022-00665-2

404. Ashton NJ, Pascoal TA, Karikari TK, et al. Plasma p-tau231: a new biomarker for incipient Alzheimer's disease pathology. *Acta Neuropathol*. 2021;141(5):709-724. doi:10.1007/S00401-021-02275-6

405. Pereira JB, Janelidze S, Smith R, et al. Plasma GFAP is an early marker of amyloid-β but not tau pathology in Alzheimer's disease. *Brain*. 2021;144(11):3505-3516. doi:10.1093/BRAIN/AWAB223

406. Cicognola C, Janelidze S, Hertze J, et al. Plasma glial fibrillary acidic protein detects Alzheimer pathology and predicts future conversion to Alzheimer dementia in patients with mild cognitive impairment. *Alzheimers Res Ther*. 2021;13(1). doi:10.1186/S13195-021-00804-9

407. W Verberk IM, Thijssen E, Koelewijn J, et al. Combination of plasma amyloid beta (1-42/1-40) and glial fibrillary acidic protein strongly associates with cerebral amyloid pathology. doi:10.1186/s13195-020-00682-7

408. Chatterjee P, Pedrini S, Stoops E, et al. Plasma glial fibrillary acidic protein is elevated in cognitively normal older adults at risk of Alzheimer's disease. doi:10.1038/s41398-020-01137-1

409. Oeckl P, Halbgebauer S, Anderl-Straub S, et al. Glial Fibrillary Acidic Protein in Serum is Increased in Alzheimer's Disease and Correlates with Cognitive Impairment. *J Alzheimers Dis*. 2019;67(2):481-488. doi:10.3233/JAD-180325

410. heller carolin, Foiani M, Moore K, et al. Plasma glial fibrillary acidic protein is raised in progranulin-associated frontotemporal dementia. *J Neurol Neurosurg Psychiatry*. 2020;91:263-270. doi:10.1136/jnnp-2019-321954

411. Benussi A, Ashton NJ, Karikari TK, et al. Serum Glial Fibrillary Acidic Protein (GFAP) Is a Marker of Disease Severity in Frontotemporal Lobar Degeneration. *J Alzheimers Dis*. 2020;77(3):1129-1141. doi:10.3233/JAD-200608

412. Czeiter E, Amrein K, Gravesteijn BY, et al. Blood biomarkers on admission in acute traumatic brain injury: Relations to severity, CT findings and care path in the CENTER-TBI study. *EBioMedicine*. 2020;56. doi:10.1016/J.EBIOM.2020.102785

413. Foerch C, Curdt I, Yan B, et al. Serum glial fibrillary acidic protein as a biomarker for intracerebral haemorrhage in patients with acute stroke. *J Neurol Neurosurg Psychiatry*. 2006;77(2):181. doi:10.1136/JNNP.2005.074823

414. Abdelhak A, Foschi M, Abu-Rumeileh S, et al. Blood GFAP as an emerging biomarker in brain and spinal cord disorders. doi:10.1038/s41582-021-00616-3

415. Janelidze S, Mattsson N, Palmqvist S, et al. Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. *Nature Medicine 2020 26:3*. 2020;26(3):379-386. doi:10.1038/s41591-020-0755-1

416. Moscoso A, Grothe MJ, Ashton NJ, et al. Time course of phosphorylated-tau181 in blood across the Alzheimer's disease spectrum. *Brain*. 2021;144(1):325-339. doi:10.1093/brain/awaa399

417. Mattsson N, Cullen NC, Andreasson U, Zetterberg H, Blennow K. Association between Longitudinal Plasma Neurofilament Light and Neurodegeneration in Patients with Alzheimer Disease. *JAMA Neurol*. 2019;76(7):791-799. doi:10.1001/jamaneurol.2019.0765

418. Marks JD, Syrjanen JA, Graff-Radford J, et al. Comparison of plasma neurofilament light and total tau as neurodegeneration markers: associations with cognitive and neuroimaging outcomes. *Alzheimers Res Ther*. 2021;13(1):1-14. doi:10.1186/S13195-021-00944-Y/TABLES/5

419. Benedet AL, Ashton NJ, Pascoal TA, et al. Plasma neurofilament light associates with Alzheimer's disease metabolic decline in amyloid-positive individuals. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*. 2019;11:679. doi:10.1016/J.DADM.2019.08.002

420. Ossenkoppele R, Smith R, Mattsson-Carlgren N, et al. Accuracy of Tau Positron Emission Tomography as a Prognostic Marker in Preclinical and Prodromal Alzheimer Disease: A Head-to-Head Comparison Against Amyloid Positron Emission Tomography and Magnetic Resonance Imaging. *JAMA Neurol*. 2021;78(8):961-971. doi:10.1001/JAMANEUROL.2021.1858

421. Giannakopoulos P, Herrmann FR, Bussière T, et al. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology*. 2003;60(9):1495-1500. doi:10.1212/01.WNL.0000063311.58879.01

422. Hedden T, Oh H, Younger AP, Patel TA. Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology*. 2013;80(14):1341. doi:10.1212/WNL.0B013E31828AB35D

423. Pascual-Lucas M, Allué JA, Sarasa L, et al. Clinical performance of an antibody-free assay for plasma Aβ42/Aβ40 to detect early alterations of Alzheimer's disease in individuals with subjective cognitive decline. *Alzheimers Res Ther*. 2023;15(1):1-14. doi:10.1186/S13195-022-01143-Z/FIGURES/5

424. Rentz DM, Amariglio RE, Becker JA, et al. Face-name Associative Memory Performance is Related To Amyloid Burden in Normal Elderly. *Neuropsychologia*. 2011;49(9):2776. doi:10.1016/J.NEUROPSYCHOLOGIA.2011.06.006

425. Rajan KB, Aggarwal NT, McAninch EA, et al. Remote Blood Biomarkers of Longitudinal Cognitive Outcomes in a Population Study. *Ann Neurol*. 2020;88(6):1065-1076. doi:10.1002/ANA.25874

426. Simpson JE, Ince PG, Lace G, et al. Astrocyte phenotype in relation to Alzheimer-type pathology in the ageing brain. *Neurobiol Aging*. 2010;31(4):578-590. doi:10.1016/J.NEUROBIOLAGING.2008.05.015

427. Asken BM, Elahi FM, La Joie R, et al. Plasma Glial Fibrillary Acidic Protein Levels Differ Along the Spectra of Amyloid Burden and Clinical Disease Stage. *J Alzheimers Dis*. 2020;78(1):265-276. doi:10.3233/JAD-200755

428. O'Connor A, Karikari TK, Poole T, et al. Plasma phospho-tau181 in presymptomatic and symptomatic familial Alzheimer's disease: a longitudinal cohort study. *Molecular Psychiatry 2020 26:10*. 2020;26(10):5967-5976. doi:10.1038/s41380-020-0838-x

429. Simrén J, Leuzy A, Karikari TK, et al. The diagnostic and prognostic capabilities of plasma biomarkers in Alzheimer's disease. *Alzheimer's & Dementia*. 2021;17(7):1145-1156. doi:10.1002/ALZ.12283

430. Moscoso A, Grothe MJ, Ashton NJ, et al. Longitudinal Associations of Blood Phosphorylated Tau181 and Neurofilament Light Chain With Neurodegeneration in Alzheimer Disease. *JAMA Neurol*. 2021;78(4):396-406. doi:10.1001/JAMANEUROL.2020.4986

431. Therriault J, Benedet AL, Pascoal TA, et al. Association of plasma P-tau181 with memory decline in non-demented adults. *Brain Commun*. 2021;3(3). doi:10.1093/BRAINCOMMS/FCAB136

432. Janelidze S, Mattsson N, Palmqvist S, et al. Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. *Nat Med*. Published online 2020. doi:10.1038/s41591-020-0755-1

433. Mattsson-Carlgren N, Janelidze S, Palmqvist S, et al. Longitudinal plasma p-tau217 is increased in early stages of Alzheimer's disease. *Brain*. 2020;143(11):3234-3241. doi:10.1093/BRAIN/AWAA286

434. Donohue MC, Sperling RA, Salmon DP, et al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol*. 2014;71(8):961-970. doi:10.1001/JAMANEUROL.2014.803

435. Mattsson-Carlgren N, Salvadó G, Ashton NJ, et al. Prediction of Longitudinal Cognitive Decline in Preclinical Alzheimer Disease Using Plasma Biomarkers. *JAMA Neurol*. Published online February 6, 2023. doi:10.1001/JAMANEUROL.2022.5272

436. Ashton NJ, Janelidze S, Al Khleifat A, et al. A multicentre validation study of the diagnosticvalue of plasma neurofilament light. *Gorka Fernández-Eulate*. 24:41. doi:10.1038/s41467-021-23620-z

437. Mielke MM, Syrjanen JA, Blennow K, et al. Plasma and CSF neurofilament light: Relation to longitudinal neuroimaging and cognitive measures. *Neurology*. 2019;93(3):e252. doi:10.1212/WNL.0000000000007767

438. Ossenkoppele R, Pichet Binette A, Groot C, et al. Amyloid and tau PET-positive cognitively unimpaired individuals are at high risk for future cognitive decline. *Nat Med*. 2022;28(11):2381-2387. doi:10.1038/S41591-022-02049-X

439. Fandos N, Pérez-Grijalba V, Pesini P, et al. Plasma amyloid β 42/40 ratios as biomarkers for amyloid β cerebral deposition in cognitively normal individuals. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*. 2017;8:179. doi:10.1016/J.DADM.2017.07.004

440. Ashton NJ, Leuzy A, Lim YM, et al. Increased plasma neurofilament light chain concentration correlates with severity of post-mortem neurofibrillary tangle pathology and neurodegeneration. *Acta Neuropathol Commun*. 2019;7(1):5. doi:10.1186/S40478-018-0649-3/FIGURES/5

441. Li Q, Joo SJ, Yeatman JD, Reinecke K. Controlling for Participants' Viewing Distance in Large-Scale, Psychophysical Online Experiments Using a Virtual

Chinrest. *Scientific Reports 2020 10:1*. 2020;10(1):1-11. doi:10.1038/s41598-019-57204-1

442. Le test de copie d'une figure complexe; contribution à l'étude de la perception et de la mémoire. Accessed March 30, 2023. https://psycnet.apa.org/record/1946-02126-001

443. Lezak MD, Howieson D. Neuropsychological assessment (5th ed.). [References]. *(2012)*. Published online 2012.

444. Reitan RM. Trail making test. Manual for administration, scoring, and interpretation. Indianapolis, IN Indiana University Press. - References - Scientific Research Publishing. Published online 1956.

445. Zeng Z, Miao C, Leung C, Shen Z. Computerizing Trail Making Test for long-term cognitive self-assessment. *International Journal of Crowd Science*. 2017;1(1):83-99. doi:10.1108/IJCS-12-2016-0002/FULL/PDF

446. Corsi PM. Human memory and the medial temporal region of the brain. Published online 1972.

447. Arce T, Mcmullen K. The Corsi Block-Tapping Test: Evaluating methodological practices with an eye towards modern digital frameworks. Published online 2021. doi:10.1016/j.chbr.2021.100099

448. Ashton NJ, Janelidze S, Mattsson-Carlgren N, et al. Differential roles of Aβ42/40, p-tau231 and p-tau217 for Alzheimer's trial selection and disease monitoring. *Nature Medicine 2022 28:12*. 2022;28(12):2555-2562. doi:10.1038/s41591-022-02074-w

449. Milà-Alomà M, Ashton NJ, Shekari M, et al. Plasma p-tau231 and p-tau217 as state markers of amyloid-β pathology in preclinical Alzheimer's disease. *Nat Med*. 2022;28(9):1797-1801. doi:10.1038/s41591-022-01925-w

450. Burnham SC, Fandos N, Fowler C, et al. Longitudinal evaluation of the natural history of amyloid-β in plasma and brain. *Brain Commun*. 2020;2(1). doi:10.1093/BRAINCOMMS/FCAA041

451. De Simone MS, Perri R, Rodini M, et al. A Lack of Practice Effects on Memory Tasks Predicts Conversion to Alzheimer Disease in Patients With Amnestic Mild Cognitive Impairment. *https://doi.org/101177/0891988720944244*. 2020;34(6):582-593. doi:10.1177/0891988720944244

452. Reed GF, Lynn F, Meade BD. Use of Coefficient of Variation in Assessing Variability of Quantitative Assays. *Clin Diagn Lab Immunol*. 2002;9(6):1235. doi:10.1128/CDLI.9.6.1235-1239.2002

453. Verberk IMW, Misdorp EO, Koelewijn J, et al. Characterization of pre-analytical sample handling effects on a panel of Alzheimer's disease–related blood-based biomarkers: Results from the Standardization of Alzheimer's Blood Biomarkers (SABB) working group. *Alzheimer's & Dementia*. 2022;18(8):1484-1497. doi:10.1002/ALZ.12510

454. Velayudhan L, Ryu SH, Raczek M, et al. Review of brief cognitive tests for patients with suspected dementia. *Int Psychogeriatr*. 2014;26(8):1247-1262. doi:10.1017/S1041610214000416

455. de Paula JJ, Malloy-Diniz LF, Romano-Silva MA. Reliability of working memory assessment in neurocognitive disorders: A study of the Digit Span and Corsi Block-Tapping tasks. *Revista Brasileira de Psiquiatria*. 2016;38(3):262-263. doi:10.1590/1516-4446-2015-1879

456. Amodio P, Wenin H, Del Piccolo F, et al. Variability of trail making test, symbol digit test and line trait test in normal people. A normative study

taking into account age-dependent decline and sociobiological variables. *Aging Clin Exp Res*. 2002;14(2):117-131. doi:10.1007/BF03324425

457. Levine AJ, Miller EN, Becker JT, Selnes OA, Cohen BA. Normative Data for Determining Significance of Test-Retest Differences on Eight Common Neuropsychological Instruments. *Clin Neuropsychol*. 2004;18(3):373-384. doi:10.1080/1385404049052420

458. Zaninotto P, Batty GD, Allerhand M, Deary IJ. Cognitive function trajectories and their determinants in older people: 8 years of follow-up in the English Longitudinal Study of Ageing. *J Epidemiol Community Health*. 2018;72(8):685-694. doi:10.1136/JECH-2017-210116

459. Bailly M, Destrieux C, Hommet C, et al. Precuneus and Cingulate Cortex Atrophy and Hypometabolism in Patients with Alzheimer's Disease and Mild Cognitive Impairment: MRI and (18)F-FDG PET Quantitative Analysis Using FreeSurfer. *Biomed Res Int*. 2015;2015. doi:10.1155/2015/583931

460. Brodaty H, Pond D, Kemp NM, et al. The GPCOG: a new screening test for dementia designed for general practice. *J Am Geriatr Soc*. 2002;50(3):530-534. doi:10.1046/J.1532-5415.2002.50122.X

461. Schlemmer M, Desrichard O. Is Medical Environment Detrimental to Memory? A Test of A White Coat Effect on Older People's Memory Performance. *Clin Gerontol*. 2018;41(1):77-81. doi:10.1080/07317115.2017.1307891

462. Öhman F, Hassenstab J, Berron D, Schöll M, Papp K V. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*. 2021;13(1). doi:10.1002/DAD2.12217

463. Kaye J, Aisen P, Amariglio R, et al. Using Digital Tools to Advance Alzheimer's Drug Trials During a Pandemic: The EU/US CTAD Task Force. *Journal of Prevention of Alzheimer's Disease*. 2021;8(4):513-519. doi:10.14283/JPAD.2021.36/FIGURES/1

464. Wu YH, Lewis M, Rigaud AS. Cognitive Function and Digital Device Use in Older Adults Attending a Memory Clinic. *https://doi.org/101177/2333721419844886*. 2019;5:233372141984488. doi:10.1177/2333721419844886

Appendix

Chapter 3.

	Metric	Set size	Delay	Set size x Delay
Group EHC	**Identification Time**	F = 572.5, ***p < 0.001** η2 = 0.641	F = 118.3, ***p < 0.001** η2 = 0.051	F = 6.0, ***p = 0.016** η2 = 0.002
	Localization Time	F = 580.2, ***p < 0.001** η2 = 0.634	F = 176.7, ***p < 0.001** η2 = 0.070	F = 2.4, p = 0.125
	Identification Accuracy	F = 664.2, ***p < 0.001** η2 = 0.605	F = 89.5, ***p < 0.001** η2 = 0.057	F = 23.9, ***p < 0.001** η2 = 0.013
	Absolute Error	F = 827.3, ***p < 0.001** η2 = 0.732	F = 64.2, ***p < 0.001** η2 = 0.022	F = 28.5, ***p < 0.001** η2 = 0.009
	Target detection	F = 1389.5, ***p < 0.001** η2 = 0.838	F = 27.6, ***p < 0.001** η2 = 0.006	F = 3.828, p = 0.052
	Misbinding	N/A	F = 6.107, ***p = 0.015** η2 = 0.038	N/A
	Guessing	F = 229.0, ***p < 0.001** η2 = 0.444	F = 50.1, ***p < 0.001** η2 = 0.036	F = 0.7, p = 0.407
	Imprecision	F = 181.6, ***p < 0.001** η2 = 0.370	F = 49.7, ***p < 0.001** η2 = 0.042	F = 2.3, p = 0.131
Group SCI	**Identification Time**	F = 125.8, ***p < 0.001** η2 = 0.592	F = 33.8, ***p < 0.001** η2 = 0.052	F = 14.7, ***p < 0.014** η2 = 0.014
	Localization Time	F = 107.9 ***p < 0.001** η2 = 0.554	F = 46.6, ***p < 0.001** η2 = 0.071	F = 14.5, ***p < 0.001** η2 = 0.013
	Identification Accuracy	F = 152.5, ***p < 0.001** η2 = 0.590	F = 20.2, ***p < 0.001** η2 = 0.041	F = 4.7, ***p = 0.035** η2 = 0.008
	Absolute Error	F = 212.8, ***p < 0.001** η2 = 0.732	F = 9.3, ***p = 0.004** η2 = 0.010	F = 4.5, ***p = 0.038** η2 = 0.004
	Target detection	F = 328.4, ***p < 0.001** η2 = 0.808	F = 2.2, p = 0.148	F = 0.000, p = 0.980
	Misbinding	N/A	F = 0.1, p = 0.703	N/A
	Guessing	F = 58.9, ***p < 0.001** η2 = 0.420	F = 6.9, ***p = 0.011** η2 = 0.020	F = 1.1, p = 0.308
	Imprecision	F = 38.5, ***p < 0.001** η2 = 0.334	F = 6.0, ***p = 0.018** η2 = 0.019	F = 0.6, p = 0.450
Group MCI	**Identification Time**	F = 114.5, ***p < 0.001** η2 = 0.546	F = 15.0, ***p < 0.001** η2 = 0.019	F = 1.5, p = 0.222
	Localization Time	F = 108.0, ***p < 0.001** η2 = 0.524	F = 14.3, ***p < 0.001** η2 = 0.022	F = 0.8, p = 0.385

	Identification Accuracy	F = 409.3, ***p < 0.001** η2 = 0.730	F = 27.1, ***p < 0.001** η2 = 0.030	F = 1.5, p = 0.218
	Absolute Error	F = 589.0, ***p < 0.001** η2 = 0.730	F = 10.6, ***p = 0.002** η2 = 0.009	F = 4.2, ***p = 0.046** η2 = 0.003
	Target detection	F = 948.7, ***p < 0.001** η2 = 0.889	F = 1.0, p = 0.309	F = 1.784, p = 0.187
	Misbinding	N/A	F = 2.8, p = 0.099	N/A
	Guessing	F = 27.0, ***p < 0.001** η2 = 0.201	F = 20.5, ***p < 0.001** η2 = 0.070	F = 0.3, p = 0.560
	Imprecision	F = 19.4, ***p < 0.001** η2 = 0.142	F = 17.9, ***p < 0.001** η2 = 0.069	F = 0.2, p = 0.618
Group AD	**Identification Time**	F = 27.1, ***p < 0.001** η2 = 0.211	F = 12.0, ***p < 0.001** η2 = 0.029	F = 0.7, p = 0.413
	Localization Time	F = 21.5, ***p < 0.001** η2 = 0.164	F = 16.9, ***p < 0.001** η2 = 0.033	F = 0.1, p = 0.748
	Identification Accuracy	F =322.2, ***p < 0.001** η2 = 0.661	F = 32.8, ***p < 0.001** η2 = 0.039	F = 0.4, p = 0.534
	Absolute Error	F = 288.6, ***p < 0.001** η2 = 0.683	F = 33.4, ***p < 0.001** η2 = 0.036	F = 3.0, p = 0.081
	Target detection	F = 816.3, ***p < 0.001** η2 = 0.871	F = 8.1, ***p = 0.006** η2 = 0.004	F = 2.798, p = 0.099
	Misbinding	N/A	F = 0.9, p = 0.341	N/A
	Guessing	F = 0.6, ***p = 0.043** η2 = 0.006	F = 26.6, ***p < 0.001** η2 = 0.069	F = 0.9, p = 0.342
	Imprecision	F = 2.0, p = 0.001 η2 = 0.020	F = 22.1, ***p < 0.001** η2 = 0.061	F = 4.3, ***p = 0.042** η2 = 0.009

Table 3.1S | Group-by-group Set size and Delay effects

	Left HV	Right HV
Identification time	t = -3.2, *p = 0.001	t = -2.5, *p = 0.012
Localization time	t = -2.8, *p = 0.004	t = -2.5, *p = 0.012
Identification accuracy	t = 3.4, *p < 0.001	t = 2.6, *p = 0.009
Absolute Error	t = -2.7, *p = 0.008	t = -2.7, *p = 0.006
Target detection	t = 6.3, *p = 0.002	t = 2.2, *p = 0.030
Misbinding	t = -2.6, *p = 0.009	t = -2.1, *p = 0.039
Guessing	t = -3.1, *p = 0.002	t = 2.6, *p = 0.012
Imprecision	t = -2.5, *p = 0.013	t = 2.6, *p = 0.002

Table 3.2S | Linear regression for left and right hippocampal volumes

	ALL	EHC/SCI	EHC/MCI	SCI/MCI	EHC/AD	SCI/AD	MCI/AD
Identification Time	F = 14.7, ***p < 0.001** η2 = 0.431	n.s.	n.s.	n.s.	t = -6.4, ***p < 0.001**, d = -2.355	t = -5.0, ***p < 0.001**, d = -2.121	t = -3.8, ***p = 0.001**, d = -1.653
Localization Time	F = 7.3, ***p < 0.001** η2 = 0.271	n.s.	n.s.	n.s.	t = -4.5, ***p < 0.001**, d = -1.748	t = -3.4, ***p = 0.006**, d = -1.470	t = -3.3, ***p = 0.007**, d = -1.391
Identification Accuracy	F = 23.9, ***p < 0.001** η2 = 0.570	n.s.	t = 3.3, ***p = 0.006,** d = 1.319	t = 2.8, ***p = 0.014**, d = 1.266	t = 8.0, ***p < 0.001**, d = 2.939	t = 6.8, ***p < 0.001**, d = 2.887	t = 3.8, ***p = 0.002**, d = 1.621
Absolute Error	F = 24.2, ***p < 0.001** η2 = 0.565	t = -2.2, ***p = 0.049**, d = -0.807	t = -4.6, ***p < 0.001**, d = -2.115	t = -2.6, ***p = 0.032**, d = -1.3	t = -8.3, ***p < 0.001**, d = -3.118	t = -5.6, ***p < 0.001**, d = -2.381	t = -2.3, ***p = 0.049**, d = -1.073
Target detection	F = 28.7, ***p < 0.001** η2 = 0.635	n.s.	t = 5.5, ***p < 0.001**, d = 2.272	t = 4.3, ***p < 0.001**, d = 1.936	t = 8.4, ***p < 0.001**, d = 3.208	t = 6.5, ***p < 0.001**, d = 2.872	t = 2.1, **p = 0.090**, d = 0.937
Misbinding	F = 11.5, ***p < 0.001** η2 = 0.405	n.s.	t = -3.9, ***p = 0.001**, d = -1.522	t = -2.7, ***p = 0.027**, d = -1.162	t = -5.3, ***p < 0.001**, d = -2.052	t = -3.8, ***p = 0.002**, d = -1.692	t = -1.2, **p = 0.439**, d = -0.530
Guessing	F = 26.8, ***p < 0.001** η2 = 0.606	n.s.	t = -5.1, ***p < 0.001**, d = -2.260	t = -3.7, ***p = 0.002**, d = -1.786	t = -8.5, ***p < 0.001**, d = -3.129	t = -6.2, ***p < 0.001**, d = -2.655	t = -1.9, **p = 0.132**, d = -0.869
Imprecision	F = 16.8, ***p < 0.001** η2 = 0.470	n.s.	t = -3.2, ***p = 0.011**, d = -1.360	n.s.	t = -6.9, ***p < 0.001**, d = -2.578	t = -5.1, ***p < 0.001**, d = -2.222	t = -2.7, ***p = 0.027**, d = -1.218

Table 3.3S | Group comparisons - longitudinal dataset

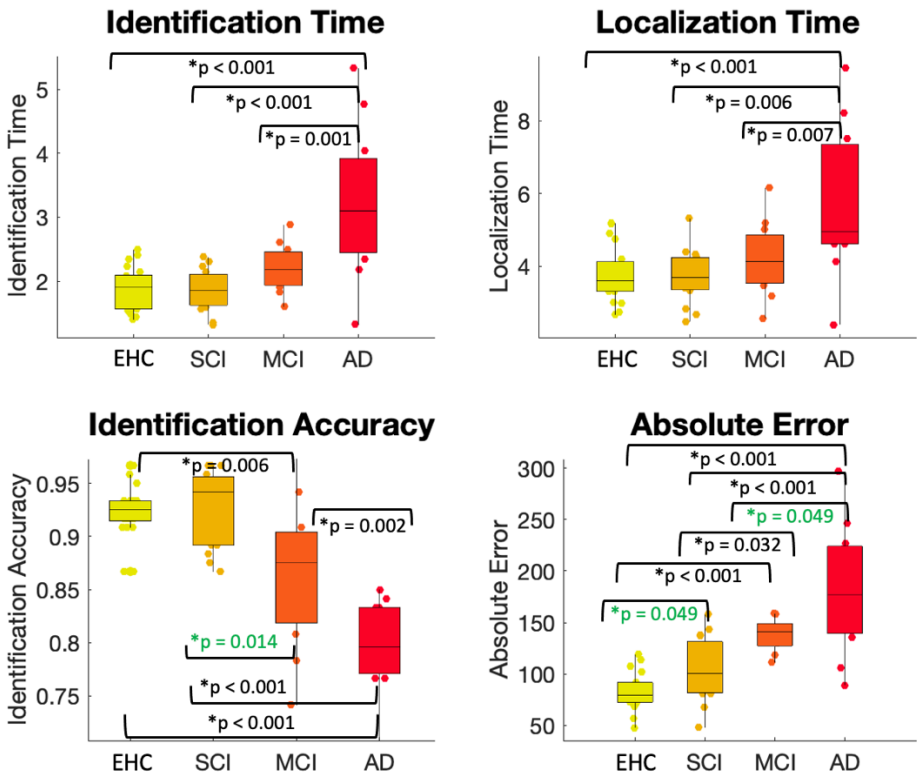


Figure 3.1S | Group comparisons - longitudinal dataset – basic metrics

In green: correlations that were not significant in the dataset of Chapter 3.1, but significant in the longitudinal dataset. In blue: correlations that were significant in the dataset of Chapter 3.1 but were not significant in the longitudinal dataset.

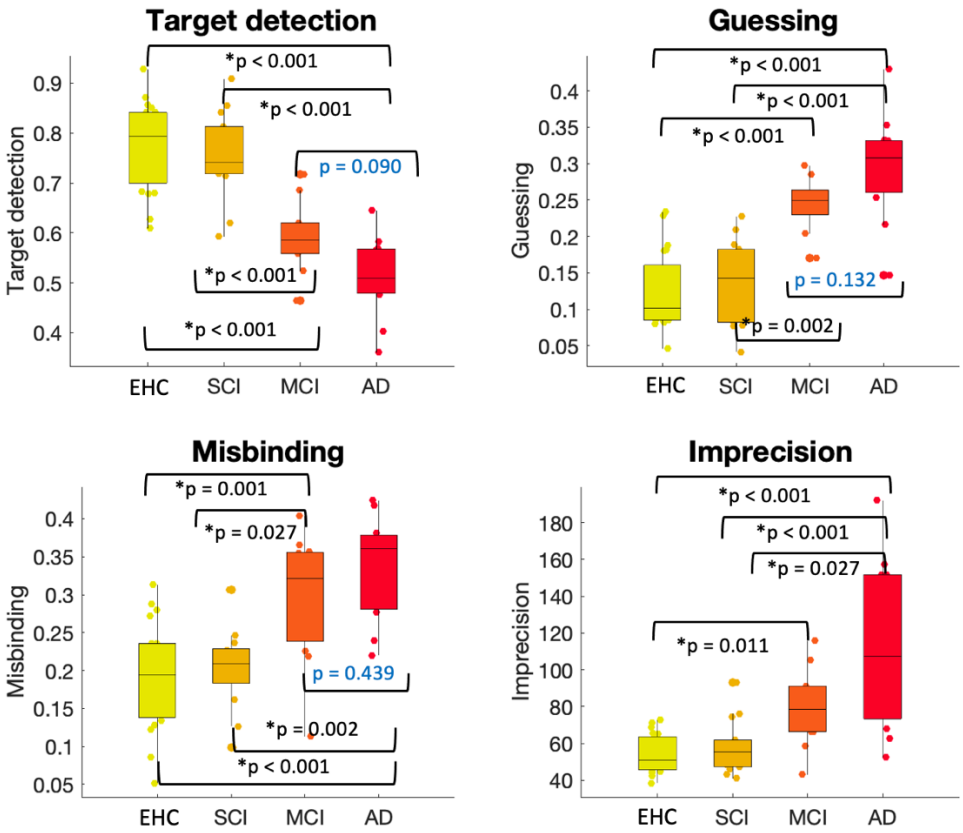


Figure 3.2S | Group comparisons - longitudinal dataset – basic metrics

	Decline in ACE scores
Identification time	F (1,45), t = 4.5, *p = 0.002, R^2 = 0.3143
Localization time	F (1,45), t = 4.6, *p =0.001, R^2 = 0.3239
Identification accuracy	F (1,45), t = -5.0, *p < 0.001, R^2 = 0.3537
Absolute Error	F (1,45), t = 5.3, *p < 0.001, R^2 = 0.3537
Target detection	F (1,45), t = -5.2, *p < 0.001, R^2 = 0.3902
Misbinding	F (1,45), t = 4.0, *p < 0.001, R^2 = 0.2706
Guessing	F (1,45), t = 4.6, *p = 0.005, R^2 = 0.3333
Imprecision	F (1,45), t = 5.5, *p < 0.001, R^2 = 0.4107

Table 3.4S | Group comparisons - longitudinal dataset - basic metrics

VSTM metrics in AD

Compared to the analysis on the whole sample AD and EHC), lower GM in the Precuneus and right Amygdala were not associated with higher Guessing and lower Target Detection (**Figure 4.1S**, top row). GM atrophy in the right posterior division of the Middle Temporal Gyrus, left Insula and left Orbitofrontal cortex didn't correlate with increased Identification Time (**Figure 4.1S**, bottom row, left).

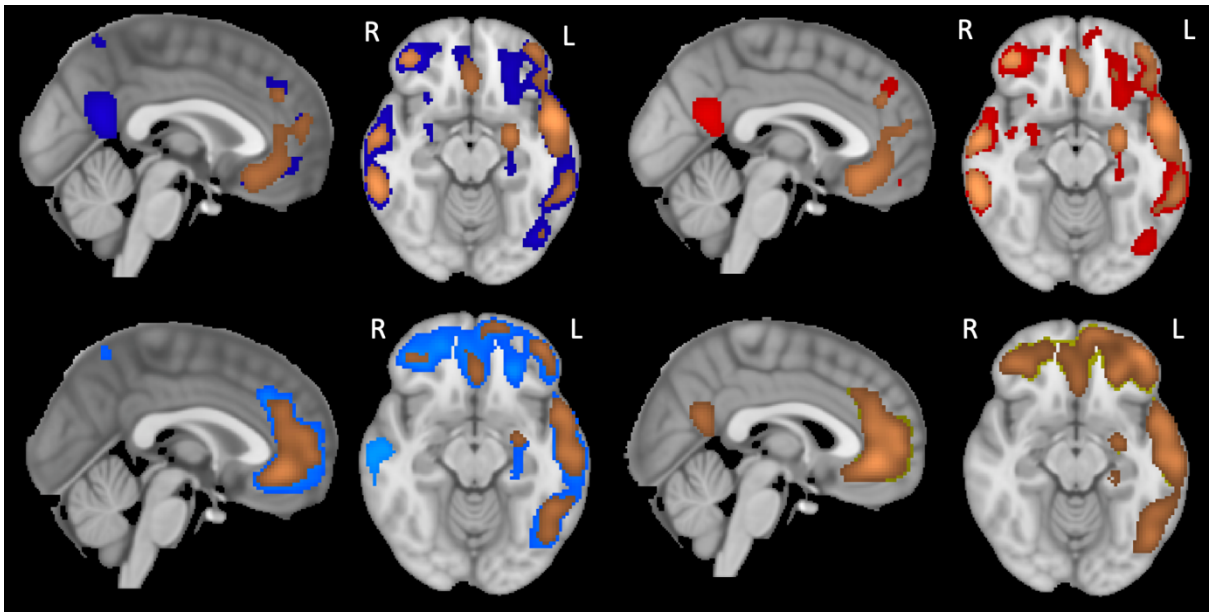


Figure 4.1S | Spatial localization - Identity metrics and Reaction Times metrics in AD

Statistically significant maps are overlaid with GM areas in AD in copper for all metrics (lower cluster size) on top. Blue = Guessing (top left), red = Target detection (top right), light blue = Identification Time (bottom left), yellow = Localization Time (bottom right). R= right, L = left.

GM atrophy in the left Parahippocampal Gyrus did not correlate with lower Identification accuracy (**Figure 4.2S**, top row). Lower GM volumes in the Posterior Cingulate Cortex and left Fusiform Gyrus were not associated respectively with increased Absolute Error and increased memory Imprecision (**Figure 4.2S**, bottom row).

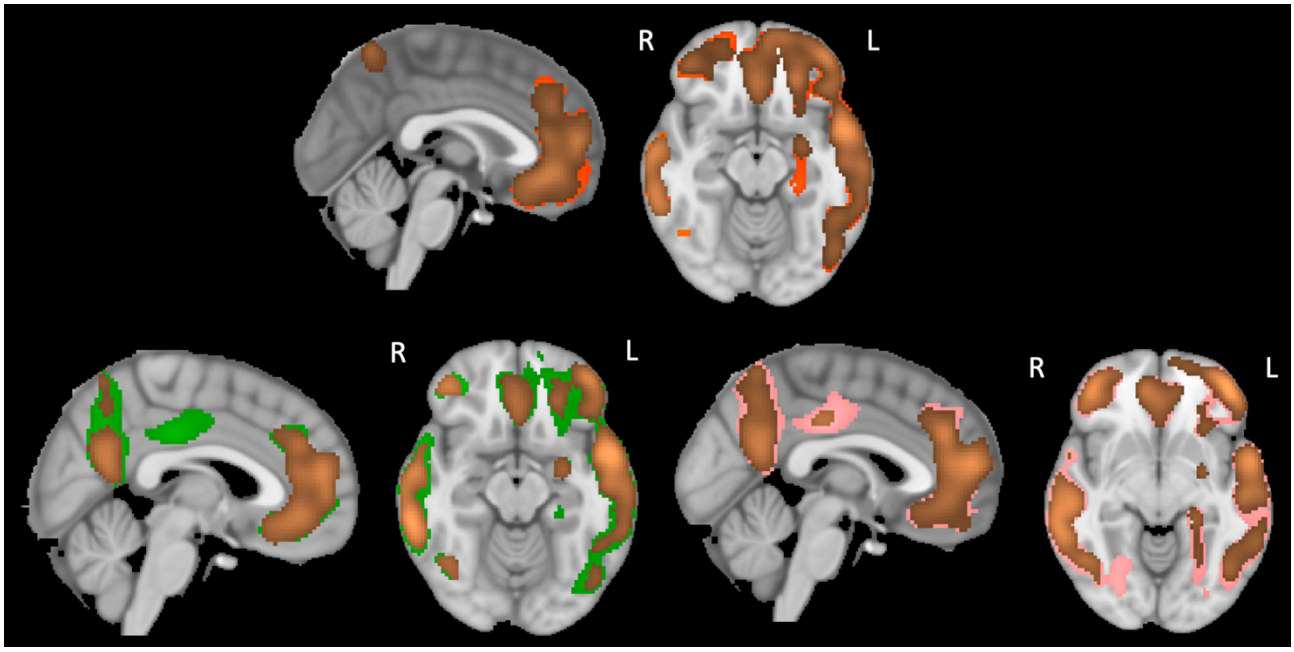


Figure 4.2S | Object Identity metric and Spatial localization - Precision metrics in AD

Statistically significant maps are overlaid with GM areas in AD in copper for all metrics (lower cluster size) on top. Orange = Identification accuracy (top row), Green = Absolute Error (bottom left), pink = memory Imprecision (bottom right), R= right, L = left.

VSTM metric	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Target detection	1	Temporal Pole (L)	10812	-48	12	-26	5.1
	2	Middle Temporal Gyrus temporooccipital division (R)	1717	62	-42	-10	4.37
	3	Frontal Pole (R)	448	42	46	-8	4.46
	4	Temporal Pole (R)	276	44	18	-26	4.11
Guessing	1	Temporal Pole (L)	11804	-50	8	-26	5.33
	2	Middle Temporal Gyrus temporooccipital division (R)	859	62	-42	-10	4.45
	3	Frontal Pole (R)	446	42	44	-10	4.39
	4	Middle Temporal Gyrus posterior division (R)	383	60	-8	-14	3.98
	5	Temporal Pole (R)	164	44	16	-24	3.95
	6	Heschl's Gyrus (R)	40	56	-12	6	3.42
	7	LOC inferior division (L)	17	-46	-70	-16	2.54
Identification time	1	Temporal Pole (L)	9660	-52	8	-26	5.24
	2	Frontal Pole (L)	7882	-28	56	8	5.73
	3	Frontal Pole (R)	660	24	46	30	3.63
	4	Hippocampus (L)	90	-22	-36	0	4

Localization time	1	Temporal Pole (L)	27911	-52	6	-26	5.65
Identification accuracy	1	Temporal Pole (L)	31041	-52	6	-26	6.43
	2	Middle Temporal Gyrus temporo-occipital division (R)	1576	62	-40	-10	4.44
	3	Thalamus/Hippocampus (L)	232	-20	-36	2	4.22
Absolute Error	1	Temporal Pole (L)	24241	-52	6	-24	5.41
	2	Middle Temporal Gyrus posterior division (R)	3286	62	-38	-14	4.65
	3	Frontal Pole (R)	716	42	46	-6	4.52
	4	Thalamus/Hippocampus (L)	81	-22	-36	0	4.08
Imprecision	1	Temporal Pole (L)	28669	-52	6	-24	5.63
	2	Middle Temporal Gyrus posterior division (R)	3162	62	-38	-12	4.33
	3	Frontal Pole (R)	1159	42	46	-6	4.23
	4	Insula (L)	137	-36	-18	2	3.48
	5	Cingulate Cortex posterior division (L)	60	0	-26	36	3.36

Table 4.1S | Correlations between VSTM metrics and GM areas in AD patients.

VSTM contribution independent of general cognition across the whole cohort

VSTM metrics showed a similar segregation pattern as seen in the previous analysis, with four different localization clusters. Areas found in this analysis included the left Frontal Pole, Paracingulate Gyrus, right Inferior Temporal Gyrus and Precuneus (**Figure 4.3S, Table 4.2S**).

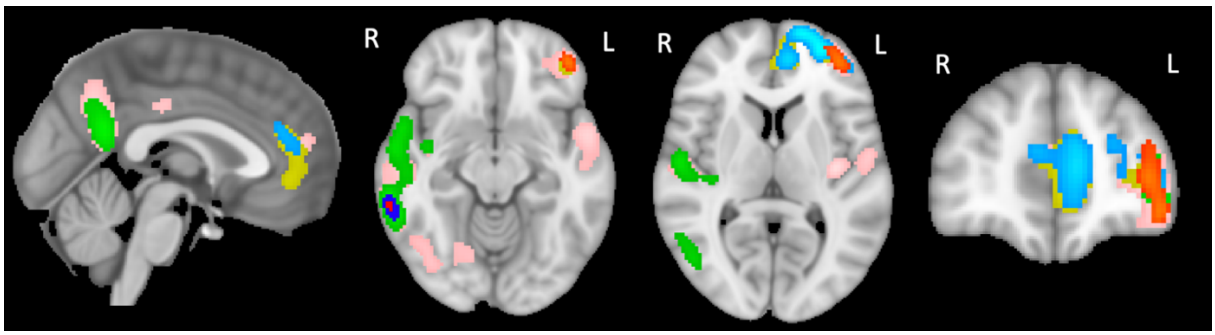


Figure 4.3S | VSTM metrics correlations with GM atrophy – independent of general cognition

Statistically significant maps are overlaid with the metrics with increasing lower cluster size at the top: red = Target detection, blue = Guessing, yellow = Localization time, light blue = Identification time, orange = Identification accuracy, green = Absolute Error, pink = memory Imprecision.

VSTM metric	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Target detection	1	Inferior Temporal Gyrus temporooccipital part (R)	20	62	-40	-14	4.29
Guessing	1	Inferior Temporal Gyrus temporooccipital part (R)	167	60	-40	-16	4.33
	2	Middle Temporal Gyrus posterior division (R)	12	54	-28	-6	3.75
Identification Time	1	Frontal Pole (L) Paracingulate Gyrus (L)	2094	-26	56	8	4.75
	2	Middle Frontal Gyrus (L)	67	-44	34	22	3.41
Localization Time	1	Paracingulate Gyrus (L) Frontal Pole (L)	2310	-8	50	8	4.64
Identification Accuracy	1	Frontal Pole (L)	463	-44	44	-8	4.04
Absolute Error	1	Middle Temporal Gyrus (R) Superior Temporal Gyrus (R) Inferior Temporal Gyrus (R) Parietal Operculum Cortex (R) Supramarginal Gyrus (R) Heschl's Gyrus (R) Insula (R) Postcentral Gyrus (R) Lateral Occipital Cortex (R)	3361	54	-26	-6	4.51
	2	Precuneus	883	4	-54	22	4.23
	3	Frontal Pole (L)	399	-44	44	-6	4.44
	4	Middle Frontal Gyrus (L)	119	-44	34	22	4.18
Imprecision	1	Middle Temporal Gyrus (R) Superior Temporal Gyrus (R) Inferior Temporal Gyrus (R) Parietal Operculum Cortex (R) Supramarginal Gyrus (R) Heschl's Gyrus (R) Insula (R) Postcentral Gyrus (R) Lateral Occipital Cortex (R)	3576	54	-24	-6	4.57
	2	Frontal Pole (L) Middle Frontal Gyrus (L)	1942	-42	46	-4	4.27
	3	Insula (L) Supramarginal Gyrus (L) Superior Temporal Gyrus (L)	1936	-36	-16	4	3.27

	Middle Temporal Gyrus (L) Temporal Pole (L) Central Opercolum Cortex (L) Parietal Opercolum Cortex (L) Heschl's Gyrus (L)					
4	Precuneus	1454	2	-54	22	4.13
5	Occipital Fusiform Gyrus (R) Lingual Gyrus (R)	227	24	-66	-8	3.48
6	Posterior Cingulate Cortex (L)	187	-10	-18	44	3.6
7	Paracingulate cortex (L)	91	-6	38	32	3.58

Table 4.2S | Correlations between VSTM metrics and GM areas in AD and EHC - independent of general cognition

Lower performances at Spatial localization - Identity Metrics (Target detection and Guessing) were associated with atrophy in a small area in the right Inferior Temporal Gyrus (**Figure 4.4S**, **Table 4.2S**).

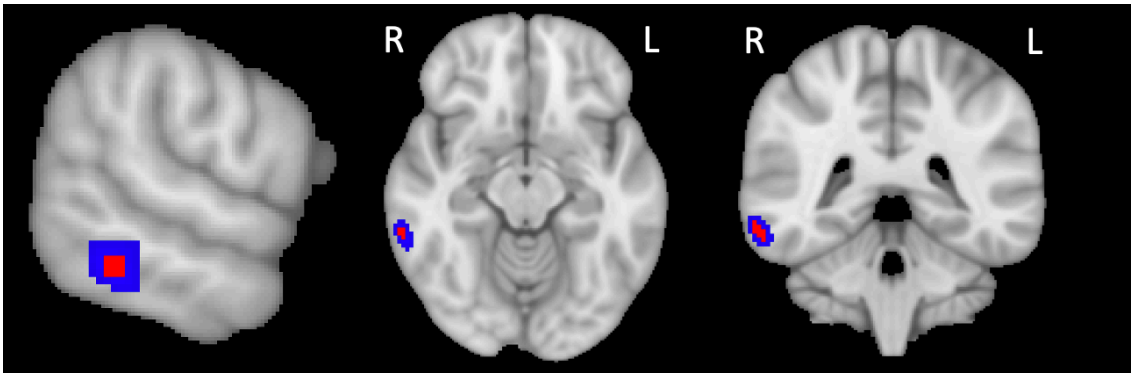


Figure 4.4S | Spatial localization – Identity metrics, independent of general cognition

Statistically significant maps are overlaid with Target detection (lower cluster size) on top. Red = Target detection, blue = Guessing. R= right, L = left.

For the Reaction Times metrics, higher Identification time and Localization time were associated with lower GM volumes in the left Frontal Pole and Paracingulate Gyrus (**Figure 4.5S**, **Table 4.2S**).

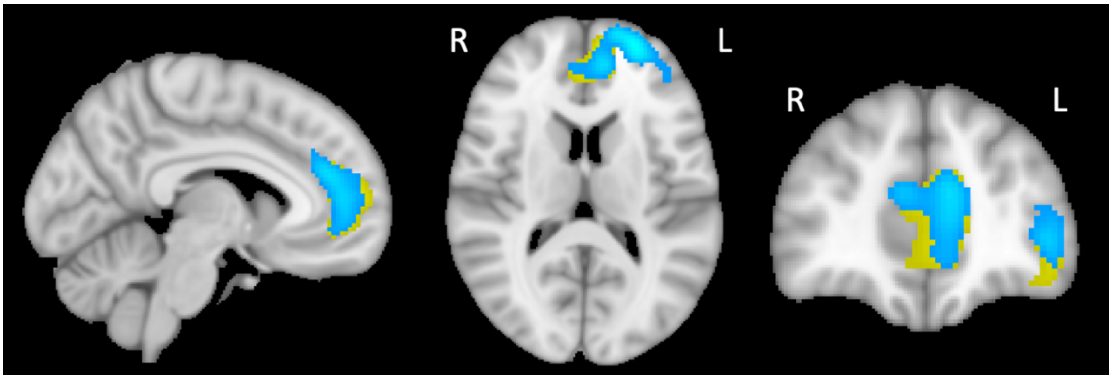


Figure 4.5S | Reaction Times Metrics, independent of general cognition

Statistically significant maps are overlaid with Identification Time (lower cluster size) on top. Light blue = Identification time, yellow = Localization time. R= right, L = left.

Lower Identification accuracy was associated with GM atrophy in the left Frontal Pole (**Figure 4.6S**, **Table 4.2S**).

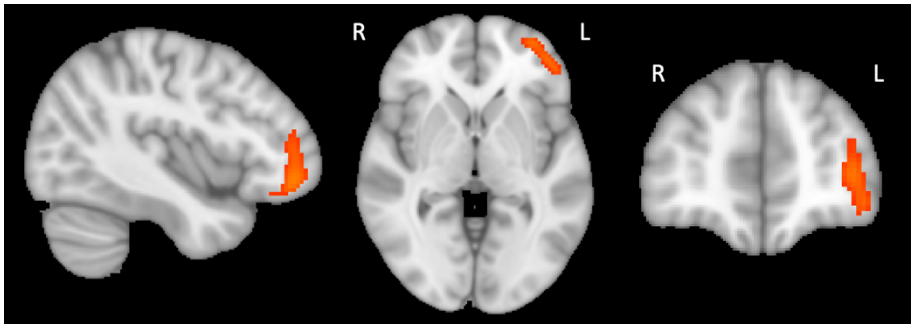


Figure 4.6S | Object Identity metric, independent of general cognition

Identification accuracy in orange. R= right, L = left.

Regarding the Spatial localization - Precision metrics, increased Absolute Error and memory Imprecision were associated with lower GM volumes in the left Frontal Pole, left Middle Frontal Gyrus, right Temporal Gyrus and Precuneus. Precision also showed an involvement of the Posterior Cingulate Cortex, Paracingulate Gyrus, left Insula, left Supramarginal Gyrus, right Occipital Fusiform Gyrus and right Lingual Gyrus (**Figure 4.7S**, **Table 4.2S**).

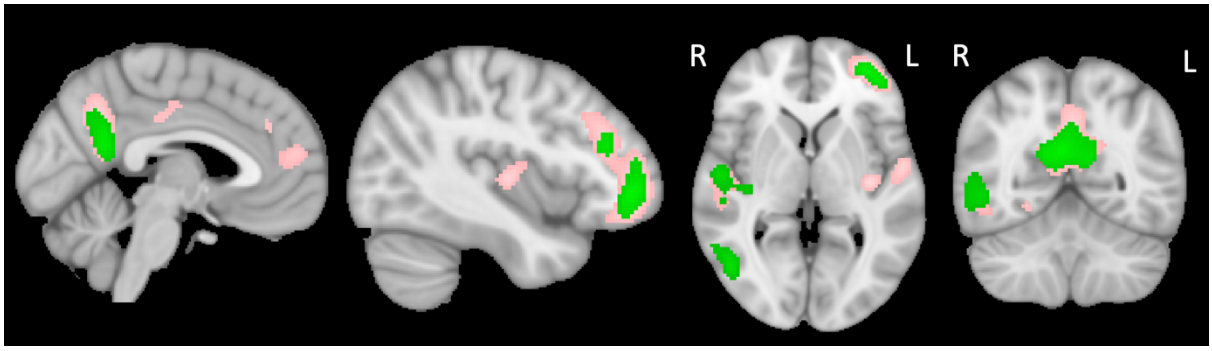


Figure 4.7S | Spatial localization - Precision metrics, independent of general cognition

Statistically significant maps are overlaid with Absolute Error (lower cluster size) on top. Green = Absolute Error, pink = memory Imprecision. R= right, L = left.

VSTM metrics independent of general cognition – in AD

Spatial localization - Identity Metrics (Target detection and Guessing) were not associated with any GM areas in the AD alone. Greater Identification time was associated with lower GM volume in the left Frontal Pole (**Figure 4.8S**, top row, and **Table 4.3S**). Higher Localization time correlated with greater atrophy in the left Frontal Pole and Paracingulate Gyrus (**Figure 4.8S**, middle row, and **Table 4.3S**). Lower Identification accuracy correlated with greater atrophy in the left Frontal Pole (**Figure 4.8S**, bottom row, and **Table 4.3S**).

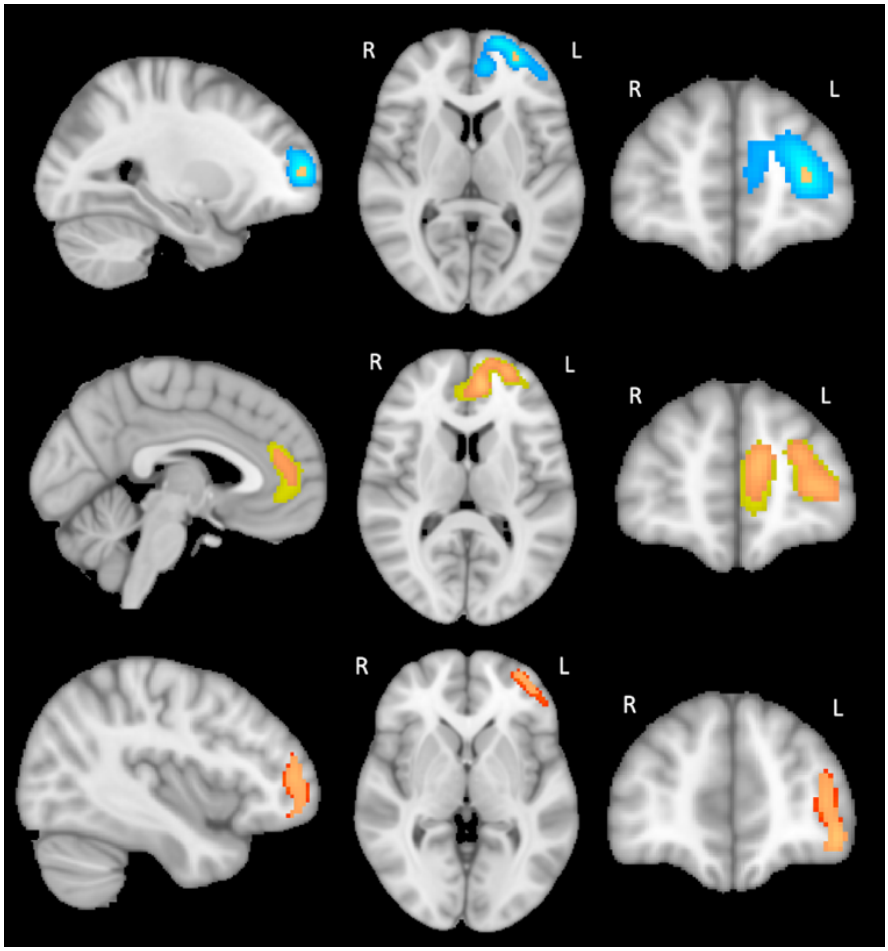


Figure 4.8S | Reaction Times and Object Identity metrics, independent of general cognition – AD

Statistically significant maps are overlaid with metrics in AD in copper (lower cluster size) on top. The background layer represents the findings in the cohort of EHC and AD: Light blue = Identification time (top row), yellow = Localization time (middle row), Identification accuracy = orange (bottom row). R= right, L = left.

For Spatial Localization - Precision metrics, in patients with AD increased Absolute Error was associated with lower GM volume in the right Middle and Inferior Temporal Gyrus, and right Supramarginal Gyrus (**Figure 4.9S**, top row, **Table 4.3S**). Higher memory Imprecision correlated with greater atrophy in the right Middle Temporal Gyrus (**Figure 4.9S**, bottom row, **Table 4.3S**).

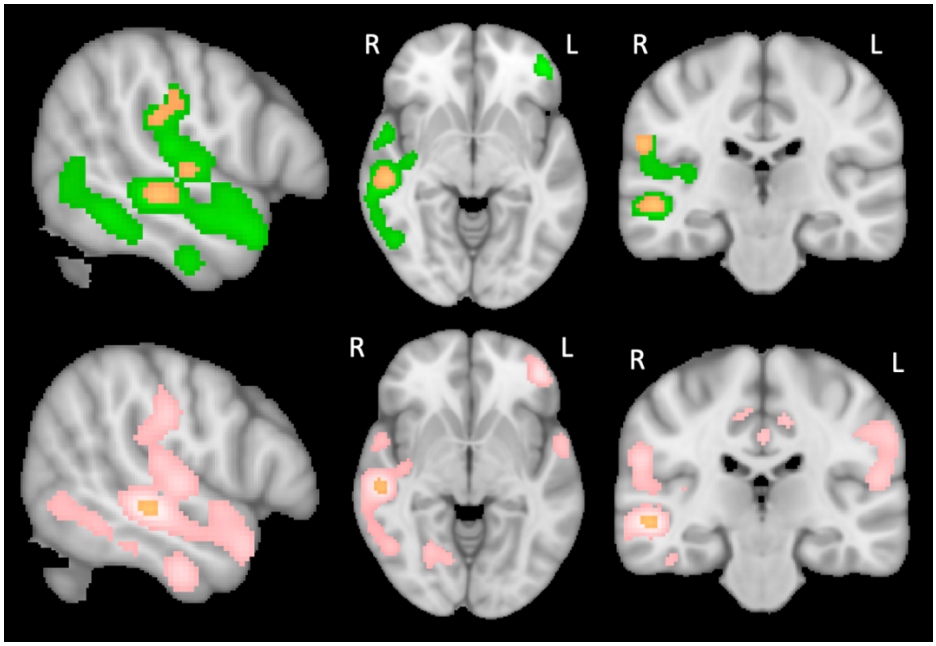


Figure 4.9S | Spatial Localization - Precision Metrics, independent of general cognition – AD

Statistically significant maps are overlaid with metrics in AD in copper (lower cluster size) on top. The background layer represents the findings in the cohort of EHC and AD: Light Green = Absolute Error (top row), pink = memory Imprecision (bottom row). R= right, L = left.

VSTM metric	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Identification Time	1	Frontal Pole (L)	12	-26	56	8	4.17
Localization Time	1	Paracingulate Gyrus Frontal Pole (L)	1333	-8	50	10	4.26
Identification accuracy	1	Frontal Pole (L)	266	-44	46	-8	4.15
Absolute Error	1	Middle Temporal Gyrus (R)	119	54	-24	-6	3.99
	2	Supramarginal Gyrus (R)	102	60	-24	24	3.91
	3	Heschl's Gyrus (R)	26	54	-14	4	3.71
	4	Inferior Temporal Gyrus posterior division (R)	25	62	-38	-16	3.5
Imprecision	1	Middle Temporal Gyrus (R)	36	54	-24	-6	4.11

Table 4.3S | Correlations between VSTM metrics and GM areas in AD - independent of general cognition

VSTM metrics independent of general cognition – in EHC

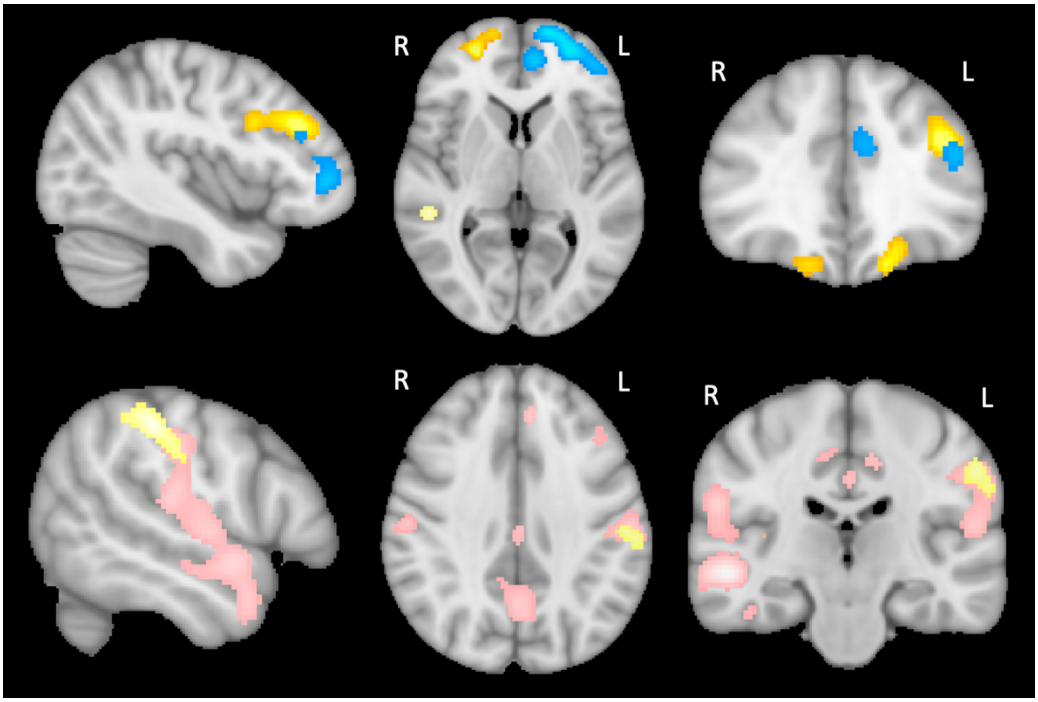


Figure 4.10S | GM atrophy and VSTM, independent of general cognition – EHC

Statistically significant maps are overlaid with EHC in light yellow for both metrics. The background layer represents the findings in the cohort of EHC and AD: Light blue = Identification time (top row) and pink = memory Imprecision (bottom row).

VSTM metric	Cluster	Atlas	Cluster Size	Coordinates (x,y,z)			Z-score
Identification Time	1	Frontal Pole (R)	959	20	48	-16	4
	2	Middle Frontal Gyrus (L)	790	-40	32	28	4.01
	3	Orbito-frontal Cortex (L)	310	-18	30	-20	4.04
	4	Middle Temporal Gyrus (R)	82	52	-36	0	5.4
	5	Middle Frontal Gyrus (R)	32	4	46	-20	3.33
Imprecision	1	Supramarginal Gyrus (L)	352	-56	-30	44	4.3

Table 4S | Correlations between VSTM metrics and GM areas in EHC

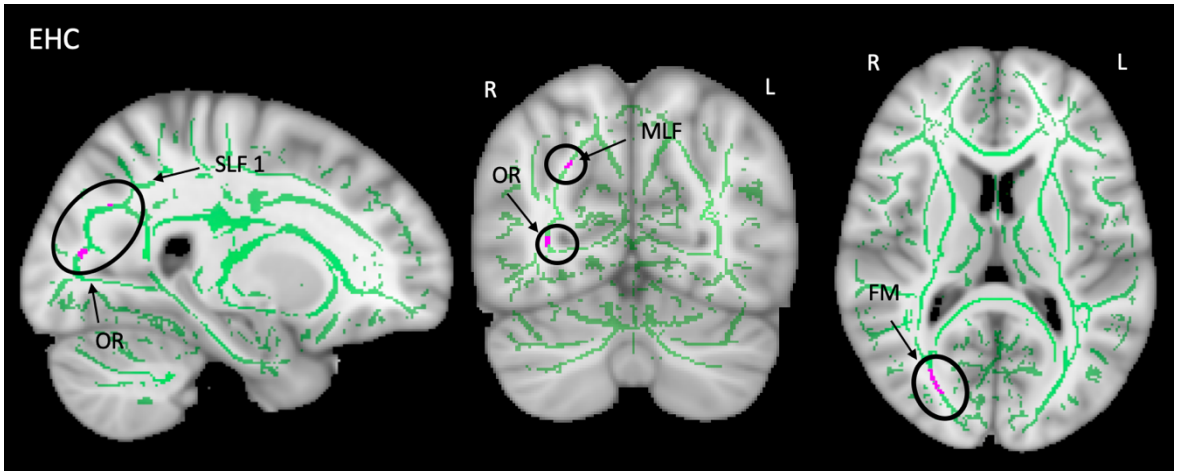


Figure 5.10 | Correlations between Identification accuracy and FA in EHC, independent of global cognition

	Cluster	White matter tract	Cluster Size	Coordinates (x,y,z)			Z-score
Identification accuracy	1	Vertical Occipital Fasciculus (R) (34%) Optic Radiation (R) (26%) Middle Longitudinal Fasciculus (R) (26%)	77	28	-74	13	3.39
	2	Forceps Major (R)	35	24	-74	36	4.75
	3	Optic radiation (R)	28	32	-67	4	3.46

Table 5.1S | Correlations between VSTM metrics and FA in EHC, independent of global cognition: Clusters and diffusivity peaks

R = right, L = left.

White matter tract	IA
Optic Radiation (R)	38.692
Middle Longitudinal Fasciculus (R)	19.023
Vertical Occipital Fasciculus (R)	15.542
Forceps Major	14.946
Inferior Fronto-occipital Fasciculus (R)	6.531
Superior Longitudinal Fasciculus 1 (R)	0.082

Table 5.2S | Correlations between VSTM metrics and FA in EHC, independent of global cognition: Tracts probability

IA = Identification accuracy. R = right, L = left.

	EHC (n= 36)	AD (n= 25)	p-value	Rank-Biserial Correlation
Age	69.23 (7.4)	69.57 (9.3)	0.538	
Gender (M/F)	15/21	9/16	0.655	
Education	15.8 (3.0)	13.9 (3.5)	* <0.001	
pTau181 (pg/ml)	2.8 (1.8)	5.2 (3.2)	* < 0.001	-0.579
GFAP (pg/ml)	119.9 (64.0)	249.9 (138.5)	* < 0.001	-0.706
NFL (pg/ml)	17.0 (7.8)	31.3 (21.7)	* < 0.001	-0.580
Aβ42/40 ratio	0.067 (0.01)	0.058 (0.01)	* < 0.001	0.530
ACE	97.8 (2.3)	66.9 (17.7)	* < 0.001	0.998
OMT - Identification time	2.77 (2.6)	3.78 (1.9)	* < 0.001	-0.420
OMT - Localization time	5.8 (2.9)	8.1 (3.5)	* < 0.001	-0.418
OMT - Identification Accuracy	0.92 (0.05)	0.81 (0.09)	* < 0.001	0.718
OMT - Absolute Error	124.5 (60.1)	234.5 (76.0)	* < 0.001	-0.773
OMT - Target detection	0.76 (0.1)	0.53 (0.2)	* < 0.001	0.801
OMT - Guessing	0.13 (0.1)	0.28 (0.1)	* < 0.001	-0.765
OMT - Misbinding	0.22 (0.1)	0.37 (0.1)	* < 0.001	-0.750
OMT - Imprecision	76.7 (46.8)	133.7 (56.3)	* < 0.001	-0.670
OIS – Immediate Object Accuracy	96.1 (5.7)	74.4 (19.9)	* < 0.001	0.828
OIS – Delayed Object Accuracy	72.1 (20.7)	30.5 (19.4)	* < 0.001	0.828
OIS – Immediate Location Error	0.97 (0.4)	3.04 (2.1)	* < 0.001	-0.747
OIS – Delayed Location Error	1.83 (1.1)	4.86 (2.8)	* < 0.001	-0.671
OIS – Immediate Semantic Accuracy	99.7 (1.5)	91.2 (12.1)	* < 0.001	0.571
OIS – Delayed Semantic Accuracy	82.3 (16.8)	42.1 (25.2)	* < 0.001	0.777
ROCF - copy	89.7 (9.8)	72.3 (26.9)	* < 0.001	0.544
ROCF - recall	84.8 (12.0)	50.1 (25.2)	* < 0.001	0.842
DSST - correct responses	34.0 (12.3)	14.6 (7.0)	* < 0.001	0.781
DSST - reaction times	3.8 (3.4)	10.2 (9.8)	* < 0.001	-0.786
TMT	35.0 (14.1)	99.3 (101.6)	* < 0.001	-0.861
CORSI	10.5 (4.2)	21.7 (7.6)	* < 0.001	-0.842

Table 6.1S | Demographics, plasma biomarkers and digital cognitive metrics for the longitudinal dataset

ACKNOWLEDGEMENTS

I hope after reading this thesis for Doctor of Philosophy, the (medical) Doctor who wrote it made you fond of (greek: 'Philos') and increased your knowledge (greek: 'Sophos') on the fine-grain cognitive changes occurring in Alzheimer's Disease. I certainly became even more passionate about pursuing wisdom than what my name - Sofia - had already gifted me with.

On this line, my parents are the first I should thank, as they have always encouraged me throughout my life. They taught me the value of hard work, and that the hardness will be always outclassed by the following joy of the reward. A special thanks goes to my mother Vallì, who has always provided me with an example of intellectual sharpness coupled with generosity and good will that I can only aspire to follow. I thank Franco for being such a steady and joyful rock, and my sister Irene for all her wise advice and deep understanding of life matters.

It is also a pleasure for me to thank Prof. Masud Husain, who is simply outstanding as supervisor, as clinician, as scientist and in terms of emotional intelligence. He has supported my growth in all possible ways in the past years, and always with such a positive and warm attitude that is hard to find. I also thank my second supervisor Prof. Clare Mackay for inspiring me to apply to Oxford in the first place just by reading her beautiful papers from afar, and for providing an excellent example of translational research in dementia with the Brain Health Clinic.

I also want to thank all the people in my lab that made it possible, especially Raquel for being such a great companion and support especially in days when we barely had time to eat between testing sessions, Rob for the filtering work, Bahaa and Michele for the useful neuroimaging tips, Sanjay for making me feel that if I ever had a question, the repository of ground truth was sitting three desks away from me, Xin-you for his many skills (from amazing coding, to eating tacos all day and still being slim, to conquering a tiger with his bare hands), Imran for sharing with me pratical ideas and his positive life attitude towards the DPhil and life, Kengo, Malin and Pablo for lifting my spirit whenever they are around. A special thanks goes to Verena for setting the bar very high on scientific rigour and skills

I have to achieve by looking at her example, and for all her warm help and company in and out of office, to Anna for making my days lighter on many levels and for being my favourite (non)neurologist, and to Sijia for all her amazing work on the digital battery, for being efficient, reliable and a lot of fun too.

My biggest thanks goes to my partner Cristian for all his useful advice and for his warm emotional support, I wouldn't have been able to do it without you.