Spatial memory deficits in mild cognitive impairment: a virtual reality study of hippocampal and entorhinal functioning

Elizabeth J. V. Harding

D.Clin.Psy. thesis (Volume 1), 2018

University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Name: Elizabeth J. V. Harding

Date: 19th June 2018

Overview

This thesis explores the topic of Alzheimer's disease (AD) and the earlier clinical state of mild cognitive impairment (MCI), with a particular focus on early identification and therapeutic intervention. The first part comprises a systematic review of the literature base examining the efficacy of transcranial direct current stimulation (tDCS; a form of non-invasive brain stimulation) in the treatment of MCI and AD. Specifically, this review focuses on the efficacy of tDCS for improving cognitive outcomes in MCI and AD patient groups. A database search identified fourteen studies that examined the relationship between tDCS and cognitive outcomes in these patient populations. The findings of these studies were summarised separately for MCI and AD patient groups. Results in both patient groups were found to be tentatively positive, however minimal research was carried out with MCI patient groups. Further, the heterogeneity of the identified research designs limited firm conclusions as to the factors associated with efficacy. Results are considered in tandem with an assessment of methodological quality. Consideration is given to the clinical implications of these findings, as well the areas that would benefit from further exploration in future research.

In the second part of this thesis, an empirical paper is presented that examines the utility of a novel spatial memory task in the identification of early AD symptomatology. A virtual reality (VR) object-location memory task (OLT) was used to assess aspects of spatial memory that are underpinned by brain regions known to be affected in the earliest stages of AD: the hippocampus and entorhinal cortex (EC). In order to assess the utility of this task in a proof-of-concept study, the OLT was administered to patients diagnosed with amnestic MCI (aMCI), a diagnosis that represents a high-risk for conversion to AD, and healthy control participants.

Alongside the OLT, a comparator battery of neuropsychological tests and a flat-screen measure of hippocampal function were administered as an index of the construct validity of the task. Results showed that there were significant group differences in performance on the OLT, and that task performance was able to predict group membership (aMCI or control) with a high degree of accuracy. Further, OLT performance was shown to be correlated with comparator measures of cognitive function. These results were interpreted as evidence for the utility of the OLT as a diagnostic measure. The implications of these findings were discussed in terms of the brain regions that the OLT may recruit, as well as the limitations of this study and how these might be addressed in future research.

The third and final part of this thesis is a critical appraisal of the research process. This offers a reflective exploration of the experiential components of the OLT and how these might compare to more traditional measures of neuropsychological assessment. The emotional challenges associated with neuropsychological assessment are discussed, making use of quotes from the author and prominent AD spokesperson, Terry Pratchett. This section also includes reflections from the researcher on the process of carrying out this research, and the learning that took place as a result.

This is a joint thesis, carried out with Adrienne Li (DClinPsy, 2018); a summary of the contributions of each author to this study is given in Appendix E.

Impact statement

The current study has a number of implications both in the domains of academic research, as well as clinical utility. In academic terms, this project represents a forward step in the use of commercially available virtual reality (VR) technologies in integrative neuroscience and clinical research. Here, we have demonstrated the successful use of novel, fully-immersive VR with a healthy older adult population, as well as with individuals who experience clinically significant cognitive difficulties. This is something that is not currently routinely practiced within academic research and therefore represents an important contribution to the methodological research base.

More specifically, the current study shows that VR paradigms can be used to assess aspects of neurological functioning that may not be routinely assessed in either pen-and-paper neuropsychological tests, nor in flat-screen computer tasks. Immersive VR paradigms, such as that employed in the current study, are uniquely placed to assess aspects of spatial learning that incorporate movement-related feedback – such as spatial learning that is supported by the entorhinal cortex (EC). Here, we have provided preliminary evidence that a VR paradigm targeting these aspects of spatial memory can be used to distinguish cognitively impaired patients from non-impaired counterparts. Future research may then build on this by demonstrating a link between performance on such measures and neuroimaging data, as well as through establishing a link between task performance and conversion to Alzheimer's disease (AD) over time. Both of these research aims would further support the concept that VR measures of spatial memory may be used to assess EC functioning that may be compromised in at-risk groups.

This research also has a number of potential clinical benefits. This work supports previous research that has demonstrated the utility of spatial learning paradigms targeting the hippocampus and entorhinal cortex as an effective tool for clinically diagnostic information. Here, we have demonstrated that such a paradigm can be used to distinguish individuals with cognitive impairment from non-impaired counterparts. It is yet to be shown that this tool can be used to distinguish AD pathology from other forms of cognitive impairment, however it does lay the groundwork for future studies that may do so. If it can be shown that tests such as these are able to distinguish underlying AD pathology from non-AD cognitive impairment, then this would lend strong support to the use of spatially-informed VR tasks in the early diagnosis of AD.

A significant implication of the current study is that it speaks to the potential integration of ecologically valid technology into clinical practice. To date, the high expense of VR software has rendered this an impractical means of neuropsychological assessment either in research or clinical practice. However, the high-quality and readily available commercial VR software opens up potential new avenues for the use of VR in clinical assessments. This is particularly the case when considered in relation to the high cost of neuropsychological assessment, which in some cases are in fact more expensive than some commercial VR equipment, while offering reduced ecological validity and, potentially, less specific focus on symptoms relevant to the early stages of AD.

Table of contents

Index of Tables

Literature Review

Empirical paper

Index of Figures

Acknowledgements

Firstly, I would like to thank the participants in this study for giving their time and energy to make this possible. It was an honour to witness their generosity and to hear their stories. A very big thank you goes to my supervisor, John King, for his incredibly thoughtful, responsive and supportive input at every stage of the research process. To my project partner, Adrienne Li, for being not just a wonderful colleague, but also a great friend since our days on first year placement. My gratitude and admiration go also to Andrea Castegnaro for his design of the computer and virtual reality tasks. From Cambridge, I would like to thank Dennis Chan for facilitating this research and allowing us to recruit from his memory clinic. I would particularly like to thank the Dream Team, David Howitt and Emma Barham, for being the linchpins in the recruitment and running of the project, as well as for their constant patience with our anxiety-filled emails. Thanks also to Zoe Adler and Volker Reisner for all their amazing help with testing. I would like to thank my clinical supervisors, Richard Rushe and Patricia McHugh, for their understanding and support in balancing my research and clinical demands. Thanks to the UCL 2015 cohort for making the process of training so much more enjoyable, and in particular to my flatmates, Lisa Masters and Melissa Barry, for going through the write-up experience alongside me and helping it to feel a far more bearable process. I would like to thank my family and friends for their support and care throughout my life in general and my training in particular. Especial thanks go to my parents for their constant love and encouragement, as well as their patient draft reading. Finally, I would like to thank the Deliveroo drivers of South West London, for never verbally judging me.

Part 1: Literature Review

Transcranial direct current stimulation (tDCS) as a therapeutic tool in Alzheimer's disease and Mild Cognitive Impairment.

Abstract

Aim: There is a clinical need for the identification of efficacious therapeutic tools for the treatment of mild cognitive impairment (MCI) and Alzheimer's disease (AD). This review aimed to assess the use of transcranial direct current stimulation (tDCS) as a treatment for MCI and AD, with a focus on cognitive outcomes.

Method: A systematic review of the literature was carried out using the databases PsychINFO, PubMed, MEDLINE and EMBASE. Data were synthesised in a narrative format and included consideration of methodological quality, assessed using the QualSyst tool. A manual search of citations and references was performed in adjunct. This process identified 1652 papers that were screened for eligibility.

Results: Fourteen papers met eligibility criteria for inclusion in the current review. Of these, three addressed the use of tDCS in MCI patients, all of which reported positive clinical outcomes on measures of cognition. A further eleven papers addressed the use of tDCS for AD patients. Findings in this regard were again largely positive, with eight studies reporting improved clinical outcomes following the use of tDCS. Three high quality studies reported negative findings, making firm conclusions difficult to draw.

Conclusions: The literature is tentatively in support of the therapeutic use of tDCS for cognitive outcomes in MCI and AD. It was difficult to draw firm conclusions regarding efficacy, or the specific factors that may contribute to beneficial therapeutic outcomes, largely due to the heterogeneity in study protocols.

1 Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder characterised by progressive deterioration in memory and cognitive functioning, with particular early deficits noted in episodic memory, alongside behavioural disturbances (e.g. e.g. McKhann et al., 1984). It is widely recognised that AD represents a rapidly increasing public health burden, with sources estimating that 1 million people in the UK will have been diagnosed by 2021 (Alzheimer's Society, 2014). Consequently, the identification of efficacious treatments for AD has become a high research and public health priority (e.g. WHO, 2012). Despite this increased focus, current therapeutic efforts have been recognised as having limited value in altering disease progression.

It is recognised that the cognitive deficits linked to AD exist on a continuum and likely begin at an earlier stage, clinically referred to as mild cognitive impairment (MCI). MCI is diagnosed when an individual presents with marked cognitive decline in one or more domains in the absence of a functional decline in daily activities (Albert et al., 2011). While it is understood that MCI is a broad diagnostic category, associated with varied aetiology, symptomatology and clinical outcomes, there remains a subsection of individuals for whom it represents a prodromal stage of AD; this is primarily thought to be those presenting with an amnestic variant of the condition (aMCI; e.g. Fischer et al., 2007). For this reason MCI and AD are increasingly viewed as two points on a continuum of impairment, associated with varying degrees of neuropathology and cognitive difficulty, but nevertheless underpinned, in many cases, by a shared pathology.

To date, the therapeutic focus for AD has centred on pharmacological interventions. This approach is based on the principle that a neurotoxic build-up of proteins are responsible for cognitive decline through the degeneration of cortical

tissue (e.g. Ramirez-Burmudez, 2012). Intercellular plaques made up of β-amyloid aggregates are thought to contribute to neuronal inflammatory processes and, it is argued, to the development of intracellular neurofibrillary tangles made up of hyperphosphorylated tau (Mattson et al., 1992). These neurofibrillary tangles are also thought to contribute to neuronal death due to the damage they inflict on the cytoskeleton. Based on this understanding, a number of drugs have been developed that seek to prevent the further spread of the neuropathological processes that underpin AD. These fall into two broad classes of drugs: acetylcholinesterase inhibitors and NMDA-receptor antagonists (Bishara, Sauer & Taylor, 2015). The former seeks to promote cholinergic transmission, which is otherwise compromised due to the loss of cholinergic neurones and reduced levels of acetylcholine, while the latter aims to reduce neuronal loss caused by excitotoxicity (i.e. neuronal death due to over-exposure to glutamate, an excitatory neurotransmitter). However, despite the promising theoretical backdrop to pharmacological intervention, in practice the literature has struggled to provide strong or consistent support for the efficacy of these drugs, with some studies reporting a success rate of as little as 0.4% (Cummings, Morstorf & Zhong, 2014). Where positive clinical gains are reported, these are often associated with small effect sizes and, as is the case with all pharmacological interventions, side-effects. While cognitive interventions, such as cognitive stimulation therapy, have offered some positive outcomes for AD patients (Spector et al., 2010), these are aimed primarily at later stage intervention and clinicians may face a number of practical challenges in their implementation (Choi & Twamley, 2013). Further, a distinction must be drawn between therapeutic efforts towards symptom management as compared to those that seek to effect change via altering the course of disease progression; while pharmacological treatments and cognitive therapies may

act toward the former goal, there remains a dearth of therapies aiming to effect change at a neuropathological level.

In this vein, there is growing support for the use of non-invasive brain stimulation techniques as an alternative intervention in neurodegenerative disorders (e.g. Hansen, 2014). Non-invasive brain stimulation is an umbrella term that may refer to multiple specific methods of intervention, primary among which are transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS). Transcranial stimulation differs from both pharmacological and cognitive interventions in that it allows for change to be affected at a direct neurological level, although the precise mechanisms of action depend on the method of stimulation used. In the case of TMS, changes in neuronal excitability are induced through the generation of a local magnetic current. The duration and intensity of the magnetic current, as well as the number of sessions an individual undergoes, are thought to regulate the extent of the impact (e.g. Guerra et al., 2011).

In contrast to TMS, tDCS is thought to induce a mild electrical current that may either increase or decrease neuronal excitability, dependent on the polarity of the current. Broadly, it is thought that anodal stimulation increases the membrane potential by a magnitude of several millivolts, while cathodal stimulation decreases it (Nitsche et al., 2008). In this way, tDCS may contribute to either increased or decreased neuronal firing in a polarity dependent manner. The mechanisms underlying the long-term after effects of this form of stimulation are yet to be fully understood, however it has been proposed that they may enhance synaptic plasticity through long-term potentiation (LTP) or long-term depression (LTD)-like processes. This idea is supported by research demonstrating that NMDA receptor agonists are successful in enhancing the synaptic effects of tDCS (Nitsche et al., 2004), while NMDA

antagonists suppress these effects (Fritsche et al., 2010); given the known role of NMDA receptor in LTP-dependent learning, this suggests that tDCS may operate via similar mechanisms. This is unlikely to constitute the whole story as many other factors, such as increased post-stimulation levels of brain derived neurotrophin factor (BDNF), are associated with tDCS (e.g. Fritsch et al., 2010). However, it does allow us to begin to construct a picture wherein tDCS application modulates cortical excitability in both the short and long term via complex neuromodulatory processes, all of which suggest that it has potential as a tool for the enhancement of learning and memory.

There is emerging evidence that tDCS can be used to successfully enhance cognitive outcomes in healthy participants; for example, Martin, Lui, Alonzo, Green & Loo (2014) showed that anodal tDCS applied to the dorsolateral prefrontal cortext (DLPFC) could be used to improve performance on a working memory cognitive training. Researchers compared performance on a dual n-back working memory task across conditions of active and sham stimulation and found that those in the active treatment condition were more accurate in their responses and showed a steeper learning curve; it should, however be noted that these improvements did not translate into post-stimulation performance improvements, as no group differences were noted during 'offline' task performance at follow-up. The finding of enhanced working memory performance associated with anodal tDCS has been supported elsewhere (e.g. Talsma, Kroese & Slagter, 2017; Gill, Shah-Basak & Hamilton, 2015). Indeed, other researchers have gone further in showing that the effects of tDCS can not only be demonstrated during online stimulation, but may also be sustained at follow-up. Meinzer et al. (2014) showed that anodal tDCS applied to left temperoparietal areas during a language acquisition paradigm enhanced performance in active as compared

to sham stimulation. The magnitude of these changes is also notable, with performance differences in recall associated with a medium-large effect size (Cohen's d = 0.73) reported at the final learning session and a medium effect retained at follow-up (Cohen's d = 0.6).

The use of transcranial stimulation as a tool for enhancing cognitive outcomes has not been limited to healthy populations, but has increasingly been applied as an intervention in the neurodegenerative disorders. As the field of transcranial stimulation has developed, a number of reviews have aimed to assess the efficacy of these techniques in the treatment of AD or MCI (e.g. Nardone et al., 2012; Gonslavez et al., 2017). Most recently, Birba et al. (2017) evaluated the literature on both TMS and tDCS in parallel for MCI patients. They concluded that results were mixed, and attributed this to the diversity of methodological approaches. Somewhat earlier, a similar combination review focussing on AD patients identified evidence of cognitive improvement associated with transcranial stimulation, however again noted the mixed nature of results and the limited conclusions that could be drawn (Elder & Taylor, 2014). Throughout these reviews a clear thread emerges in which the outcomes appear tentatively positive, however firm conclusions are hampered by the limited research base and diversity of methodological approaches. Further, despite these attempts to synthesise the literature base on transcranial stimulation, it is consistently noted that the broad methodological applications of either TMS or tDCS individually can hamper the possibility of drawing clear conclusions regarding efficacy. There is therefore a strong argument for a more refined approach that evaluates the individual contribution of TMS and tDCS as independent techniques. In keeping with this, a recent review and meta-analysis has examined the unique contribution of repetitive TMS (rTMS) to cognitive outcomes in AD, including only methodologically rigorous RCTs in the

analysis (Cheng et al., 2017). The authors concluded that there was good evidence that high-frequency rTMS improved outcomes for individuals with mild to moderate AD. Given the clear clinical and research imperative to identify efficacious treatment avenues for AD, alongside the current stagnation in the field of pharmacological research, this review aims to examine the current state of play as regards the use of tDCS in MCI and AD. The cognitive deficits that develop throughout disease progression are the key clinical characteristic of these conditions and the focus of this review will therefore be on clinical outcomes related to cognition and memory. Through examining both the prodromal and clinical stages of the disease it is hoped that any stage-dependent effects, such as can be observed in the use of pharmacological interventions, may be identified. The current review therefore aims to address the following questions:

(1) Is tDCS an effective method of improving cognitive outcomes for individuals with MCI?

(2) Is tDCS an effective method of improving cognitive outcomes for individuals with AD?

2 Method

2.1 Inclusion criteria

2.1.1 Participants

Studies were included if patients had received a primary diagnosis of AD or MCI, as assessed by scores on a validated screening tool, such as the Mini Mental State Examination (MMSE; Folstein et al., 1975) or Addenbrooke's Cognitive Examination-Revised (ACE-R; Mioshi et al, 2006), or by nationally recognised

diagnostic criteria. In the literature a distinction is often drawn between multi-domain MCI and amnestic MCI (aMCI), a presentation characterised by primary memory impairment; the latter is thought to be at highest risk of progression to a diagnosis of AD (e.g. Fischer et al., 2007). While there is an argument for inclusion of studies that focus only on the highest risk individuals (i.e. with a diagnosis of aMCI), in practice there is at present a limited research base in this area. Inclusion only of these studies would limit significantly the conclusions that could be drawn about the efficacy of tDCS at earlier stages of disease progression; therefore at this stage a decision was made to retain all studies where MCI was a primary diagnosis.

2.1.2 Cognitive outcomes

In order to be included in this review, the outcome of interest was required to relate to a domain of cognition and memory. Outcomes could refer to either one specific cognitive domain (e.g. visual memory), or to a broader assessment of cognition (e.g. the ACE-R). While this limits the conclusions that can be made about efficacy as regards specific cognitive processes, it does allow for a broader picture of efficacy. Given that AD symptomatology, although relating primarily to memory, can result in a broad spectrum of impairment, it was felt that wider inclusion would represent a more holistic view of the efficacy of tDCS.

2.1.3 Measures

Outcome measures included a quantitative measure of cognition.

2.1.4 Transcranial direct current stimulation

Studies were eligible for inclusion if the primary intervention was tDCS, although were also included if other interventions, such as cognitive training or medication, were used in adjunct. No parameters were set for the duration, location or intensity of tDCS, as this represents a further refining of literature that may at this

stage be premature. However, the inclusion was limited to a standardised tDCS protocol; studies including slow-wave oscillation tDCS were excluded on the basis that they operate via different hypothesised mechanisms (i.e. facilitation of sleep-dependent learning only, which may not be directly comparable to other forms of tDCS).

Where tDCS had been used as a tool to evaluate other aspects of cognition or neural processes, these studies were excluded; only those for which tDCS was applied as a therapeutic intervention were deemed appropriate for inclusion.

2.2 Search strategy

2.2.1 Electronic search

An electronic search of the following databases was carried out: PsychINFO, MEDLINE, PubMed and EMBASE. The databases were searched using the terms "Alzheimer's disease" OR "AD" OR "dementia" AND ("tDCS" OR "transcranial direct current stimulation" OR "TMS" OR "transcranial magnetic stimulation" OR "deep brain stimulation" OR "DBS" AND ("treatment" OR "treat*" OR "therap*" OR "therapeutics")). These search terms were designed to be broad and did not reference cognitive outcomes, as the potential terminology used would be too wide ranging.

Search limits were applied to remove all papers that were not written in English, as it was not possible to obtain support with translation. Additionally, any studies on non-human subjects were excluded via the filter function.

2.2.2 Data sorting

An initial search identified 1652 papers from the above search terms. A screen of titles and abstracts was then applied by the researcher in order to exclude those studies that did not meet inclusion criteria on this basis.

From this, 29 studies were assessed as examining an intervention for tDCS on cognitive outcomes in AD or MCI. Following an examination of the full text, studies were excluded based on the following: not published in a peer reviewed journal (n = 9); they reporting on a mixed (i.e. including diagnoses other than AD or MCI) patient population (n = 4); use of a sleep-based tDCS protocol (n = 1); and describing only a prospective study protocol (n = 1). Fourteen studies were therefore eligible for inclusion in the following review.

2.3 Quality assessment

All papers that met the inclusion criteria were subject to a formal assessment of quality using the Qualsyst tool for quantitative research papers (Kmet, Lee & Cook, 2004). This tool evaluates fourteen methodological and reporting factors including, but not limited to, the clarity of research aims, methods, sample size and selection, as well as the use of blinding and randomisation. In brief, this tool operates a scoring system ranging from 0-2, with a higher score indicating a greater quality of design and reporting. Where it was felt that a criterion was not applicable to a given study, it is possible to rate this as 'not applicable'. Subsequent to rating each criterion a summary score was given, ranging between 0 and 1. If a criterion was deemed non-applicable then it was possible to adjust for this while tabulating the final summary score. In this way, only the quality of relevant criterion would contribute to the summary score. This is important when considering findings that may span a diverse range of research methodology; for example, both RCTs and case study designs may provide clinically informative information, and yet it would be a challenge to apply all criterion equally to both. This allowed for the inclusion and evaluation of a broader range of clinically relevant research. In addition, the QualSyst tool has been shown to have an inter-rater

reliability ranging between 73-100% in terms of for quantitative research studies (Kmet, Lee & Cook, 2004), making it a valid tool for quality assessment.

This tool was selected on the basis that it had been designed specifically to assess clinical outcomes studies and as such was well-suited for the format of studies included in this review. Where studies reported multiple research aims (e.g. assessing both neuroimaging and cognitive outcomes), quality was assessed based on the criteria as they applied to the measurement of cognitive outcomes only, as these were the focus of the current review.

2.4 Data synthesis

The focus of this review was on change pre-and post- intervention. It was therefore considered that the outcomes would be helpfully synthesised meta-analytically; however, heterogeneity in outcome data across studies precluded this possibility. Consequently, the data have been presented in a narrative format. The results will first summarise the quality of the papers included, before moving on to a consideration of the literature as it applies to MCI and AD separately. This is a helpful distinction to draw, as studies invariably intervened with only one of these patient groups, as well as being in keeping with the rationale that interventions may have differential efficacy at different stages of disease progression.

The analysis of quality assessment will not include an in-depth assessment of each paper included; rather, the literature base to date will be summarised according to the relative strengths and weaknesses of the field, as well as providing a brief summary of those studies that were ranked among the highest or lowest quality in the current review. In this way, we aim to inform the understanding of the current state of the field more generally, as well as to highlight the studies for which particular

attention – whether positive or negative - is warranted. Further consideration of quality assessment will be included into a summary of the clinical outcomes.

3 Results

3.1 Quality assessment (QA)

Overall, the studies included in the current review ranged in methodological quality between 42% and 93%, therefore representing a broad spectrum of study quality. Individual scores for each criterion, as well as summary scores for each study, can be seen in Table 1. In order to examine this information in closer detail, we will begin first with a summary of those studies that occupied both the highest and lowest ends of the spectrum in terms of methodological quality, before going on to consider the criteria for which consistently high or low scores were obtained across studies.

The current review identified Andrade et al. (2016), Bystad et al. (2017), Murugaraja et al. (2017) and Marceglia et al. (2016) as the lowest scoring studies in the field, all of which received a score ranging between 42% and 54%. With the exception of Murugaraja et al., this reflects the case-study design of the studies. For this reason, all were necessarily rated as not fulfilling the criterion of appropriate sample size, nor did they report any attempt blinding for either subject or researcher. In the case of Andrade et al., further issues arose due to the confounding factor of cognitive training administered alongside tDCS; no attempts to control for the potential impact of this on clinical outcome were reported; consequently, it is difficult to disentangle the effects of tDCS from those of the cognitive intervention.

In contrast, the highest scoring study was also from Bystad et al. (2016), who received a score of 93% for their double-blind and sham controlled trial. The methodological strengths of this research included the combined use of both subject

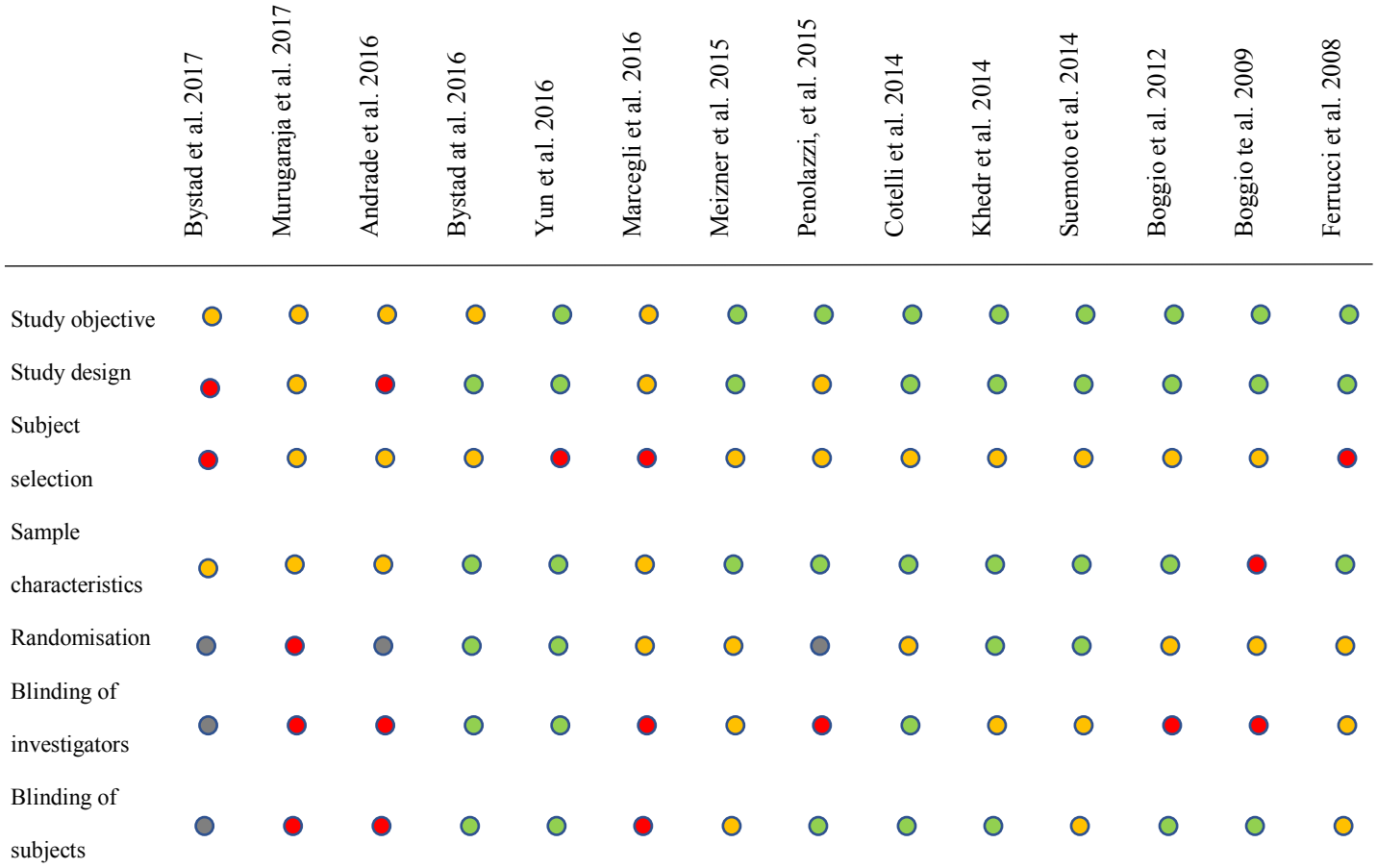
and researcher blinding, comprehensively described randomisation procedures and thorough reporting of results. Here, the only criterion for which a 'partial' score was given related to the reporting of outcomes and an incomplete description of control group selection. Sharing a score of 89%, Cottelli et al. (2014), Suemoto et al. (2014) and Khedr et al. (2014) were those studies for whom the next highest ratings were given. All three reports received a high proportion of 'full' ratings, with weaknesses again identified primarily limited to the use of an opportunity sample for patient selection.

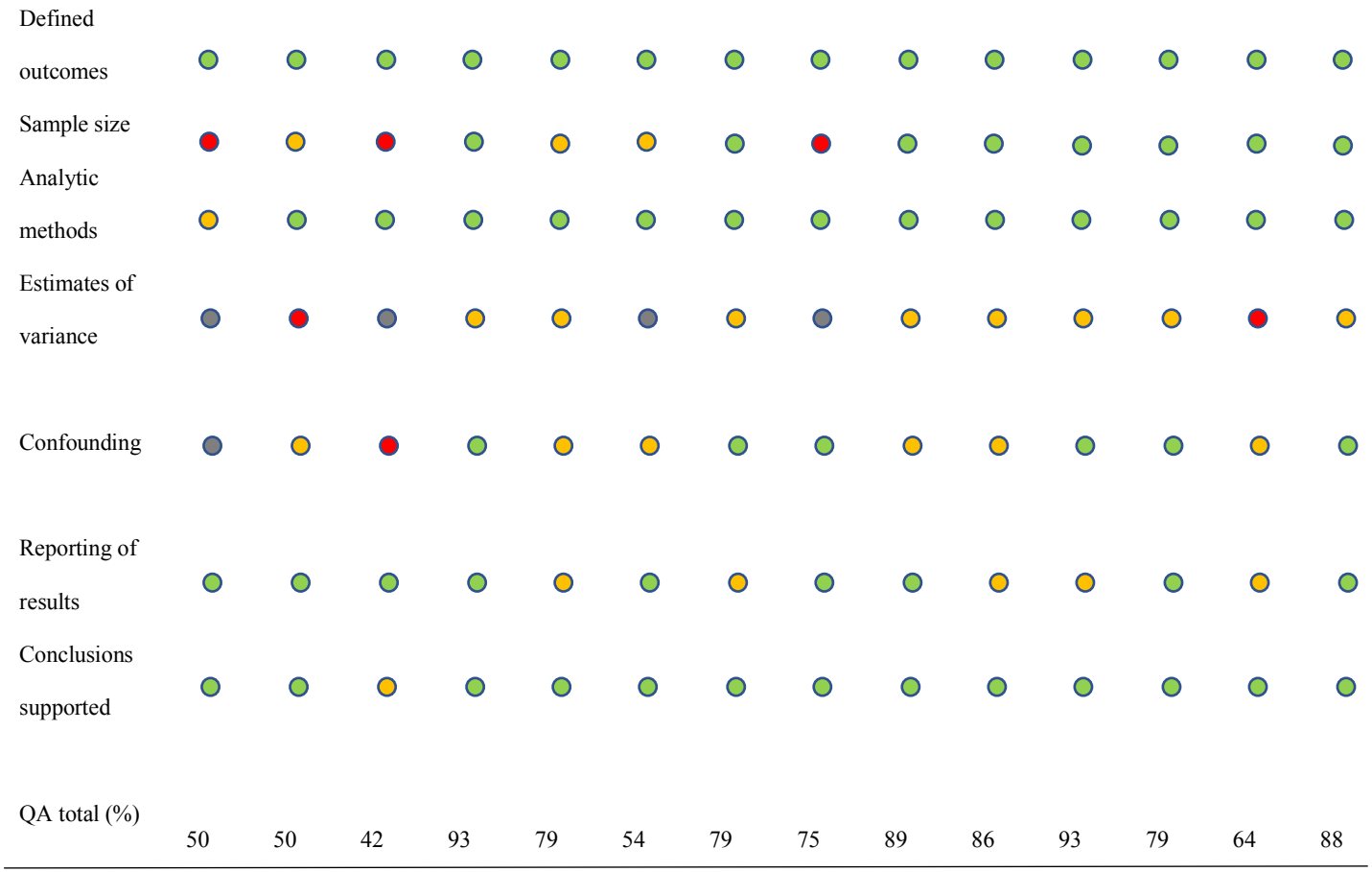
Across the studies reported, consistent high scores were obtained in terms of the use of appropriate and objective outcome measures. There was no instance of a study reporting the use of subjective or non-replicable data, despite the variation in the nature of clinical outcomes (i.e. use of both neuropsychological batteries and individualised measures of cognition). Further, there was a consistent high standard in the interpretation of the data. Further strengths included the use of appropriate analytical methods across studies, with only Bystad et al., (2017) receiving a partial score; this was assigned due to their use of only descriptive data to report on the case study. It is, therefore, in the thorough analysis of robust clinical data that the field to date finds it strengths.

Finally, consistent weaknesses were identified in the areas of patient recruitment and blinding of investigators, with only three studies receiving 'full' scores in the latter instance and no studies scoring above a 'partial' fulfilment in the former. These are both areas that have important implications for the interpretation of findings. The lack of thoroughly outlined recruitment procedures means that it is difficult to eliminate the possibility of any bias in participant selection at this stage, nor are we able to draw firm conclusions about the generalisability of results outside

of a given study. The lack of adequately described researcher blinding protocols similarly introduces the possibility of bias in the administration process and interpretation of outcomes. These areas therefore represent important weaknesses in the field to date.

Table 1: Quality assessment (QA) ratings according to study





Key: ● = Criterion fulfilled in full; ● = criterion partially fulfilled; ● = criterion not fulfilled; ● = criterion not applicable.

3.2 Studies with MCI samples

The earliest identified study of tDCS in an MCI population was carried out in 2015 by Meinzer and colleagues (QA 79%), who studied semantic word-retrieval following stimulation of the inferior frontal gyrus (IFG). Importantly, this paper recruited participants who had been diagnosed with amnestic MCI only; this allowed them to draw more specific conclusions regarding the therapeutic utility of tDCS for individuals with this diagnostic subtype. The authors reported that anodal stimulation improved the performance of patients on a computerised word-retrieval task, stating that post-stimulation performance was comparable with that of control participants. This finding is notable as it is based on a single session stimulation protocol, with intensity of stimulation relatively lower in both duration and intensity than many other studies. Improvements in measures of cognition are therefore linked, in the authors' view, to short-term modulation of membrane potentials; this is contrasted to longer-lasting changes that are thought to be linked to protein synthesis following multiple stimulation sessions. The main difficulty in the interpretation of these results stems from incomplete reporting on the processes of randomisation and blinding. Despite this issue this study was assessed as being methodologically robust, lending support to the conclusions drawn.

Subsequently, Yun et al. (2016) (QA 79%) investigated the effect of both anodal and cathodal tDCS on memory function, as assessed using the Multifactorial Memory Questionnaire (MMQ), a self-report measure of subjective memory function (Troyer & Rich, 2002). Stimulation was applied to left dorsolateral prefrontal cortex (left DLPFC), an area associated with long-term memory. Following three weeks' stimulation, applied thrice weekly, the authors reported significant improvements in both the contentment and ability subscales of the MMQ relative to sham conditions;

this difference was not observed for the strategy subscale. Based on these findings, the authors conclude that multiple tDCS sessions do produce improvements in transient memory performance. It should, however, be considered that these conclusions are based on a self-reported measure of ability. While it may be argued that the use of a subjective measure of primary outcomes does align with the main presenting problem of MCI – i.e. subjective memory complaints – the parallel use of an objective or informant report of cognition would lend support to the conclusions drawn. This conceptual limitation aside, the work of Yun and colleagues received a high rating of methodological quality, indicating that this is a methodologically robust piece of research.

The most recent research was carried out by Murugaraja et al. (2017) (QA 50%), who attempted to adapt the study of tDCS and cognition to an Indian setting. For this reason, cognitive outcomes were assessed via the Picture Memory Impairment Test (PMIT), a measure adapted by the authors in order to assess immediate and delayed visual recall while accounting for the variability in language and educational background of patients. The cortical target was again the left DLPFC, and anodal stimulation was applied. It was reported that patients showed improvement in immediate and delayed recall as measured by the PMIT, and that this improvement in delayed recall was sustained at one month follow-up. This is the only study with MCI patients to report follow up data, and it is promising that this shows evidence for sustained improvements after the tDCS protocol has finished. It should, however, be noted that this study received one of the lowest ratings of quality in this report. There were a number of contributory factors to this low score, significantly among which was the lack of blinding of either researchers or participants. In the absence of blinding it is difficult to draw firm conclusions regarding the true meaning of these results.

Further difficulties were noted in the reporting of participant characteristics and recruitment procedures. It is unclear how patients were identified for inclusion in the study, nor is there any further information regarding the clinical or personal characteristics of the participants; for example, there is no reference to any concurrent medication use and the authors provide only the mean age of participants. As a consequence, it is difficult to paint a clear picture of the patient group involved in this study and it would therefore be inappropriate to generalise these findings to anyone outside of the included participants. Nevertheless, this study does represent an impressive attempt to adapt and evidence a complex intervention within a culturally and clinically diverse patient group, and makes an interesting contribution to the current report.

3.2.1 Summary of findings: MCI

There are relatively few studies that have examined the use of tDCS within an MCI patient group. Of the three identified studies, all reported positive cognitive outcomes. The standard of methodological quality among these was relatively high, with two out of three papers scoring above 75% when assessed using the QualSyst tool. Cortical targets were the left DLPFC and the IFG, therefore conclusions regarding efficacy can only be made as regards stimulation of these areas. There was variability in the manner of outcomes measures, with two papers making use of a measure developed by the authors themselves (Meinzer et al., 2015; Murugaraja et al., 2017) and one reporting subjective memory improvement (Yun et al. 2016). Only one study (Meinzer et al., 2015) restricted their work to the amnestic subtype of MCI (aMCI), which is most closely linked to future onset of AD. Overall, there a positive picture is painted for the use of tDCS in an MCI population.

3.3 Studies with AD samples

A seminal paper examining the use of tDCS to treat cognitive symptomatology in AD was published by Ferruci et al. in 2008 (QA 88%). In this article, both anodal and cathodal stimulation was used to stimulate temperoparietal (TP) areas, citing hypoactivity in AD populations as a rationale for selecting this cortical target. Outcomes were assessed via a word-recognition test (WRT) that the authors adapted from the Alzheimer's Disease Assessment Scale-Cognitive version (ADAS-Cog). Performance across groups was compared both to a sham stimulation condition and on a visual attention task; these conditions were included to allow conclusions to be made about whether TP stimulation induces task-specific effects on recognition memory, as opposed to broader cognitive changes (e.g. increased attention). They reported that only anodal stimulation resulted in improved recognition memory; those who received cathodal stimulation performed worse on the WRT and performance was unchanged in the sham condition. No stimulation-related changes were observed in the visual attention task. These results were based on a single stimulation session and suggest promising implications for the use of tDCS in AD. The authors conclude that tDCS has therapeutic benefit for the cognitive symptoms of AD, although do not rule out the possibility that the results may be linked to changes in attentional processes. While the authors report both patient and researcher blinding in addition to randomisation, little information regarding these processes is provided. This suggests that the findings are likely to be robust, however some caution should be held when interpreting the data as we are unable to thoroughly assess these aspects of the protocol.

Several studies in this area have been carried out by Boggio and colleagues, who were among the earliest to investigate the use of tDCS in AD. Their first study, published in 2009 (Boggio et al., 2009) (QA 64%), applied tDCS to both the left

DLPFC and left temporal cortex. The cognitive outcomes under consideration were attention, working memory (WM) and visual recognition memory (VRT), assessed respectively via the Stroop task, digit span, and a computerised task of recognition memory designed by the experimenters. They reported that relative to sham stimulation, active anodal tDCS in both areas enhanced VRT performance, however had no such impact was observed in relation to WM or attentional processes. From this, it was concluded that tDCS has a specific effect on task facilitation for VRM, with broader non-specific effects of attention ruled out.

This research group then went on in 2012 to assess the long-term effects of repeated stimulation, using a protocol refined from this initial study. Boggio et al. (2012) (QA 79%) extended their protocol to five sessions of anodal tDCS to the temporal lobes. Outcomes were again assessed using the VRM paradigm as a primary outcome, however the ADAS-Cog, MMSE and Visual Attention Task (VAT) were included as secondary measures of cognition. A follow-up period of four weeks was included to assess whether cognitive changes were sustained over time. This study replicated the initial finding of VRM improvements following anodal tDCS, and further showed that these improvements did persist four weeks post- stimulation. In this more recent study, the quality of Boggio and colleagues' work improved in terms of methodological quality, placing it among the most robust studies included in this review and increases one's confidence in the findings. However, it should be highlighted that some important concerns remain regarding the lack of investigator blinding, which do in some ways limit the conclusions drawn.

Subsequent to these initial findings, Suemoto et al. (2014) (QA 93%) went on to study the impact of lDLPFC stimulation on individuals diagnosed with AD. While the primary focus of this study was on apathy, ADAS-Cog outcomes were included as

Table 2

Summary of results

Author	Patient group	Design	Mean age of active group	Gender	N	Medication controlled	Stimulation site	Type of stimulation	Current in mA (duration)	Number of sessions	Control	Measures of cognition	Improvement in any outcome measure
Bystad, M. et al. (2017)		CS	60		1		Left temporal lobe/right frontal		2 (30)	Daily 8 months	None	RBANS	
Murugaraja, V. et al. (2017)		PP	59.6		11		Left DLPFC		2 (20)	5	None	Picture memory impairment test (PMIT)	
Andrade, S. et al. (2016)		CS	73		1		DLPFC		2 (30)	10	None	ADAS-Cog	
Bystad, M. et al. (2016)		PG	70		25		Left temporal lobe		2 (30)	1	Sham	CVLT-II, MMSE, TMT A & B	
Yun, K. et al. (2016)		R, DB	74.75		16		DLPFC		2 (30)	9	Sham	Multifactorial Memory	

Marceglia, S. (2016)	PG	75.4		7		Bilateral TP		1.5 (15)	2 (1 anodal, 1 cathodal)	None	Questionnaire (MMQ) Word recognition task	
Meinzer, M. (2015)	DB, CO	67.44		18		Left ventral IFG		1 (20)	1	Sham and HCs	Semantic word retreival	
Penolazzi, B. (2015)	CS	60		1		Left DLPFC		2 (20)	10	Sham	WRT, VWMT, PFT, CPT	
Cotelli, M. (2014)	PG	76.6		36		Left DLPFC		2 (25)	10	Sham	Face-name association task (FNAT)	
Khedr, E. (2014)	PG	69.7		34		Left DLPFC		2 (25)	10	Sham	MMSE, WAIS-III	
Suemoto, C. (2014)	DB, PG	80.5		40		Left DLPFC		2 (20)	6	Sham	ADAS-Cog	

Boggio, P. et al. (2012)	CO, WS	77.5/80.6*		15		Temporal cortex		2 (30)	5	Sham	Visual recognition memory, ADAS-Cog	
Boggio, P. et al. (2009)	CO, WS	79.1		10		Left DLPFC and left temporal cortex		2 (30)	1 per region	Sham	Stroop, digit span, Recognition memory	
Ferrucci, R. (2008)	CO, WS	75.2		10		Temporoparietal		1.5 (15)	1 each	Sham	Word recognition and visual attention	

Key: Patient group: = AD; = MCI

Study design: CO = cross-over; DB = double-blind; WS = within-subjects; CS = case study; PG = parallel groups; R = randomised; PP = Pre/Post only

Sex: = female; = male

Medication: = controlled for; = no control reported

Type of stimulation: = anodal; = cathodal

Result: = improvement in outcomes; = no improvement in outcomes

*ages reported separately across test sites

a secondary measure. Suemoto and colleagues applied a blinding protocol, with patients randomly allocated to receive either active anodal or sham stimulation. In contrast to earlier studies, no significant effects on cognitive outcomes were reported in this instance. The authors consider the negative findings as potentially relating to disease-stage factors, particularly that disease progression may have been too advanced to observe changes with a limited number of stimulation sessions. In terms of quality, Suemoto's work represents a very high level of methodological rigor. It is therefore important to highlight that these negative findings bear significant weight when interpreting this body of literature.

In the same year, Khedr et al. (2014) (QA 86%) conducted a large-scale study on the effects of repeated tDCS over a longer-term follow-up period. Following on from earlier studies, the stimulation site was LDLPFC, however cognitive outcomes were assessed via the MMSE and performance on the WAIS-III. In contrast to other protocols, cathodal stimulation was included alongside anodal and sham as an intervention method, with patients allocated randomly to each condition. The authors reported that both forms of stimulation were associated with improvements on MMSE scores that continued at both one and two month follow-up. WAIS scores were analysed as performance, full-scale IQ and verbal IQ; the only significant improvement was noted for performance IQ, where cathodal stimulation showed greater improvements than sham. It was concluded that the WAIS may not be suitable for cognitive assessment in patients with dementia, and for this reason no observable effects on cognition were observed on this measure. Along with Suemoto and colleagues, this study was assessed as a high quality paper. Any concerns in this regard were limited to minor issues relating to reporting and unlikely to reflect any serious concerns in study design. Therefore, a good deal of confidence can be placed in these results.

In an alteration from existing research paradigms, Cotelli et al. (2014) (QA 89%) included a condition of memory training into the traditional stimulation protocol. They coupled tDCS, applied over the lDLPFC, with a computerised training programme aimed at improving performance on the Face-Name Association memory Task (FNAT), which was the primary cognitive outcome in this study. The training programme was individualised based on the baseline performance of each patient and centred around an errorless learning method. Performance on the FNAT in this group was compared to a control group who took part in cognitive training alongside sham stimulation. As a further control, a motor (i.e. non-cognitive) training condition was included in which patients underwent the active tDCS protocol while also taking part in a motor training programme. The authors reported that improvements were observed following cognitive training, regardless of stimulation condition; improvements did not differ significantly across active stimulation and placebo conditions. This was the case both at the end of treatment and at long-term follow-up (2 and 4 months post-stimulation). Again, this study was rated at 89% in terms of methodological quality and therefore is considered to be among the most robust in this literature base. While it could be considered that the efficacy of the cognitive training programme was such that any stimulation-related changes were comparably non-significant (i.e. non-additive), it remains that in this context tDCS did not offer any measurable therapeutic gains over and above cognitive training.

In a further work from the Ferrucci et al. research team, a more recent study has examined the link between tDCS and performance on a word recognition test (WRT). Marceglia et al. (2016) showed that in a single stimulation session, anodal tDCS applied to temperoparietal areas could induce performance improvements on the WRT relative to sham stimulation. Again, the authors report a directional effect whereby anodal stimulation was associated with improvements in WRT, while

cathodal stimulation was associated with worse WRT outcomes, although these findings were non-significant in the latter instance. Given the relatively lower intensity of stimulation used in this study, in which only a single active session was administered and at lower amplitude than many other studies, this is a notable finding. One of the methodological difficulties with this study was the lack of blinding of either patients or experimenters, as well as the use of a convenience sample; these factors introduce the potential for biased results.

Multiple of the more recent studies in this field have been carried out by Bystad et al.'s research team. The earliest of these (Bystad et al., 2016) (QA 93%) was a randomised sham-controlled trial with double-blinding applied throughout the intervention period; this was among the highest quality studies included in this review. In this instance, the primary outcome measure was the California Verbal Learning Test (CVLT-II), used as an index of verbal learning. However, this study reported no improvement on any subscale of the CVLT-II, or on any secondary outcome measures, which included the MMSE and TMT. The authors attribute their null findings to several possible factors, including the use of a fixed stimulation protocol, as opposed to individualised localisation for cortical targets. Further, it was noted that the patient group showed advanced disease progression and low baseline performance rates, potentially indicating that the neuroplasticity thought to be induced by tDCS may have been lacking.

3.3.1 Case studies

More recently, several researchers have reported the effects of tDCS in single case studies. The earliest of such reports comes from Penolazzi et al. (2015) (QA 75%), who administered both an active and sham protocol to a man in his sixties, thereby placing him in the position of a control for himself. Both pre- and post-stimulation the patient engaged in ten daily sessions of cognitive training, comprising the WRT, a

verbal working memory task (VWMT), a phonemic fluency task (PFT) and a continuous performance task (CPT). Stimulation was applied to the lDLPFC in a series of ten sessions, in combination with daily cognitive training. A two-month interlude was then given before then repeating an identical cycle, however with sham stimulation in lieu of active intervention. The authors reported that the VWMT was the only aspect of the training battery to show a treatment-related effect. It is important to note that in this study although the patient was blind to the condition, researchers were not, thereby introducing the possibility of bias. Despite this, and the lack of generalisability inherent in a case-report design, Penolazzi et al.'s report remains the highest quality case study to be included in this review, with clear attempts made to mitigate the impact of common limitations, such as the lack of control or the reporting of purely descriptive data.

A second such case study comes from Andrade and colleagues (2016) (QA 42%), who similarly administered lDLPFC stimulation alongside cognitive training. In this instance, however, the ADAS-Cog was used as a primary measure of cognition. The authors reported improvement on several subscales of the ADAS-Cog, most notably in the 'commands' and 'spoken language ability' domains, where post-stimulation performance was 100% improved relative to baseline. Other notable, although less substantial improvements, were observed in the domains of word recall and orientation. While a promising result, this should be interpreted with caution in light of the lack of control conditions. Further, it is reported that the patient was also receiving both medication and cognitive training with a neuropsychologist at the time of the intervention. While this is reported, no attempts are made to control for the potential therapeutic effects that either would have on cognitive outcomes. For this reason, alongside other issues with study design and reporting, this study was assessed

as being the lowest quality in the present review. It is therefore important to interpret these findings with caution.

In 2017 the research group of Bysted and colleagues (QA 50%) went on to apply a long-term stimulation protocol in a case-study design, with a neuropsychological assessment battery carried out at baseline, midpoint and at the end of the intervention. The research team trained the patient, a man in his sixties, to self-administer anodal stimulation to the left temporal lobe on a daily basis for a period of eight months; this was by far the longest period of intervention reported in any identified study. The repeatable battery for the assessment of neuropsychological status (RBANS) was used in order to monitor cognitive functioning throughout the intervention. The authors report stabilisation in cognitive functioning over the study period, as well as noting improvements in immediate and delayed recall of 39% and 23% respectively. As this research uses a case-study design, there are inherently limitations in generalisability; however, there were additional difficulties with the reporting, with limited information provided as to the nature of patient selection and a lack of clarity around study design. For these reasons results should be interpreted with due caution.

3.3.1 Summary of findings: AD

There are mixed results as to the efficacy of tDCS as a therapeutic tool in AD. While the majority of studies (8 out of 11) reported improvements in cognition following tDCS stimulation, 3 studies reported no effect. Improved outcomes were reported both on measures designed to target specific cognitive process (e.g. semantic word-retreival), as well as on battery-based measures of cognition (e.g. ADAS-Cog). Cortical targets for stimulation included the dorsolateral prefrontal cortex, the inferior frontal gyrus, the left temporal lobe and the temperoparietal lobe. Efficacy was reported following tDCS in all of these areas. The only studies to report a long-term follow-up of outcomes were Boggio et al. (2012), Khedr et al. (2014) and Cotelli et al. (2014). All reported stability in performance at follow-up; i.e. any immediate improvements were sustained over time, or null effects (Cotelli) remained the same at follow-up.

4 Discussion

4.1 Therapeutic efficacy of tDCS

The application of transcranial stimulation as a therapeutic tool for MCI and AD is relatively recent, with research into the efficacy of tDCS beginning a little over a decade ago. Since that time a substantive body of research has emerged, with the field growing rapidly. The majority of studies reported positive outcomes on a variety of cognitive measures, including both clinical diagnostic tools (e.g. the ADAS-Cog, MMSE) and more targeted tools developed for research purposes (e.g. adapted word-retrieval tasks). However, the findings of the current review are not unanimously positive in regards to the therapeutic potential of tDCS. It is clear that tDCS is a complex neuromodulatory technique, with factors such as the polarity, intensity and

duration of stimulation to be considered, as well as the cortical target and cognitive outcome of interest. With this in mind, the present review aimed firstly to synthesise what is known about its efficacy in both AD and its prodromal stage, before considering how these factors may play a role in influencing differential outcomes.

4.1.1 MCI

There are relatively fewer studies that focus on the prodromal stage of disease progression, with only three meeting criteria for the current review. Of those studies that are available, methodological quality was high in only two out of the three reports. Therefore despite the positive result reported in these studies, it is too soon to draw any firm conclusions regarding efficacy for the use of tDCS in an MCI population. Caution should be exercised both due to the small number of papers on which the conclusions are based, as well as the limitations inherent in the papers reported. The studies themselves rely on relatively small sample sizes, with no study exceeding eighteen participants. When considered in combination with the use of opportunity sampling, which was a limitation in all reported studies, this raises questions about the generalisability of findings.

Further, it is important to closely examine the nature of the cognitive outcomes reported. Two of these studies reported improvement based on an outcome measure developed by the researchers themselves (Meinzer et al., 2015; Murugaraja et al., 2017). While this does not mean that they do not assess a valid cognitive construct, it remains that these studies lack an independent assessment of the primary outcome measure and findings are not comparable to other studies. This is problematic, as in order to develop a coherent narrative within the literature base, it is necessary to ensure some way of comparing outcomes across studies. The third study, conducted by Yun and colleagues (2014), also made use of a potentially problematic outcome measure in that improvement is measured by subjective patient report. As noted above, this may

be theoretically aligned with the subjective nature of MCI, however for firm conclusions to be drawn it would be important to consider this in conjunction with an objective measure of cognition.

Given that an MCI diagnosis is associated with diverse clinical outcomes (e.g. Abner et al., 2017), it is important to note that this is a potential limitation of the literature as it stands. To date, only one study has attempted to refine the patient population to the amnestic variant of MCI (Meinzer et al., 2015), which is most likely to be associated with progression to AD. While it is promising that the literature in the MCI field to date indicates that tDCS is associated with positive outcomes, it is important to bear in mind that while the MCI subtype remains undefined conclusions in this regard are limited. For example, it is known that patients with MCI due to Parkinson's disease can also benefit from tDCS (e.g. Manetti et al., 2016).

4.1.2 AD

When considering the literature for AD, the picture becomes somewhat more mixed. Most studies reported positive outcomes, with the evidence base broadly in support of the use of tDCS to enhance cognitive outcomes (e.g. Khedr et al., 2014; Boggio et al. 2012); however, there were some notable exceptions. Bystad et al. (2016), Cotelli et al. (2014) and Suemoto et al. (2014) all reported no effect of tDCS on the cognitive outcome of interest. The reasons postulated for these negative findings varied, including a speculation that there may be dwindling effects of plasticity in the more severe stages of disease progression, highlighting the importance of considering disease stage when evaluating cognitive outcomes. A further technical issue, highlighted by Bystad and colleagues (2016), is the use of fixed vs. individualised stimulation protocols; the latter refers to the use of computational modelling to take into account factors such as skull density, allowing individualised placement of electrodes. It is true that all those studies reporting negative findings also made use of

a fixed protocol, however there were a number of studies that employed this protocol and yet reported findings in support of tDCS. For this reason, any of these explanations alone are unlikely to fully account for these negative findings.

It is also important to note that those studies reporting negative findings are also among the highest in terms methodological quality. An association between methodological rigour and the likelihood of reporting positive results is notable and should be given consideration, as it is important not to rule out the possibility that this may in part reflect a true negative outcome for the use of tDCS in an AD population. It is a well-reported phenomenon that studies reporting positive findings have a higher rate of publication (i.e. publication bias) and that this may skew a literature base. While it is not possible to conclude that this is the case on the basis of the present review, it is important to highlight this difference and to bear this in mind for future studies; it also again highlights the research imperative for further high quality studies in the field.

It is also worth noting that the choice of cortical target did not appear to be related to the efficacy of tDCS on cognitive outcomes; those studies reporting negative findings chose to target the temporal cortex and the dorsolateral prefrontal cortex, both areas that were associated with cognitive improvement elsewhere in the literature. This does, however, highlight a challenge in synthesising the literature, in that there was a notable lack of consistency in the choice of stimulation site across studies. This may potentially be appropriate given that there is also a great deal of heterogeneity in the cognitive outcome of interest; it is of both theoretical and clinical importance that the cortical targets of tDCS are chosen according to their hypothesised link to a targeted cognitive outcome. However, the disparity in both outcomes and stimulation sites does result in a literature base that makes it hard to conclude with any certainty that tDCS can consistently demonstrate therapeutic utility. In order to move forward, it may be

important to consider the identification of key outcomes of interest; from this point, it would then be easier to draw specific conclusions regarding tDCS efficacy. This may be a target for a future review, when the research base has expanded sufficiently to allow the evaluation of separate cognitive processes.

There are a number of technical considerations that must be made when considering the efficacy of tDCS in this population. First among these is the question of stimulation parameters; i.e. the amplitude, duration and frequency at which tDCS is applied. Within the studies reviewed here each of these parameters have been varied in some way, with some studies employing relatively low amplitude stimulation (e.g. 1mA), or using only single stimulation sessions. While there is not sufficient evidence at this stage to state that any one application protocol is superior to another, it is the case that neither a lower amplitude nor less frequent stimulation appeared to be associated with poorer outcomes in any of the studies reviewed here.

There does, however, appear to be a consensus emerging in the literature that anodal stimulation is used preferentially over cathodal, with only three studies attempting to include the latter condition (Ferrucci et al., 2008; Marceglia et al., 2016; Khedr et al., 2014). This may have been based on the initial finding by Ferrucci and colleagues that cathodal stimulation was associated with poorer performance on cognitive outcomes; this is consistent with the hypothesis that anodal stimulation is associated with increased neuronal excitability. Hover, it is worth noting that in the case of Khedr and colleagues this was not the case, with both anodal and cathodal stimulation associated with improved performance on the MMSE. Therefore while it can be concluded that the precedent in the literature to date for the use of anodal stimulation, the evidence for cathodal stimulation is not conclusively negative.

A conceptual issue that warrants consideration is how cognition and memory are defined in this context. Cotelli and colleagues speculated that more successful

outcomes may be observed for isolated cognitive processes, as opposed to outcomes related to associative learning (e.g. the Face-Name Association Test). It is true that this was the only study to examine associative mechanisms, with the majority of studies assessing outcomes via either a single aspect of memory, or a clinical neuropsychological measure (e.g. ADAS-Cog, RBANS). In this instance it is likely that a more widely distributed neural network would be activated, and that any short-term neural changes induced in one location may therefore not be sufficient to induce network-wide improvement. Further research into associative outcomes is warranted before any conclusions can be made in this regard.

Only a handful of the studies in the current review have included information regarding long-term follow-up (Boggio et al., 2012; Khedr et al., 2014; Cotelli et al., 2014; Murugaraja et al., 2017). This is important because in order to have true clinical utility, it would be necessary for the effects of tDCS to outlast the stimulation session by some time; if this were not the case, then it would be hard to justify the cost of continued intervention. Of those that did report follow-ups, the outcomes have appeared to remain stable over time. Interestingly, this is the case both for single stimulation sessions as well as repeated stimulation protocols. Therefore, while the small number of studies including follow-up information means that it is premature to reach firm conclusions, the literature appears tentatively positive in this regard.

Previous reviews considering the therapeutic benefits of transcranial stimulation have similarly highlighted the mixed nature of the literature, particularly in relation to factors such as the stimulation parameters, site of stimulation and outcomes of interest (Birba et al., 2017). Given the time elapsed since previous reviews of the transcranial stimulation literature, as well as the more refined focus on only tDCS, it was hoped that a certain degree of clarity may have emerged in this regard. This has not been the case and it remains a challenge to the evidence base that there

are such a variety of factors that may influence the apparent efficacy of this intervention.

It was notable that throughout the studies no commentary was made in regard to effect sizes. Further, due to either incomplete reporting, the use of non-parametric data or alternative estimates of variance (e.g. the reporting of standard error as opposed to standard deviation), it was not possible to calculate effect sizes for the majority of studies reported. This is relevant as it limits the conclusions that can be drawn regarding the magnitude of clinical effect that we might expect to observe when applied to real-world clinical settings. Further, it limits the extent to which this can be incorporated into considerations of sample size when designing future studies; for the purposes of the QualSyst tool, adequate sample size can be implied by the finding of significant between group differences. While this is a technically accurate assessment it remains that the lack of consideration of power or effect size does not stand out as an evident finding across studies. As a result, the magnitude of potential clinical benefits cannot be assessed at this time. However, were tDCS to be adopted into routine clinical practice for AD and MCI it would represent a new class of therapeutic intervention, in which the aim is to promote plasticity through non-invasive neuromodulation. The implication of this is the potential for clinically significant gains in cognitive function following an initial diagnosis of MCI or AD; this would be highly clinically significant in the context of a diagnosis for which the aim is currently only to maintain cognitive function and prevent further decline.

4.2 Clinical recommendations

The overarching finding of this review is that tDCS appears to be a safe and potentially effective therapeutic tool for the treatment of cognitive symptoms in MCI and AD. This should be interpreted with the caveat that several high quality studies

report null findings, therefore efficacy cannot at this stage be said to be conclusively demonstrated. It is, however, important to note that the safety of tDCS was not explicitly assessed within this review and was not routinely examined as part of the study design, with studies inconsistently reporting patient feedback on any adverse effects. Based on the literature to date it appears that the strongest evidence base is in relation to the use of anodal stimulation, with potential stimulation sites largely considered to be the left dorsolateral prefrontal cortex, as well as temporal and temperoparietal areas. The choice of stimulation site should be selected based on a consideration of the targeted cognitive outcome. Following on from this, it should be noted that research to date only supports the use of tDCS in relation to isolated cognitive processes, as opposed to associative processes that may depend on the contribution of multiple cortical systems. In the majority of studies reviewed, tDCS was used as an adjunct to pharmacological interventions. There is therefore no evidence to support the use of tDCS as a standalone treatment option; it remains to be seen whether this tDCS can be used more successfully when applied in conjunction with cognitive training. Currently, it is a challenge to make further recommendations as to the most successful manner in which to implement this clinically due to the heterogeneity of the research base. Before further conclusions regarding clinical utility can be drawn, research must first move towards elucidating the factors that contribute to successful versus unsuccessful clinical outcomes.

4.3 Research recommendations

The difficulty in drawing any firm conclusions from the clinical evidence to date is likely to be a reflection of the disparate nature of the research base. In order to move forward from this challenge, which has also plagued previous attempts to review the literature on transcranial stimulation, it is necessary to consider the current

strengths and weaknesses of the literature, both in terms of methodological quality and conceptual backdrop. From this, it is possible to inform the next steps in constructing methodologically sound and clinically informative research studies.

A consistent methodological weakness identified across studies was the use of opportunity sampling; many studies reported either minimal recruitment from a single clinic site or, in some instances, no descriptors were given regarding the process of participant selection. Further, it was noted that the lowest quality studies in the current review were among the most recent. This is an unusual finding and appears to reflect an increased movement towards the use of case study designs. If the research base wishes to move forward in quality then it would be necessary to include thoughtfully designed multi-site studies that aim to include a patient sample representative of the MCI/AD population.

The current review focussed on the effect of tDCS on cognitive outcomes in MCI and AD, with outcomes assessed primarily through performance on neuropsychological measures. While it was not the focus of the current review, it is of note that measures of daily living and functional performance were not assessed in the studies included here. In order to more fully demonstrate clinical utility it would be necessary to assess whether any observed cognitive improvements successfully translated into improvement on functional outcomes; for example, the Alzheimer's Disease Cooperative Study activities of daily living inventory (Galasko et al.,1997) is a tool that assesses functional capacity and could helpfully be incorporated alongside cognitive outcomes. Future research in this field would benefit from incorporating tools to assess functional outcomes and quality of life, alongside cognitive improvement.

The primary conceptual issue that the literature must address is the diversity of approaches taken to the use of tDCS in this patient group. There is little cross-study

consistency in the cognitive outcomes of interest, nor is there a firm consensus reached on the most effective stimulation parameters (i.e. cortical target, number of sessions or duration of stimulation). It would be helpful in this regard to be led by clinical need, with a hypothesis-driven approach to addressing the core clinical symptomatology of AD and MCI (e.g. long-term memory deficits). Further refinement would also be beneficial in the MCI literature in particular, where the potential heterogeneity of the patient population hampers our ability to form firm conclusions about efficacy of tDCS for prodromal AD. Increased focus on an amnestic MCI patient group would address this concern. Finally, studies would benefit from a movement toward the use of long-term follow-ups as a standard practice in this area of research. For a strong case to be made for the use of tDCS in future clinical practice, it will be necessary to demonstrate the long-term implications of this intervention.

4.4 Limitations

There were a number of limitations to the current review that may impact on interpretation of the results. Firstly, the results included are restricted to those papers that were written in the English language. While this has not prevented the inclusion of several studies that were conducted in non-English speaking settings (e.g. Murugaraja et al., 2017), it may nevertheless represent a Western skew of the evidence base. Further, only those papers that were included in a peer-reviewed journal have been included. This was set as a criterion for inclusion in order to ensure the quality of the papers assessed; a variety of study designs were included in order to attempt to present the broadest possible range of clinical evidence, however it is possible that a wider evidence base may be lacking due to this omission. Finally, due to limited resources it was not possible to have either the literature screen or the quality assessment verified by a second assessor. It is therefore possible that there is some bias

in the screening and assessment of papers which would be helpfully addressed by the inclusion of a second researcher or independent assessor.

4.5 Conclusions

The literature to date is cautiously in support of the use of tDCS as a therapeutic tool in MCI and AD patients. The studies included present largely positive results in terms of cognitive outcomes, which do not appear to be linked to the duration or intensity of the stimulation protocol. There are some notable exceptions to this, with several high-quality studies reporting null findings. For this reason, it remains necessary to take a tentative perspective on the clinical use of tDCS. One key difficulty in interpreting these findings is the heterogeneity of the literature, with a wide range of stimulation protocols and cognitive outcomes assessed. Further research is needed to refine those factors that contribute to successful clinical outcomes, as well as defining the cognitive difficulties that may be most helpfully addressed by the use of tDCS.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Garmst, A., Holtzman, D., Jagust, W. J., Petersen, R.C., Snyder, P. J., Carrillo, M.C., Thies, B., & Phelps, C.H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, *7*(3), 270-279.

Alzheimer's Society (2014). Dementia UK: The Second Edition. Alzheimer's Society: London

Andrade, Suellen Marinho, Mendonça, Camila Teresa Ponce Leon de, Pereira, Thobias Cavalcanti Laurindo, Fernandez-Calvo, Bernardino, Araújo, Regina Coely Neves, & Alves, Nelson Torro. (2016). Adjuvant transcranial direct current stimulation for treating Alzheimer's disease: A case study. *Dementia & Neuropsychologia*, *10*(2), 156-159.

Birba, A., Ibáñez, A., Sedeño, L., Ferrari, J., García, A. M., & Zimerman, M. (2017). Non-Invasive Brain Stimulation: A New Strategy in Mild Cognitive Impairment?. *Frontiers in aging neuroscience*, *9*, 16.

Bishara, D., Sauer, J., & Taylor, D. (2015). The pharmacological management of Alzheimer's disease. *Progress in Neurology and Psychiatry*, *19*(4), 9-16.

Boggio, P. S., Khoury, L. P., Martins, D. C., Martins, O. E., De Macedo, E. C., & Fregni, F. (2009). Temporal cortex direct current stimulation enhances performance on a visual recognition memory task in Alzheimer disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *80*(4), 444-447.

Boggio, P. S., Ferrucci, R., Mameli, F., Martins, D., Martins, O., Vergari, M., Tadini, L., Scarpini, E., Fregni, F., & Priori, A. (2012). Prolonged visual memory enhancement after direct current stimulation in Alzheimer's disease. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, *5*(3), 223-230.

Bystad, M., Grønli, O., Rasmussen, I. D., Gundersen, N., Nordvang, L., Wang-Iversen, H., & Aslaksen, P. M. (2016). Transcranial direct current stimulation as a memory enhancer in patients with Alzheimer's disease: a randomized, placebo-controlled trial. *Alzheimer's research & therapy*, *8*(1), 13.

Bystad, M., Rasmussen, I. D., Grønli, O., & Aslaksen, P. M. (2017). Can 8 months of daily tDCS application slow the cognitive decline in Alzheimer's disease? A case study. *Neurocase*, *23*(2), 146-148.

Cheng, C. P. W., Wong, C. S. M., Lee, K. K., Chan, A. P. K., Yeung, J. W. F., & Chan, W. C. (2018). Effects of repetitive transcranial magnetic stimulation on improvement of cognition in elderly patients with cognitive impairment: a systematic review and meta-analysis. *International journal of geriatric psychiatry*, *33*(1), e1-e13.

Choi, J., & Twamley, E. W. (2013). Cognitive rehabilitation therapies for Alzheimer's disease: a review of methods to improve treatment engagement and self-efficacy. *Neuropsychology review*, *23*(1), 48-62.

Cotelli, M., Manenti, R., Brambilla, M., Petesi, M., Rosini, S., Ferrari, C., Zanetti, O., & Miniussi, C. (2014). Anodal tDCS during face-name associations memory training in Alzheimer's patients. *Frontiers in aging neuroscience*, *6*, 38.

Cummings, J. L., Morstorf, T., & Zhong, K. (2014). Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimer's research & therapy*, *6*(4), 37.

Elder, G. J., & Taylor, J. P. (2014). Transcranial magnetic stimulation and transcranial direct current stimulation: treatments for cognitive and neuropsychiatric symptoms in the neurodegenerative dementias?. *Alzheimer's research & therapy*, *6*(5), 74.

Ferrucci, R., Mameli, F., Guidi, I., Mrakic-Sposta, S., Vergari, M., Marceglia, S., Cogiamanian, F., Barbieri, S., Scarpini, E., & Priori, A. (2008). Transcranial direct current stimulation improves recognition memory in Alzheimer disease. *Neurology*, *71*(7), 493-498.

Fischer, P., Jungwirth, S., Zehetmayer, S., Weissgram, S., Hoenigschnabl, S., Gelpi, E., Karmpla, W., & Tragl, K. H. (2007). Conversion from subtypes of mild cognitive impairment to Alzheimer dementia. *Neurology*, *68*(4), 288-291.

Folstein, M. F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.

Fritsch, B., Reis, J., Martinowich, K., Schambra, H. M., Ji, Y., Cohen, L. G., & Lu, B. (2010). Direct current stimulation promotes BDNF-dependent synaptic plasticity: potential implications for motor learning. *Neuron*, *66*(2), 198-204.

Gill, J., Shah-Basak, P. P., & Hamilton, R. (2015). It's the thought that counts: examining the task-dependent effects of transcranial direct current stimulation on executive function. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, *8*(2), 253-259.

Galasko, D., Bennett, D., Sano, M., Ernesto, C., Thomas, R., Grundman, M., & Ferris, S. (1997). An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. *Alzheimer disease and associated disorders*.

Gonsalvez, I., Baror, R., Fried, P., Santarnecchi, E., & Pascual-Leone, A. (2017). Therapeutic noninvasive brain stimulation in Alzheimer's disease. *Curr Alzheimer's Res*, *14*(4), 362-376.

Guerra, A., Assenza, F., Bressi, F., Scrascia, F., Del Duca, M., Ursini, F., Vollaro, S., Trotta, L., Tombini, M., Chisari, C. & Ferreri, F. (2011). Transcranial

magnetic stimulation studies in Alzheimer's disease. *International Journal of Alzheimer's Disease*, *2011*.

Hansen, N. (2012). Action mechanisms of transcranial direct current stimulation in Alzheimer's disease and memory loss. *Frontiers in psychiatry*, *3*, 48.

Hansen, N. (2014). Brain stimulation for combating Alzheimer's disease. *Frontiers in neurology*, *5*, 80.

Khedr, E. M., Gamal, N. F. E., El-Fetoh, N. A., Khalifa, H., Ahmed, E. M., Ali, A. M., Noaman, M., El-Baki, A., & Karim, A. A. (2014). A double-blind randomized clinical trial on the efficacy of cortical direct current stimulation for the treatment of Alzheimer's disease. *Frontiers in aging neuroscience*, *6*, 275.

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

Manenti, R., Brambilla, M., Benussi, A., Rosini, S., Cobelli, C., Ferrari, C., Petesi, M., Orizio, I., Padovani, A., Borroni, B., & Cotelli, M. (2016). Mild cognitive impairment in Parkinson's disease is improved by transcranial direct current stimulation combined with physical therapy. *Movement Disorders*, *31*(5), 715-724.

Marceglia, S., Mrakic-Sposta, S., Rosa, M., Ferrucci, R., Mameli, F., Vergari, M., Arlotti, M., Ruggiero, F., Scarpini, E., Galimberti, D., Barbieri, S. & Priori,

A. (2016). Transcranial direct current stimulation modulates cortical neuronal activity in Alzheimer's disease. *Frontiers in neuroscience*, *10*, 134.

Martin, D. M., Liu, R., Alonzo, A., Green, M., Player, M. J., Sachdev, P., & Loo, C. K. (2013). Can transcranial direct current stimulation enhance outcomes from cognitive training? A randomized controlled trial in healthy participants. *International Journal of Neuropsychopharmacology*, *16*(9), 1927-1936.

Mattson, M. P., Cheng, B., Davis, D., Bryant, K., Lieberburg, I., & Rydel, R. E. (1992). beta-Amyloid peptides destabilize calcium homeostasis and render human cortical neurons vulnerable to excitotoxicity. *Journal of Neuroscience, 12*, 376–389

Meinzer, M., Jähnigen, S., Copland, D. A., Darkow, R., Grittner, U., Avirame, K., Rodriguez, A., Lindenberg, R. & Flöel, A. (2014). Transcranial direct current stimulation over multiple days improves learning and maintenance of a novel vocabulary. *Cortex*, *50*, 137-147.

Meinzer, M., Lindenberg, R., Phan, M. T., Ulm, L., Volk, C., & Flöel, A. (2015). Transcranial direct current stimulation in mild cognitive impairment: behavioral effects and neural mechanisms. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, *11*(9), 1032-1040.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test

battery for dementia screening. *International journal of geriatric psychiatry*, *21*(11), 1078-1085.

Murugaraja, V., Shivakumar, V., Sinha, P., Venkatasubramanian, G., & Sivakumar, P. T. (2017). Transcranial direct current stimulation for mild cognitive impairment. *Journal of Geriatric Mental Health*, *4*(2), 106.

Nardone, R., Bergmann, J., Christova, M., Caleri, F., Tezzon, F., Ladurner, G., Trinka, E. & Golaszewski, S. (2012). Effect of transcranial brain stimulation for the treatment of Alzheimer disease: a review. *International Journal of Alzheimer's Disease*, *2012,*

Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., Paulus, W., Hummel, F. Boggio, P.S., Fregni, F. & Pascual-Leone, A. (2008). Transcranial direct current stimulation: state of the art 2008. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, *1*(3), 206-223.

Nitsche, M. A., Jaussi, W., Liebetanz, D., Lang, N., Tergau, F., & Paulus, W. (2004). Consolidation of human motor cortical neuroplasticity by D-cycloserine. *Neuropsychopharmacology*, *29*(8), 1573-1578.

Penolazzi, B., Bergamaschi, S., Pastore, M., Villani, D., Sartori, G., & Mondini, S. (2015). Transcranial direct current stimulation and cognitive training in the

rehabilitation of Alzheimer disease: a case study. *Neuropsychological rehabilitation*, *25*(6), 799-817.

Ramirez-Bermudez, J. (2012). Alzheimer's disease: critical notes on the history of a medical concept. *Archives of medical research*, *43*(8), 595-599.

Spector, A., Orrell, M., & Woods, B. (2010). Cognitive Stimulation Therapy (CST): effects on different areas of cognitive function for people with dementia. *International journal of geriatric psychiatry*, *25*(12), 1253-1258.

Suemoto, C. K., Apolinario, D., Nakamura-Palacios, E. M., Lopes, L., Leite, R. E. P., Sales, M. C., Nitrini, R., Brucki, S.M., Morillo, L. S., Magaldi, R.M., & Fregni, F. (2014). Effects of a non-focal plasticity protocol on apathy in moderate Alzheimer's disease: a randomized, double-blind, sham-controlled trial. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, *7*(2), 308-313.

Talsma, L. J., Kroese, H. A., & Slagter, H. A. (2017). Boosting cognition: effects of multiple-session transcranial direct current stimulation on working memory. *Journal of cognitive neuroscience*, *29*(4), 755-768.

Yun, K., Song, I. U., & Chung, Y. A. (2016). Changes in cerebral glucose metabolism after 3 weeks of noninvasive electrical stimulation of mild cognitive impairment patients. *Alzheimer's research & therapy*, *8*(1), 49.

World Health Organization. (2012). Dementia: a public health priority. *Dementia*, 112.

Part 2: Empirical Paper

Spatial memory deficits in mild cognitive impairment: a virtual reality study of hippocampal and entorhinal functioning

Abstract

Introduction: Neurophysiological research has identified that the hippocampus and entorhinal cortex (EC) are the brain regions affected in the earliest stages of Alzheimer's disease (AD). These areas play a key role in spatial memory, notably the allocentric location of an individual within an environment, as well as the location of objects within an environment. Given that a diagnosis of amnestic mild cognitive impairment (aMCI) represents a high-risk for conversion to AD, this study used a virtual reality (VR) task to establish whether a spatial memory task sensitive to EC and hippocampal function could be used as an early diagnostic measure for AD by establishing a proof-of-concept in this (aMCI) patient group.

Method: Twenty individuals diagnosed with aMCI and twenty-two age-matched healthy controls completed a VR study of object-location memory – the object-location task (OLT). This was followed by a desktop computer assessment of object recognition memory, as well as object-environment recognition. Alongside this, a comparator battery of neuropsychological tests sensitive to early AD was administered in order to assess construct validity of the OLT.

Results: Performance on the VR component of the OLT was poorer in aMCI patients compared to healthy controls. A model containing OLT response accuracy, a flat-screen measure of hippocampal function and premorbid IQ was highly significant and able to predict the patient status of participants (aMCI or healthy control) with 95.1% accuracy. Performance on object-only, but not object-context recognition was poorer in the aMCI group relative to controls. There were significant correlations between OLT performance and the neuropsychological testing battery for aMCI patients but not control subjects.

Conclusion: The OLT is a useful tool in distinguishing aMCI from healthy controls and has the potential to form the basis of a sensitive early diagnostic measure for AD. Performance on the VR component of the OLT relates to other known neuropsychological measures used diagnostically in AD, supporting the construct validity of the task. However, the OLT did not relate to a flat-screen measure of hippocampal functioning, potentially suggesting greater EC involvement in this task. Implications of these findings are considered and suggestions for future research are made.

1 Introduction

Alzheimer's disease (AD) is the most prevalent presentation of the dementias and, due to the effects of an aging population, represents a growing public health concern. It is estimated that 1 million people will receive this diagnosis by 2021 (Alzheimer's Society, 2014) and that the number of people diagnosed will double every twenty years (Prince et al., 2015). It is statistics such as this that have prompted a number of global health initiatives to name the identification and treatment of AD as a high public health priority (e.g. WHO, 2012). AD represents approximately two thirds of dementia diagnoses and is clinically distinguished from other forms of dementia due to the early symptomatic hallmarks of episodic memory loss and topographic disorientation (e.g. McKhann et al., 1984; Guariglia & Nitrini, 2009). A clinical diagnosis requires that individuals present with a gradual onset of symptoms that worsen over time and interfere with functioning (McKhann et al., 2011). While anterograde memory loss is considered the most prominent feature of AD, it is understood that this is often accompanied by impairment in language, executive function, visuospatial reasoning and semantic knowledge that, at later stages, progresses to global cognitive impairment (Salmon and Bondi, 1999).

Important links have now been made between the neuropathology of AD and its associated cognitive deficits. From a neurological perspective, AD is distinguished from other forms of dementia by the distinct neuropathological profile of β-amyloid plaques (Aβ) and intracellular tangles of the hyper-phosphorylated neurofibrillary protein, tau. Together, these plaques and tangles are understood to contribute to neural cell death through inflammatory processes and neuronal destruction (DeLaGarza, 2003). These neuropathological processes do not appear to occur in a uniform manner across brain regions, but rather are likely to begin in the medial temporal lobe, an area

associated with episodic memory (e.g. Scoville & Milner, 1957; Vargha-Khadem et al., 1997; Squire et al., 2004); specifically, these changes are thought to progress from a starting point in the entorhinal cortex (EC) to the hippocampus proper (Braak & Braak, 1991).

The link between the cognitive and behavioural symptomatology of AD and the neuropathological changes that underlie them is now more clearly understood, however, this has failed to translate into the development of more efficacious drug treatments. For many years the primary treatment options available for AD have been pharmacological, with anticholinesterase inhibitors and NMDA receptor antagonists as the two primary options (Bishara, Sauer & Taylor, 2015). However, not only are there debates as to the degree of clinical benefit associated with pharmacological interventions, but these interventions seek only to prevent or reduce the rate of further cognitive decline, rather than offering any remediation of symptoms already present, which reflect neural damage which may be irreversible. Non-pharmacological treatment options include cognitive training paradigms, such as cognitive stimulation therapy (CST; Spector et al., 2008) or, more recently, there has been tentative support for the use of non-invasive transcranial stimulation (i.e. transcranial magnetic or transcranial direct current stimulation) (for a recent review, see Birba et al., 2017 or Gonslavez et al., 2017). This latter avenue is in its infancy and it remains to be seen how and in what ways it may be incorporated into routine clinical practice. Given that current clinical interventions offer only support with maintaining cognitive function and reducing further decline, there is a clear imperative for the early identification of and intervention in AD.

It is now recognized that AD is preceded by a prodromal phase of mild cognitive impairment (MCI), in which individuals are symptomatic but do not yet meet diagnostic criteria for dementia. MCI is diagnosed when an individual shows

impairment in one or more cognitive domains in the presence of spared functional abilities (Albert et al., 2011). However, a significant clinical challenge remains in that MCI is a heterogenous diagnosis that includes many individuals who do not progress to develop AD. For this reason, a diagnosis of MCI alone is not sufficient to determine the likely endpoint or time-course of cognitive impairment; rather, efforts are now being made to identify those features of early cognitive decline that are likely to progress to a full diagnosis of AD. The term amnestic MCI (aMCI) has been coined to refer to specific presentations of MCI in which memory loss is the predominant early symptom and this patient group is broadly held to be at highest risk of conversion to AD; for example, Fischer et al. (2007) showed a 48% conversion rate to AD among aMCI patients compared to a 26% conversion rate in non-aMCI patients over a 30 month period. While this suggests higher conversation rates in this group, this is clearly not conclusively predictive of progression to AD. To date traditional neuropsychological tests are not able to reliably differentiate MCI due to AD (MCIad) from MCI due to other causes. Studies attempting to differentiate MCIad from non-AD MCI have provided a number of neuropsychological predictors, including (but not limited to) the Trail Making Test-B (Ewers et al., 2012); the Digit Symbol Test subtest of the Weschler Adult Intelligence Scale-Revised (Tabert et al., 2006); and the Orientation and Memory Subtests of the Cambridge Cognitive Examination (CAMCOG; Conde-Sala et al., 2012). While these studies indicate that neuropsychological measures are informative in the diagnosis of MCIad, there remains a lack of consensus as to clear neuropsychological predictors of conversion to AD. While recent advances have made it possible to diagnose MCIad through the use of amyloid and tau biomarkers that are detectable in cerebrospinal fluid (CSF), these methods are both invasive and costly, rendering them impracticable as a routine feature of diagnostic testing.

A promising new source of diagnostically informative research has come from rodent studies that are elucidating the contribution of the hippocampus and associated regions to spatial and episodic memories. The discovery of cells in the rodent cornu ammonis subfield (CA1) that fire in a spatially determined manner based on an organism's location within its environment – termed place cells (O'Keefe & Dostrovsky, 1971) - has led to the hypothesis that the hippocampus supports the encoding of allocentric spatial frameworks (i.e. information that is encoded relative to the environment, as opposed to idiothetic frameworks, which are encoded relative to the individual). In this way, the hippocampus is thought to form the neural basis of a 'cognitive map' (O'Keefe & Nadel, 1978). In support of this hypothesis, researchers have since identified cells in the neighbouring medial entorhinal cortex (MEC) that fire in a regular hexagonal pattern spanning an environment, referred to as 'grid cells' (Hafting et al., 2005). This has been accompanied by the finding of cells in the presubiculum that encode an animal's heading within space, known as 'head direction cells' (Ranck, 1984; Taube et al. 1995). A proliferation of both animal and human studies has continued to elucidate the factors that contribute to the differential involvement of these neural areas in spatial learning and memory. One key avenue in this research has been the role of landmarks and environmental boundaries in supporting the cognitive map. Doeller, King and Burgess (2008) used fMRI to demonstrate the selective activation of the hippocampus under conditions of environmental boundary-based learning during an object-location memory task. Participants navigated within a circular arena that contained a single landmark and were required to learn and identify the location of everyday objects relative to either the landmark or arena boundaries. The authors report that the hippocampus was activated when object-location learning took place relative to the arena boundary, but not when learning was relative to the landmark. Here, we see that the hippocampus

acts as a neural substrate for the encoding of allocentric spatial information, and that this encoding incorporates salient environmental features. On the basis of these findings, more recent thinking has put forward that the role of the hippocampus in humans is likely to be in binding event representations to their spatial and temporal context – the 'what', 'where' and 'when' of memory formations (Burgess, Maguire & O'Keefe, 2010).

These findings are important as they provide a clear link between the role of hippocampus, the neuropathology and the clinical presentation of AD, thereby providing a potential framework in which hippocampal pathology can be used to inform diagnostically sensitive measures. In keeping with this, research has demonstrated that measures of allocentric spatial memory tasks are sensitive to hippocampal dysfunction associated with AD. The Four Mountains Task (Hartley et al., 2007) presents images of landscapes (specifically, mountains) from either the same or shifted-viewpoint, hypothesising that hippocampal dysfunction should selectively impair image recognition under conditions of shifted-viewpoint, which requires allocentric representations of spatial relations. This has indeed been found to be the case, with research now showing that not only does performance on the Four Mountains Task successfully distinguish dementia due to AD from frontotemporal dementia (FTD; Bird et al., 2010; Pengas et al., 2010) – a condition associated with non-hippocampal neuropathology – but it can also distinguish individuals with MCIad from non-AD MCI (determined by CSF biomarkers) with 100% sensitivity and 90% specificity (Moodley et al., 2015). It should be noted that these results derive from small samples and may not translate to large-scale community studies; in addition, it was noted that specificity was reduced in non-UK samples. However, it remains that this is strong evidence that measures informed by neurophysiological research can be successfully used to design diagnostically informative measures for AD.

While this has proved an informative and promising research avenue, it remains that the hippocampus proper is not the first area to become effected by AD pathology; rather, this process begins in the EC (Braak & Braak, 1991), an area that provides important input into the hippocampus. It follows that a more complete understanding of the role of the EC may guide the design of measures that are sensitive to an earlier, potentially preclinical stage of AD-related neuropathology. In this vein, attempts are now being made to disentangle the relative contributions of different sub-regions of the EC. It is widely held that grid cell activity in the MEC supports a form of idiothetic (i.e. self-referential) navigation referred to as path integration – i.e. navigation back to a start point through integration of vestibular, proprioceptive and visual information (McNaughton et al., 2006). In line with this idea, there is evidence that path integration is impaired both in the context of normal aging (Harris & Wolbers, 2012) and, relevantly, in individuals diagnosed with both AD and MCI (Mokrisova et al., 2016). Further, there is evidence that the EC grid representation is impaired in healthy individuals who carry a high genetic risk for AD, despite the absence of any cognitive decline. Kunz et al. (2015) demonstrated that carriers of the *APOE*-ε4 gene - a known risk factor for development of AD - showed reduced grid-cell-like activation during an object-location task, despite equivalent spatial memory performance. When considered together these results indicate that there may be early changes in the EC that precede both hippocampal impairment and clinically distinguishable cognitive decline.

In contrast, the lateral entorhinal cortex (LEC) has been viewed as relatively insensitive to spatial changes, but rather associated with encoding of objects within an environment (e.g. Hargreaves et al., 2005). This has for some time been viewed as evidence of a functional distinction between the MEC and LEC, with these areas

hypothesised to provide spatial and non-spatial input respectively (Knierim et al., 2014). However, emerging evidence now suggests that this distinction is not clear, and that the role of the MEC and LEC may be far more complex. Recent theories suggest that while the MEC alone is responsible for the processing of idiothetic information, the contribution of both LEC and MEC is required for alloethetic navigation (Save & Sargolini, 2017). Further, it is thought that factors such as the behaviour of the animal and the complexity of the environment may modulate the level of MEC and LEC involvement. Studies making use of a free exploration paradigm have shown that MEC lesions impair an animal's ability to identify spatial changes (e.g. object relocation) within an environment, while LEC lesions impair recognition of changes that are both spatial and non-spatial (i.e. object identity; Van Cauter et al., 2013). Deshmukh and Knierim (2011) showed that cells in the LEC exhibited spatially dependent firing only in the presence of objects. Further, it has been shown that lesions of the LEC impair the ability to recognise item-context associations, but not item recognition per se (Wilson et al., 2013). Here, we can see an emerging picture that the MEC and LEC contribute to constructing conjunctive representations of item, context and location that are likely integrated downstream in the hippocampus.

Research into the neural underpinnings of spatial memory represents a promising avenue in the assessment of AD, with translation from rodent neurophysiological studies to use with human subjects greatly facilitated by the advent of virtual reality (VR) systems. VR draws on a tradition of computerised testing and refers to systems or tasks that enable individuals to retain a sense of physical presence in a computerised 'virtual' environment. Garcia-Betances et al. (2015) describe the differing levels of immersion that can be obtained via a range of technological platforms, such as 3D mounted head displays, and note that the recent advances in VR technologies offer a potential tool in both the assessment of AD. VR paradigms offer

multiple advantages over pen-and-paper measures of neuropsychological function, primary among which are the increased ecological validity of immersive and engaging technologies; this is particularly the case in patient groups, such as AD patients, for whom sustained attention may be an issue. There is evidence to support the concept that performance on VR measures is comparable to real-world performance, with Cushman, Stein and Duffy (2008) demonstrating that navigational patterns in individuals with MCI and early stage AD was closely related across the two paradigms. Further, there is a precedent to show that VR technology can be successfully used as a screening measure for MCI. Zygouris et al. (2015) showed that a VR task in which participants were required to navigate around a virtual supermarket was successful in distinguishing MCI patients from age-matched healthy controls with a classification accuracy of 87%.

Despite the increasing use of VR with AD populations, there remains arguments around the ecological validity of the medium – in which older adults, with limited exposure to immersive gaming technologies, are asked to envision themselves 'within' a desktop scene. Further, it is clear from examination of the role of the EC that this incorporates complex non-visual information, such as vestibular and proprioceptive feedback, that are not available in flat-screen mediums. These limitations are applied primarily to the use of VR utilising lower levels of immersion, with minimal integration of multisensory feedback. However, a recent review highlights that research to date does not frequently make use of VR technologies with a high level of immersion (Garcia-Betances, Waldmeyer, Fico and Cabera-Umpierrez, 2015). This is likely due to the fact that, until recently, commercial VR was associated with high expense and low quality; however, recent advances have seen the introduction of widely available and relatively low-cost commercial systems, which increases their potential utility within research and, perhaps, clinical practice.

Given that the EC is an area affected by the earliest stages of AD neuropathology, it follows that individuals with early stage AD are likely to show selective deficits in tasks designed to tap specific aspects of spatial memory, relative to non-affected individuals. The current research seeks to draw on literature supporting the role of the MEC and LEC in the neural encoding of the location of objects, as well as the role of the LEC in conjunctive object-context associations, in order to develop a measure of spatial memory sensitive to EC pathology. In addition, the current study includes parallels with previous flat-screen VR tasks known to recruit the hippocampal-dependent skill of locating an object relative to environmental boundary cues (Doeller, King & Burgess, 2008).

This study draws on literature suggesting that an aMCI diagnosis represents a high risk for progression to AD in order to establish a proof of concept in an at-risk patient group. Here, we use immersive VR technology that allows participants to navigate freely within a virtual environment as a platform for a novel measure of spatial memory that aims to assess hippocampal and EC function. This offers a number of advantages over desktop measures of spatial memory, including increased ecological validity and greater task engagement; this is in addition to the incorporation of vestibular and proprioceptive feedback that is known to underlie EC encoding. On this basis, we established an object-location memory task that required participants to recall the location of objects within a virtual environment. Based on the known role of the hippocampus in allocentric spatial memory, particularly when reliant on distal boundary cues (Doeller, King & Burgess, 2008), and the role of the EC in object-location representations, this task was intended to recruit both regions affected in early stage AD. In addition, participants were asked to complete desktop assessment of both object and object-context recognition memory. A comparator neuropsychological testing battery was included, against which performance on the OLT was assessed.

Alongside pen-and-paper diagnostic measures, performance on the OLT was compared to the Four Mountains Task (4MT; Hartley et al., 2007), as a known measure of hippocampal dysfunction that is sensitive to MCI due to underlying AD (Moodley et al., 2015).

Research questions

(1) Is an immersive virtual reality test of object-location memory able to distinguish individuals with amnestic MCI (aMCI) from healthy controls?
(2) Do patients with an aMCI diagnosis show reduced performance relative to controls in the recognition of object-context associations, in the context of preserved object-only recognition?
(3) Does performance on a virtual reality task of object-location memory correlate with traditional neuropsychological measures of episodic memory?

2 Methods

2.1 Design

This study made use of a quasi-experimental between-subjects design in which all subjects completed both the immersive VR and neuropsychological battery. The between groups factor was patient status ('MCI' or 'HC') and within groups measures were the performance indices on the VR task, each pen-and-paper test within the neuropsychological battery and the Four Mountains (4MT) test (Hartley et al., 2007).

2.2 Participants

MCI patients were recruited from a hospital-based memory clinic attached to the research centre. All participants within the patient group were diagnosed by a neurologist with amnestic MCI (aMCI) according to the Peterson criteria (Petersen,

2004), which requires that individuals present with: (1) informant corroborated memory complaints; (2) objective memory impairment relative to others of the same age; (3) preserved general cognitive function; (4) intact functional activities; and (5) the absence of dementia. Participants received an initial memory screening comprising the Addenbrookes Cognitive Examination-Revised (ACE-R; Mioshi et al., 2006) and a Mini Mental State Examination (MMSE; Folstein, Folstein & McHugh, 1975) in order to determine objective memory disturbance. During their attendance at the memory clinic, the attending neurologist discussed the research with patients. Those individuals who were interested in taking part in ongoing research then completed a consent form allowing members of the research team to contact them in the future.

Age-matched healthy controls were recruited via the Join Dementia Research initiative developed by the National Institute for Health Research (NIHR) or were spouses of patients. The MMSE and ACE-R were also administered to control participants within six months of their participation in the study.

Recruitment for both patients and control groups was carried out by research staff, including a Research Nurse and doctoral student. Interested individuals were contacted via telephone, at which time the study was explained to them verbally and an information sheet was also sent out via post (see Appendix B). Participants were given opportunities to raise any questions or concerns about the research at this stage, as well as prior to their participation in the study. Inclusion criteria across both patient and control groups were as follows: (1) Capacity to consent to participation; (2) fluent English-speakers; (3) no psychiatric, neurological or substance misuse difficulties or learning disability that would interfere with capacity to participate in the neuropsychological or VR testing; (4) no sensory or motor difficulties that would

interfere with capacity to participate in neuropsychological or VR testing. All participants could claim reimbursement for travel expenses.

2.3 CSF biomarkers

As part of the diagnostic procedure, a subsection of the patient sample also underwent lumbar punctures to ascertain the presence of β-amyloid or tau biomarkers in the cerebrospinal fluid (CSF). Where this procedure was undertaken, the outcome was made available for inclusion in the current study. A positive CSF biomarker status indicates the presence of Aβ and tau biomarkers, suggesting that cognitive impairment is due to underlying AD pathology. CSF status was determined according to criteria outlined elsewhere (Shaw et al., 2009), and procedures were carried out by a qualified nurse using ELISA assay kits (Innotest, Innogenetics, Ghent, Belgium).

2.4 Ethical considerations

Ethical approval for this research was given by the Cambridge South NHS Research Ethics Committee (REC reference 16/EE/0215; see Appendix B for confirmation letter) and was carried out in line with the Declaration of Helsinki (WMA, 2013). UCL researchers were granted Visiting Researcher status in April 2017 (see Appendix C for letter granting Visiting Researcher status).

2.5 Equipment

The immersive VR component of the object-location task (OLT) was carried out on a commercially available VR system, the HTC Vive. The HTC Vive hardware comprised a headset with resolution 1080 x 1200 pixels per eye, with a refresh rate of 90 Hz. The location of participants in the room was tracked using a system called

Lighthouse, which uses head position and the location of two handheld controllers to locate individuals within a rectangular space. Testing rooms were set-up with tracking equipment arranged at the perimeter and programmed to form a 3.5 x 3.5 metre space in which participants could navigate freely throughout the task. In order to improve comfort, minimise distractions and increase safety, a wireless backpack containing a laptop was used to run the task. The backpack was MSI VR One, incorporating an Intel^{R} core i7 processor and weighing 3.6kg. The OLT was then controlled remotely by researchers through a laptop connected to the backpack via a virtual private network (VPN). The OLT was programmed by a UCL PhD student with a computing background; they made use of Unity gaming software (San Francisco, California, USA).

Desktop aspects of the OLT were programmed in the Cogent 2000 Matlab toolbox (Wellcome Department of Neuroimaging Neuroscience: http://vislab.ucl.ac.uk/cogent) presented on MATLAB software on a 15 inch MacBook Pro, with screen resolution 1440 x 900. Responses were given via a Bluetooth keyboard.

2.6 Magnetic Resonance Imaging (MRI) scanning

A subsection of the aMCI group also underwent volumetric MRI scanning as part of their diagnostic work-up. These scans were conducted on-site at the clinical and research centre using 3T Siemens Prisma scanners. These scans took place within six months of testing and the imaging data obtained comprise the body of a joint thesis with Adrienne Li (2018); therefore the results of this will not be discussed further here.

2.7 Object-location task (OLT)

2.7.1 iVR component of the OLT

The immersive VR component of the task was made up of four active trials and one practice trial, with each trial divided into a 'learning' phase and a 'recall' phase. During the 'learning' phase, participants were required to first attempt to memorise the locations of everyday objects within a virtual world, while the 'recall' phase required them to indicate where they believed each object had been located. The order of iVR trials is shown diagrammatically in Figure 1. Each trial contained three objects and was set within a different virtual environment. The virtual environments were enclosed by a square wall, ensuring that navigation relied only on distal landmark cues (see Figure 2 for examples of environments). Outside of this wall, a range of distal landmarks were arranged that participants were instructed to use to assist in locating themselves within the environment. No landmark cues were present within the walled portion of the environment.

Upon entering the virtual environment for the 'learning' phase, participants were given sixty seconds to familiarise themselves with the environment, after which a first object was presented. Objects were presented sequentially, with only one object present at any time. Participants were presented with each object three times (i.e. 3 objects x 3 presentations), in a randomised order, and were cued to return to the centre of the environment between each 'round' of presentations. All objects were presented on a pedestal; in this way, it was ensured that all objects occupied a standardised area of the environment floor. After each object had been presented three times, participants were returned to a neutral grey 'waiting room' where they were given further instructions by the researcher.

In the 'recall' phase, participants were given a handheld controller that allowed them to produce a simulated pedestal, identical to the ones on which objects had been placed, which they could position anywhere inside the boundary. The participant was then placed back into the centre of the environment and asked

to indicate where they believed each object had been located. Only one object location was presented at each time, with a small image cueing the target object remaining present in the corner of the screen throughout, to ensure that performance was not affected by difficulty in recalling the task. Participants indicated each choice twice, with objects again presented in a randomised order. The primary outcome measure from this was the displacement error of responses (i.e. the distance between the identified and actual object location), taken in centimetres; lower displacement error therefore indicated higher accuracy and better performance on the task.

2.5.2 Desktop component of the OLT

The flat screen measure of the OLT comprised two separate components: (1) simple object recognition; and (2) object-context recognition. Participants were, in the first instance, asked to identify whether an object on the screen was either 'old' (i.e. had previously been seen in the iVR OLT) or 'new' (i.e. not presented in the iVR OLT). Images were presented one at a time, with each object present on the screen for 30 seconds. Participants were required to indicate their response within this time limit, or else it was recorded as an incorrect response. Where 'new' objects were presented, these were foils – i.e. a different visual image of a concept that had been presented in the iVR OLT. For example, a foil might include either a previously seen 'real' duck, or a rubber duck that had not been previously presented (Figure 3 shows examples of 'old' objects and 'new' foils).

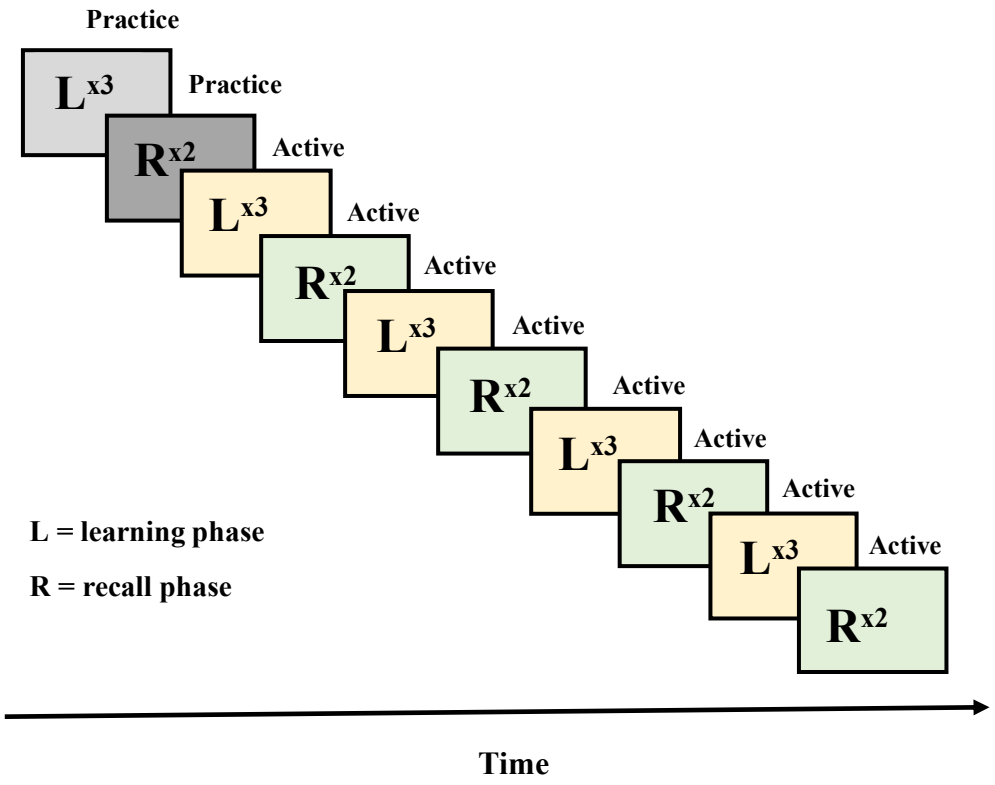


Figure 1: Order of presentation for VR component of the OLT

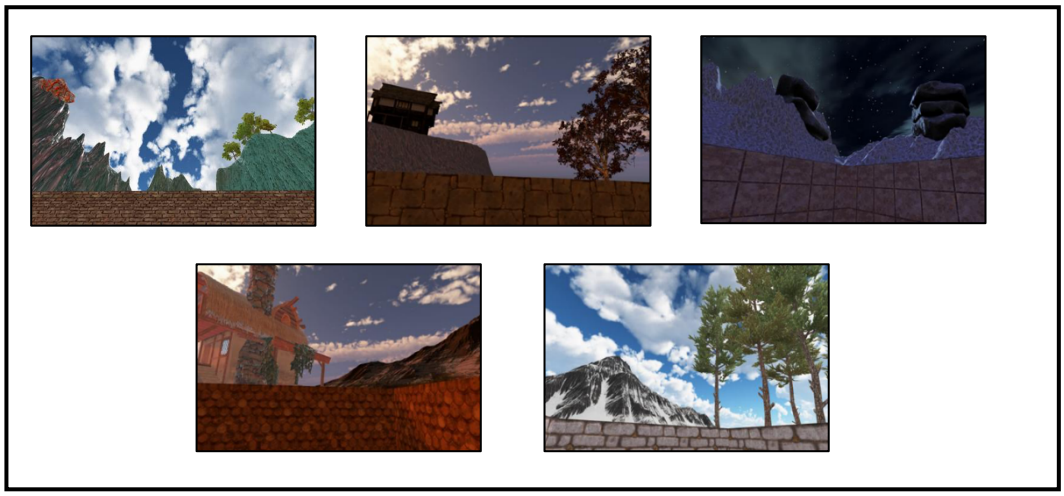


Figure 2: Still images of the five virtual environments as seen from the perspective of participants.

In the instance that an 'old' object had been presented, the participant was then asked to identify which environment it had been seen in; this occurred regardless of whether an object was correctly identified as 'old'. Still images of all four potential environments were presented on the screen alongside one another, appearing in a random order (see Figure 4). Participants were asked to identify the environment that the object had been seen, again with a 30 second time limit on responses. An image of the target object remained on the screen alongside the pictures of the environment to minimise the demands of the task and to ensure that poor performance was due to object-environment associations, as opposed to difficulty recalling the object. No feedback was given on performance throughout these trials, however the experimental trials were preceded by a practice round, in which performance feedback was given.

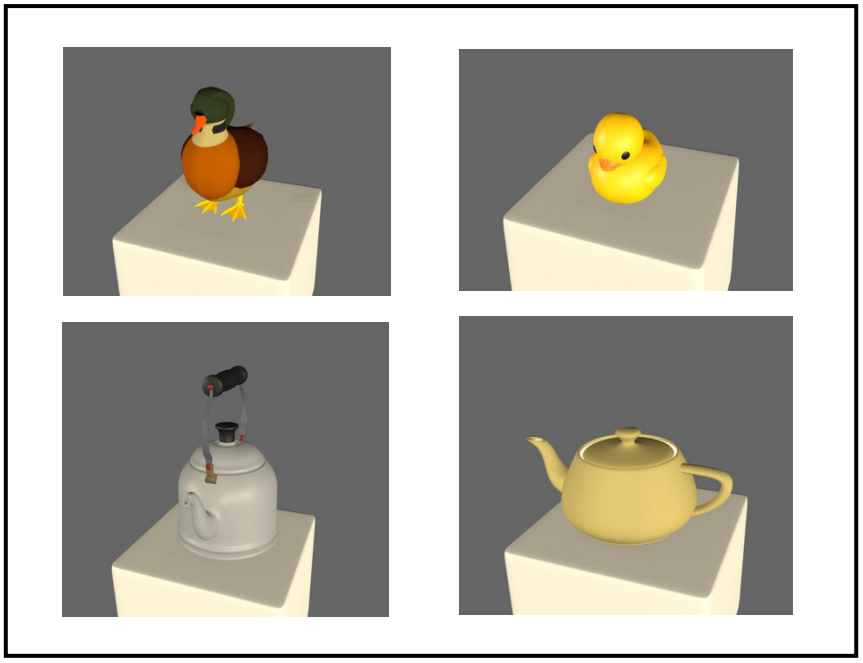


Figure 3: Objects used in the OLT, as presented in the desktop task. Original images presented on the left, foils (decoy images) presented on the right.

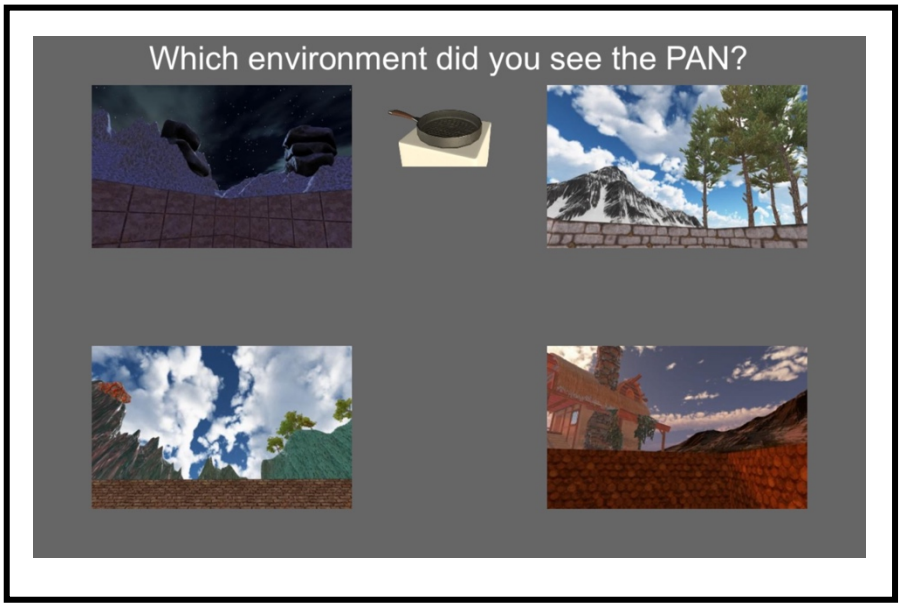


Figure 4: Presentation of the object-context recognition condition in the desktop OLT.

2.6 Neuropsychological measures

A neuropsychological test battery comprising measures currently validated as a screening tool for AD and MCI were assembled. This comprised the Rey-Osterrieth Complex Figure Test (RCFT), as a measure of visual recognition and recall; Trail Making Test B (TMT-B), as a measure of executive function (Reitan & Wolfson, 2004; Digit Symbol Test (DST; (Weschler, 1981); Free and Cued Selective Reminding Test (FCSRT), as an index of verbal recall under free and cued conditions (Buschke, 1984).The Four Mountains Test (4MT) was also included in this battery as a measure of hippocampal-dependent spatial learning (Hartley et al., 2007). Finally, the National Adult Reading Test (NART) was administered as a measure of premorbid IQ (Nelson & Willison, 1991).

These measures have all been demonstrated to be sensitive diagnostic indicators of the early stages of AD. The FCSRT has been shown to identify the early stages of AD (Grober, Sanders and Lipton, 2010) and distinguish this from frontotemporal dementia (FTD) (Lemos, Duro and Santana, 2014). The possible

outcomes from the FCSRT include both free and total recall, with the latter representing performance when memory was cued. In the current study we report free and total recall (iFR) and total recall (iTR), as well as corresponding measures for the delayed recall condition (dFR and dTR, respectively).

The RCFT is used as an index of immediate and delayed visual recall, with points scored based on the accuracy with which participants are able to replicate an image based on their recall; lower scores on this measure are therefore indicative of poorer performance. Visual memory indexed via the RCFT has been shown elsewhere to be impaired in mild AD (Kasai et al., 2006).

The TMT-B was used as a measure of executive functioning. The primary outcome measure from this is the time taken in seconds to complete the task, with higher scores therefore indicating poorer performance. This has been shown to have significant diagnostic accuracy in predicting conversion from MCI to AD, particularly when considered alongside EC volume and CSF biomarker status (Ewers et al., 2012).

The DSST requires the participant to match symbols with a corresponding number within a set time limit, with higher numbers representing a greater completion rate and, therefore, better performance. This has been shown to be a predictor of time to AD conversion in MCI patients (Tabert et al., 2006).

The 4MT task was designed as a measure of hippocampal-dependent spatial memory, with participants required to identify images of mountains from shifted viewpoints, thereby tapping allocentric spatial memory. The primary outcome from this is the number of correctly identified scenes, with lower scores thereby indicating poorer performance. As highlighted above, the 4MT has been shown to be a sensitive and specific predictor of conversion to AD (Moodley et al., 2015). This was carried out on a small, handheld computer tablet.

2.7 Testing procedure

2.7.1 Structure of testing

All testing took place on site in the Institute of Public Health (IPH) or the MRC Cognition and Brain Sciences Unit (CBSU), both of which belong to the University of Cambridge. The duration of testing was approximately 150 minutes, although this was dependent on individual performance and could range between 90 and 180 minutes. Regular breaks were offered throughout to ensure participant comfort, as well as protecting data quality from fatigue-related effects. Testing was carried out in three main stages: (1) an immersive VR task; (2) a desktop computer task, based on participation in part 1; and (3) a pen-and-paper neuropsychological assessment. Due to the dependence of (2) on participation in (1), it was necessary that these tasks were performed in that order. Consequently, each participant began with either the VR task or the neuropsychological battery in a counterbalanced manner to eliminate carry-over effects.

At the outset, the format of the testing session and purpose of the study was explained to all participants, who again were given the opportunity to read through the information sheet and ask any questions of the researchers. All participants were made aware that they could withdraw at any point. Subsequent to this, participants provided written informed consent for participation in the study (Appendix D).

2.7.2 VR OLT procedure and instructions

Prior to beginning the VR task, participants were fitted with a backpack and headset and given time to become familiar with the equipment. At this point, verbal instructions were again given by the researcher and a practice trial was initiated. Prior to beginning the VR task, participants were made aware

that this would be followed by a computerised task in which they were asked to identify the objects and environments that they encounter in the VR. During the 'learning' phase of the VR task, participants were informed of the task set up and structure, including the number of objects and manner of presentation. Once this had been completed, they were removed from the environment and returned to the 'waiting room'. Participants were then given handheld controllers and instructed on their use, as well as allocating time to practice with these while still in the 'waiting room'. Participants were returned to the virtual environment and began the 'recall' phase only when both they and the researcher felt that they had mastered the use of the controller.

Researchers gave minimal further instructions after this point, except to indicate any transition between phases and environments. In addition, participants were prompted at the beginning of each trial to spend some time familiarising themselves with the environment and learning the landmarks in each new VR environment. If a participant was evidently struggling to recall earlier instructions or appeared to have difficulty engaging in any aspect of the task, the researcher would then provide additional support and instructions until it was felt that the participant could manage without this.

2.7.3 Desktop OLT procedure and instructions

After completion of the iVR component of the OLT, participants moved to a table-top computer. They were given task instructions verbally and via on screen instructions. Responses were given via a keyboard, and relevant keys highlighted both by visual cues and verbal instructions from the researcher.

2.7.4 Neuropsychological battery

The neuropsychological battery was administered by a single researcher over one session that lasted approximately 60 minutes. Breaks were offered if necessary.

2.8 Sample size

A sensitivity analysis showed that for alpha = 0.05 and beta = 0.2 the study was powered to detect only large effects given this sample size. While this is not ideal, prior studies using spatial tasks had found them to be highly discriminant in similar samples and it was decided that the study would be valuable despite the potential to be somewhat underpowered.

2.9 Data analysis

All data were analysed using the Statistical Package for the Social Sciences (SPSS) version 25 (IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp).

2.9.1 Group comparisons

In order to ensure the comparability of groups, demographic information - namely age at testing and years of education - was compared across groups using independent samples t-tests. Comparisons across groups were also made in relation to performance on the OLT, again making use of independent samples t-tests; the primary OLT outcome variable was displacement error (centimetres), however object recognition (% accuracy) and object-environment identification (% accuracy) were also examined. In all instances 'group' (i.e. MCI or control) was used as the independent variable, with the relevant measures reported as a dependent variable. Where assumptions of normality were violated the appropriate non-parametric statistics were reported. Given that CSF biomarker

status was not available for the majority of participants, descriptive data will be provided that examines OLT and 4MT performance according to patient biomarker status.

At this stage it is necessary to note two technical considerations in analysis of the data. Firstly, the initial six patients completed a version of the OLT including four objects per environment; this was later reduced due to feedback on the difficulty of the task. It was not possible to remove these individuals from the analysis, nor to make a statistical comparison to the performance of individuals who completed the three object version of the OLT as either analysis would be underpowered. The data were examined to determine the presence of any outliers or extreme scores in this group that would warrant removal from the analysis. Secondly, seven data points were missing from the conditions for object and object-environment recognition, meaning that only thirteen patients were included in this analysis; this was due to technical difficulties in data recording that affected earlier participants. It is assumed that this would affect the data in a non-systematic manner.

2.9.2 Relationship to neuropsychological testing battery

The relationship between OLT displacement error, as the primary outcome measure, and performance on the neuropsychological measures will be examined using Pearson correlations. RCFT immediate and delayed recall, FCSRT immediate and delayed free recall, DSST, TMT-B and the 4MT will all be examined in relation to their association with OLT displacement error. Raw scores for each outcome were used, due to the limited range of participant ages. Due to the high number of tests performed, Bonferroni corrections were applied.

2.9.3 Prediction of patient status

Binomial logistic regression was used in order to evaluate the ability of each variable to correctly classify patient status (either 'MCI' or 'control') on the basis of OLT performance. In order to assess the independent contribution of variables to the prediction of patient status, variables were added sequentially over three stages of modelling. In the first stages, predictors of individual resilience, including premorbid IQ and education level (indexed by the variables 'NART' and 'Years Education') were added to the model. This was then followed by 4MT as a hippocampal-dependent measure of spatial learning. Lastly, the variable 'OLT displacement error' was added as a measure of entorhinal-dependent spatial learning. At each stage all variables were assessed to determine whether they made a statistically significant contribution to prediction, with variables that become non-significant removed in the final model. The final model, including chi-square goodness of fit statistic, the adjusted R^2 and odds ratios will be reported, alongside the contribution of each variable to prediction of patient status.

A final model was constructed, again using binomial logistic regression, in which the ability of the comparator neuropsychological testing battery to correctly classify patient status (again either 'MCI' or 'control'). The initial stage predictors of resilience factors - i.e. 'NART' and 'Years Education' – were added to the model at a first step, in the same way as the OLT model. At the next stage, the variables 'RCFT immediate', 'RCFT delayed', 'FCSRT immediate free recall', 'FCSRT delayed free recall', 'TMT' and 'DST' were then added to the model. Again, the final model statistics and independent contributions of each variable will be reported. This will be used as a comparator against which the utility of the OLT/4MT model can be assessed.

3 Results

3.1 Demographics

Overall, twenty MCI patients (males = 13) and twenty-two healthy controls (males = 9) took part in this study. There were no significant differences between the MCI and HC groups in relation to age or years of education (see Table 1 for group means and associated *t* values). Due to the high number of comparisons, Bonferroni corrections were applied to comparisons for the neuropsychological testing battery and differences assessed against a more stringent *p* value of .004 (.05/11).

3.2 OLT performance

Mann-Whitney U tests revealed a significantly greater error rate of object-location displacement error for MCI patients when compared to controls ($U = 85.0$, $p = .001$). For the second part of the task, thirteen patients and twenty-one controls were included in the analysis. Where participants were not included, this was due to technical difficulties in data recording that prevented the analysis of this aspect of the OLT. There was an observable difference in object-recognition accuracy, with the patient group demonstrating lower accuracy than controls ($U = 69.0$, $p = .015$). Independent t-tests revealed no differences between MCI and control groups in terms of object-environment accuracy ($t(32) = .779$, $p > .05$; 95% CI = -8.24, 18.09).

Table 1

Demographic information and neuropsychological test performance across groups

	MCI (n=20)	**Control (n=22)**	
	Mean (SD)	***Mean (SD)***	*p value*
Age	68.35 (9.65)	65.31 (7.57)	.262
Years Education	15.15 (4.09)	15.59 (3.80)	.611
ACE	85.85 (8.92)	97.91 (2.93)	<.001*
MMSE	27.6 (2.46)	29.91 (0.30)	<.001*
NART	13.4 (9.61)	5.64 (3.02)	.014
RCFT copy	32.75 (3.97)	35.96 (.21)	<.001*
RCFT immediate	13.88 (10.70)	21.66 (8.15)	.011
RCFT delayed	12.83 (11.71)	21.17 (8.15)	.012
TMT-B	145 (65.70)	70.33 (20.12)	<.001*
DST	49.50 (13.74)	66.77 (10.44)	<.001*
FCSRT iFR	21.05 (13.25)	35.55 (5.09)	<.001*
FCSRT iTR	41.15 (9.54)	47.82 (.50)	<.001*
FCSRT dFR	8.00 (6.01)	13.96 (1.46)	<.001*
FCSRT dTR	13.45 (3.97)	16.00 (.00)	.001*
Four Mountains	7.60 (3.32)	10.50 (2.04)	.003*

**significant difference between groups following Bonferroni correction*

Table 2

Performance on OLT task

	MCI	Control
	Mean (SD)	***Mean (SD)***
OLT displacement error	90.62 (53.98)	38.18 (22.28)
Object recognition (%)	82.45 (13.31)	91.47 (6.78)
Object-environment recognition (%)	41.83 (18.41)	36.91 (17.59)

3.3 Biomarker status

Overall, four participants were identified as positive for CSF biomarkers of amyloid and tau. A further four individuals within the patient group were identified as negative according to CSF biomarkers. The remaining twelve participants were of unknown biomarker status. Figure 5 shows a descriptive breakdown of OLT displacement error and 4MT performance according to CSF biomarker status.

3.4 Relationship to neuropsychological measures

In the case of the patient group, significant negative relationships were shown between OLT displacement error and performance on RCFT immediate ($r(18)=-.855$, $p<.001$) and delayed recall ($r(17)=-.843$, $p<.001$); FCSRT immediate ($r(18)=-.882$, $p<.001$) and delayed free recall ($r(18)=-.893$, $p<.001$); and DST ($r(18)=-.5941$, $p=.006$). No significant relationship was observed between OLT displacement error and performance on the NART, 4MT or the TMT-B ($p>.05$ in all cases). In contrast, for the control group there was no significant relationship observed between OLT displacement error and any comparator neuropsychological test ($p>.05$ in all cases).

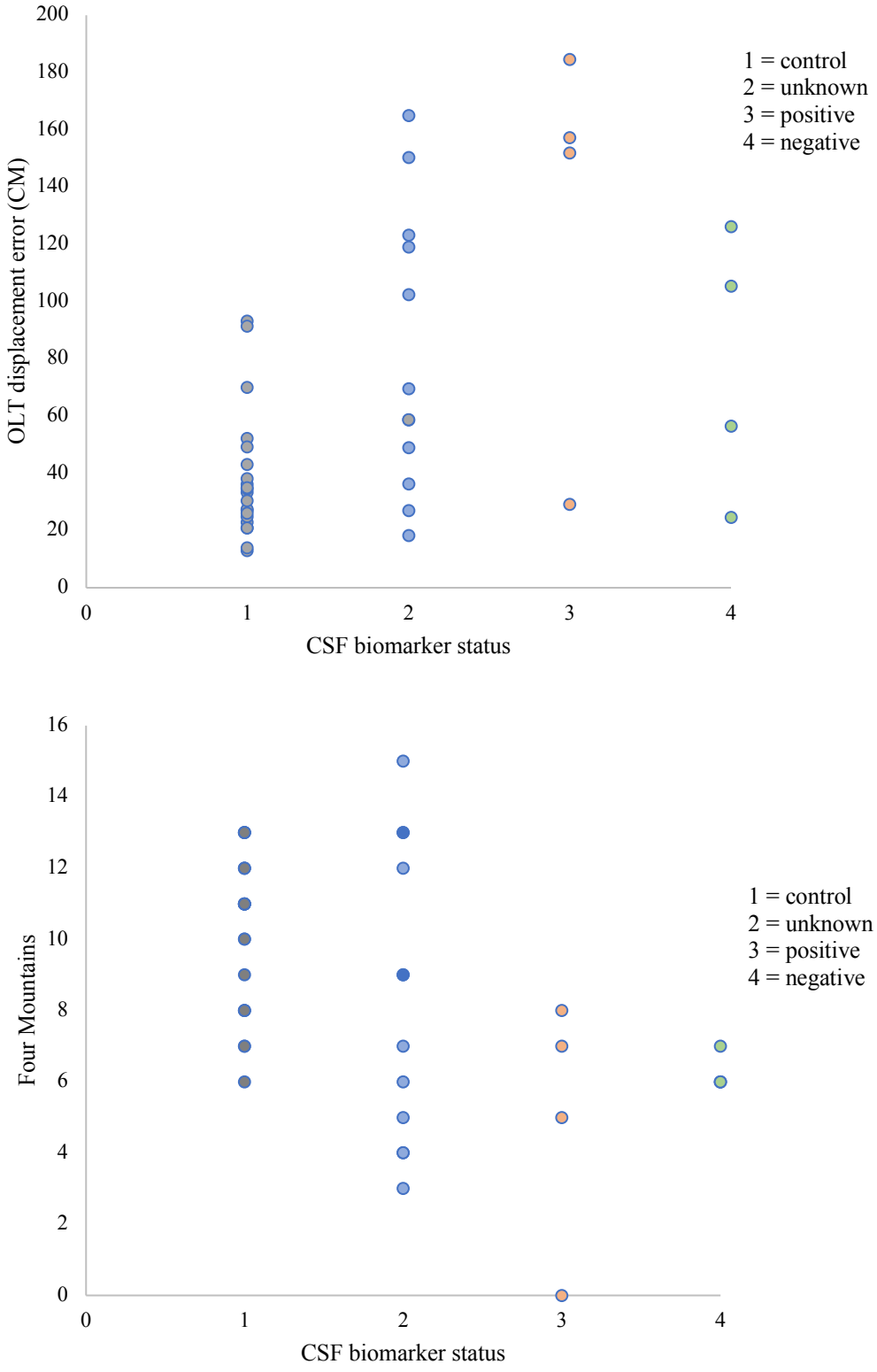


Figure 5: Graphs showing performance on the OLT and 4MT according to biomarker status

3.5 Prediction of patient status

3.5.1 OLT model

Model one The extent to which 'NART' score and 'Years Education' predicted patient status were added in the first stage of analysis, as these were characterised as resilience factors in line with ideas of cognitive reserve. A model containing these variables was significantly better at predicting patient status than a model that included only the constant ($X^2(1)=12.315$, $p<.001$) and explained 34.6% of the variance (Nagelkerke R^2). Only 'NART' contributed significantly to the predication of patient status. The variable 'Years Education' did not contribute significantly to the model ($p>.05$) and was therefore removed at this stage. Overall, this model was able to correctly predict group membership in 70.7% of cases.

Model two At the next stage of modelling the variable '4MT' score was added to the model to assess its independent contribution to the prediction of patient status. The addition of this variable significantly improved the model's predictive power compared to model one ($X^2(2)=19.377$, $p<.001$) and explained 50.2% of the variance (Nagelkerke R^2). This model was able to correctly predict group membership in 75.6% of cases, which can be further broken down to 81% and 70% of controls and patients respectively.

Model three Finally, the contribution of the variable 'OLT Displacement Error' was added to form a last stage of the model. This stage of the model was significantly better at predicting patient status than model two ($X^2(3)=43.182$, $p<.001$) and explained 86.8% of the variance (Nagelkerke R^2). The classification accuracy of the model was improved to 95.1%, with 95.2% and 95.0% of

controls and patients correctly classified respectively. Coefficients, odds ratios and 95% confidence intervals for all variables included in the final model are presented in Table 3.

Table 3

Coefficients, odds ratios and associated statistics for variables included in the final OLT model

Variable	B (SE)	Significance	Exp(B) (95% CI)
NART	.392 (.161)	.015	1.480 (1.079, 2.030)
4MT	-.981 (.434)	.024	.375 (.160, .878)
OLT Displacement Error	.092 (.039)	.018	1.097 (1.016, 1.184)

3.5.2 Comparator model

The variables 'RCFT immediate', 'RCFT delayed', 'FCSRT immediate free recall', 'FCSRT delayed free recall', 'DST' and 'TMT-B' were added to a second binomial logistic regression, with 'Group' (either MCI or control) as a dependent measure. The first stage of modelling, in which the variables 'NART' and 'Years education' were added to the model were identical to the first stage of the OLT model; this is outlined above and therefore will not be repeated here. The addition of these variables significantly improved the model relative to the inclusion of only premorbid IQ and education level ($X^2(3)=39.213$, $p<.001$) however only the variables 'TMT' and 'FCSRT immediate free recall' (FCSRT iFR) were retained for inclusion; all other variables were removed at this stage.

The final comparator model therefore included the variables 'NART', 'TMT' and 'FCSRT immediate free recall' and was able to account for 81% of variance in the data. Further, this model correctly classified 92.9% of cases; this constituted 100% correct identification of control and 85% of MCI cases.

Table 4

Coefficients, odds ratios and associated statistics for variables included in the final comparator model

Variable	B (SE)	Significance	Exp(B) (95% CI)
NART	.147 (.138)	.288	1.159 (.883, 1.520)
TMT	.050 (.025)	.042	1.051 (1.002, 1.103)
FCSRT iFR	-.168 (.089)	.060	.845 (0.709, 1.007)

4 Discussion

4.1 Main findings

This study aimed to assess the utility of an object-location iVR task as an assessment tool for amnestic MCI. Here, we demonstrate that the OLT does have potential diagnostic use for distinguishing patients with aMCI from age-matched healthy controls. Patients demonstrated significantly higher rates of OLT displacement error relative to control participants, suggesting that this task does assess aspects of object-location memory that are selectively impaired in individuals with aMCI. As aMCI patients are a group at high risk of progression to AD, this supports the further investigation of the OLT as a potential early diagnostic measure. Further, this study

demonstrates that the OLT, when used in combination with a measure of hippocampal function and premorbid IQ, is able to predict patient status with a high degree of accuracy, correctly classifying patients in 95% of cases. In diagnostic terms, this represents a high degree of sensitivity to early stage deterioration in cognitive dysfunction associated with an aMCI diagnosis.

This measure was designed to assess hippocampal and EC pathology, on the basis that these are regions known to be affected by the earliest stages of AD pathology. Here we see evidence that the OLT is correlated with other neuropsychological measures known to be sensitive to early AD, although importantly was not related to performance on the 4MT task, a known measure of hippocampal function. Not only did the two measures not correlate, but they each independently contributed to a predictive model of patient status. Together, this suggests that the OLT is sensitive to cognitive dysfunction associated with aMCI, however is distinct from difficulties assessed via the 4MT. It is potentially the case that in the model predicting patient status hippocampal dysfunction was accounted for by the 4MT, while the remaining variance in spatial memory due to EC dysfunction was then accounted for by OLT performance. However, this does not account for the lack of correlation between the two measures. Rather, this provides tentative support for the concept that the OLT taps into an earlier stage of disease progression, at which point hippocampal pathology may not yet be evident; i.e. the OLT may rely more heavily on object-location memory supported by the EC. This was not an expected result based on the initial design of the OLT, which has some structural similarities to the hippocampal-dependent task used in Doeller, King and Burgess (2008). One important difference between this and the OLT is the use of fully immersive VR that allows for free movement within the task, therefore incorporating self-motion information into the process of object-location memory in a manner that would not have been possible in

Doller et al. It is potentially the case that this resulted in greater reliance on the EC in this instance. However, in the absence of neuroimaging data this remains a speculative interpretation of this finding, and it is not possible to conclusively demonstrate that this was the case.

Despite prior evidence that LEC lesions do not impair object recognition under non-associative conditions (Wilson et al., 2013), in this study we found there was significantly worse performance in aMCI patients relative to control participants. It is, however, important to note that in this instance object recognition was high across both groups, exceeding 80% accuracy even amongst aMCI patients. Nevertheless, the finding of reduced performance was not predicted in this group. These findings are most likely attributable to the heterogeneous nature of the participant group; only four individuals were confirmed as CSF positive for AD, with a further three identified as negative for CSF negative for amyloid and tau and the remained unconfirmed in their CSF status. While it would be expected that object recognition would remain unimpaired if pathology were limited to the EC – as is thought to be the case in the preclinical and prodromal stages of AD (Braak & Braak, 1991) - MCI in this instance is potentially (and, in some cases, certainly) due to non-AD pathology and therefore associated with a potentially diverse range of brain regions. It may be expected that were the data and sample size sufficient to allow comparisons across aetiologies, there may be a dissociation between those who are CSF positive and negative. However, it is also necessary to consider that this hypothesis is formulated from rodent data and as such may need further refining in order to successfully translate to human studies.

In contrast, both MCI and control groups performed equally poorly when required to link objects to the environments in which they were presented and no group differences were observed. The poor performance is potentially related to the loss of data in this patient group, which may have rendered the analysis underpowered to

detect meaningful changes in this condition. Alternatively, it is possible that this aspect of the task was too complex, resulting in floor effects across groups. Indeed, it has been established in rodent studies that the role of the LEC and MEC in binding object and spatial representations may vary in line with the complexity of the task, with factors such as the number and diversity of objects playing a role in modulating performance (Save & Sargolini, 2017). Broadly, it seems that a higher number of objects increases cognitive demand and, consequently, requires greater LEC involvement (Save & Sargolini, 2017; Ku et al., 2017). This latter part of the task requires participants to hold in mind the object and context across all tasks, totalling twelve objects and four environments. While this was designed as a measure of recognition memory for spatial and non-spatial conjunctions, it appears that in this instance the cognitive load involved may be too demanding, resulting in poor performance across both patients and controls.

It was not possible to comment on performance according to CSF biomarker status as the proportion of individuals with known biomarker status was not sufficient to identify any statistical differences across groups. Descriptive breakdown of the current results according to biomarker status showed some indication of higher OLT error rates in the CSF positive group relative to controls or those with CSF negative or unknown biomarkers. There is some variation in this, with one individual notably scoring among the lowest in the sample. In contrast, 4MT data did not show any clear differentiation according to CSF status; this is not in line with previous findings (Hartley et al., 2007) which have indicated a clearer breakdown in this patient group. This may be due to the small sample size in this study precluding the observation of clear distinctions according to CSF biomarker status.

In addition to demonstrating that the OLT is a sensitive measure of cognitive functioning, this study aimed to assess the construct validity of this task through

comparators to a neuropsychological testing battery comprising the current 'gold standard' of AD assessment measures. This included measure of verbal (FCSRT) and visual (RCFT) immediate and delayed recall, as well as executive function (TMT-B). A strong negative relationship was found between performance on the OLT and measures of both immediate and delayed verbal and visual recall, indicating that poor performance on the OLT was related to poorer performance on standard neuropsychological measures. Further, this observation is in line with the clinical symptomatology of aMCI and AD as specifically associated with memory impairments. Interestingly, no relationship was observed between TMT and the OLT, which would again be clinically consistent as TMT is designed as a measure of executive function, a cognitive area less associated with decline in the earlier stages of AD. Further, the comparator model demonstrated somewhat lower accuracy in the prediction of patient status than did the model making use of spatial measures (i.e. the OLT and 4MT); this was particularly the case for the identification of the MCI group, with the comparator model showing lower accuracy in the identification of patients relative to controls (85% and 100% respectively). This is of note and tentatively suggests some advantages of the OLT over traditional neuropsychological measures; however, the difference between the predictive power of these models is small and the heterogeneous nature of the patient group precludes any conclusive statements in this regard. In order to comment on this with greater certainty it would be important to assess predictive power as regards the rates of transition to AD; it may be expected that the spatial model would show higher predictive power for those individuals who do go on to transition to AD.

4.2 Acceptability of task

Given the novel nature of the task, which makes use of VR technology incorporating free movement tracking with visual immersion, it is helpful to briefly reflect on the usability of the OLT and VR environment to users with cognitive impairment. There were no instances in which patients failed to complete the OLT or withdrew consent during testing; nor, indeed were there significant outliers in which performance was much worse that the group mean. This, along with qualitative patient feedback, provides support for the acceptability of the OLT in this patient group, as well as the use of immersive VR technologies in this patient group. This is in line with previous research that supports the use of VR in MCI and AD patient groups (Cushman, Stein & Duffy, 2008). Given that the role of the EC is uniquely characterised by the integration of information from multiple sensory modalities (e.g. McNaughton et al., 2006), as well as the ability to integrate the spatial context of object recognition (Deshmukh & Knierim, 2011), this is a patient group for which immersive technologies may present important diagnostic potential, as the combination of immersive visual experiences with motion tracking software offers a unique opportunity to tap into the multisensory aspects of spatial cognition.

4.3 Limitations

Despite demonstrating concurrent validity, it remains to be seen whether the OLT will demonstrate predictive validity as regards AD. For practical reasons it was not possible to recruit exclusively CSF biomarker positive participants and the current sample therefore comprises a heterogeneous patient group. Further, the current results are taken at only one time point, precluding any follow-up of the conversion rate to AD in this sample. It is therefore not possible at this stage to comment conclusively on the specificity of the OLT as a diagnostic tool for MCI due to AD. In addition, the absence of concurrent neuroimaging data means that we cannot state with certainty

that the OLT in its current format specifically indexes EC functioning, rather this can only be inferred from knowledge regarding the neuropathology of MCI patients and its relation to other neuropsychological assessment tools.

It is important to acknowledge that the control group demonstrated significantly higher scores on the NART, suggesting higher IQ in this group relative to patients. This may have arisen due to bias in the sampling, with a volunteer sample from a university town associated with a high likelihood of presenting with above average IQ; this is in contrast to a patient sample, which is drawn from a wider demographic area. Due to time constraints and practical limitations in the availability of volunteers, it was not possible at this time to recruit an alternative control group. In order to control for this group difference, years of education and NART error rates were included as a first stage in the modelling of the data; this is aimed at reducing the impact of this difference on any final assessment of OLT utility. However, this is only a partial solution to the difficulty and the sensitivity of the OLT to patient status may be reduced were this pre-existing difference not present (Miller and Chapman, 2001). It is therefore important that the current results be interpreted with caution, and future studies should aim to replicate these findings with a control group that is matched according to performance on the NART, alongside other demographic factors.

The number of objects in each trial of the OLT was reduced during the course of testing, thereby introducing variation in complexity amongst patient groups; this represented a challenge in the current analysis. Despite piloting of the task in healthy samples prior to the outset of testing, issues such as this were not possible to identify prior to testing with a patient group precisely because they are issues specific to individuals with cognitive impairment; while four objects was acceptable in a non-impaired sample, it presented a greater challenge to an MCI cohort and therefore could not be identified through piloting. Given the limited availability of clinical volunteers

and the high demands that testing places on them, it was not deemed feasible or appropriate to recruit this group solely for the purposes of piloting. While examination of the data did not identify extreme scores in those who completed the four-object version, it remains that this additional variation must be considered when interpreting the current results.

Related to this, it is important to acknowledge that this study did not explicitly assess feasibility or acceptability of either the task itself or the VR set-up. While some information – particularly regarding drop-out rates and participant debrief discussions – was collected, this was on an informal basis. Adaptations to the protocol (i.e. changing number of objects per round) were based on this informal feedback. For this reason, it is not possible at this stage to make conclusive statements regarding the feasibility of the OLT in routine clinical practice. In order to build on the findings of the current work, it would be important to carry-out a formal assessment of feasibility and acceptability.

Bowen et al. (2009) set out protocols for determining the appropriateness of an intervention, including assessments of acceptability and any changes to protocol. The use of pre- and post- change surveys may be informative in determining the effect of changes in the OLT protocol, such as the alteration in object number described above. As regards acceptability to the target population, Bowen et al. propose that at the initial stages of design a small-scale focus group with the intended population – in this case MCI patients – may be carried out to more fully understand patient perceptions of the task. At these early stages, it would therefore be important to incorporate this form of feedback into future studies involving the OLT. At later stages, were the OLT to move towards clinical use, this may then progress to an RCT comparison of experiences of memory assessment via the OLT relative to assessment via routine neuropsychological testing.

A final limitation concerns that nature of the participant group – including both patients and controls - which was recruited from a single location within a university town. It is therefore likely that this groups represents a highly educated sample that likely does not represent a diverse sample in terms of either culture or socioeconomic status. This is particularly important given the known cross-cultural variations in performance on neuropsychological assessment tools (e.g. Ardila, 1995) as well as the observed cultural differences in specificity in the 4MT (Moodley et al., 2015).

4.4 Further work

The current study sought to demonstrate a proof of concept of the utility of an immersive VR measure of object location memory in distinguishing aMCI patients from non-cognitively impaired counterparts. In order to go further and demonstrate that the OLT is specific to EC dysfunction thought to be the hallmark of preclinical AD, a number of further steps must be taken. Importantly, it must be demonstrated that performance on the OLT is able to distinguish individuals with underlying AD pathology from individuals with cognitive impairment deriving from other aetiologies. This may be achieved in a number of ways, including the comparison of performance across CSF biomarker positive and biomarker negative patients. Alternatively (or, ideally, in parallel) a longitudinal examination of AD conversion rates over time on the basis of OLT performance would lend further support to the predictive utility of the OLT.

Further, it is assumed that the OLT assesses abilities thought to be supported by EC function and, consequently, sensitive to EC dysfunction. While demonstrating the predictive power of OLT is an important step in supporting this conclusion, confirmatory neuroimaging studies would make an important contribution to this conclusion. It is possible that other neural regions subserving spatial memory (e.g. the

hippocampus) may be more heavily involved in this task. While the lack of relation to the 4MT in the current study suggests that this is unlikely to be the case, it would be important to confirm this with complimentary neuroimaging data.

Finally, the uniformly low scores in the identification of object-environment conjunctions suggest that future incarnations of the OLT would benefit from reducing the complexity level in this part of the desktop task. This may be achieved by introducing a short prompt at the end of each environment encouraging participants to recall which objects had been recalled in that round. Alternatively, decreasing the similarity between environments may support participants in distinguishing between them when viewing a still screen image.

4.5 Conclusions

Overall, this study provides support for the use of VR technology as offering diagnostic potential in MCI patient groups. These results offer a proof-of concept that a task assessing object-location memory can successfully distinguish between aMCI patients and controls with a high degree of sensitivity. What is more, a model including measures of object-location memory and hippocampal function, when used in combination with premorbid functioning, was more successful than a model of pen-and-paper neuropsychological measures in predicting patient status. These results would be supported by future work aligning object-location performance to neuroimaging data in order to assess the involvement of the EC in the current task. The specificity of this measure in identifying individuals at high risk of conversion to AD also remains to be seen and should be assessed in future research.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Garmst, A., Holtzman, D., Jagust, W. J., Petersen, R.C., Snyder, P. J., Carrillo, M.C., Thies, B., & Phelps, C.H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, *7*(3), 270-279.

Alzheimer's Society (2014). Dementia UK: The Second Edition. Alzheimer's Society: London.

Ardila, A. (1995). Directions of Research in Cross-Cultural Neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17*,143- 150.

Birba, A., Ibáñez, A., Sedeño, L., Ferrari, J., García, A. M., & Zimerman, M. (2017). Non-Invasive Brain Stimulation: A New Strategy in Mild Cognitive Impairment?. *Frontiers in aging neuroscience*, *9*, 16.

Bird, C. M., Chan, D., Hartley, T., Pijnenburg, Y. A., Rossor, M. N., & Burgess, N. (2010). Topographical short-term memory differentiates Alzheimer's disease from frontotemporal lobar degeneration. *Hippocampus*, *20*(10), 1154–1169.

Bishara, D., Sauer, J., & Taylor, D. (2015). The pharmacological management of Alzheimer's disease. *Progress in Neurology and Psychiatry*, *19*(4), 9-16.

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... & Fernandez, M. (2009). How we design feasibility studies. *American journal of preventive medicine*, *36*(5), 452-457.

Braak H, Braak E (1991) Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica* 82, 239-259.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Conde-Sala, J. L., Garre-Olmo, J., Vilalta-Franch, J., Llinàs-Reglà, J., Turró-Garriga, O., Lozano-Gallego, M., Hernández-Ferrándiz, M., Pericot-Nierga, I., & López-Pousa, S. (2012). Predictors of cognitive decline in Alzheimer's disease and mild cognitive impairment using the CAMCOG: a five-year follow-up. *International Psychogeriatrics*, *24* (6), 948–958.

model of early Alzheimer's disease. *Neuron,* 73, 685-97

Cushman, L. A., Stein, K., & Duffy, C. J. (2008). Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. *Neurology*, *71*(12), 888-895.

Delagarza, V. W. (2003). Pharmacologic treatment of Alzheimer's disease: an update. *American family physician*, *68*(7), 1365-1372.

Deshmukh, S. S., & Knierim, J. J. (2011). Representation of non-spatial and spatial information in the lateral entorhinal cortex. *Frontiers in behavioral neuroscience*, *5*, 69.

Doeller C, Barry C, Burgess N (2010) Evidence for grid cells in a human memory network. *Nature, 463*, 657-661

Doeller, C. F., King, J. A., & Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Sciences*, *105*(15), 5915-5920.

Ewers, M., Walsh, C., Trojanowski, J. Q., Shaw, L. M., Petersen, R. C., Jack, C. R., Feldman, H. H., Bode, A.L., Alexander, G. E., Scheltens, P., Vellas, B., Dubois, B., Weiner, M., Hampel, H., in collaboration with the North American Alzheimer's Disease Neuroimaging Initiative (ADNI). (2012). Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiology of Aging*, *33*(7), 1203–1214.

Fischer, P., Jungwirth, S., Zehetmayer, S., Weissgram, S., Hoenigschnabl, S., Gelpi, E., Krampla, W., & Tragl, K. H. (2007). Conversion from subtypes of mild cognitive impairment to Alzheimer dementia. *Neurology*, *68*(4), 288-291.

Folstein, M. F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.

García-Betances, R. I., Arredondo Waldmeyer, M. T., Fico, G., & Cabrera-Umpiérrez, M. F. (2015). A succinct overview of virtual reality technology use in Alzheimer's disease. *Frontiers in aging neuroscience*, *7*, 80.

Gonsalvez, I., Baror, R., Fried, P., Santarnecchi, E., & Pascual-Leone, A. (2017). Therapeutic noninvasive brain stimulation in Alzheimer's disease. *Curr Alzheimer's Res*, *14*(4), 362-376.

Guariglia, C. C., & Nitrini, R. (2009). Topographical disorientation in Alzheimer's disease. *Arquivos de neuro-psiquiatria*, *67*(4), 967-972.

Hafting T, Fyhn M, Molden S, Moser M, Moser EI. (2005). Micro- structure of a spatial map in the entorhinal cortex. *Nature*, 436, 801–806.

Hargreaves, E. L., Rao, G., Lee, I., & Knierim, J. J. (2005). Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *science*, *308*(5729), 1792-1794.

Harris, M. A., & Wolbers, T. (2012). Ageing effects on path integration and landmark navigation. *Hippocampus*, *22*(8), 1770–1780.

Hartley, T., Bird, C. M., Chan, D., Cipolotti, L., Husain, M., Vargha-Khadem, F., & Burgess, N. (2007). The hippocampus is required for short-term topographical memory in humans. *Hippocampus*, *17*(1), 34-48.

Knierim, J. J., Neunuebel, J. P., & Deshmukh, S. S. (2014). Functional correlates of the lateral and medial entorhinal cortex: objects, path integration and local–global reference frames. *Phil. Trans. R. Soc. B*, *369*(1635), 20130369.

Ku, S. P., Nakamura, N. H., Maingret, N., Mahnke, L., Yoshida, M., and Sauvage, M. M. (2017). Regional specific evidence for memory-load dependent activity in the dorsal subiculum and the lateral entorhinal cortex. *Frontiers in Systems Neuroscience*, 11:51. doi: 10.3389/fnsys.2017.00051

McKhann G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E.M. .(1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human servIces task force on Alzheimer's disease. *Neurology,* 34(7):939–44.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's

disease. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, *7*(3), 263-269.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M. B. (2006). Path integration and the neural basis of the "cognitive map." *National Review Neuroscience*, *7*(8), 663–78.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International journal of geriatric psychiatry*, *21*(11), 1078-1085.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of abnormal psychology*, *110*(1), 40.

Mokrisova, I., Laczo, J., Andel, R., Gazova, I., Vyhnalek, M., Nedelska, Z., Levcik, D., Cerman, J., Vicek K., & Hort, J. (2016). Real-space path integration is impaired in Alzheimer's disease and mild cognitive impairment. *Behavioural Brain Research*, *307*(March), 150–158.

Moodley, K., Minati, L., Contarino, V., Prioni, S., Wood, R., Cooper, R., D'Incerti, L., Tagliavini, F., & Chan, D. (2015). Diagnostic differentiation of mild cognitive impairment due to Alzheimer's disease using a hippocampus-dependent test of spatial memory. *Hippocampus*, *25*(8), 939–951.

O'Keefe J, Dostrovsky J. (1971). The hippocampus as a spatial map: preliminary evidence from unit activity in the freely- moving rat. *Brain Research*, 34,171–5.

O'Keefe J, Nadel L. (1978). The Hippocampus as a Cognitive Map. Oxford, UK: Oxford University Press.

Pengas, G., Patterson, K., Arnold, R. J., Bird, C. M., Burgess, N., & Nestor, P. J. (2010). Lost and found: Bespoke memory testing for Alzheimer's disease and semantic dementia. *Journal of Alzheimer's Disease*, *21*(4), 1347–1365.

Prince, M., Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., & Prina, M. (2015). World Alzheimer Report 2015. The Global Impact of Dementia. Alzheimer's Disease International. *Alzheimer's Disease International (ADI), London*.

Ranck, J.B. Jr (1984). Head direction cells in the deep cell layer of dorsal presubiculum in freely moving rats. *Society for Neuroscience Abstracts, 10*, 176.12.

Salmon DP, Bondi MW. *Neuropsychology of Alzheimer's disease*. In: Terry RD, Katzman R, Bick KL, Sisodia SS, editors. Alzheimer disease. 2nd ed. Lippincott Williams & Wilkins; Philadelphia: 1999. pp. 39–56

Save, E., & Sargolini, F. (2017). Disentangling the role of the MEC and LEC in the processing of spatial and non-spatial information: contribution of lesion studies. *Frontiers in systems neuroscience*, *11*, 81.

Scoville, W. B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry, 20*, 11–21.

Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., Dean, R., Siemers, E., Potter, W., Lee, V.M., & Tojanowski, J.Q.. (2009). Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Annals of neurology*, *65*(4), 403-413.

Spector, A., Woods, B., & Orrell, M. (2008). Cognitive stimulation for the treatment of Alzheimer's disease. *Expert Review of Neurotherapeutics*, *8*(5), 751-757.

Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Reiew. Neuroscience, 27*, 279-306.

Tabert, M.H., Manly, J.J., Liu, X., Pelton, G.H., Rosenblum, S., Jacobs, M., Zamora, D., Goodkind, M., Bell, K., Stern, Y., & Devanand, D.P. (2006). Neuropsychological Prediction of Conversion to Alzheimer Disease in Patients With Mild Cognitive Impairment. *Arch Gen Psychiatry,* 63(8):916-924.

Taube J.S., Muller R.U., Ranck J.B., Jr. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats, I: description and quantitative analysis. *Journal of Neuroscience, 10*, 420–35.

Vargha-Khadem, F., Gadian, D.G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277, 376–380.

Wilson, D. I., Watanabe, S., Milner, H., and Ainge, J. A. (2013). Lateral entorhinal cortex is necessary for associative but not nonassociative recognition memory. *Hippocampus, 23*, 1280–1290. doi: 10.1002/hipo.22165

Van Cauter, T., Camon, J., Alvernhe, A., Elduayen, C., Sargolini, F., & Save, E. (2012). Distinct roles of medial and lateral entorhinal cortex in spatial cognition. *Cerebral Cortex*, *23*(2), 451-459.

Zygouris, S., Giakoumis, D., Votis, K., Doumpoulakis, S., Ntovas, K., Segkouli, S., Karagiannidis, C., Tzovaras, D., & Tsolaki, M. (2015). Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. *Journal of Alzheimer's Disease*, *44*(4), 1333-1347.

Part 3: Critical Appraisal

1 Overview

This critical appraisal of the research process explores the experiential aspects of carrying out this study. It will begin with a consideration of the emotional aspects of neuropsychological assessments for individuals with cognitive impairment and how this may relate to their experience of testing both in traditional formats and also with the object-location task (OLT) specifically, and virtual reality (VR) more generally. Throughout, pertinent points will be illustrated by quotes from the late, great Terry Pratchett, being both a prominent Alzheimer's spokesperson and one of my favourite authors. The second part of these reflections will concern the practical challenges that arose throughout the process, with a focus on the challenges associated with multisite collaboration and the position of visiting researcher. Finally, this appraisal will end with a reflection on the role of the clinician in establishing clinically-informative research frameworks, and how this has shaped my own sense of what is possible in dual clinician-researcher roles.

2 Experiential aspects of testing

The late Terry Pratchett, when reflecting on his diagnosis, wrote that "*Alzheimer's is me, unwinding, losing trust in myself, a butt of my own jokes and on bad days capable of playing hunt the slipper by myself and losing…. It steals you from yourself*" (Pratchett, 2015) Here, there is a poignant message of the intensely personal nature of Alzheimer's disease (AD) and the impact that it can have on the personhood of those who receive this diagnosis. Prior to beginning this research, I was allocated a six-month placement within a neuropsychology service, where my main role was providing one-off cognitive assessments, most often to individuals who had either suffered from strokes or were going through the process of dementia testing. Far from

the detached, objective process that neuropsychological screening had seemed during clinical teaching, I soon learned that this was one of the most highly emotive clinical settings that I had yet encountered. When pausing to reflect on this, it is understandable that this may well be the case. As Pratchett notes, one's identity is often highly entwined with our own perceived abilities, as well as the memories and experiences that constitute our life. The public perception of dementia in general, and Alzheimer's disease (AD) in particular, is that it gradually erodes not only one's cognitive capacity, but the autobiographical memories of a life. It conjures images of dependence – the 'second childhood' – and a reduced capacity to engage in the activities that constitute a meaningful engagement with life. In keeping with this concept, Steeman, Casterlé, Godderis and Grypdonck (2006) reviewed qualitative research into the experience of early stage memory difficulties and found that there was a high association with threat towards the sense of self-determination and meaning within society; not only this, but they noted an associated emotional burden of fear and uncertainty. Cognitive testing is unique in its ability to rapidly confront people with difficulties that they may have previously been unaware of having or, in some cases, may greatly fear that they are developing. It is little wonder then that cognitive testing is a highly emotional process for both examinee and examiner. What is a source of confusion, however, is a lack of consideration of this in relation to neuropsychological test construction.

In the design and write-up of this study I have followed the tradition of the dementia literature by framing the advantages of virtual reality (VR) use in relatively concrete terms: the cost and usability of the software; its utility in tapping into aspects of neurological functioning that can't be accessed through desktop technologies; or its tolerability to patients, demonstrated through task performance and drop-out rates. These factors are undoubtedly important when considering the use of immersive VR technologies in the assessment of MCI and AD, however they do not constitute the

whole picture. Equally important is the emotional experience of undergoing such an assessment. Despite this, it was not possible to include any formal assessment of the subjective experience of patients completing the OLT. This was primarily determined by the pre-existing ethical arrangements of the study, which had been set-up prior to our joining the research team and did not incorporate any formal assessment of the subjective experience of testing – e.g. a post-study follow-up questionnaire or qualitative interview. Nonetheless, it would have represented an important contribution to the research process and is an area of cognitive testing which is consistently overlooked.

This reflects a consistent theme amongst the dementia literature, as well as in the current study, in which the medical model is undoubtedly the dominant conceptual framework. The use of diagnostic testing is spoken about primarily in terms of the validity and utility of the assessment tools, as well as the predictive power of the tools. These are, of course, relevant factors to consider and yet little room is given to discussion of the experiential process of being assessed. Here I turn again to the words of Pratchett, who wrote that "*It occurred to me that at one point it was like I had two diseases – one was Alzheimer's and the other was knowing I had Alzheimer's. There were times when I thought I'd have been much happier not knowing, just accepting that I'd lost brain cells and one day they'd probably grow back or whatever*" (Pratchett, 2015). This provides us with a beautiful summary of the sense that the assessment and diagnosis of AD itself can be so painful as to be another form of disease. This is in line with qualitative assessments of neuropsychological testing in AD populations, where it is highlighted that the process of assessment can, for many, feel intricately tied up in a sense of self-worth. Tolhurst (2015) carried out interviews with men who had dementia diagnoses to ascertain their sense of the diagnostic process; he found that the sense of testing could be intimately linked with a sense of

self, with decline in test performance associated with emotional challenges. Interestingly, the author notes that this difficulty wasn't highlighted in regard to less detached assessment processes, such as neurological scans, suggesting that it is not the concept of AD or the confirmation of decline per se that presents an emotional challenge. Rather, the author links this to the social, relational and identity focussed nature of cognitive testing. It is suggested that this manner of confrontation with deficit, which occurs within a relationship and may be perceived to challenge the intellectual validity of the respondent, may be felt to more acutely undermine personhood than do other forms of deficit assessment (Tolhurst, 2015).

While it was not possible to gather formal feedback on the experience of the OLT, it remains that myself and my fellow researchers were actively involved in all aspects of data collection – including both OLT and pen and paper 'traditional' neuropsychological tests. In this way, I was able to observe the impact and the experience that the OLT appeared to have from an observer perspective, as well as based on informal feedback post-testing. While it's of course not possible to represent all the views given – of which there were many – it is possible to reflect on the overall sense of the participant experience. While it was often the case that participants found the pen and paper tests highly challenging, often eliciting high levels of self-criticism, this was reported much less frequently following completion of the OLT. While it is of course the case that for some individuals the OLT felt challenging and caused some anxiety about performance, this appeared to be far less common response than was the case in testing. Indeed, there were many instances in which participants reported finding the task enjoyable despite being aware of the challenges that it posed. This was reflected in the structure of the recruitment process, in which a single pool of participants had been invited to attend multiple VR based testing sessions, the OLT being the most recent of these. The fact that individuals had not only returned having

attended at least one previous testing session, but in many cases expressed willingness to return for future studies, suggests that there is something about the VR set-up that may feel less threatening to those who engage with it. This sense of positive emotional engagement with VR technology has been supported elsewhere in the literature, with one study finding that older adults reported a range of positive emotional experiences associated with the VR experience; these included references to novelty and escapism associated with the format, as well as a high degree of engagement and even excitement (Roberts, De Shuter, Franks & Radina, 2018).

A common criticism that examinees often level at traditional neuropsychological assessment is the lack of relevance that it bears to their everyday life. Both when working as a clinician and throughout this research process I have seen many cases in which people respond to tests by noting that this is something that they would always have struggled with, and do not feel that their difficulties are represented by the highly abstract tasks that make up these assessments. Again, this issue was noted by Tolhurst (2015) in interviews with male dementia patients and their carers, who reported a sense that neither their aptitudes or difficulties were adequately represented during the process of cognitive testing, and the associated frustration that this could cause. This is little wonder when we consider how devoid from real life challenges these tests can be; when asked to connect numbers and letters, draw complex images from memory or count backwards in multiples of seven, it is only a select few who would feel that these skills comprise a significant part of their day to day experiences. However, the challenge has often been how to capture isolated cognitive functions in a manner that is both relatively specific to said function, standardisable and replicable, while also maintaining a sense of real-world relevance (or, in research terms, ecological validity). Here, perhaps VR can offer some solution. Pratchett wrote that "*fanstasy is an exercise bicycle for the mind. It might not take you*

anywhere, but it tones up the muscles that can." (Pratchett & Briggs, 2013). It is not hard then to draw an analogy between imaginative fantasy and the 'fantasy' of a virtual world; while it may not be a direct correspondent to the real-life struggles that AD patients may face, perhaps it can open up a portal to the muscles that those everyday tasks require, as it were. In the case of the present study it would not be true to say that views on the task were uniformly positive, however there was no instance during testing when I experienced an individual reporting that the testing process did not feel that it measured a difficulty that felt relevant to their real-life experiences; indeed, misplacing objects feels perhaps one of the most common cognitive lapses for us all.

3 Multisite collaboration and the role of visiting researcher

I count myself very fortunate to have stumbled upon a project that was so highly set-up at the outset that a number of the challenges common to clinical research, such as establishing clinical links and applying for NHS ethics, were hurdles that my project partner and I were able to avoid. The trade-off that was made for these shortcuts was that all of our research was conducted off-site, in a testing centre approximately an hour and a half away from our home institution. In addition, the design of the research was in many ways constrained by the pre-existing structures already in place, notably the ethical approval, as well as the need to balance our demands with the priorities and needs of an entirely different research team. Altogether, this project was tied in with the research needs of both myself, my project partner and supervisor, as well as an additional two PhD projects and two MSc projects. This is alongside the overarching research aims of the senior research team, for whom volume of research data was a clear priority and this study was one amongst many that they were hoping to run within a fairly short time period. As is inevitable in such cases, each of these priorities was running on a different time scale and therefore decisions as to how and

when many aspects of the research would be carried out did not run to a time-scale that was ideal for myself and my partner. In addition, I had to make peace with a lower level of control over the research process than I would typically be comfortable with. Here, I would like to take the time to reflect on a one such issue that we encountered, and the learning I've taken from them.

One of the key challenges that arose was in the time scale and specifics of recruitment, which was handled largely by a research nurse affiliated with the off-site research team. The designation of an individual specifically to support with recruitment was a massive asset to this study, and their hard work undoubtedly saved myself and my project partner a great deal of time in calling and administration. An additional factor in the decision to have recruitment handled on-site was that our visiting researcher status granted us access to buildings, data and testing facilities, however did not grant computer access. This greatly restricted the involvement that we were able to have in the screening of participants, given that patient data was held securely on an on-site server. Due to pressures to recruit in an accelerated time-frame and a lack of access to patient data, there was an initial lack of clarity as regards the nature of the participants that we were recruiting. While it was always agreed that our research would focus on an MCI patient cohort, it was initially thought that this may focus specifically on those individuals who had also tested as positive for cerebrospinal fluid (CSF) biomarkers of β-amyloid and neurofibrillary tau, giving our data greater specificity to an AD diagnosis. During initial discussions with senior research investigators it had been thought that this would have been a possibility due to the high number of potential patients from which we could recruit. However, when we then broached this issue with those responsible for recruitment it emerged that this was not possible nor had this been a factor in those patients who were recruited for the

study. While this did not in essence change the rationale or execution of the study, it did underline for me how easy it was for miscommunications to occur that had the potential to alter our research in a significant way. This difficulty was resolved, with communication between, myself, my project partner and our supervisor, as well as through discussions that took place.

From this incident, I have taken a number of points. Firstly, I found that large research teams come with a number of advantages, but also their fair share of challenges. While there are undoubted benefits in terms of the broader range of thinking and potential to divide the work load, to name but a few, communication amongst team members becomes far more complex. On one level, there were communication challenges involved in co-ordinating a complex research set-up, involving multiple sites and a large amount of technical equipment, with my fellow researchers with whom I had no day-to-day contact. This was also evident in our assumptions that the information that had been given to us regarding patient recruitment had also been conveyed to all members of the off-site research team. Given the large number of people and locations involved, it became difficult to arrange that all team members were in the same place at the same time, and this led to issues of miscommunication where such assumptions were made. At another level, there are often differing priorities and aims between more senior researchers and staff, such as myself, working at the ground level. While for myself and my colleagues the aim was more clearly invested in the practical running of this individual study, it is necessary for more senior colleagues to focus on broader aims, with this study being only one amongst many that are being planned. Therefore practical details, such as was the issue with participant recruitment, can at times become lost through miscommunication. On reflection, I do feel that this was an issue that was noted at an early stage and significant efforts were made to increase regular effective communication. However, in the future

I would have this in mind before engaging in any multi-site research project and ensure that from the outset a clear method for consistent communication across all team members is in place, with clear structures for including colleagues at all levels of the research project.

Secondly, I found that I as a researcher can find it a challenge to feel in some ways out of control within the research process. At every stage of this study my own research was dependent on the support of multiple colleagues; this was true in terms of the design of the computer and VR tasks, the technical set-up of the equipment and the recruitment of participants. Further, all data were stored on-site at the testing location, meaning that I did not have access to hard copies of the data in between testing session. In all of my previous research experience, I have taken on the role of research assistant; in this role, I have had a high level of control over the hard copies of the data and have worked primarily within single-site projects and a close research team (both in terms of professional relationships and physical proximity). This was therefore a new experience for me and one that required some adjusting. I found this a particular challenge during the process of write-up, where acquiring many aspects of the data – e.g. demographic information, data collected outside of the testing session – would require my requesting this from a PhD student on-site at the testing centre, who had access to both hard copies of data and the computer databases that I was unable to reference. Of course, they were incredibly helpful and responsive, which meant that no major difficulties came up in this regard. However, I did realise within myself that this lack of control feels very uncomfortable, and my desire to check and verify various aspects of data collection had to be restrained more so than is usually the case. This was an important opportunity for me to sit with the discomfort of feeling that control was shared amongst a number of team members; through this I have come to understand the importance of sharing equal responsibility and of relinquishing a

need to feel so highly involved in every aspect of the research process. Further, this encouraged me during the process of write-up to shift my focus from the smaller details of project management to the broader sense of the narrative and over-arching aims of the project; for example, rather than attempting to re-check and re-analyse the data multiple times in order to understand an unexpected result, I instead took a step back and thought about how this might make sense in the context of our particular patient group, and what alternative theoretical perspective might say about this issue. In this way, I feel that I have learned the personal and academic benefits of sharing involvement in research in a meaningful way, and have attempted to take a step back from a need to control or manage all aspects of a project.

4 The clinical research process

A final brief space here must go to my observations of the recruitment set-up that allowed this research to take place. As is highlighted in the methods section of my empirical paper, our patient sample was recruited directly through their attendance at the memory clinic, with all participants initially approached via their attending clinician. As part of the pre-recruitment and testing stage of research myself and my project partner sat in on several clinic sessions, as the senior researcher in our guest institution was also, in most cases, the clinician who gave the MCI diagnoses and approached individuals about research. Here I was impressed at the dedication that this clinician showed both to the clinical aspects of their role, but their ability to think about the broader potential to use this dual-position for the increasing research opportunities. Of course, there are difficulties with this manner of recruitment, primary among which is the fact that the process of diagnosis is a very sensitive one, and that individuals can feel grateful to and obligated to comply with the wishes of their clinician when asked if they are willing to participate in research, or may not have had time to consider the

request in greater detail. This then raises some questions about whether individuals are truly comfortable with consenting to taking part in research. However, what we did find throughout recruitment was that those individuals who did not, after reflection, wish to take part did feel able to give this feedback when subsequently approached by the research team and therefore did fairly quickly appear to withdraw from the research.

What was evident was that many of the barriers often encountered in carrying out clinically relevant research, such as finding and establishing relationships with disconnected clinical teams, did not feel to be an issue in this instance. Rather, there was an impressive amount of joined-up thinking about the clinical and research reality of the individuals in the study; this was evident both in small, practical ways, such as the ease of access to diagnostic information, but also in the broader process issues such as the familiarity of conversations between participants and the research team. Further, because of the close links between the research team, participants and clinical team, there was a clear consideration of each participants personal situation when planning the testing session. This felt particularly salient for me as an individual on the cusp of qualification, considering not only how I would like to place myself as regards the clinical settings that I work within, but also how I would want to balance this with potential future work in research. This encouraged me to think that there is an important and vital role for clinicians in actively participating in and establishing frameworks for the routine carrying out of clinically-relevant research.

5 Conclusions

The issues considered here highlight the clinical and practical learning that can be taken from even highly technical or neuroscience-focussed theses. Despite the veneer of objectivity and detachment that can surround neuropsychological

assessment, there is undoubtedly a high degree of emotional content involved in both the process and the outcome of diagnosis. This appraisal considered how this was relevant in the context of the current research, as well as how this may apply more broadly to technological advances in AD assessment. It also put forward personal learning for myself in terms of the benefits and challenges of multisite collaboration within research, how this might intersect with my own experiences and preferences as a researcher, and the ways in which I might take this forward in the future. I also reflect on the ways in which the research process has informed my views on the importance of dual clinician and researcher roles, and the unique position that they offer in terms of increasing research opportunities and promoting clinical thinking within research teams. Finally, I would like to end with a concluding quote from Terry Pratchett, which aptly summarises the process of learning through difficulties and mistakes:

"That was always the dream, wasn't it? 'I wish I'd known then what I know now'? But when you got older you found out that you NOW wasn't YOU then. You then was a twerp. You then was what you had to be to start out on the rocky road of becoming you now, and one of the rocky patches on that road was being a twerp." (Pratchett, 2002).

References

Pratchett, T., (2015, 15 March). 'A butt of my own jokes': Terry Pratchett on the disease that finally claimed him. *The Guardian.* Retrieved from https://www.theguardian.com/books/2015/mar/15/a-butt-of-my-own-jokes-terry-pratchett-on-the-disease-that-finally-claimed-him

Pratchett, T. (2002). *Night Watch: A Discworld Novel*. London: Doubleday.

Pratchett., T., Briggs, S., & Simonetti, M. (2013). Turtle Recall: The Discworld Companion. Gollancz.

Roberts, A. R., Schutter, B. D., Franks, K., & Radina, M. E. (2018). Older adults' experiences with audiovisual virtual reality: perceived usefulness and other factors influencing technology acceptance. *Clinical gerontologist*, 1-7.

Steeman, E., Casterlé, D., Dierckx, B., Godderis, J., & Grypdonck, M. (2006). Living with early-stage dementia: A review of qualitative studies. *Journal of advanced nursing*, *54*(6), 722-738.

Tolhurst, E. (2015). The Experiential Impacts of Cognitive Function Tests upon Men with Dementia and Their Carers. *The Journal of New Writing in Health and Social Care, 2*(1), 1-13.

Appendix A
Participant Information Sheet (PIS)

UNIVERSITY OF CAMBRIDGE

Cambridge University Hospitals NHS
NHS Foundation Trust

Department of Clinical Neurosciences

Dr Dennis Chan PhD MD FRCP

Herchel Smith Building for Brain and Mind Sciences
Forvie Site, Robinson Way
Cambridge CB2 0SZ
Tel: 0044 (0)1223 760697
Fax: 0044 (0)1223 336 581

PARTICIPANT INFORMATION SHEET

Study title: VIrtual Reality Testing of Entorhinal Cortex and Hippocampal function in Alzheimer's Disease (VIRTECH-AD)

IRAS ID: 193437

Part 1 - Background:

We would like to invite you to take part in a research study that uses virtual reality/augmented reality (VR/AR) technology to test spatial navigation (getting from A to B) and spatial memory (remembering the location of objects) in people with mild cognitive impairment

You have been asked to participate in this research because your memory clinic specialists have diagnosed you with mild cognitive impairment.

Before you decide to take part we would like you to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. You can talk to others about the study if you wish or ask one of the researchers for further information.

What is the purpose of the study?

Spatial navigation and memory are abilities controlled by regions of the brain called the entorhinal cortex and the hippocampus, which are the first brain regions to be affected by Alzheimer's disease (AD). As such, when people are in the earliest stages of AD, these abilities will be impaired and if we can identify this impairment will help detect AD in its very earliest stages. At present, when people such as yourself are

diagnosed with mild cognitive impairment on clinical grounds, it is not possible to know whether this may be due to AD unless specialised and invasive tests (amyloid-PET brain scanning or lumbar puncture studies) are undertaken. It would therefore be greatly desirable to have new tests that in the future can help doctors understand whether a person's mild cognitive impairment is due to AD.

The problem is that it is not practical to test people's spatial navigation and memory in real world settings such as parks or streets. However, current wearable VR/AR technology provides a solution to this problem. VR headsets (see Figure 1, left) give the wearer the experience of being immersed within a virtual environment (Figure 1, right), within which they can move around using hand-held controls. By comparison, AR headsets generate artificial objects and scenes (holographic images) superimposed on the real world (hence the term "augmented reality"). While this does not have the advantage of VR in creating a fully immersive simulated environment, the ability to perceive the real world as well as the simulated images may prove less disorientating.

Together, these novel technologies provide previously unavailable opportunities to test spatial navigation and memory in a way that may be applied in clinical practice.

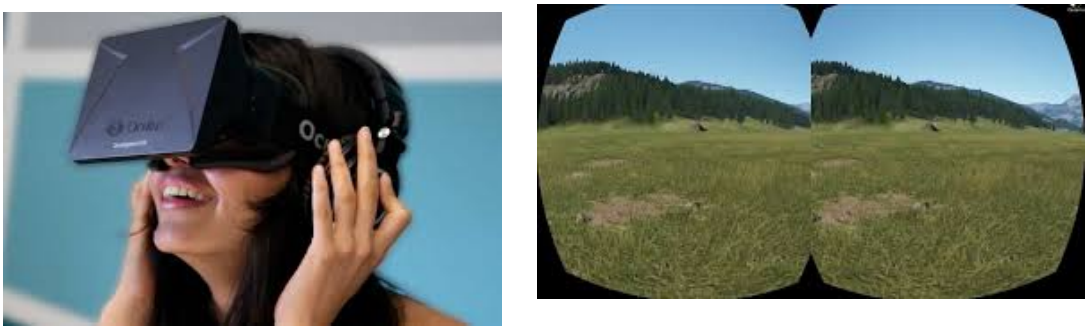


Figure 1. Left: a wearable VR headset. Right: an example of the immersive virtual environment visualised by the headset wearer.

However, to date, VR/AR has not been used for testing spatial navigation or memory in people with early AD. To do so, we need to undertake studies to investigate which VR/AR technology can be used to help diagnose early AD and also to obtain feedback on whether people experience any problems with these tests, such as nausea or a sense of disorientation while moving around within the simulated environment using the VR/AR headsets. Any such feedback will be used to adjust the test in order to prevent these issues in the future.

Do I have to take part?

Participation in this study is entirely voluntary. We will describe this study to you and go through this information sheet, which will be given to you. Should you decide to take part, we will ask you to sign a consent form. You are free to withdraw from the study at any time, with no need to give a reason.

What does participating involve?

This study requires you to wear a VR or AR headset. At the beginning of the study you will be invited to use a VR headset (the HTC Vive, currently available commercially). As the study proceeds, the study team will have access to an AR headset (the Microsoft Hololens, not yet available commercially) and you will be asked if you would like also to be tested using this headset.

It takes less than ten minutes to get used to the headsets. You will be asked to explore a simulated scene which a number of everyday objects (such as a vase or a ball) are placed at specific locations. Once you have familiarised yourself with this scene and with the VR or AR equipment, you will be tested on your ability to navigate within the simulated scene and your memory for the location of these everyday objects.

You will be tested once with the VR headset and once with the AR headset, if you agree to return at a later date for testing with the latter. Each test will take around 30 minutes.

At the end of the testing period we will ask you a few questions about your experience during the task, and in particular whether you experienced any discomfort or disorientation during the task.

Who will have access to my medical records?

Your medical records will be reviewed by the study team before you start the study to check your eligibility for the study.

In addition to the study team, your records may be reviewed as part of an auditing process carried out by the R & D department. These audits serve to uphold rules relating to good clinical practice.

What happens to the study data?

Research data will be stored for 10 years. All personal identifiable data will be stored on NHS computers held in lockable offices and password-protected. These data will be held for 10 years before being destroyed.

Will I be told about the results from this study?

Yes. We will endeavour to inform you of the results of tests. Furthermore our aim is to present the study results at public-patient meetings and you will be invited to attend these meetings. These presentations will describe the overall study results rather than results on individuals and will not refer to you, or any other participant, personally or in any other way that would compromise your anonymity.

In the event that my condition deteriorates and my mental capacity is lost, what will happen to my personal data?

Assuming that you gave your consent at the beginning of the study, the research team would retain the study data collected and continue to use it confidentially in connection with the purposes for which consent is being sought.

Expenses and Payments

We will be able to contribute £20 to the cost of travel to our research site.

What are the possible disadvantages and risks of taking part?

There are no major disadvantages or risks associated with participation. However, some users may find VR headsets makes them feel sick and/or disorientated. It is important that if you experience any of these problems that you make the researcher aware so they can stop the test.

What are the possible benefits of taking part?

There are no immediate benefits to you. However, if the study is successful in its aims, then it will help us diagnose AD in its very earliest stages.

What if there is a problem?

If you have a concern about any aspect of the study or wish to make a complaint, you can speak to one of the researchers who will do their best to answer your questions or address your complaint, or by contacting Dr Dennis Chan in writing, or by telephone 01223 760696. If you wish to express concerns or complaints to someone outside the research team then you are advised to contact the Patient Advisory and Liaison Service (PALS), Addenbrooke's Hospital, Cambridge.

Confidentiality – who will have access to the data?

If you join this study your personal information and the data we collect will be stored on a secure network and only the research team will have access to it. Since we work with other researchers worldwide in the study of AD, it is possible that these researchers may analyse *anonymised* data arising from this study, abiding by the terms of formal collaborations between the University of Cambridge and other academic institutions. In such instances your personal details will not be shared.

What will happen to the results of the research study?

The results of the study will be published in scientific journals. Participant data will be anonymised so that it is not possible for you to be identified in published articles.

Who is organising and funding the research?

This study is organised by Dr Dennis Chan and funded by the Medical Research Council and by Alzheimer's Research UK. The study results will be analysed by Dr Chan's research group.

Contact details

If you have any queries, please contact Dr Dennis Chan in writing, or by telephone 01223 760696.

Appendix B
Ethical Approval Letter

NHS
Health Research Authority

East of England - Cambridge South Research Ethics Committee

The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Please note: This is the favourable opinion of the REC only and does not allow the amendment to be implemented at NHS sites in England until the outcome of the HRA assessment has been confirmed.

16 February 2018

Mr David Howell
Herchel Smith Building
Robinson Way
Cambridge
CB2 0SZ

Dear Mr Howell

Study title:	**VIrtual Reality Testing of Entorhinal Cortex and Hippocampal function in early Alzheimer's disease (VIRTECH-AD)**
REC reference:	**16/EE/0215**
Amendment number:	**SA3**
Amendment date:	**18 January 2018**
IRAS project ID:	**193437**

The above amendment was reviewed at the meeting of the Sub-Committee held in correspondence on 09 February 2018

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Appendix C
Letter confirming visiting researcher status

Cambridge University Hospitals NHS
NHS Foundation Trust

Research and Development Department
Box 277
Addenbrooke's Hospital
Hills Road
Cambridge
CB2 0QQ

Ms Elizabeth Harding
Visiting Researcher
Dept of Clinical Neurosciences
Forvie Site

R&D Manager: Stephen Kelleher
stephen.kelleher@addenbrookes.nhs.uk
HR Manager: Debbie Richards
01223 274660
deborah.richards@addenbrookes.nhs.uk
HR Advisor: Gayle Lindsay
01223 348496
gayle.lindsay@addenbrookes.nhs.uk

18th April 2017

Dear Elizabeth

Letter of access for research – A093863 - VIrtual Reality Testing of Entorhinal Cortex and Hippocampal function in ageing and mild cognitive

This letter confirms your right of access to conduct research through Cambridge University Hospitals NHS Foundation Trust for the purpose and on the terms and conditions set out below. This right of access commences on **31st March 2017** and ends on **30th September 2018** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permission to conduct the project and you have provided the Trust's R&D department with written evidence that you have completed GCP training from an EU institution before you start your research.

The information supplied about your role in research at Cambridge University Hospitals NHS Foundation Trust has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to Cambridge University Hospitals NHS Foundation Trust premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through Cambridge University Hospitals NHS Foundation Trust, you will remain accountable to your place of work **University of Cambridge** but you are required to follow the reasonable instructions of **Dr Dennis Chan and Rodney Laing** in this NHS organisation or those given on their behalf in relation to the terms of this right of access.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any

Appendix D
Participant consent forms for control and MCI participants

7th October 2016
Version 2

UNIVERSITY OF CAMBRIDGE

Herchel Smith Building for Brain and Mind Sciences
Forvie Site, Robinson Way
Cambridge CB2 0SZ
Tel: 0044 (0)1223 760697
Fax: 0044 (0)1223 336 581

Department of Clinical Neurosciences

Dr Dennis Chan PhD MD FRCP

PARTICIPANT CONSENT FORM

Control Participants

Study title: VIrtual Reality Testing of Entorhinal Cortex and Hippocampal function (VIRTECH).

Principal Investigator: Dr Dennis Chan, University Lecturer and Honorary Consultant in Clinical Neurosciences.

Please tick the boxes:

1. I confirm that I have read and understood the information sheet version 5 dated 7th October 2016 for the above study and have been given a copy to keep.	
2. I have had the opportunity to ask questions about the study and have received answers to my questions.	
3. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	
4. I understand that my personal details and information about me that is gathered for this research study will be held on a secure, confidential computerised database that is only accessible to members of the research group. This is in accordance with the Data Protection Act 1998.	
5. I confirm that I have had sufficient time to consider whether or not I want to be included in the study.	
6. I give consent for additional virtual reality based tests of spatial navigation and memory to be performed, as outlined in the participant information leaflet.	
7. I understand that none of my results will be given to me and that I will not benefit financially from taking part	
8. I understand that the research data collected from the study will only be share within our research group and I will not be identified if the results are published.	
9. I agree to take part in the above study.	

7th October 2016
Version 2

CONSENT FORM (CONFIDENTIAL)

Title of project: **VIrtual Reality Testing of Entorhinal Cortex and Hippocampal function (VIRTECH)**

Principal Investigator: Dr Dennis Chan, University Lecturer and Honorary Consultant in Clinical Neurosciences

Participant Identification Number:

Project ID Number:

_____	_____	_____
Name of subject	Date	Signature
_____	_____	_____
Name of Person taking consent (if different from researcher)	Date	Signature
_____	_____	_____
Researcher (to be contacted (in the event of any problems)	Date	Signature

Comments or concerns during the study.

If you have any comments or concerns you may discuss these with the investigator.

If you wish to go further and complain about any aspect of the way you have been approached or treated during the course of the study, you should write or get in touch with the Complaints Manager, University of Cambridge University.

One form for Participant

One form to be kept as part of the study documentation

14 June 2016
Version 2

UNIVERSITY OF CAMBRIDGE

Cambridge University Hospitals NHS
NHS Foundation Trust

Department of Clinical Neurosciences
Dr Dennis Chan PhD MD FRCP

Herchel Smith Building for Brain and Mind Sciences
Forvie Site, Robinson Way
Cambridge CB2 0SZ
Tel: 0044 (0)1223 760697
Fax: 0044 (0)1223 336 581

PARTICIPANT CONSENT FORM

Study title: Virtual Reality Testing of Entorhinal Cortex and Hippocampal function in Alzheimer's Disease (VIRTECH-AD).

Principal Investigator : Dr Dennis Chan, University Lecturer and Honorary Consultant in Clinical Neurosciences

Please tick the boxes:

1. I confirm that I have read and understood the information sheet version 2 dated 14 June 2016 for the above study and have been given a copy to keep.	
2. I have had the opportunity to ask questions about the study and have received answers to my questions.	
3. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	
4. I understand that my personal details and information about me that is gathered for this research study will be held on a secure, confidential computerised database that is only accessible to members of the research group. This is in accordance with the Data Protection Act 1998.	
5. I confirm that I have had sufficient time to consider whether or not I want to be included in the study.	
6. I give consent for the study tests of spatial navigation and memory to be performed, as outlined in the Participant Information Sheet.	
7. I understand that my medical records and the research data collected from the study may be looked at by the sponsor, by regulatory authorities or by the NHS Trust, as part of auditing policies.	

8. I understand that the research data collected from the study will be analysed on an anonymous basis.	
9. I understand that I will not receive any financial benefit from any intellectual property arising from this study.	
10. I agree to take part in the above study.	

CONSENT FORM (CONFIDENTIAL)

Title of project: **Virtual Reality Testing of Entorhinal Cortex and Hippocampal function in Alzheimer's Disease (VIRTECH-AD).**

Principal Investigator: Dr Dennis Chan, University Lecturer and Honorary Consultant in Clinical Neurosciences

Patient Identification Number:

Project R&D Number:

_____	_____	_____
Name of participant	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature

One form for Participant

One form to be kept as part of the study documentation

Appendix E
Statement regarding joint thesis contributions

Testing was carried out jointly by Elizabeth Harding (EH) and project partner, Adrienne Li (AL; DClinPsy student, UCL). Other researchers who contributed to the testing of patients were David Howitt (DH; PhD student, University of Cambridge), Emma Barham (EB; Research Nurse, University of Cambridge), Zoe Adler (ZA; MSc student, University of Cambridge) and Volker Reisner (VR; MSc student, University of Cambridge).

AL's DClinPsy thesis also makes use of OLT data, however the focus is on the relation to neuroimaging MRI measures.

EH and AL carried out virtual reality testing for all of the aMCI patient group and approximately half of control participants.

Neuropsychological testing was carried out by EH, AL, EB and ZA, with EB and ZA carrying out the majority of testing.

Scoring and analysis of neuropsychological testing data was carried out solely by EH.

Scheduling of participants was carried out by EB and DH.

Lumbar punctures, from which CSF data was gathered, was carried out by EB.

MRI scanning of the aMCI patient group was scheduled by DH, and analysed by AL, DH, ZA and VR.