Longitudinal profiling of Mild Cognitive Impairment subtypes

By Shannon Z. Klekociuk

BPsych (Hons.), B.P.A

School of Psychology

Submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy

University of Tasmania, October 2013
Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in this thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed………………………………………………………………………

Date………………………………
Statement of Authority of Access

This thesis may be made available for loan and limited copying and communication in accordance with the copyright Act 1968

Signed…………………………………………………………………….

Date…………………………….
Statement regarding published work contained in this theses

The publishers of the papers comprising Chapters 4, 5, 6, 7 and 8 hold the copyright for that content, and access to the material should be sought from the respective journals. The remaining non-published content of the thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Signed…………………………………………………………………….

Date………………………………
Statement of Co-authorship

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Name and School = Shannon Klekociuk, School of Psychology
Name and institution = Dr Mathew J. Summers, School of Psychology (UTAS) & Wicking Dementia Research and Education Centre, School of Medicine (UTAS)
Name and institution = Prof. Jeff Summers, School of Psychology (UTAS) & Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom.
Name and institution = Prof. James C. Vickers, Wicking Dementia Research and Education Centre, School of Medicine (UTAS)

Author details and their roles:

Paper 1, The self-fulfilling prophecy of episodic memory impairment in mild cognitive impairment (MCI): Do episodic memory deficits identified at classification remain evident when later examined with different memory tests?
Located in chapter 4
Candidate was the primary author (70%) and with author 2 contributed to the idea, its formalisation and development, refinement and presentation (30%)

Paper 2, Lowered performance in working memory and attentional sub-processes are most prominent in multi-domain amnestic mild cognitive impairment subtypes
Located in chapter 5
Candidate was the primary author (70%) and with author 2 contributed to the idea, its formalisation and development, refinement and presentation (30%)

Paper 3, Exploring the validity of Mild Cognitive Impairment (MCI) subtypes: Multiple-domain amnestic MCI is the only identifiable subtype at longitudinal follow up
Located in chapter 6
Candidate was the primary author (70%) and with author 2 contributed to the idea, its formalisation and development, refinement and presentation (30%)

Located in chapter 7
Candidate was the primary author (70%) and with author 2 contributed to the idea, its formalisation and development, refinement and presentation (30%)

Paper 5, Reducing false positive MCI diagnoses with neuropsychological assessment
Located in chapter 8
Candidate was the primary author (60%) and with author 2 contributed to the idea, its formalisation and development, refinement and presentation (30%)
Author 3 (5%) and 4 (5%) assisted with refinement and presentation
We the undersigned agree with the above stated “proportion of work undertaken” for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed: ___________________ ______________________

Dr Mathew Summers
Supervisor
School Of Medicine
University of Tasmania

Prof. Jenn Scott
Head of Discipline (Psychology)
School of Medicine
University of Tasmania

Date: ________________
Statement of Ethical Conduct

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government’s Office of the Gene Technology Regulator, and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

Signed………………………………………………………………………………

Date……………………………..
Publications from thesis

Peer reviewed articles


Papers currently accepted for publication pending revisions:


Abstracts


2012:


Abstract

Mild Cognitive Impairment (MCI) was originally conceptualized as a condition that manifested prior to the onset of clinical dementia, particularly Alzheimer’s disease. However, longitudinal studies show that MCI has an unstable course and may lead to various outcomes including dementia, but also stability of cognitive deficits or recovery to age appropriate levels of functioning. As a result, the status of MCI as a genuine diagnostic entity remains questionable. The aim of the present thesis was to examine the validity of the MCI concept by tracking groups of individuals classified into one of the MCI subtypes and to monitor their neuropsychological profiles over time. To avoid previous criticisms of circularity, participants were classified as MCI on a neuropsychological test battery and then reassessed longitudinally using an alternate battery of neuropsychological tests. At each stage of testing, participants were assessed on a comprehensive neuropsychological test battery tapping the cognitive domains implicated in MCI. Findings from this thesis indicate that multiple domain amnestic MCI may be the most valid subtype of MCI due to consistently poor performance over time on a range of neuropsychological measures. Results also demonstrate that those who are likely to remain on the MCI spectrum can be differentiated from healthy older adults using reliable and valid measures of sustained attention, semantic memory, verbal episodic memory, visual and verbal working memory, selective attention and strategy use. Despite these findings, evidence from this thesis indicates that existing MCI clinical criteria lack sufficient sensitivity and specificity. Although the concept of MCI remains useful, it cannot be considered a clinical diagnostic entity. Future research should prioritize the observation of those presenting with a multiple domain amnestic profile as these individuals may have the poorest prognosis. Further, studies must utilize comprehensive testing protocols to increase the sensitivity and specificity of identifying those with genuine subclinical impairments.
Acknowledgements

To my supervisor, Mat: I cannot thank you enough for your support during my candidature. I have so much respect for the way that you work as a researcher, as an academic, and as a psychologist. You were a big motivator for me to undertake this journey and I’ve had a great time working with you (even during those tough times!). You’ve always given me your time so generously; for personal chats and endless discussions about research, academia, and teaching. And while I did have to put up with your constant barrage of poor dad jokes and digs about my technology choices, I’d say that it was all worth it. I am forever grateful for your honesty, guidance, and friendship.

I would also like to thank my office buddies Mai and Megan for always being there to chat, to whinge about marking, or provide me with creative ways to procrastinate. To my co-supervisors, Jeff and James, thank you for your support. The psychology lads; Pete, Matt, and Douglas: You’ve taught me so much about the crazy world of academia, I am eternally grateful. Thanks for your encouragement throughout my PhD, your support with teaching, as well as your regular political counsel over coffee and cake. Thanks must also go to Anthea for her admin support.

To my beautiful family: You have always been so supportive of my endeavours and this huge undertaking was no exception. Thank you for always being there for me. To my mum Merril, and sister Hannah; you are the two most special women in my life. I could not have conquered this beast without you! To my dearest partner, Dave, who put up with me during this crazy time; thank you for your kindness, your patience, and enduring love.
My sincere thanks must go to UTAS, the Australian Government, and the Wicking Dementia Research and Education Centre for their financial support during my candidature. Of course, none of this would have been possible had it not been for the dedicated participants and family members involved. I have learnt so much about aging gracefully (and disgracefully!). I will remember our sessions with great fondness, thank you.
Table of Contents

Declaration of Originality ................................................................. i
Statement of Authority of Access ......................................................... ii
Statement regarding published work contained in this theses ......................... iii
Statement of Co-authorship ............................................................... iv
Statement of Ethical Conduct .......................................................... vi
Publications from thesis .................................................................. vii
Papers and Poster Presented at Conferences ........................................... ix
Abstract ......................................................................................... x
Acknowledgements .......................................................................... xi
List of Figures .................................................................................. xviii
List of Tables .................................................................................... xix
Chapter 1 ......................................................................................... 1
Introduction ..................................................................................... 1

Alzheimer’s Disease: Neuropathology and Neuropsychology ...................... 1
  Neuropathology of Early AD ......................................................... 2
  Episodic Memory in Early AD ........................................................ 3
  Attention in Early AD ................................................................. 4
  Working Memory in Early AD ...................................................... 5
  Executive Function in Early AD .................................................... 6
  Semantic Memory in Early AD ...................................................... 6

Pre-dementia Syndromes .................................................................... 7

Mild Cognitive Impairment ............................................................... 8
  MCI Criteria .................................................................................. 9
  Revised criteria ............................................................................ 9
  National Institute on Aging recommendations .................................. 10

The Neuropathological and Neuropsychological Profile of MCI ................. 12
  Neuropathology of MCI ............................................................. 12
  Episodic Memory Function in MCI .............................................. 13
  Attention function in MCI .......................................................... 15
  Working Memory Function in MCI ............................................. 17
  Executive Function in MCI ......................................................... 17
  Semantic Memory Function in MCI .......................................... 19

Genetics and MCI ........................................................................... 19
Current Issues in MCI Research .................................................................20
  The Validity of the MCI Concept and MCI Subtypes ........................21
  The Stability of MCI ........................................................................22
  Lack of Comprehensive Neuropsychological Assessment among MCI cohorts ... 31
  Circular Reasoning ........................................................................33

Aim of the Present Thesis and Outline of Chapters ..................................34

Chapter 2 ..................................................................................................38
General Methodology ..............................................................................38
  Recruitment ..........................................................................................39
  Pre-screening .......................................................................................39
  Screening .............................................................................................40
  Participants .........................................................................................40
  Materials ............................................................................................45
  Procedure ...........................................................................................53
  Experimental stages ..........................................................................54
  Participants .........................................................................................55
  Materials ............................................................................................55
  Procedure ...........................................................................................65

Chapter 3 ..................................................................................................66
Screening Assessment .............................................................................66
  Statistical analysis ............................................................................67
  Results ...............................................................................................68
  Discussion ..........................................................................................76
  a-MCI ...............................................................................................76
  na-MCI .............................................................................................77
  a-MCI+ .............................................................................................78
  Control ...............................................................................................78

Chapter 4 ..................................................................................................79
The self-fulfilling prophecy of episodic memory impairment in mild cognitive impairment (MCI): Do episodic memory deficits identified at classification remain evident when later examined with different memory tests? ....... 79
  Abstract .............................................................................................80
  Introduction .......................................................................................81
  Method .............................................................................................83
  Study population ...............................................................................83
  Materials ...........................................................................................86
Lowered performance in working memory and attentional sub-processes are most prominent in multi-domain amnestic mild cognitive impairment subtypes

Abstract

Introduction

Method

Study population
Materials
Procedure

Results

Discussion

Chapter 6

Exploring the validity of Mild Cognitive Impairment (MCI) subtypes: Multiple-domain amnestic MCI is the only identifiable subtype at longitudinal follow up

Abstract

Introduction

Method

Participants
Materials
Procedure

Results

Data analysis
Demographics
Verbal and Visual Episodic Memory
Attention Measures
Immediate Memory Measures
Working Memory Measures
Language

Discussion

Chapter 7

The learning profile of persistent Mild Cognitive Impairment (MCI): A potential diagnostic marker of persistent amnestic MCI
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>145</td>
</tr>
<tr>
<td>Method</td>
<td>148</td>
</tr>
<tr>
<td>Participants</td>
<td>148</td>
</tr>
<tr>
<td>Materials</td>
<td>153</td>
</tr>
<tr>
<td>Procedure</td>
<td>154</td>
</tr>
<tr>
<td>Results</td>
<td>155</td>
</tr>
<tr>
<td>Data analysis</td>
<td>155</td>
</tr>
<tr>
<td>Discussion</td>
<td>160</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>164</td>
</tr>
<tr>
<td>Reducing false positive MCI diagnoses with neuropsychological assessment</td>
<td>164</td>
</tr>
<tr>
<td>Method</td>
<td>169</td>
</tr>
<tr>
<td>Participants</td>
<td>169</td>
</tr>
<tr>
<td>Materials</td>
<td>174</td>
</tr>
<tr>
<td>DNA extraction</td>
<td>175</td>
</tr>
<tr>
<td>PCR</td>
<td>175</td>
</tr>
<tr>
<td>Gel electrophoresis</td>
<td>176</td>
</tr>
<tr>
<td>Procedure</td>
<td>176</td>
</tr>
<tr>
<td>Results</td>
<td>177</td>
</tr>
<tr>
<td>Discussion</td>
<td>180</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>184</td>
</tr>
<tr>
<td>Discussion</td>
<td>184</td>
</tr>
<tr>
<td>Chapter 4 Discussion</td>
<td>187</td>
</tr>
<tr>
<td>Chapter 4 Key findings</td>
<td>187</td>
</tr>
<tr>
<td>Chapter 5 Discussion</td>
<td>188</td>
</tr>
<tr>
<td>Chapter 5 Key findings</td>
<td>189</td>
</tr>
<tr>
<td>Chapter 6 Discussion</td>
<td>190</td>
</tr>
<tr>
<td>Chapter 6 Key findings</td>
<td>191</td>
</tr>
<tr>
<td>Chapter 7 Discussion</td>
<td>192</td>
</tr>
<tr>
<td>Chapter 7 Key findings</td>
<td>192</td>
</tr>
<tr>
<td>Chapter 8 Discussion</td>
<td>193</td>
</tr>
<tr>
<td>Chapter 8 Key findings</td>
<td>194</td>
</tr>
<tr>
<td>General discussion</td>
<td>195</td>
</tr>
<tr>
<td>MCI is an unstable condition and should not be considered a diagnostic entity</td>
<td>195</td>
</tr>
<tr>
<td>The a-MCI+ profile is a better representation of the MCI condition</td>
<td>197</td>
</tr>
<tr>
<td>Limitations</td>
<td>199</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. Diagnostic algorithm for MCI subtypes .............................................................. 10
Figure 2. Display screen for Paired Associates Learning (PAL) ..................................... 59
Figure 3. Display screen for Simple Reaction Time (SRTI) ........................................ 60
Figure 4. Display screen for Choice Reaction Time (CRTI) ........................................ 60
Figure 5. Display screen for Matching to Sample (MTS) ............................................. 61
Figure 6. Display screen for Rapid Visual Information Processing (RVP) ................... 62
Figure 7. Display screen for Spatial Span (SSP) .......................................................... 63
Figure 8. Display screen for Spatial Working Memory (SWM) .................................... 64
Figure 9. Group differences visual episodic memory (mean ± SEM) ......................... 91
Figure 10. Group differences verbal episodic memory (mean ± SEM) ....................... 92
Figure 11. Mean spatial span length for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to a-MCI group at $p < .05$) ......................... 107
Figure 12. Mean forward Digit Span length forward and backward for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to control and a-MCI group at $p < .05$) ........................................... 108
Figure 13. Mean letter-number sequencing for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to a-MCI and control groups at $p < .05$) ........... 109
Figure 14. Mean spatial working memory total errors for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to a-MCI at $p < .05$) ....................... 110
Figure 15. Mean RVP mean latency for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to control group at $p < .05$) ....................................... 112
Figure 16. Mean RVP A’ for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to a-MCI and control groups at $p < .05$) ....................... 113
Figure 17. Participant flow chart .................................................................................. 124
Figure 18. Participant progression chart ...................................................................... 149
Figure 19. Flow chart of classification of participants at each assessment phase ......... 152
Figure 20. Learning curves for MCI subgroups on the Rey Auditory Verbal Learning Test ......................................................................................................................... 158
Figure 21. Participant progression chart ...................................................................... 171
Figure 22. Stability classification procedure ................................................................. 174
List of Tables

Table 1. Testing Protocols and Rates of Reversion across Studies Tracking MCI .................. 25
Table 2. Screening Neuropsychological Test Battery .............................................................. 44
Table 3. Experimental Psychological and Neuropsychological Test Battery ....................... 57
Table 4. Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores at Screening, Time 1, and Time 2 .......................................................... 69
Table 5. Group Differences on Screening Neuropsychological Assessment ...................... 74
Table 6. Screening Test Battery Used for MCI Classification ............................................... 85
Table 7. Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores ................................................................................................................ 89
Table 8. Test Batteries for Screening and Working Memory and Attention Assessment ...... 103
Table 9. Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores ................................................................................................................ 106
Table 10. Non-significant Group Differences in Attention Performance ......................... 111
Table 11. Screening and Experimental (Time 1 and Time 2) Test Batteries ......................... 126
Table 12. Group Mean Scores for Demographic and Screening Measure ..................... 130
Table 13. Episodic Memory Performance in MCI subtypes at Time 1 and Time 2 .......... 132
Table 14. Performance on Measures of Attention in MCI Subtypes at Time 1 and Time 2 ... 134
Table 15. Performance on Immediate Memory and Working Memory Measures in MCI Subtypes at Time 1 and Time 2 .......................................................... 136
Table 16. Test Batteries for Screening and Follow-up Testing ............................................ 151
Table 17. Classification of Participants at Screening, Time 1, and Time 2 ........................ 152
Table 18. Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores ................................................................................................................ 156
Table 19. Effect Sizes of Mean Differences of Persistent a-MCI and Healthy Control, and Persistent na-MCI and Healthy Control across RAVLT trials ......................... 159
Table 20. Test Batteries for Screening and Follow-up Testing ........................................ 172
Table 21. Demographic Information and APOE E4 Status ............................................... 179
Table 22. Outcome Classification Predicted from Time 1 Neuropsychological Scores ...... 179
Table 23. Sensitivity and Specificity of MCI Diagnosis Using Standard Diagnostic Procedures (Winblad et al., 2004) and DFA analysis .......................................................... 180
Table 24. Summary of the Main Results of the Present Thesis ........................................ 185
Chapter 1

Introduction
Alzheimer’s disease (AD) is a progressive cortical dementia that typically presents with deficits to episodic memory followed by widespread impairment to multiple cognitive domains such as learning, attention, decision making, and language function. It is the leading cause of dementia in older adults accounting for approximately 40% of all cases in the USA (Alzheimer's Association, 2009) and approximately 50% of all cases in Australia (Access Economics, 2009). In Australia, there has yet to be a national study of dementia prevalence that utilises clinical diagnostic data. Current estimates and projections have relied on published epidemiological studies and meta-analyses. To date, the most prominent study of dementia prevalence was undertaken by Deloitte Access Economics (2011) for Alzheimer’s Australia. In their report, Deloitte Access Economics estimated that there were 266,574 Australians living with dementia in 2011, with forward projections estimating that this would rise to 553,285 by 2030 and 924,624 by 2050. In a previous report, it was estimated that delaying the onset of dementia by only 5 months would save $6.6 billion, with a delay in onset of 5 years resulting in a saving of $67.5 billion in national health care costs by 2040 (Access Economics, 2009). In light of the increase in average lifespan across the world as well as the aging of the post-World War II “baby boomer” generation in Western countries, an emerging dementia crisis has been identified. This concern has resulted in an increased effort to identify those at the early stages of pathological cognitive decline. The assumption being that intervention strategies will be most effective at the earliest possible stages of the neurodegenerative process (Petersen & Negash, 2008).

**Alzheimer’s Disease: Neuropathology and Neuropsychology**

Alzheimer’s disease (AD) is progressive cortical dementia that primarily manifests as an amnestic syndrome with prolific disturbances to episodic memory (Perry & Hodges, 1999).
As AD progresses, disruption to higher cognitive functions such as language, attention, visuospatial abilities and executive functions become more evident (Perry & Hodges, 1999; Zillmer, Spiers, & Culbertson, 2008). The follow sections describe the neuropathological changes associated with early AD and the clinical symptoms that occur as a result.

**Neuropathology of Early AD**

In early AD, neuropathological changes primarily involve medial temporal lobe (MTL) structures, predominantly the hippocampus and entorhinal complex, but also regions in the frontal lobes (Dickerson & Sperling, 2008). Neurons in these areas are particularly susceptible to the accumulation of neurofibrillary plaques, neuritic tangles, as well as neuronal atrophy (Braak & Braak, 1991; de Leon et al., 2004). Neurofibrillary tangles are disproportionately represented in temporoparietal regions; the hippocampal complex, particularly the entorhinal complex, as well as the periamygdaloid cortex and corticomedial region of the amygdala (de Leon et al., 2004). Plaques tend to concentrate in the frontal and temporal regions, particularly in the hippocampus (Zillmer et al., 2008). MTL structures play a vital role in mnemonic function, receiving afferent projections from the hippocampus as well as providing the hippocampus with critical input from various association cortices (Dickerson & Sperling, 2008). Therefore, the sequence of MTL neuropathology is seen to account for the hallmark memory deficits in early AD. Upon examination, AD brains show generalised neurodegeneration in association cortices of the frontal, temporal, and parietal lobes with additional subcortical atrophy evident in the hippocampus and amygdala (Zillmer et al., 2008).
Episodic Memory in Early AD

Episodic memory refers to memory for spatial and temporal information about a particular event (Zec, 1993). Early stage AD is primarily characterised by anterograde amnesia reflecting an underlying difficulty encoding, consolidating, and retrieving new episodic information (Zec, 1993). Conceptually, this represents a failure in transferring information from short term memory to long term memory. Individuals with early AD also experience temporally graded episodic retrograde amnesia although this is considered to be less severe than the deficits to anterograde amnesia (Hodges, 2006). Episodic memory deficits in early AD can affect both visual and verbal domains, and be evident across different conditions e.g. free recall, cued recall, and recognition (Backman, Small, & Fratiglioni, 2001). When assessed on neuropsychological measures, early AD individuals demonstrate difficulty recalling information beyond their immediate memory span (Zillmer et al., 2008). Further, individuals at this stage do not tend to benefit from repeated presentation of stimuli indicating a deficit to encoding and consolidation (Storey, Kinsella, & Slavin, 2001; Zillmer et al., 2008). This is particularly evident on list learning tasks where early AD individuals demonstrate a flattened learning curve despite the repeated nature of the task (Grober & Kawas, 1997). Performance on clinical tests of delayed free recall are effective at identifying early AD cases given the prominent deficit to new learning (Salmon & Bondi, 1999; Storey et al., 2001). However, many studies have revealed that deficits to non-memory domains (e.g. attention, working memory etc.) may appear earlier than initially expected (Bondi et al., 2008). Some research indicates that working memory deficits may underpin episodic memory difficulties in early AD (Germano & Kinsella, 2005). Other research suggests that deficits to attentional control are a prominent feature of early AD and may be useful in identifying those who are destined for dementia (Rapp & Reischies, 2005).
Attention in Early AD

Research examining attentional function delineate between three broad processes: Selective attention; divided attention; and sustained attention. Selective attention refers to the ability to screen out irrelevant stimuli and is often subdivided into the following sub processes; detecting relevant stimuli, filtering irrelevant stimuli, and inhibiting irrelevant stimuli from interfering with an appropriate response (Perry & Hodges, 1999). Divided attention refers to the ability to undertake cognitive tasks simultaneously, whereas sustained attention refers to the ability to maintain attention over an extended period of time (Perry & Hodges, 1999). In terms of early AD, there is ongoing controversy regarding the exact mechanisms that are impaired and the stage at which this impairment becomes apparent (Perry, Watson, & Hodges, 2000).

Some research suggests that deficits to selective attention in early AD are a result of difficulty inhibiting prepotent response. Studies have found that early stage AD is associated with difficulty inhibiting the automatic reading of the word rather than the colour on the Stroop task (Perry et al., 2000) as well as impaired responding to distracter items on choice reaction time tasks (Baddeley, Baddeley, Bucks, & Wilcock, 2001). Research examining divided attention suggests that early AD is associated with significant impairments to the ability to undertake multiple tasks or to simultaneously attend to two or more stimuli (Baddeley et al., 2001; Belleville, Chertkow, & Gauthier, 2007). However, the assessment of divided attention using dual task paradigms has been criticised for their reliance on other (non-attentional) executive processes (Perry & Hodges, 1999). Perry et al. (2000) argue that many studies fail to account for disease severity, and that this has led to ambiguity regarding the time at which divided attention deficits become evident. In their study using a dual task paradigm, Perry and colleagues found that divided attention was impaired in a middle stage
AD group but not early stage AD group relative to controls. In terms of sustained attention, there is limited research detailing the integrity of this process in early AD. It has been suggested that sustained attention remains relatively intact during the earliest stages, however there is some evidence that early AD individuals have diminished accuracy and increased RT on complex sustained attention tasks (Rizzio, Anderson, Dawson, Myers, & Ball, 2000).

**Working Memory in Early AD**

According to Baddeley and Hitch’s (1974) original model, working memory is an interactive system consisting of an attentional control component, the central executive, which is assisted by two slave systems, the phonological loop and the visuospatial sketchpad (Baddeley, 1998). The central executive component is an attentional system with a limited storage capacity that is responsible for the coordination and prioritisation of mental processes (Baddeley, 2002). The visuospatial sketchpad and the phonological loop act as temporary stores of visual and verbal information respectively (Baddeley, 1998). The model makes a distinction between the passive storage capacity of the slave systems versus the active processing capacity of the central executive (Germano & Kinsella, 2005). Evidence of impaired central executive function has been found in research showing that AD individuals have difficulty orchestrating two concurrent tasks (e.g. Belleville, Peretz, & Malenfant, 1996). In terms of the slave components of working memory, the general conclusion from studies is that the phonological loop and visuospatial sketchpad remain relatively intact until middle stage AD (Baddeley, Bressi, Della Sala, Logie, & Spinnler, 1991; Cherry, Buckwalter, & Henderson, 1996).
Executive Function in Early AD

Attention and working memory processes are sometimes subsumed under the label of executive functioning. Typically, the domain of executive function is conceptualised as multiple higher order processes associated with goal oriented behaviours (McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). For example, planning, sequencing, cognitive flexibility, decision making, judgement, rule application, set shifting, and abstract thinking, are all considered executive processes (Brandt et al., 2009). In terms of executive functioning in early AD, evidence suggests that individuals experience difficulty with tasks that make demands on mental tracking (e.g. repeating the alphabet in reverse); and also flexibility and self-monitoring (e.g. card sorting tasks) (Baudic et al., 2006). Storey, Kinsella and Slavin (2001) suggest that subtle changes to executive function may underpin the impairments to everyday activities evident in early stage AD, such as becoming confused in novel situations or having trouble adapting to changes in routine.

Semantic Memory in Early AD

While episodic memory is recognised as the major domain of memory disturbance in early AD, semantic memory may also be impaired (Zillmer et al., 2008). Semantic memory can be defined as the memory for facts, figures, and names which all require access to word knowledge (Storandt, 2008). Research indicates that individuals in early AD demonstrate a reduced capacity to recall well established information such as the name of everyday objects (Storandt, 2008). Early AD individuals may also experience difficulty maintaining and understanding semantic relationships between words and what they represent (Lezak, 1995). Reading of lower frequency or orthographically irregular words can also be disrupted (Storey
et al., 2001). The exact stage at which semantic difficulties become clinically evident is less clear with some research suggesting that it is observable after the breakdown of episodic memory and attention (Perry et al., 2000) whereas others suggest that it occurs at the same time as episodic memory dysfunction (Mickes et al., 2007).

**Pre-dementia Syndromes**

Although there is yet to be a treatment capable of reversing or delaying the progression of AD, researchers agree that such a treatment would be most effective during or even before mild symptomology manifests (Dickerson & Sperling, 2008; Zillmer et al., 2008). Subsequently, a major research effort has been dedicated to characterising the pre-clinical syndrome of AD. The concept of a pre-clinical syndrome is not new. The first description of such a state came from Kral (1962) who introduced the term *benign senescent forgetfulness* as a way of capturing mild memory dysfunction in older adults (Levy, 1994). This was followed by *age associated memory impairment* (AAMI) from the National Institute of Mental Health Workgroup in 1986, which described a purely amnestic syndrome characterised by performance ≥ 1SD below the mean on an objective memory measure (K. Ritchie & Touchon, 2000). AAMI was criticised for its amnestic focus and for using normative data based on young adults (Petersen & Negash, 2008). The ability of AAMI to adequately discriminate between those experiencing normal versus pathological cognitive decline was also questioned. *Age associated cognitive decline* was proposed by the International Psychogeriatric Association (Levy, 1994) in attempt to overcome the caveats of AAMI. This concept referred to performance that was ≥ 1SD below the mean for age based norms across a range of domains such as memory, attention, or language (Levy, 1994). A
similar concept emerged from the Canadian Study of Health and Aging and was referred to as *Cognitive Impairment-No Dementia* (CIND) (Tuokko & Frerichs, 2000). CIND is defined by the presence of multiple cognitive impairments that do not fulfil the diagnostic criteria for clinical dementia. Whilst various labels have been used to define this transitional phase, the term Mild Cognitive Impairment (MCI) has come to dominate the literature (Tuokko & Hultsch, 2006).

**Mild Cognitive Impairment**

The MCI profile was initially identified from a series of MAYO Clinic studies attempting to identify predictive risk factors for AD (Petersen et al., 1997; Petersen et al., 1999; Petersen, Stevens, et al., 2001). Originally, MCI was conceptualised as a purely amnestic condition based on the MAYO Clinic studies findings that episodic memory compromise was prominent among those who eventually converted to AD (Petersen et al., 1997). Petersen and colleagues subsequently proposed the following criteria for MCI: (i) subjective report of memory impairment (preferably substantiated by an informant); (ii) objective memory impairment compared to age appropriate norms; (iii) no significant impairment to activities of daily living; (iii) normal general cognition; and (iv) an absence of clinical dementia (Petersen & Morris, 2005). Petersen et al. (1999) claimed that MCI was an adequate clinical entity in terms of its ability to highlight those individuals who were at most risk of future cognitive decline, particularly those likely to transition to AD.
MCI Criteria

Since the initial MAYO Clinic characterisation MCI criteria have evolved to accommodate the findings from additional research. Primarily, changes to the MCI criteria have been driven by research demonstrating evidence of widespread heterogeneity among MCI cohorts in terms of progression and outcome (A. J. Mitchell & Shiri-Feshki, 2009). The subsequent sections will examine the criterion changes implemented by Winblad et al. (2004) which form the current framework for MCI classification. This will be followed by a discussion regarding recent recommendations for MCI criteria as outlined by the working group from the Alzheimer’s Association National Institute on Aging (Albert et al., 2011).

Revised criteria (Winblad et al., 2004)

In 2004, The International Working Group on Mild Cognitive Impairment extended the MCI concept to include various subtypes (Winblad et al., 2004) (see Figure 1). These subtypes reflect differential patterns of impairments to specific cognitive domains thought to indicate an underlying aetiology. This conceptual shift was partly influenced by evidence suggesting that memory deficits, although a major feature of AD, may not be the earliest features and that other non-memory domains (e.g. attention, working memory) may be affected first (Perry & Hodges, 1999). It was also influenced by research indicating that individuals may transition to non-AD dementias (Busse, Hensel, Gühne, Angermeyer, & Riedel-Heller, 2006). Two major adjustments were made to the original criteria:

(1) criterion (ii) was altered to include individuals who present with memory and/or non-memory deficits;
(2) The wording of criterion (i) was changed to capture individuals reporting a cognitive complaint rather than just those concerned with their memory.

Theoretically, individuals may be classified as either amnestic MCI (isolated memory impairment), non-amnestic MCI (isolated non-memory impairment), multiple-domain amnestic MCI (memory and non-memory impairment) or multiple domain non-amnestic MCI (multiple non-memory impairments) (Winblad et al., 2004). It has been suggested that amnestic variants, particularly multi-domain amnestic, are more likely to transition to AD, whereas non-amnestic variants may progress to non-AD type dementias (Fischer et al., 2007).

![Diagram of Diagnostic Algorithm for MCI Subtypes](image)

**Figure 1.** Diagnostic algorithm for MCI subtypes

**National Institute on Aging recommendations (Albert et al., 2011)**

The National Institute on Aging and Alzheimer’s Association recently published a series of recommendations for revisions to the Winblad et al. (2004) criteria based on the growing body of MCI research (Albert et al., 2011). The working party outlined a set of *core clinical criteria* with the aim of being sufficiently broad to be used across a variety of clinical
settings. The recommendations also propose separate *clinical research criteria* which incorporate the use of biomarkers in the identification of MCI due to AD. It is important to note that these recommendations use the term MCI to refer to a prodromal phase of AD. The aim of such criteria is to capture those individuals whose likely pathophysiology is AD, although the authors do recognised that AD pathophysiology may occur in conjunction with other aetiologies. The recommendations by Albert et al. primarily reflect changes to the criteria regarding subjective and objective cognitive impairment. The subjective complaint criterion has been reworded from a subjective complaint of cognitive decline to “*a concern regarding a change in cognition*” (Albert et al., 2011, p. 271). This can be derived from a subjective judgement made by the individual; somebody who knows the individual well; or a clinical practitioner who has observed the individual. Although it is not explicitly stated, it is presumed that the term *change* refers to a decline in functioning relative to prior ability. With respect to objective cognitive performance, Albert et al. suggest that individuals should exhibit a lowered performance (relative to their age and education history) in at least one area of cognitive functioning. In contrast to the Winblad et al. criterion, which defined evidence of objective decline occurring over time, the Albert et al. revisions only require evidence of a least one impaired domain at a single time point. The recommended revisions by Albert et al. go further than previous criteria in terms of outlining memory and non-memory domains that should be assessed in potential MCI cases. However, Albert and colleagues do not specify the number of domains that should be assessed nor do they state the optimum number of tests that should be used. As with previous iterations, these recommendations state that criteria are guidelines only, and that clinical expertise is required to make a diagnosis of MCI (Albert et al., 2011; Winblad et al., 2004).
The Neuropathological and Neuropsychological Profile of MCI

On the premise that MCI represents a genuine precursor to AD, researchers have used the early AD profile as a framework to characterise the syndrome of MCI. Specifically, evidence regarding the neuropathology and neuropsychology of early AD has formed the basis for understanding MCI. The next sections outline the neuropathological changes that occur in MCI, followed by a detailed discussion of the neuropsychological features of MCI.

Neuropathology of MCI

Given the primary role of medial temporal lobe structures in early AD pathology, these same structures have been of interest to researchers examining MCI (Machulda et al., 2009). Currently, there are small number of fMRI studies that have investigated hippocampal activation in MCI, although results remain inconsistent (Dickerson & Sperling, 2008). Some studies reveal hypoactivation of the hippocampus during episodic memory tasks (Machulda et al., 2009), whilst others have reported significant hyperactivation (Johnson et al., 2004). Dickerson and Sperling suggest that the variability in findings is related to the severity of clinical impairment. Specifically, hyperactivation of MTL structures may occur in the early stages of MCI, acting as a compensatory mechanism to impending AD neuropathology. As memory deficits progress and hippocampal atrophy increases, hypoactivation may occur as a result of an inability to activate this region during learning (Dickerson & Sperling).

Recent evidence emphasising the role frontal regions in dementia related pathologies has initiated further research in MCI. Machulda et al. (2009) used fMRI to show that structural changes to frontal regions are evident at the MCI stage. During an encoding task, individuals
with multiple-domain amnestic MCI displayed reduced activation across temporoparietal regions, as well as in a small area in the right frontal lobe. In contrast, when performing a recognition task, there was evidence of reduced activation predominantly in temporoparietal regions. Machulda and colleagues interpreted the combination of diminished activation across temporoparietal and frontal regions as being consistent with the multiple domain amnestic MCI profile, which is characterised by deficits to memory and non-memory functions. In comparison, the non-amnestic MCI group displayed diminished activation primarily in frontal regions during encoding and recognition, reflecting their predominant deficit to non-memory domains (Machulda et al.).

**Episodic Memory Function in MCI**

The original series of clinical studies tracking the cognitive function of older adults identified episodic memory impairment as a core feature of those who transitioned to AD (Petersen et al., 1997; Petersen et al., 1999). Subsequently, episodic memory dysfunction became a core feature of the MCI syndrome. Currently, this conceptualisation is referred to as the single domain amnestic MCI subtype (Winblad et al., 2004). Given that the amnestic profile is thought to be the manifestation of emerging AD pathology, a large amount of research has been dedicated to exploring this subtype (Alexopoulos, Grimmer, Perneczky, Domes, & Kurz, 2006; Petersen & Negash, 2008). The general finding from research studies is that amnestic MCI groups perform between AD and healthy controls on tests of episodic memory (Arnaiz & Almkvist, 2003); and that episodic memory performance may be an important indicator of those who are likely to transition to AD (Ahmed, Mitchell, Arnold, Nestor, & Hodges, 2008; Albert et al., 2011; Albert, Moss, Blacker, Tanzi, & McArdle, 2007). While such patterns have emerged, comparisons between studies are complicated by
the variety of measures that are used (Arnaiz & Almkvist, 2003). In particularly, concerns have been raised in terms of studies that rely on tests of verbal episodic memory.

Many of the early MCI studies utilised measures of verbal memory to document amnestic functioning on the premise that verbal measures have greater predictive power in terms of identifying those at risk of AD (Flicker, Ferris, & Reisberg, 1991). This approach has been challenged by evidence showing that individuals may present with isolated deficits to visual memory (Bondi et al., 2008). Alladi, Arnold, Mitchell, Nestor, and Hodges (2006) examined the use of visual and verbal episodic memory measures when classifying amnestic MCI. Whilst many of the participants in this study demonstrated impairment to both verbal and visual memory, a sizeable minority demonstrated an isolated visual impairment. As a result, Alladi and colleagues argue that quantifying memory via verbal memory performance only runs the risk of missing individuals who would otherwise meet the criteria for MCI. Recent recommendations to MCI criteria have suggested some visual based tests which may be useful in identifying memory impairment in MCI, however it appears that a classification of MCI can be made without an assessment of visual memory.

Deficits to learning have also been of interest to researchers examining MCI given evidence of their appearance in early AD (Germano & Kinsella, 2005; Grober & Kawas, 1997). Ribeiro et al. (2007) found that the rate of word trial learning in MCI was higher than AD participants but significantly lower than the control group, which aligns with the notion of MCI as a precursor to AD. Rabin et al. (2009) found that a standard list learning task was more accurate in terms of identifying MCI cases compared to a prose recall task. Rabin and colleagues suggest that list learning may increase the accuracy of classification because the onus resides with the examinee in terms of encoding. This is supported by other evidence
suggesting that poor learning performance in MCI is due to ineffective strategy use at the encoding stage (Price et al., 2010; Ribeiro et al., 2007). Other enquiries into learning and MCI have suggested that the combination of trial based learning and delayed recall measures enhances the identification of MCI cases likely to develop dementia. Chang et al. (2010) examined rates of progression from MCI to AD in a group of participants with learning and/or retention deficits on a clinical list learning task. At two year follow-up, the MCI cases with combined learning and retention deficits were at the highest risk of transitioning to AD. These results imply that learning is also compromised in MCI and may assist in isolating MCI cases that are most at risk of developing dementia, in particularly AD.

While episodic memory dysfunction remains a key feature of amnestic MCI subtypes (Winblad et al., 2004), research examining non-amnestic functioning is emerging. This is largely due to evidence suggesting that deficits to non-memory domains (e.g. attention, working memory) may be detectable earlier than previously thought and may represent a key feature of those who are destined for clinical dementia (Backman, Jones, Berger, Laukka, & Small, 2004; Grober et al., 2008; Rapp & Reischies, 2005; Summers & Saunders, 2012). In addition, there is a growing body of research indicating that isolated amnestic deficits are rare, particularly if individuals are assessed on a comprehensive neuropsychological test battery (Alladi et al., 2006; Summers & Saunders, 2012).

**Attention function in MCI**

Attention function has been of interest to MCI researchers given evidence that this domain is vulnerable to decline in the early stages of AD (Perry & Hodges, 1999). However, a majority of this research focuses on attentional performance in those who eventually convert to AD.
Further, some studies do not use discrete measures of attentional processing, instead relying on broad screening measures to imply the integrity of attention function as a whole. Of the limited research that has comprehensively examined attention in MCI, it is apparent that aspects of attentional processing are compromised.

Okonkwo, Wadley, Ball, Vance, and Crowe (2008) found that MCI participants performed poorly on measures of simple sustained attention, selective attention and divided attention. It was noted that performance was particularly impaired in terms of divided attention, leading Okonkwo and colleagues to suggest that deficits to divided attention may be an important feature of MCI. Similar conclusions were drawn from a longitudinal study tracking the neuropsychological profile of MCI subtypes. Saunders and Summers (2011) revealed that amnestic and non-amnestic subtypes developed increasingly poorer performance on a five-choice reaction time task at 20 months follow up, despite performing in the normal range at baseline. From this, Saunders and Summers suggest that impairments to divided attention in amnestic and non-amnestic subtypes may be an important feature of those who eventually decline. Recent findings from the same group indicate that non-amnestic MCI and amnestic MCI groups also display a deficit to complex sustained attention (Summers & Saunders, 2012). Using discriminant function analysis, Summers and Saunders found that a measure of sustained attention was the strongest predictor variable in terms of differentiating between healthy controls, MCI, recovered MCI and AD cases. These results suggest that sustained attention may be more vulnerable at the MCI stage than initially expected (Summers & Saunders, 2012).
**Working Memory Function in MCI**

While there is evidence indicating working memory deficits in early AD, less is clear about these deficits in MCI. A small number of studies have attempted research in this area but tend to focus purely on the amnestic variant of MCI. Belleville et al. (2007) examined central executive functioning in individuals who met the criteria for amnestic MCI and multi-domain amnestic MCI in comparison to an AD and control group. Participants were assessed on their ability to manipulate, select and inhibit information, as well as their capacity to simultaneously coordinate two or more activities; all tasks that propose to tax the central executive (Belleville et al., 2007). The results revealed that deficits were most severe in the AD group, yet both MCI groups displayed a clinically significant impairment to central executive functioning. In a more recent study, Saunders and Summers (2011) assessed the longitudinal performance of amnestic and non-amnestic subtypes on working memory measures tapping short term visual storage and manipulation. Both MCI groups showed impaired but stable performances on these tasks across a 20 month period indicating that working memory is compromised at the MCI stage.

**Executive Function in MCI**

Research investigating executive functioning suggests that there are several deficits to this domain at the MCI stage and that such deficits may be useful in predicting future cognitive decline (Albert, Moss, Tanzi, & Jones, 2001; Rapp & Reischies, 2005). Brandt et al. (2009) found that difficulty with planning/problem solving and working memory was evident across all MCI subtypes, but particularly prominent in multiple domain subtypes. Brandt and colleagues suggest that these results indicate poorer prognosis for multiple domain variants.
In another study, Bisiacchi et al. (2008) examined cognitive flexibility, inhibition of prepotent response, use of strategy, and task switching in amnestic MCI and AD participants. Consistent with their classification, the amnestic MCI group performance was comparable to controls in terms of executive functioning. However, in terms of memory performance, the amnestic group were comparable to the AD group. Based on this finding, Bisiacchi and colleagues suggest that executive deficits may represent a key feature of the decline from MCI to clinical dementia. Despite these findings, there is some evidence that suggests that executive functioning may be less useful in predicting outcome in MCI.

In a recent longitudinal study, Aretouli, Tsilidis, and Brandt (2013) examined the ability of executive measures to predict dementia from MCI. At a four year follow up, several executive functions were individually associated with cognitive decline in MCI, including inhibition of prepotent response, set shifting, flexibility, planning, decision making and working memory. However, findings revealed that age, informant reports of functioning, and performance on a dementia severity scale was better at discriminating those who declined to AD than measures of executive functioning. While these findings suggest that executive measures may have limited predictive utility above demographic information and a dementia screening tool, the classification protocol may have influenced this outcome. Prior to classification, participants were not assessed on clinical measure of non-memory function, meaning that the severity of deficits in some individuals could have been underestimated. Potentially, some individuals included in the MCI cohort may have been manifesting signs of early dementia that were not identified via the amnestic centred classification protocol.
Semantic Memory Function in MCI

The degree of semantic memory dysfunction at the MCI stage is unclear. Some studies report that amnestic MCI and AD profiles demonstrate similar patterns of semantic memory breakdown (Dudas, Clague, Thompson, Graham, & Hodges, 2005; Murphy, Rich, & Troyer, 2006). Other studies report that only amnestic MCI is associated with reduced naming capacity compared to healthy controls, but that performance in this domain is relatively stable over time (N. L. J. Saunders & Summers, 2011). Although the evidence regarding semantic disturbance in MCI is mixed, studies attempting to map the trajectory of MCI should include an examination of semantic memory given that it has been previously linked to the development of AD (Murphy et al., 2006).

Genetics and MCI

Apolipoprotein E (ApoE) is considered a risk factor gene for the development of AD (Albert, 1996). ApoE plays a role in transporting cholesterol in the blood and is linked to reduced clearance of amyloid peptide from the brain, which may increase the build-up of neuritic plaques and neurofibrillary tangles in AD (Bondi, Salmon, Galasko, Thomas, & Thal, 1999). ApoE is composed of three alleles, ε2, ε3, and ε4, the latter of which has been associated with the development of AD (Albert, 1996; Albert et al., 2001; Almkvist & Tallberg, 2009; Corder et al., 1993; A. M. Saunders et al., 1993).

If the presence of an ε4 allele increases the risk of developing AD, and MCI represents a genuine prodromal phase of AD, then it is logical to suggest that ε4 status is also a risk factor for MCI. However, evidence linking ε4 status with MCI is inconsistent. Some studies report
that amnestic MCI groups are composed of a significantly greater proportion of $\varepsilon 4$ carriers compared to non-amnestic or healthy individuals (Aggarwal et al., 2005; Sasaki et al., 2009) whereas others do not (Bangen et al., 2010). Similar inconsistencies are apparent in studies that examine the association between $\varepsilon 4$ status and conversion to dementia. Some studies reveal that having at least one $\varepsilon 4$ allele is a strong predictor of AD (Petersen, Smith, Ivnik, & et al., 1995) whereas others report no association between $\varepsilon 4$ status and conversion to AD (Amieva et al., 2004). In one study, it was found that $\varepsilon 4$ status combined with performance on clinical tests of memory and executive functioning were the best predictors of AD (Fleisher et al., 2007). However, removal of $\varepsilon 4$ status as a predictor variable did not significantly reduce the predictive accuracy of the model. In a more recent study, Brainerd et al. (2013) revealed that $\varepsilon 4$ status is a risk factor for developing MCI but not for transitioning from MCI to AD. Brainerd et al. found that while $\varepsilon 4$ status increased risk of transitioning to MCI, only a proportion of those MCI cases progressed to AD. That $\varepsilon 4$ status was not a predictor of AD suggest that a significant proportion of the MCI cases were not at the early stage of AD (Brainerd et al., 2013). Farlow et al. (2004) suggest that knowing the $\varepsilon 4$ status of MCI individuals may be helpful in terms of indicating likely risk of future cognitive decline, but less helpful in cases where individuals were displaying early signs of AD.

Current Issues in MCI Research

Extensive variation among the literature in terms of progression to AD has resulted in scrutiny of the MCI concept (A. J. Mitchell & Shiri-Feshki, 2009). The discrepancies between MCI studies have largely been attributed to the way MCI criteria have been operationalised in clinical and research settings, e.g. sampling procedures (e.g. community versus clinical); the use of different cut off scores to establish impairment, and variation in
terms of the cognitive domains assessed and the neuropsychological measures used to make those assessments. (Panza et al., 2007; Winblad et al., 2004). However, such claims ignore fundamental issues with regard to MCI as a construct, such as the validity of the MCI concept and MCI subtypes; stability of the MCI syndrome; the lack of comprehensive testing in MCI studies; and the potential issue of circular reasoning in MCI research.

The Validity of the MCI Concept and MCI Subtypes

One of the major criticisms of the MCI concept relates to the validity of the classification or diagnosis of MCI and MCI subtypes (Summers & Saunders, 2012). As previously mentioned, the MCI concept was initially developed from a series of MAYO Clinic studies aimed at identifying risk factors for AD. Subsequent revisions to the MCI concept to include subtypes was a result of a growing body of evidence indicating heterogeneity of outcome rather than extensive clinical research identifying these subtype profiles as pre-dementia syndromes.

There have many post hoc attempts to validate MCI subtypes, however, differences across these studies complicate comparisons. Some studies use broad screening measures to infer the integrity of non-memory domains (e.g. attention, language function etc.), which may influence the accuracy of MCI subtype classification (Artero, Petersen, Touchon, & Ritchie, 2006; Busse et al., 2006; Fischer et al., 2007). In addition, many studies only examine amnestic variants of MCI (Kramer et al., 2006), which has resulted in considerably less understanding about the non-amnestic subtypes. Other research has translated statistically derived profiles of neuropsychological functioning in broadly defined MCI as supporting evidence for the validity of these subtypes (Delano-Wood et al., 2009; Libon et al., 2010). Overall, many MCI studies are epidemiological investigations of risk factors; studies of
prevalence rates; or investigations focused on dementia conversion. Very few studies represent genuine attempts to establish the validity of the MCI subtypes as discreet diagnostic entities.

The utility of sub-classifying MCI was perceived to be its ability to account for the heterogeneity of pre-dementia syndromes and to avoid the inconsistencies that have plagued MCI research (Panza et al., 2007). From the limited amount of research that has examined these newly proposed subtypes, it has been argued that amnestic variants, particularly multi-domain, are more likely to transition to AD, whereas non-amnestic variants may progress to non-AD type dementias (e.g. Fischer et al., 2007). However, this latter issue is somewhat contentious given that non-amnestic variants have received little empirical attention (Lopez et al., 2006). Petersen and Negash (2008) maintain that MCI subtypes are helpful in denoting the specific underlying disease process in MCI, particularly in cases of amnestic MCI. However, this remains questionable given the lack of evidence validating the presence of MCI subtypes as well as a growing body of evidence suggesting that MCI subtypes are inherently unstable in terms of trajectory and outcome (Han et al., 2012; Summers & Saunders, 2012).

The Stability of MCI

Studies continue to demonstrate a diverse range of outcomes associated with the MCI profile including stability; deterioration to AD; deterioration to other types of dementia; and return to age appropriate levels of functioning (Busse et al., 2006; Luis et al., 2004; N. L. J. Saunders & Summers, 2011; Summers & Saunders, 2012). Some researchers claim that variation in outcome is related to factors such as the source of sample recruitment (community sample
versus memory clinic sample) (Petersen & Bennett, 2005; Petersen & Morris, 2005). This argument is logically problematic. Clinic and population studies utilizing the same diagnostic criteria should reveal the same rates of conversion from an initial MCI diagnosis, despite any differences in the prevalence rates of MCI in their broader population sample. That is, while the prevalence/incidence of MCI is likely to be higher in a memory clinic sample than in a community sample; the rate of conversion from MCI to dementia must be similar across both samples if the diagnosis of MCI is clinically accurate. At present, of greatest concern to diagnostic validity are instances where individuals are seen to recover to age appropriate levels of cognitive functioning. Table 1 outlines some of the longitudinal studies tracking MCI over time and the rates at which MCI cases returned to healthy cognitive status at follow up. There is considerable variation in terms of the rates at which MCI cases revert to normal status over time. However, evidence of recovery of function does not align with the definition of MCI as a genuine prodrome of a neurodegenerative disease. Rather, recovery from initial symptoms suggests that transient factors (e.g. fatigue, anxiety etc) may be responsible for cognitive difficulties. However, if clinical diagnostic criteria demonstrate that significant proportions of individuals revert to healthy status, then major questions arise in terms of the sensitivity and specificity of these criteria.

In a recent study, Han et al. (2012) examined the longitudinal trajectory of individuals classified as one of the four MCI subtypes (single domain amnestic and non-amnestic; multiple domain amnestic and non-amnestic). Han and colleagues tracked 140 participants over 18 months to evaluate the stability of MCI over time as well as the predictive validity of each subtype. In terms of conversion, 19 cases transitioned to dementia at follow up. AD was the most common diagnosis (13) followed by vascular dementia (5) and frontotemporal dementia (1). Conversion was more frequent in individuals diagnosed with multiple domains
impaired (amnestic or non-amnestic) than single domains impaired (amnestic or non-amnestic) at baseline. However, reversion was more common in cases classified as single domain impaired (amnestic or non-amnestic). In total, 40 cases no longer met the criteria for an MCI classification at follow up. Between group instability was observed with approximately 20% of non-amnestic cases shifting to an amnestic subtype and vice versa. Despite a proportion of individuals converting, the between group instability and high rate of reversion lead Han and colleagues to conclude that MCI is a heterogeneous condition that is highly unstable.

Instability within the MCI condition is likely to have major consequences for early intervention programs. Specifically, if samples are enriched with individuals who will eventually return to normal levels of functioning, the true effectiveness of a treatment may be overestimated (J. Mitchell, Arnold, Dawson, Nestor, & Hodges, 2009). Further, the identification of accurate clinical diagnostic biomarkers will not be possible if MCI groups contain proportions of individuals who will eventually revert to healthy status.
Table 1.

**Testing Protocols and Rates of Reversion across Studies Tracking MCI**

<table>
<thead>
<tr>
<th>Study authors</th>
<th>Description of sample</th>
<th>Classification criteria</th>
<th>Classification assessment</th>
<th>Follow up assessment</th>
<th>Average follow up duration</th>
<th>Rates of reversion at follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed et al. (2008)</td>
<td>Clinical sample; mean age 68.6 years (SD=5.4); United Kingdom</td>
<td>Petersen criteria, 1.5SD below norms; amnestic MCI</td>
<td>Screening and Comprehensive (memory)</td>
<td>Comprehensive</td>
<td>12 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Albert et al. (2007)</td>
<td>Community sample; ≥65 years; United States</td>
<td>MCI=CDR of .05; did not specify subtypes</td>
<td>Screening</td>
<td>Comprehensive</td>
<td>36-48 months</td>
<td>30% of the original sample were excluded due to reversion</td>
</tr>
<tr>
<td>Alexopoulous et al. (2006)</td>
<td>Clinical sample; &gt;55 years; Germany</td>
<td>Petersen criteria; impairment level not specified; amnestic, non-annestic, and multiple domain MCI</td>
<td>Screening</td>
<td>Screening (same as baseline)</td>
<td>42 months</td>
<td>18%</td>
</tr>
<tr>
<td>Aretouli et al. (2013)</td>
<td>Clinical sample; &gt;55 years; United States</td>
<td>Petersen criteria; 1.5SD age and education norms; single and multiple, amnestic and non-annestic MCI</td>
<td>Comprehensive (memory)</td>
<td>Comprehensive (non-memory)</td>
<td>48 months</td>
<td>8%</td>
</tr>
<tr>
<td>Study authors</td>
<td>Description of sample</td>
<td>Classification criteria</td>
<td>Classification assessment</td>
<td>Follow up assessment</td>
<td>Average follow up duration</td>
<td>Rates of reversion at follow up</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------</td>
<td>-------------------------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Artero et al. (2008)</td>
<td>Community sample; ≥65 years; France</td>
<td>Adjusted Petersen criteria (subjective change in cognition, change, may have difficulty with ADL’s); lowest quartile for age and education norms; subtypes not specified</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>48 months</td>
<td>37%</td>
</tr>
<tr>
<td>Bennett et al. (2002)</td>
<td>Community sample; mean age 78.6 years (SD=6.8); United States</td>
<td>MCI defined by clinical judgement</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>53 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Bickel et al. (2006)</td>
<td>Clinical sample; 65-85 years; Germany</td>
<td>Winblad; impairment &lt;1.96SD below norm sample; all MCI subtypes</td>
<td>Screening</td>
<td>Screening</td>
<td>3.5 months</td>
<td>65.60%</td>
</tr>
<tr>
<td>Brodaty et al. (2012)</td>
<td>Community sample; 70-90 years; Australia</td>
<td>Winblad criteria, 1.5SD below age and education norms, all MCI subtypes</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>24 months</td>
<td>28.20%</td>
</tr>
<tr>
<td>Busse et al. (2006)</td>
<td>Community sample; ≥75 years; Germany</td>
<td>Petersen criteria; 1 &amp; 1.5SD below, all MCI subtypes</td>
<td>Screening</td>
<td>Screening (same as baseline)</td>
<td>52 months</td>
<td>20%</td>
</tr>
<tr>
<td>De Jager &amp; Budge (2005)</td>
<td>Community sample; 60-90 years; United Kingdom</td>
<td>1.5SD below mean for sample; amnestic MCI</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>48 months</td>
<td>13.40%</td>
</tr>
<tr>
<td>Study authors</td>
<td>Description of sample</td>
<td>Classification criteria</td>
<td>Classification assessment</td>
<td>Follow up assessment</td>
<td>Average follow up duration</td>
<td>Rates of reversion at follow up</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------</td>
<td>---------------------------------</td>
<td>----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>De Rotrou et al. (2005)</td>
<td>Clinical sample; mean age 70 years ($SD=4$); France</td>
<td>1.5$SD$ below normative sample; MCI subtypes not specified</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>12 months</td>
<td>48%</td>
</tr>
<tr>
<td>De Routrou et al. (2005)</td>
<td>Clinical sample; 70 years ($SD=4$); France</td>
<td>1.5$SD$ below norm for sample; MCI subtypes not specified</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>12 months</td>
<td>48%</td>
</tr>
<tr>
<td>Fischer et al. (2007)</td>
<td>Community sample; mean age 75.7 years ($SD=0.45$); Austria</td>
<td>1.5 $SD$ below age adjusted mean of sample; amnestic and non-amnestic MCI</td>
<td>Screening</td>
<td>Screening (same as baseline)</td>
<td>30 months</td>
<td>9.50%</td>
</tr>
<tr>
<td>Fleisher et al. (2007)</td>
<td>Clinical sample; 55-90 years; United States</td>
<td>Petersen criteria; 1.5-2$SD$ below education norms; amnestic MCI</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>36 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Forlenza et al. (2009)</td>
<td>Clinical sample; mean age 68.5 years ($SD=6.1$); Brazil</td>
<td>Petersen criteria; 1.5$SD$ below age and education norms; all MCI subtypes.</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>31 months</td>
<td>Amnestic 22.5%; non-amnestic 21%; 15% multiple domain</td>
</tr>
<tr>
<td>Ganguli et al. (2004)</td>
<td>Community sample; ≥65 years; United States</td>
<td>1.5$SD$ below mean for sample; amnestic MCI</td>
<td>Screening</td>
<td>Screening (same as baseline)</td>
<td>120 months</td>
<td>55%</td>
</tr>
<tr>
<td>Griffiths et al. (2006)</td>
<td>Clinical sample; MCI participants mean age 68.47 ($SD=8.65$); United States</td>
<td>Petersen criteria; 1.5$SD$ below age and education equivalent controls; at least 1 follow up diagnosis of amnestic MCI required</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>24 months</td>
<td>0%</td>
</tr>
<tr>
<td>Study authors</td>
<td>Description of sample</td>
<td>Classification criteria</td>
<td>Classification assessment</td>
<td>Follow up assessment</td>
<td>Average follow up duration</td>
<td>Rates of reversion at follow up</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>---------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Han et al. (2012)</td>
<td>Community sample; ≥65 year; Korean</td>
<td>Winblad criteria; &lt;1.5SD below age and education norms; all MCI subtypes</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>18 months</td>
<td>28.57%</td>
</tr>
<tr>
<td>Ishikawa &amp; Ikeda (2007)</td>
<td>Community sample; ≥65 years; Japan</td>
<td>Memory impairment = 0/3 or 1/3 on MMSE; amnestic MCI</td>
<td>Screening</td>
<td>Screening</td>
<td>60 months</td>
<td>38.50%</td>
</tr>
<tr>
<td>Kryscio et al. (2006)</td>
<td>Community sample; ≥60 years; United States</td>
<td>1.5SD below age; amnestic and multiple domain amnestic</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>12 months</td>
<td>52.50%</td>
</tr>
<tr>
<td>Loewenstein et al. (2007)</td>
<td>Clinical and community; mean age 77.66 years (SD=5.8); United States</td>
<td>1.5 SD below age and education norms, amnestic, non-amnestic and multiple domain MCI</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>12 months</td>
<td>amnestic 7.7%; non amnestic 16.7%; multiple domain 6.7%</td>
</tr>
<tr>
<td>Luis et al. (2004)</td>
<td>Clinical sample; mean age 72.1 (SD=7.9) Canada</td>
<td>Petersen criteria; cut off not specified; amnestic MCI</td>
<td>Screening</td>
<td>Screening (same as baseline)</td>
<td>28 months</td>
<td>4.50%</td>
</tr>
<tr>
<td>Mitchell et al. (2009)</td>
<td>Clinical sample; ≥ 50 years; United Kingdom</td>
<td>Petersen criteria; 1.28SD below sample controls; amnestic, non-amnestic and multiple domain MCI</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>24 months</td>
<td>amnestic 18%; non amnestic 70%; multiple domain 5%</td>
</tr>
<tr>
<td>Study authors</td>
<td>Description of sample</td>
<td>Classification criteria</td>
<td>Classification assessment</td>
<td>Follow up assessment</td>
<td>Average follow up duration</td>
<td>Rates of reversion at follow up</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td>-----------------------------</td>
<td>----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Perri et al. (2007)</td>
<td>Clinical sample; mean age 70.1 years (SD=7.2); Italy</td>
<td>Petersen criteria; impairment on three memory measures; amnestic MCI</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>24 months</td>
<td>only followed up those who converted or maintained MCI status</td>
</tr>
<tr>
<td>Petersen et al. (1999)</td>
<td>Community sample; MCI mean age 80.9 (SD=1); United States</td>
<td>Petersen criteria; 1.5SD below age and education norms; amnestic MCI</td>
<td>Comprehensive</td>
<td>Abbreviated battery not specified</td>
<td>120 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Rabin et al. (2009)</td>
<td>Community and clinical sample; ≥ 60 years; United States</td>
<td>Petersen criteria; 1.5SD below age, amnestic MCI</td>
<td>Comprehensive (memory)</td>
<td>Comprehensive (memory)</td>
<td>48 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Rasquin et al. (2005)</td>
<td>Clinical sample; ≥ 55 years; The Netherlands</td>
<td>No subjective memory impairment, 1.28SD (Z scores) below local norms, amnestic, non- amnestic, and multiple domain MCI</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>24 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Ravaglia et al. (2006)</td>
<td>Clinical sample; ≥ 60 years; Italy</td>
<td>Petersen criteria; 1.5SD below age and education norms; amnestic, non- amnestic, and multiple domain amnestic MCI</td>
<td>Comprehensive</td>
<td>Comprehensive (same as baseline)</td>
<td>36 months</td>
<td>4%</td>
</tr>
<tr>
<td>Study authors</td>
<td>Description of sample</td>
<td>Classification criteria</td>
<td>Classification assessment</td>
<td>Follow up assessment</td>
<td>Average follow up duration</td>
<td>Rates of reversion at follow up</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Ritchie et al. (2001)</td>
<td>Community sample; ≥ 60 years; France</td>
<td>Petersen criteria; 1.5SD below age and education norms amnestic MCI</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>12 months</td>
<td>92.6%</td>
</tr>
<tr>
<td>Sachdev et al. (2013)</td>
<td>Community sample; 70-90 years; Australia</td>
<td>Winblad criteria, 1.5SD below age and education norms, all MCI subtypes</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>24 months</td>
<td>29.5%</td>
</tr>
<tr>
<td>Summers &amp; Saunders (2012)</td>
<td>Community sample; ≥ 60 years; Australia</td>
<td>Petersen criteria; 1.28 SD below age and education norms; amnestic and non-amnestic MCI</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>20 months</td>
<td>24.7%</td>
</tr>
<tr>
<td>Tabert et al. (2006)</td>
<td>Clinical sample; ≥ 40 years; United States</td>
<td>Petersen criteria; 1.5 SD below age and education norms; all MCI subtypes</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>34.6 months</td>
<td>Not reported</td>
</tr>
<tr>
<td>Yaffe et al. (2006)</td>
<td>Clinical sample; mean age 72.9 ± 9.3 years; United States</td>
<td>Clinical judgement, amnestic and multiple domain MCI</td>
<td>Comprehensive</td>
<td>Comprehensive</td>
<td>36 months</td>
<td>7%</td>
</tr>
</tbody>
</table>

*Comprehensive = discreet measures of neuropsychological functions; Screening = broad screening measure*
Lack of Comprehensive Neuropsychological Assessment among MCI cohorts

As MCI criteria do not specify which neuropsychological measures should be utilised, there is considerable variability between studies in terms of the type of tests that are used, as well as the number of tests that are used. In some studies, researchers utilise measures with good reliability and validity, and adopt measures that are widely used in clinical neuropsychology. Others rely on brief screening measures (e.g. MMSE; ACE; CDR) to classify MCI cases as well as to track cognitive function over time. Further, some studies utilise reliable and valid neuropsychological measures to assess episodic memory function (e.g. list learning tasks, paragraph recall) but concurrently utilize broad screening measures to assess non-memory function (e.g. Alexopoulos et al., 2006; Artero et al., 2006; Brodaty et al., 2012; Busse et al., 2006; Fischer et al., 2007; Gavett et al., 2009; Jungwirth, Zehetmayer, Hinterberger, Tragl, & Fischer, 2012; Loewenstein, Acevedo, Agron, & Duara, 2007; Lonie, Herrmann, Donaghey, & Ebmeier, 2008). Table 1 outlines a number of longitudinal studies that have utilized comprehensive measures of discreet neuropsychological functions (comprehensive) at either classification or follow up stages, and those that have used broad screening measures (screening) at either classification or follow up stages.

The use of screening measures as indicators of domain specific function (e.g. memory, attention) may lead to inaccurate MCI classifications. Screening measures such as the MMSE have established sensitivity for detecting severity of dementia, however their reliability and validity for detecting subclinical (non-demented) impairment is questionable. The MMSE has been criticised for lacking adequate sensitivity to comprehensively assess non-memory function in MCI (Alladi et al., 2006). This criticism is supported by studies demonstrating that individuals with subclinical memory impairments frequently demonstrate additional non-
memory deficits when assessed on a comprehensive neuropsychological test battery utilising sensitive and reliable measures of non-memory domains (Belleville et al., 1996; Lonie et al., 2008; K. Ritchie & Touchon, 2000; N. L. J. Saunders & Summers, 2010). Kramer et al. (2006) found that more than 50% of their amnestic MCI sample performed at least one standard deviation lower than the control group on various non-memory tasks. Recently, Saunders and Summers (2010) found that 83% of their amnestic MCI participants displayed an impairment to either attention and/or working memory, suggesting that a majority of amnestic MCI individuals would be more appropriately classified as multiple domain amnestic MCI.

There are two major implications of these findings. First, if studies fail to comprehensively assess non-memory function, then subtype classifications may be inaccurate resulting in increased heterogeneity of subtype diagnosis and outcome. Second, conclusions regarding the trajectory of subtypes may also be inaccurate due to this increased heterogeneity of subtype diagnosis. Han et al. (2012) claim that studies failing to comprehensively assess non-memory function may result in multiple-domain amnestic cases being misclassified as single domain amnestic. This may have led to erroneous claims about the trajectory of the single domain amnestic subtype (Han et al., 2012). Further, early studies of single-domain amnestic MCI may be contaminated with cases of multiple-domain amnestic MCI and may have excluded individuals who would fulfil the criteria for single domain non-amnestic MCI (A. J. Mitchell & Shiri-Feshki, 2009). Recent recommendations suggest that MCI studies should include an assessment of memory and non-memory function (Albert et al., 2011), although insufficient detail is provided regarding the assessment of non-memory function. Given the growing body of evidence demonstrating enhanced sensitivity and specificity in MCI classification using comprehensive testing (Arnaiz & Almkvist, 2003; Bozoki, Giordani, Heidebrink, Berent, &
Foster, 2001; Summers & Saunders, 2012), MCI studies should prioritize extensive cognitive testing in assessment protocols.

**Circular Reasoning**

A potential issue that has yet to be formally investigated in MCI is that of circular reasoning. It is not uncommon for studies to identify MCI cases with a neuropsychological test battery that is identical to the test battery used to track cognitive change over time (see Table 1). Some claim that by using the same test battery there is a lack of independence between the MCI classification and follow-up neuropsychological assessments (Tuokko & Frerichs, 2000). This may lead to a self-fulfilling prophecy whereby individuals maintain their MCI status due to the circularity created by using the same test batteries (Burns & Zaudig, 2002; K. Ritchie & Touchon, 2000). Some researchers make attempts to avoid circularity by using separate test batteries to classify MCI and monitor cognitive changes over time. While this may reduce the potential impact of circular reasoning, the extent to which it presents a genuine concern for MCI researchers remains uncertain. However, if circularity is an issue, conclusions about the neuropsychological functioning of MCI cohorts, particularly in terms of predicting future cognitive decline, may be erroneous.
Aim of the Present Thesis and Outline of Chapters

The overall aim of the present research was to observe the longitudinal neuropsychological profile of MCI, particularly in terms of examining the validity of MCI subtypes as discreet clinical syndromes. A series of studies were designed utilising comprehensive neuropsychological assessment at each stage of testing. This was considered a priority given previous evidence indicating increased sensitivity and specificity of the MCI classification using comprehensive test protocols. In addition, the studies in this thesis were designed to reduce the potential issue of circular logic by using an experimental test battery that was predominantly distinct from the test battery used to identify MCI cases at screening.

The main aim of chapter three was to examine the group performances of each of the MCI subtypes after the initial screening assessment. It was considered necessary to cross check the performance of each subtype to ensure that group performances aligned with published MCI profiles. For example, evidence of reduced memory performance from the amnestic variants compared to non-amnestic variants would indicate accurate classification within these subtypes. Although this chapter presents an analysis of neuropsychological performance, the discussion about group differences was restricted to cross checking subtype profiles as any conclusions beyond this would be confounded by circularity. Chapter three includes all neuropsychological data from screening as well as demographic data from time 1 and time 2.

Chapter four represents an attempt to investigate the potential influence of circular reasoning in MCI. This was achieved by examining episodic memory performance at time 1 which utilised a different set of memory measures to those used at screening. Of primary interest was whether MCI groups would continue to demonstrate similar patterns of memory disturbance a time 1. This chapter presents data pertaining to the memory performance of
MCI subtypes at time 1 only. Chapter four as presented has been published in a peer reviewed journal:


The aim of chapter five was to investigate attention and working memory function of MCI subtypes at time 1. Given evidence that working memory and attention function may be impaired at the MCI stage, it was of interest to examine the function of MCI subtypes within these domains. Further, previous studies may be confounded by circularity and may have used broad screening measures to assess non-memory function at the classification stage. Chapter five includes demographic data from time 1 as well as data from the working memory and attention measures at time 1. Chapter five as presented has been published in a peer reviewed journal:


Chapter six is an investigation into the longitudinal neuropsychological profile of MCI subtypes. This chapter reveals the status of each neuropsychological domain implicated in the MCI profile. Specifically, this chapter discusses which domains are showing evidence of
stability; recovery; and decline among MCI subtypes over a 20 month period. The stability of each subtype is also examined with regards to the consistency of subtype profiles over time. This chapter includes subtype demographic information as well as data from all neuropsychological measures at time 1 and time 2. Chapter six as presented has been published in a peer reviewed journal:


The aim of chapter seven was to reclassify individuals according to stability over time and investigate the learning profile of these stable groups. As data from chapter six indicated instability across MCI subtype groups, it was of interest to examine what could be a feature of MCI subtypes with established longitudinal stability. After re-classifying MCI participants according to the stability of their subtype over time, time 2 list learning performance was analysed to reveal the learning profiles of stable MCI subtypes. Chapter seven provides learning trial data from each group at time 1 as well as demographic information. Chapter seven as presented has been published in a peer reviewed journal:

The main objective of chapter eight was to identify which domains of neuropsychological function would be able to differentiate between those who maintained MCI status over time versus those who did not. Participants were reclassified as either MCI or unimpaired according to the stability of their MCI classification across the 20 month time period. Performance across multiple cognitive measures was then analysed to uncover which domains were able to predict membership to MCI or unimpaired status. Chapter eight presents data from time 1 and group demographics including ApoE ε4 status. Chapter eight as presented has been accepted for publication with revisions:


Chapter nine presents an overview of each chapter and the key findings. This is followed by a general discussion and an examination of the limitations and implications of this body of work.
Chapter 2

General Methodology
Recruitment

Participants in this study were recruited via advertisements placed in local media (television and radio) as well as general medical practices seeking older adults (≥60 years) with subjective cognitive complaints (see Appendix A). Control participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment.

Pre-screening

A total of 247 participants responded to the advertising and were pre-screened via telephone interview for their suitability according to the following criteria:

(1) Subjective report of declining memory, attention/concentration, speech or language

(2) Preserved general cognition;

(3) Preserved activities of daily living;

(4) Subjective impairments and areas of preservation corroborated by an informant (e.g. partner, family member, or friend);

(5) Absence of any significant medical, neurological or psychological illness;

(6) Absence of any sensory impairment or major impairment to hand mobility.

Of the 286 respondents, 207 successfully completed telephone screening while 79 were excluded for various medical, neurological, and psychological reasons. In addition, seven participants pulled out of the study prior to screening assessment; physical illness (n =4); spouse illness (n =1); personal issues at the time of testing (n =2). A total of 200 participants were maintained at the pre-screening stage.
Screening

Participants

Of the 200 participants who completed screening assessment, 31 were excluded from data analysis and further assessment. Seventeen of those displayed clinical levels of anxiety (HADS A >8); one displayed clinical levels of depression and anxiety (HADS A and HADS D >8); one individual had a diagnosis of Autism Spectrum Disorder (not disclosed at telephone screening); one individual had a diagnosis of PTSD (not disclosed at telephone screening); two individuals had suffered a cerebrovascular accident just prior to their assessment; two individuals had a significant visual impairment that may have compromised their test performance; two individuals were ill at time of testing which may have compromised their test performance; one individual was experiencing chronic pain which may have compromised their test performance; one individual had recently performed in a different study with a similar test battery; and three were referred for further testing due to clinically significant cognitive deficits with a pattern suggestive of possible early-stage AD. A total of 169 participants (66 males & 103 females) were retained from the screening assessment.

The screening assessment was comprised of a brief interview as well as a comprehensive neuropsychological evaluation spanning multiple cognitive domains (see Table 2). The brief interview was primarily used to establish preservation of activities of daily living. Participants were asked whether they had any difficulties with everyday activities such as cooking, cleaning, shopping, driving etc. Although functional status was confirmed by an informant (e.g. friend or family member), this approach is limited by its self report nature.
The decision to adopt this approach was largely based on the fact that there is no consensus as to the most appropriate way to assess activities of daily living in MCI.

The neuropsychological assessment at screening was required so that participants could be classified according to published MCI criteria (Winblad et al., 2004). Neuropsychological cut-off scores were set at the 10th percentile (<1.28 SD) to ensure consistency with recent research (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012) and to avoid claims that typical cut-off scores of 1.5 SD are too severe (A. J. Mitchell & Shiri-Feshki, 2009). The current study adopted the following criteria to sub classify individuals into those with amnestic deficits (a-MCI); those with non-amnestic deficits (e.g. attention, language function; na-MCI & na-MCI+); and those with a combination of amnestic and non amnestic deficits (a-MCI+).

**Amnestic MCI (a-MCI)**

1. Subjective memory complaint at screening (confirmed by an informant);
2. Evidence of objective memory impairment (visual or verbal); as defined by performances <10th percentile for age based norms;
3. No evidence of an impairment to a non-memory cognitive domain (attention/executive function, working memory, or language processing); as defined by a performance >10th percentile for age based norms;
4. No evidence of dementia (DRS-2 AEMSS score ≥9); and
5. Preserved activities of daily living.
Multidomain amnestic MCI (a-MCI+)

1. Subjective memory complaint at screening (confirmed by an informant);
2. Evidence of one or more objective memory impairment (visual or verbal); as defined by performances <10th percentile for age based norms;
3. Evidence of one or more objective impairment to a non-memory cognitive domain (attention/executive function, working memory, or language processing); as defined by a performance <10th percentile for age based norms;
4. No evidence of dementia (DRS-2 AEMSS score ≥9); and
5. Preserved activities of daily living.

Non-amnestic MCI (na-MCI)

1. Subjective memory complaint at screening (confirmed by an informant);
2. No evidence of objective memory impairment (visual or verbal); as defined by performances >10th percentile for age based norms;
3. Objective evidence of a single impairment to a non-memory cognitive domain (attention/executive function, working memory, or language processing); as defined by a performance <10th percentile for age based norms; and
4. No evidence of dementia (DRS-2 AEMSS score ≥9); and
5. Preserved activities of daily living.

Multidomain non-amnestic MCI (na-MCI+)

1. Subjective memory complaint at screening (confirmed by an informant);
2. No evidence of objective memory impairment (visual or verbal); as defined by performances >10th percentile for age based norms;
3. Objective evidence of more than one impairment to a non-memory cognitive domain (attention/executive function, working memory, or language processing); as defined by a performance <10\textsuperscript{th} percentile for age based norms; and

4. No evidence of dementia (DRS-2 AEMSS score ≥9); and

5. Preserved activities of daily living.

**Control**

1. No history of subjective memory complaint;

2. No evidence of objective memory impairment (visual or verbal); as defined by performances >10\textsuperscript{th} percentile for age based norms;

3. No evidence of an impairment to a non-memory cognitive domain (attention/executive function, working memory, or language processing); as defined by a performance >10\textsuperscript{th} percentile for age based norms;

4. No evidence of dementia (DRS-2 AEMSS score ≥9); and

5. Preserved activities of daily living.

The 169 screening participants met the MCI sub classification criteria for the following groups: a-MCI (n=24); a-MCI+ (n=27); na-MCI (n=24); na-MCI+ (n=5); and control (n=89). Given the small number of participants who met the criteria for na-MCI+, it was not possible to separately analyse this subgroup, consequently they were collapsed into the na-MCI group for statistical analysis.
**Table 2.**

*Screening Neuropsychological Test Battery*

<table>
<thead>
<tr>
<th>Test</th>
<th>Abbrev.</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dementia Rating Scale</td>
<td>DRS-2</td>
<td>Dementia screen</td>
</tr>
<tr>
<td>Wechsler Test of Adult Reading</td>
<td>WTAR</td>
<td>Estimate of premorbid IQ</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale</td>
<td>HADS</td>
<td>Clinical depression/anxiety screen</td>
</tr>
<tr>
<td>Rey Complex Figure Test</td>
<td>RCOF</td>
<td>Immediate and 30 min delayed recall of visual memory</td>
</tr>
<tr>
<td>Logical Memory</td>
<td>LM†</td>
<td>Immediate and 30 min delayed recall of verbal information</td>
</tr>
<tr>
<td>Digit Span</td>
<td>DSP†</td>
<td>Verbal short term memory span</td>
</tr>
<tr>
<td>Spatial Span</td>
<td>SSP</td>
<td>Visual short term memory span</td>
</tr>
<tr>
<td>Letter Number Sequencing</td>
<td>LNS†</td>
<td>Verbal working memory capacity</td>
</tr>
<tr>
<td>Trail Making Test</td>
<td>TMT</td>
<td>Attentional control</td>
</tr>
<tr>
<td>Digit Symbol Coding</td>
<td>DSC†</td>
<td>Executive function</td>
</tr>
<tr>
<td>Stroop Test (Victoria version)</td>
<td>Stroop</td>
<td>Executive function</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>VOC†</td>
<td>Language</td>
</tr>
</tbody>
</table>

† = WAIS-III Subtest
Materials

Tests were administered to verify the absence of dementia (DRS-2); detect clinically significant anxiety and/or depression (HADS); and ensure all groups were relatively homogeneous with regard to premorbid intelligence (WTAR). The neuropsychological tests used at screening were chosen on the basis of good reliability and validity as well as clinical utility. It was also ensured that the test battery assessed all of the domains implicated in MCI and early AD e.g. visual and verbal episodic memory, attention, language functioning, working memory, and executive processing (Albert et al., 2011). While MCI criteria do not specify a set of tests considered to be most appropriate (Albert et al., 2011; Winblad et al., 2004) there is growing evidence that comprehensive test batteries spanning multiple cognitive domains have greater sensitivity and specificity in detecting MCI (Arnaiz & Almkvist, 2003; Bozoki et al., 2001; Summers & Saunders, 2012).

The Dementia Rating Scale, 2nd edition (Jurica, Leitten, & Mattis, 2001) is specifically designed to measure cognitive functioning in patients with dementia. The 38-item instrument has been found to be valid and effective in discriminating between levels of dementia severity as well as screening individuals for possible dementia (Johnson-Greene, 2004; Miller & Pliskin, 2006). The DRS-2 has established clinical utility and validity in diagnosing AD (Jurica et al., 2001). Interpretation of scores is based on adjusted age and education MOANS (Mayo Older American Normative Studies) scale scores (AEMSS). To be retained in the present study, participants needed to display a DRS-2 AEMSS score ≥9, which is above the cut off for clinical dementia and is consistent with MCI criteria of intact general cognition (Jurica et al., 2001).
The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) was designed to identify individuals who may be experiencing clinical levels of anxiety and/or depression. The HADS is a self rating questionnaire divided into an anxiety subscale, and a depression subscale. Each subscale contains seven questions (e.g. I feel tense or ‘wound up’, I look forward with enjoyment to things) and participants respond by circling the response that best fits how they have been feeling in the past week (e.g. most of the time, a lot of the time, from time to time, occasionally, not at all). In terms of the present study, the HADS was used as a flagging tool to identify individuals with high levels of anxiety and/or depression rather than as a diagnostic tool. It is well established that anxious and depressive symptomology can have a profound and detrimental effect on cognitive performance (Lezak, Howieson, & Loring, 2004). In a meta-analysis of 747 studies, Bjelland, Dahl, Haug and Neckelmann (2002) reported that the HADS performed consistently well in terms of its ability to detect the spectrum of symptoms in anxiety and depressive disorders in both clinical and non-clinical settings. Bjelland et al. concluded that while more comprehensive tests should be employed to aid diagnosis, the HADS is a very effective screening tool to highlight individuals with probable or possible anxiety or depression.

The Wechsler Test of Adult Reading (WTAR, 2001) is an estimate of premorbid intellectual functioning in adults. The WTAR was designed to be relatively insensitive to neurological deterioration in order to estimate intellectual functioning prior to cognitive decline (Mathias, Bowden, & Barrett-Woodgridge, 2007). The WTAR has excellent reliability and validity and is currently the only premorbid IQ assessment to be co-normed with the WAIS-III (Wechsler, 1997a) making it the preferred tool for assessing premorbid intellectual ability (Mathias et al., 2007; WTAR, 2001).
Verbal Episodic Memory

Logical Memory (LM; Wechsler, 1997b) is a subtest of the Wechsler Memory Scale and assesses immediate and delayed verbal recall. For the immediate recall trial, participants are read a short story (story A) after which they are required to verbally recall as much of the story as possible. They are then read a different story (story B) and asked to recall as much of that story as possible. Story B is then repeated and again the participant asked to recall what they can, the expectation being that the participant will remember a greater amount of detail having heard the story a second time. After the immediate trial, participants are given a verbal warning that they will be assessed on their delayed recall later in the testing session. Approximately 30 minutes later, participants are tested on their ability to freely recall elements of story A and B. At both trials, participants receive a point for each story unit and or each thematic unit they recall. For example, a participant would receive one thematic point for recalling that Story B was about a man, but if they remembered that it was about a man whose name was David then they would receive one point for a story unit as well as one point for a thematic unit. As the thematic information is more general than the story units, thematic scores were not used in the following study. The percentage of each participant’s retention was calculated by dividing the delayed recall total with the immediate recall total. A low score on the immediate recall trials suggests some short term memory difficulty while a low score on the delayed recall trials suggests a weakness in episodic memory and learning. The logical memory subtest is one of the most frequently utilised tests of episodic memory in MCI and AD populations.
**Visual Episodic Memory**

The Rey Complex Figure Test (RCFT; Lezak et al., 2004) is an assessment of visual memory and visuo-spatial construction. Participants are shown a large figure comprising of 18 shapes and elements and instructed to copy the figure taking care to be as accurate as possible. Approximately five minutes later, the participant is asked to recall the figure from memory without warning. A maximum of 5 minutes is allowed to complete the figure copy but there is no time limit for the recall trial. There are both quantitative and qualitative approaches to scoring the RCFT. The original Osterrieth (1944, as cited in Lezak et al., 2004) method uses a 36-point system where participants receive points for accuracy and location. Through this system, participants score 2 points for each element that is both drawn and positioned correctly; 1 point for each element that is drawn correctly but misplaced; 1 point for each element that is recognisable although distorted or incomplete but is placed in the correct position; and half of a point for each element that is recognisable although distorted or incomplete and misplaced. A maximum of 36 points can be achieved on both the copy and recall trials. Retention score can be calculated to mediate the effect of copy trial errors on recall [(recall/copy) x 100 = % retention]. As the role of the RCFT in the present study was to highlight visual recall function, the Osterrieth system as mentioned above was used instead of other qualitative systems that focus on spatial organisation and picture integrity. The RCFT has a long clinical and experimental history demonstrating sensitivity to a range of dementia syndromes including AD, as well medial temporal lobe damage, the main brain region implicated in AD pathology (Berry, Allen, & Schmitt, 1991).
Working Memory

The *Letter-Number Sequencing* (LNS; Wechsler, 1997a) test is a subtest of the Wechsler Adult Intelligence Scale, 3rd edition. LNS requires participants to repeat a random string of verbally presented letters and numbers in a specific order (R. G. Morris, 1994). Numbers in ascending order must be recalled first followed by the letters in alphabetical order. Letter-number sequencing is largely an assessment of the central executive component of working memory as participants are required to undertake considerable information manipulation prior to recall (Cherry et al., 1996).

Short Term Memory

*Digit Span* (DSP; Wechsler, 1997a) is a subtest of the Wechsler Adult Intelligence Scale, 3rd edition. Digit span requires participants to repeat a sequence of verbally presented numbers in the same order originally presented (forward) and in the reverse order originally presented (backward). This test is used to assess forward and backward span lengths (Cherry et al., 1996; Lezak et al., 2004). Evidence suggests that verbal short term span may be compromised in early AD and MCI (Belleville et al., 2007).

The *Spatial Span* (SSP; Wechsler, 1997b) task is a subtest of the Wechsler Memory Scale, 3rd edition and is an assessment of visual short term memory capacity. It consists of a forward trial that measures spatial capacity, and a backwards trial that measures visuospatial sketchpad function as well as central executive integrity. Using a three dimensional board with randomly spaced square blocks attached, the experimenter uses their finger to tap a
specific sequence of blocks at one second time intervals. For the forward trial, the participant is then required to tap the blocks in the same presentation order. For the backwards trial, the participant is required to tap the sequence in reverse order. There are two trials for each span length for forward SSP and backward SSP, beginning at a span of two blocks and finishing at a span of nine blocks. Each trial begins with the sequence of two blocks and is continued until the participant is unable to successfully complete two attempts at any span length. One point is awarded for each correct sequence. Evidence suggests that visual short term memory span is compromised in early AD (Carlesimo, Fadda, Lorusso, & Caltagirone, 1994; Huntley & Howard, 2010) and may also be impaired at the MCI stage (N. L. J. Saunders & Summers, 2011).

Attention

The Trail Making Test (TMT) (Lezak et al., 2004) is a dual trial assessment of divided attention and speed of processing. The first trial (Trails A) requires participants to use a pencil to sequentially join circled numbers (1-25) that are randomly distributed on an A4 page. The second trial (Trails B) requires participants to use a pencil to alternately join circled numbers (1-13) and letters (A-L) that are randomly distributed on an A4 page. A practice trial is administered before each test trial to ensure that the participant is able to complete the task. Trails A is administered prior to Trails B. Participants are scored on the amount of time it takes to complete each test trial. If a mistake is made (e.g. a number is missed in the sequence), the participant is asked to correct their mistake before continuing resulting in an increase in time taken to complete the task. Although both Trails A and Trails B make demands of perceptual tracking and processing speed, Trails B makes demands on
divided attention as the participant is required to process two streams of information simultaneously. The TMT test has a long clinical history and is one of the five most commonly used tests for neuropsychological assessment (Rabin, Barr, & Burton, 2005). Further, it has been found to be sensitive to a range of neurological disorders and is recommended as a useful assessment for neurologically challenged populations (Strauss, Sherman, & Spreen, 2006).

**Executive Function**

*Digit Symbol Coding* (DSC; Wechsler, 1997a) is a psychomotor task tapping selective and sustained attention as well as incidental learning. It is a speed based task that involves copying symbols that represent numbers. During the task, participants refer to a key that displays the numbers one to nine and their respective symbol. While they are not required to memorise the symbols or the numbers that they represent, it is common for some incidental learning to take place. Participants are told to complete the task as quickly as possible and are given a single point for each correct symbol they draw in 120 seconds. Recent evidence suggests that selective and sustained attention may be compromised in MCI (Rapp & Reischies, 2005; N. L. J. Saunders & Summers, 2010, 2011).

*The Stroop Test* (Stroop; Lezak et al., 2004) is a measure of impulse control and inhibition of prepotent response. The 24-item Victoria version consists of three stages, each of which requires the participant to read out aloud the colour of the items on a single stimulus card. For the first stage, participants are presented with a card containing groups of dots that are coloured red, yellow, green, or blue. Each group consists of three dots and there are exactly 6
groups of each colour pseudo randomly distributed in a four by six table. From the left to right, the participant must read aloud the colour of each group of dots as fast as they can. If a colour is read incorrectly, the participant is required to correct the mistake before continuing as quickly as possible. The second stage involves the naming of the colour of neutral words (e.g. ‘when’, ‘and’) printed in red, yellow, green or blue. The participant is required to read the colour of each neutral word out aloud as quickly as possible. The third stage involves the naming of colour words (e.g. ‘blue’, ‘red’) that are printed in an incongruent colour (e.g. ‘blue’, ‘red’). At this stage, the participant has to read the colour of the word by relying on the visual content of each word whilst ignoring the verbal content. At each stage, time taken to complete the task is recorded as well as the number of mistakes made. The level of interference is calculated by finding the ratio of the colour incongruent trial with the dot trial. The interference score allows the examination of a participant’s performance whilst controlling for processing speed. The Victoria version of the Stroop task was chosen due its shorter administration time compared to other Stroop versions. It was also chosen because difficulties with response inhibition using longer versions are often evident within the first 40 items (Strauss et al., 2006). The shorter task time of the Victoria Version also limits the amount of task practice and may be more ideal at detecting deficits to response inhibition (Strauss et al., 2006). Compared to healthy adults, both AD and MCI populations perform poorly on the stroop task indicating difficulty with response inhibition in these populations (Bélanger, Belleville, & Gauthier, 2010).
Language function

The *Vocabulary* subtest (Voc; Wechsler, 1997a) requires participants to present oral definitions of various words (e.g. ‘Breakfast’, ‘Designate’, ‘Evolve’). Responses are rated in terms of their quality with a maximum of two points for each correct response; one point is awarded for responses that are close to correct; and zero points are awarded for an incorrect or colloquial response. There are a total of 30 words in the Vocabulary list making a maximum of 60 points possible. The test is administered until the end of the list is complete or until the participant receives six consecutive scores of zero. Each word is presented in a booklet and read out aloud by the experimenter. Recent recommendations for MCI criteria suggest that language function should be assessed in potential MCI cases (Albert et al., 2011).

Procedure

Prior to the screening stage, each participant provided informed consent in alignment with the Human Research Ethics Committee (Tasmania) Network and NHMRC Human Research Guidelines. Individual testing sessions were held at the School of Psychology building at the Launceston campus of the University of Tasmania. Informed consent was obtained in writing prior to testing. Each test was conducted in a well lit, well ventilated room and took approximately 1.5-2 hours including rest breaks. The screening test battery was administered in the following order: HADS, WTAR, DRS, DSC, LMI, RCOF (copy), DSP, SSP, RCOF (recall), Trails A & B, LMII, Stroop, and VOC. In order to limit negative practice effects, participants were given a 10 minute break during the middle of the testing sequence.
Following testing, each participant received written feedback on their performance in relation to validated normative data.

**Experimental stages**

Following screening, two longitudinal assessments (time 1 and time 2) were undertaken to track the progression of MCI subtypes. The objective of time 1 and time 2 was to examine subtype performance within the domains of episodic memory, short term memory, working memory, attention/executive control, and language functioning. Some researchers have criticized longitudinal studies that utilise the same tests to classify MCI as well as evaluate cognitive performance (Bangen et al., 2010; de Jager & Budge, 2005; L. J. Ritchie & Tuokko, 2010; Tuokko & Frerichs, 2000). The current study attempted to avoid the argument of circular logic in classification by using a test battery for follow up assessment that was predominantly different from the battery used at the screening stage. All of the tests were chosen on the basis of good reliability and validity as well as their clinical utility and experimental history. Where possible, tests were also chosen on their equivalency with screening tests.
Participants

Due to the high number of healthy controls at screening ($n=89$), 34 participants were excluded at time 1 as they were surplus to requirements. An additional five control participants did not wish to continue with the study due to personal reasons. The 50 remaining control participants were chosen on the basis of matching to the age and gender of MCI participants. In addition, two participants were excluded due to clinical levels of anxiety (HADS A $>8$); one participant relocated interstate; one joined an alternative study with a competing test battery; and four withdrew due to chronic health issues. In total, the data of 122 (48 male, 74 female) participants were included in the time 1 analysis: a-MCI ($n=23$); na-MCI ($n=25$); a-MCI+ ($n=24$); control ($n=50$). At time 2 assessment, four participants had withdrawn from the study (1 deceased; 2 health reasons; and 1 relocated interstate). A total of 118 participants (46 male and 72 female) were included in the time 2 analysis: a-MCI ($n=23$); na-MCI ($n=24$); a-MCI+ ($n=22$); control ($n=49$).

Materials

As with the screening stage, the experimental stage aimed to tap domains that are considered sensitive to the early cognitive changes in MCI (e.g. attention, episodic memory, working memory, semantic memory, and executive function). Whilst a comprehensive protocol was used, other domains (e.g. motor function, praxes) were not assessed. This decision was made for practical reasons related to the scale of the study, and also to avoid having an unnecessarily long test battery for older adults with mild cognitive difficulties. Consequently, the profiling within the present study is limited to the chosen cognitive domains and tests.
Further, a majority of the tests used in the experimental stage were taken from the Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition Ltd, 2011). The CANTAB is a computerised tool that consists of six groups of tests: screening tests; visual memory tests; executive function, working memory and planning tests; attention tests; semantic/verbal memory tests; and decision-making and response control tests. The decision to adopt several CANTAB subtests was based on a number of reasons: Firstly, the CANTAB has the advantage of being an easy to administer, non-verbal tool that gives immediate feedback, with some sub tests allowing repeat testing situations via parallel forms. In addition, the clinical validity and utility of the CANTAB has been well established, not only in detecting deficits in early AD but also in distinguishing between different neurodegenerative conditions (Robbins et al., 1998; Sahakian & Owen, 1992). The tests chosen for the experimental phase are presented in Table 3 and discussed in further detail below.
### Table 3.
*Experimental Psychological and Neuropsychological Test Battery*

<table>
<thead>
<tr>
<th>Test</th>
<th>Abbrev.</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dementia Rating Scale</td>
<td>DRS-2</td>
<td>Dementia screen</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale</td>
<td>HADS</td>
<td>Clinical depression/anxiety</td>
</tr>
<tr>
<td>Rey Auditory Verbal Learning test</td>
<td>RAVLT</td>
<td>Learning, immediate and delayed recall of verbal memory</td>
</tr>
<tr>
<td>Paired Associates Learning</td>
<td>PAL†</td>
<td>Learning and recall visual information</td>
</tr>
<tr>
<td>Reaction time</td>
<td>RTI†</td>
<td>Simple sustained and divided attention</td>
</tr>
<tr>
<td>Match to samples</td>
<td>MTS†</td>
<td>Selective attention</td>
</tr>
<tr>
<td>Rapid visual processing</td>
<td>RVP†</td>
<td>Complex sustained attention</td>
</tr>
<tr>
<td>Spatial Span</td>
<td>SSP†</td>
<td>Visual short term memory span</td>
</tr>
<tr>
<td>Digit Span</td>
<td>DSP*</td>
<td>Verbal short term memory span</td>
</tr>
<tr>
<td>Spatial working memory</td>
<td>SWM†</td>
<td>Visual working memory capacity</td>
</tr>
<tr>
<td>Letter Number Sequencing</td>
<td>LNS*</td>
<td>Verbal working memory capacity</td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td>BNT</td>
<td>Capacity to name common and uncommon objects</td>
</tr>
</tbody>
</table>

† = CANTAB subtest; *WAIS-III subtest
Episodic Memory

*The Rey Auditory Verbal Learning Test* (RAVLT; Strauss et al., 2006) is a verbal assessment of episodic memory and learning. The RAVLT consists of five consecutive learning trials of an auditory presentation of 15 item word list (list A). Following each learning trial, participants are required to recall as many of the 15 words in any order. After the fifth learning trial, a distracter list of 15 new words (list B) is presented followed by a recall trial. Following recall of the list B words, the participant is then required to recall as many words as possible from list A. Finally, recognition of list A words is assessed using a printed array of 50 words in which the original 15 list A words are embedded. The number of words recalled correctly is recorded for each learning and recall trial. The RAVLT has been found to be reliable in distinguishing between healthy controls and individuals with AD, as well as differentiating between various neurodegenerative conditions (Tierney et al., 1994).

*The Paired Associates Learning test* (PAL) (Cambridge Cognition Ltd, 2011) is a subtest of the Cambridge Neuropsychological Test Automated Battery (CANTAB). The PAL is a visual measure of episodic memory and learning and is sensitive to medial temporal lobe function (Cambridge Cognition Ltd, 2011). The PAL has demonstrated an ability to accurately discriminate between individuals with AD and healthy controls as well as the capacity to predict future cognitive decline (Swainson et al., 2001). During the PAL, participants are presented with six white boxes that open up one at a time in random order. At trial one, the computer reveals one pattern hidden in one of the six boxes (see Figure 2). At the end of the presentation sequence, the pattern appears in the centre of the screen and the participant is required to recall the location of the pattern by touching the appropriate box. Correct detection of each pattern within the allocated ten attempts allows the participant to move on
to additional trials where two, three, six, and eight patterns are hidden respectively. Failure to recall the correct location of each pattern after 10 attempts at any trial level results in termination of the test.

![Display screen for Paired Associates Learning (PAL).](image)

**Figure 2.** Display screen for Paired Associates Learning (PAL).

**Attention**

*Reaction Time* (RTI; Cambridge Cognition Ltd, 2011) is a subtest of the CANTAB that assesses simple sustained attention (SRTI) and divided attention (CRTI). In SRTI, participants are required to tap the centre of the touch screen where a yellow spot has appeared (see *Figure 3*). Similarly, in CRTI participants are required to touch the screen when they see a yellow spot in one of five circles (see *Figure 4*). The assessment stage of both SRTI and CRTI uses a response pad which records RT in milliseconds. Participants must hold down the right button and release it when they see the yellow spot appear inside the single circle (SRTI) or in one of the five circles (CRTI). Upon releasing their finger,
participants must immediately touch the appropriate circle on the screen with the same hand. The practice and assessment blocks for STRI and CRTI involve 18 and 40 trials respectively.

**Figure 3.** Display screen for Simple Reaction Time (SRTI)

**Figure 4.** Display screen for Choice Reaction Time (CRTI).

*Matching to Sample* (MTS; Cambridge Cognition Ltd, 2011) is a subtest of the CANTAB that taps selective attention by assessing speed and accuracy in matching patterns during a
visual search paradigm. Upon holding down the right response button, a central red box opens to reveal a pattern. Two seconds later, 1, 2, 4, or 8 patterns are revealed in the white boxes located around the red box (see Figure 5). The participant must match the stimulus pattern in the red box to one of the choice stimuli patterns in the white boxes by releasing their finger from the button pad and touching the correct pattern. If the chosen pattern is incorrect, a red cross appears over it with the word ‘WRONG’ to indicate the participant has chosen an incorrect stimulus match. Participants are then prompted by the experimenter to make another choice. There are 12 trials at each number of choice stimuli (1, 2, 4, and 8), with trial blocks appearing in a pseudo random order.

Figure 5. Display screen for Matching to Sample (MTS)

Rapid Visual Information Processing (RVP; Cambridge Cognition Ltd, 2011) is a subtest of the CANTAB that assess sustained attention. During RVP, participants watch a white box in the centre of the screen that presents single digits (from 2 to 9) pseudo randomly in a continual stream at 100 digits per minute (see Figure 6). In the first trial, participants are
directed to respond with the press pad when they detect the sequence, 3-5-7. In the second trial, participants are directed to respond with the press pad when they detect the appearance of the following sequences: 3-5-7; 2-4-6; and 4-6-8. During both trials, no feedback is given with regards to accuracy. Target sequences are presented at a rate of 16 per two minutes.

**Figure 6.** Display screen for Rapid Visual Information Processing (RVP).

**Short term memory**

*Spatial Span* (SSP; Cambridge Cognition Ltd, 2011) is a subtest of the CANTAB which is a computerised equivalent of the Corsi Block task, assessing visuospatial short term memory (Lezak et al., 2004). During this test, nine white boxes appear on the screen, several of which change colour, one at a time and in random order (see *Figure 7*). Participants are required to touch the boxes in the same order as originally presented by the computer. The test begins with a sequence of two boxes changing colour which increases by one additional box as
participants pick the correct sequence. The test is discontinued when participants fail three trials at any level.

![Spatial Span Display Screen](image)

**Figure 7.** Display screen for Spatial Span (SSP)

The Digit Span (DSP; Wechsler, 1997a) from the WAIS-III was chosen as an additional short term memory task in both the screening and experimental test batteries.

**Working memory**

Spatial Working Memory (SWM; Cambridge Cognition Ltd, 2011) is a CANTAB subtest used to assess working memory capacity and strategy, indicating central executive functioning (Chase, Clark, Sahakian, Bullmore, & Robbins, 2008). This test places constant demand on participant’s ability to manipulate and revise their working memory capacity and is subsequently sensitive to frontal (but not temporal) lobe impairment (Cambridge Cognition
Ldt, 2011). In SWM, participants are required to locate a number of blue tokens hidden inside an array of boxes on the computer screen (see Figure 8). Only one token is hidden at any one time and the computer only uses each box once to hide a token. Participants must subsequently avoid returning to boxes that have previously produced tokens. Participants are required to complete four separate trials of four levels of difficulty involving three, four, six, and eight tokens respectively. Both colour and position of boxes is altered on each trial in order to avoid the use of the same sequence. Two performance measures were selected from this test: SWM strategy indicates whether the individual adopts an effective systematic search pattern that is repeated at each level of the task (Chase et al., 2008); and SWM total errors denotes the number of boxes that were opened despite yielding a blue token on a previous attempt.

![Figure 8. Display screen for Spatial Working Memory (SWM)](image)

*Figure 8.* Display screen for Spatial Working Memory (SWM)

*Letter-Number Sequencing* (LNS; Wechsler, 1997a) from the WAIS-III was chosen as an additional working memory task in both the screening and experimental test batteries.
Language

The *Boston Naming Test* (Kaplan, Goodglass, & Weintraub, 1983) is a standardised test measuring language retrieval from long term semantic memory. The BNT consists of 60 line drawings of familiar and unfamiliar objects, which require correct initial identification for a positive score (Lezak et al., 2004). The BNT has been found to detect mild deficits to naming and word retrieval in neurodegenerative disorders such as AD (Barker-Collo, 2007) and is often used in MCI research. The popularity of the BNT as an assessment of verbal function is a testament to its excellent reliability and validity (Strauss et al., 2006).

Procedure

Individual testing sessions were held at the School of Psychology building at the Launceston campus of the University of Tasmania. Informed consent was obtained in writing prior to testing. Each test was conducted in a well lit, well ventilated room and took approximately 1.5-2 hours including rest breaks. At time 1, the test battery was administered in the following order: DRS-2; BNT; RTI; DSP; SSP; LNS; PAL; HADS; RAVLT; RVP; SWM; MTS. At time 2, the test battery was administered in the following order: DRS-2; PAL; RAVLT; SRTI; CRTI; DSP; RVP; LNS; MTS; SSP; HADS; SWM; BNT. To limit negative practice effects, participants were given a 10 minute break during the middle of the testing sequence. The CANTAB was administered on an IBM laptop connected to a 19-inch touch screen and response pad. Participants sat approximately 50cm from the touch screen with the response pad positioned 15cm from the touch screen. Following testing, each participant was given written feedback outlying their performance in relation to validated normative data.
Chapter 3

Screening Assessment
The main aim for the screening assessment was to classify individuals according to current published criteria (Winblad et al., 2004). To ensure that the classification procedure resulted in groups that aligned conceptually with MCI subtypes, group performances across cognitive domains were examined. For example, amnestic subtypes should demonstrate poorer performance on tests of memory function compared to non-amnestic variants as the amnestic subtype is defined by the presence of impaired memory function. Any interpretation of group performances beyond confirmation of subtype would be circular in logic and therefore erroneous. Outcome measures for this analysis were total scores (raw) e.g. items correct, time. On tests with multiple outcome measures (e.g. LM, stroop), outcome measures were chosen on the basis of their sensitivity according to previous research.

**Statistical analysis**

All statistical analyses were completed using SPSS for Windows (Version 19.0). As the tests used in the present study are sensitive to individual differences in demographic (e.g. gender, age, IQ) (Strauss et al., 2006), a series of preliminary one way ANOVAs were completed to examine any differences on these variables. MANOVA was used to control for the inflated Type 1 error rate that occurs when assessing similar cognitive domains. Significant multivariate results were followed with one-way ANOVAs and Games-Howell procedure due to unequal sample sizes (Howell, 2002).
Results

Demographic data is presented in Table 4. There were no significant differences between the groups in terms of age; depression score (HADS D); or number of years of formal education. The a-MCI+ group demonstrated a significantly lower estimated premorbid IQ than all other groups. Despite these differences, each group recorded an average estimated premorbid IQ. The a-MCI+ group also recorded a higher mean anxiety score (HADS A) than the na-MCI and control groups. Despite this difference, anxiety levels were below the clinically significant threshold (>8). The control group also scored significantly higher on the dementia screening scale (DRS-2) compared to all other groups. It is important to note that none of the groups had a mean DRS-2 score of clinical significance (all ≥9). A 2 x 4 chi square revealed no significant differences in gender ratio between the four groups (Χ² (3) = 2.58, p =.461).
Table 4.

*Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores at Screening, Time 1, and Time 2.*

<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI</th>
<th>na-MCI</th>
<th>a-MCI+</th>
<th>Control</th>
<th>p.</th>
<th>Post-hoc (at p &lt; .05)</th>
<th>Effect size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>Screening</td>
<td>25</td>
<td>30</td>
<td>26</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time 1</td>
<td>23</td>
<td>26</td>
<td>23</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>22</td>
<td>25</td>
<td>22</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Screening</td>
<td>69.68 (7.63)</td>
<td>69.30 (5.86)</td>
<td>69.15 (6.50)</td>
<td>70.84 (6.37)</td>
<td>.525</td>
<td>.013</td>
<td>.208</td>
</tr>
<tr>
<td></td>
<td>Time 1</td>
<td>70.61 (7.99)</td>
<td>70.58 (5.97)</td>
<td>69.26 (6.56)</td>
<td>72.66 (6.52)</td>
<td>.201</td>
<td>.038</td>
<td>.404</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>69.82 (8.09)</td>
<td>69.36 (5.50)</td>
<td>68.82 (6.75)</td>
<td>71.82 (6.38)</td>
<td>.240</td>
<td>.036</td>
<td>.369</td>
</tr>
<tr>
<td>Education</td>
<td>Screening</td>
<td>14.28 (3.16)</td>
<td>14.87 (3.38)</td>
<td>12.38 (3.43)</td>
<td>13.94 (3.50)</td>
<td>.054</td>
<td>.045</td>
<td>.631</td>
</tr>
<tr>
<td></td>
<td>Time 1</td>
<td>14.43 (3.16)</td>
<td>14.92 (3.52)</td>
<td>12.48 (3.53)</td>
<td>14.20 (3.74)</td>
<td>.098</td>
<td>.052</td>
<td>.535</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>14.64 (3.08)</td>
<td>14.72 (3.43)</td>
<td>12.32 (3.52)</td>
<td>14.06 (3.64)</td>
<td>.079</td>
<td>.058</td>
<td>.571</td>
</tr>
<tr>
<td>Measure</td>
<td></td>
<td>a-MCI</td>
<td>na-MCI</td>
<td>a-MCI+</td>
<td>Control</td>
<td>p.</td>
<td>Post-hoc (at p.&lt;.05)</td>
<td>Effect size</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
<td>----------------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>WTAR (est. FSIQ)</strong></td>
<td>Screening</td>
<td>110.24 (5.29)</td>
<td>109.43 (5.75)</td>
<td>103.00 (7.75)</td>
<td>109.74 (6.29)</td>
<td>&lt;.001</td>
<td>a-MCI+&lt;All</td>
<td>.134</td>
</tr>
<tr>
<td></td>
<td>Time 1</td>
<td>110.22 (5.42)</td>
<td>109.88 (5.20)</td>
<td>103.43 (8.15)</td>
<td>110.38 (5.76)</td>
<td>&lt;.001</td>
<td>a-MCI+&lt;All</td>
<td>.163</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>110.27 (5.54)</td>
<td>109.72 (5.23)</td>
<td>103.55 (8.33)</td>
<td>110.67 (5.43)</td>
<td>&lt;.001</td>
<td>A+&lt;A,NA,C</td>
<td>.168</td>
</tr>
<tr>
<td><strong>DRS-2 (AEMSS)</strong></td>
<td>Screening</td>
<td>10.44 (1.78)</td>
<td>10.40 (2.24)</td>
<td>9.27 (1.87)</td>
<td>11.72 (2.30)</td>
<td>&lt;.001</td>
<td>Control&gt; All</td>
<td>.154</td>
</tr>
<tr>
<td></td>
<td>Time 1</td>
<td>10.91 (2.17)</td>
<td>11.58 (1.86)</td>
<td>9.96 (1.97)</td>
<td>12.54 (2.14)</td>
<td>&lt;.001</td>
<td>a-MCI+&lt;na-MCI, C; a-MCI &lt; C</td>
<td>.189</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>10.77 (2.56)</td>
<td>11.52 (2.18)</td>
<td>10.233 (2.11)</td>
<td>12.55 (2.08)</td>
<td>.994</td>
<td></td>
<td>.003</td>
</tr>
<tr>
<td><strong>HADS A</strong></td>
<td>Screening</td>
<td>4.92 (3.01)</td>
<td>4.77 (2.00)</td>
<td>6.65 (2.83)</td>
<td>4.35 (2.78)</td>
<td>.003</td>
<td>a-MCI+&gt;na-MCI, Control</td>
<td>.081</td>
</tr>
<tr>
<td></td>
<td>Time 1</td>
<td>5.39 (2.79)</td>
<td>5.08 (2.65)</td>
<td>6.87 (3.79)</td>
<td>4.72 (2.56)</td>
<td>.034</td>
<td>Insufficient power</td>
<td>.071</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>6.00 (2.64)</td>
<td>4.76 (2.55)</td>
<td>6.45 (3.41)</td>
<td>4.88 (2.99)</td>
<td>.317</td>
<td></td>
<td>.030</td>
</tr>
<tr>
<td>Measure</td>
<td>a-MCI (Mean (SD))</td>
<td>na-MCI (Mean (SD))</td>
<td>a-MCI+ (Mean (SD))</td>
<td>Control (Mean (SD))</td>
<td>p.</td>
<td>Post-hoc (at p.&lt;.05)</td>
<td>Effect size (n_p^2)</td>
<td>Power</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-----</td>
<td>----------------------</td>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>HADS D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td>2.48 (2.63)</td>
<td>2.60 (1.71)</td>
<td>2.65 (1.77)</td>
<td>2.51 (1.91)</td>
<td>.985</td>
<td></td>
<td>.001</td>
<td>.059</td>
</tr>
<tr>
<td>Time 1</td>
<td>3.04 (2.35)</td>
<td>2.46 (2.01)</td>
<td>3.48 (2.47)</td>
<td>2.72 (2.33)</td>
<td>.429</td>
<td></td>
<td>.023</td>
<td>.249</td>
</tr>
<tr>
<td>Time 2</td>
<td>3.64 (2.80)</td>
<td>2.24 (1.48)</td>
<td>3.32 (2.26)</td>
<td>2.82 (2.44)</td>
<td>.495</td>
<td></td>
<td>.021</td>
<td>.219</td>
</tr>
<tr>
<td>Gender (male)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td>40.0%</td>
<td>26.7%</td>
<td>38.5%</td>
<td>43.2%</td>
<td>.461</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>39.1%</td>
<td>26.9%</td>
<td>39.1%</td>
<td>46.0%</td>
<td>.456</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>40.9%</td>
<td>24.0%</td>
<td>40.9%</td>
<td>44.9%</td>
<td>.369</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

WTAR = Wechsler Test of Adult Reading; est FSIQ = estimated Full Scale Intelligence Quotient; DRS-2 = Dementia Rating Scale-2 (Age and Education corrected); HADS = Hospital Anxiety and Depression Scale (A=Anxiety score; D = Depression score).
Analysis by MANOVA revealed significant group differences in visual and verbal episodic memory function (LMI recall total; LMII recall total, RCFT recall) (Pillai’s trace = .425, $F_{(9, 495)} = 9.071, p < .001$, power $= 1.00$ $n_p^2 = .142$). Post hoc analysis indicated that the a-MCI and a-MCI+ groups had significantly poorer immediate paragraph recall (LMI recall total) and delayed paragraph recall (LMII recall total) than the control and na-MCI groups. A similar pattern of results was evident in terms of visual episodic memory function with both the a-MCI+ and a-MCI groups demonstrating significantly poorer visual delayed recall (RCFT recall) than the na-MCI and control groups.

Analysis by MANOVA revealed significant group differences in immediate memory function (DSP forward, DSP backward, SSP forward, SSP backward) (Pillai’s trace = .226, $F_{(12, 492)} = 3.334, p < .001$, power $= .997$ $n_p^2 = .075$). In terms of immediate verbal span, the a-MCI+ group had a significantly shorter forward and backward digit span than the a-MCI and control groups. Group differences were also detected on SSP, with the a-MCI+ and na-MCI groups having a significantly shorter forward and backward visual spans than the a-MCI and control groups.

Analysis by MANOVA revealed significant group differences in attention/executive function (DSC, Trails A and B, Stroop A and C) (Pillai’s trace = .330, $F_{(15, 489)} = 4.028, p < .001$, power $= 1.00$ $n_p^2 = .110$). Performance on the DSC tasks suggests that the a-MCI+ group was significantly poorer at utilising selective and sustained attention compared to the a-MCI and control groups. In terms of attentional speed (Trails A), the a-MCI+ group were significantly slower compared to the a-MCI group. Further, the a-MCI+ were significantly slower than the control group when utilizing divided attention (Trails B). Group differences were also
observed on Stroop C where the a-MCI+ and na-MCI groups displayed greater difficulty in terms of inhibiting prepotent response compared to the a-MCI and control groups.

One way ANOVAs were used to assess group performances on language function (Vocab); visuospatial construction (RCFT-copy); and verbal working memory function (LNS) (see Table 5). Group differences were detected in language function with the a-MCI+ group scoring significantly poorer on the test of vocabulary compared to the na-MCI and control groups. Performance on visuospatial construction (RCFT-copy) was also significantly poorer in the a-MCI+ group compared to the na-MCI and control groups. In terms of verbal working memory function, the a-MCI+ group had a significantly shorter letter-number span compared to the a-MCI and control groups; and the na-MCI had a significantly shorter letter number span compared to the a-MCI group.
<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI Mean (SD)</th>
<th>na-MCI Mean (SD)</th>
<th>a-MCI+ Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>p.</th>
<th>Post-hoc</th>
<th>Effect size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>25</td>
<td>30</td>
<td>26</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMI recall total</td>
<td>30.04 (9.69)</td>
<td>38.30 (6.98)</td>
<td>28.50 (10.17)</td>
<td>39.89 (7.75)</td>
<td>&lt;.001</td>
<td>A, A+&lt;N,C</td>
<td>.248</td>
<td>1.00</td>
</tr>
<tr>
<td>LMII recall total</td>
<td>14.52 (7.14)</td>
<td>22.69 (8.44)</td>
<td>14.16 (8.67)</td>
<td>23.73 (5.41)</td>
<td>&lt;.001</td>
<td>A, A+&lt;N,C</td>
<td>.280</td>
<td>1.00</td>
</tr>
<tr>
<td>RCFT (recall)</td>
<td>13.28 (6.39)</td>
<td>18.75 (5.56)</td>
<td>10.48 (5.95)</td>
<td>18.74 (5.00)</td>
<td>&lt;.001</td>
<td>A, A+&lt;N,C</td>
<td>.266</td>
<td>1.00</td>
</tr>
<tr>
<td>DSP forward</td>
<td>10.64 (2.16)</td>
<td>9.43 (2.27)</td>
<td>8.65 (1.52)</td>
<td>10.26 (2.14)</td>
<td>.001</td>
<td>A+, A&lt;C</td>
<td>.091</td>
<td>.936</td>
</tr>
<tr>
<td>DSP backward</td>
<td>6.80 (2.26)</td>
<td>6.43 (2.79)</td>
<td>5.23 (1.75)</td>
<td>6.92 (2.26)</td>
<td>.012</td>
<td>A+, A&lt;C</td>
<td>.064</td>
<td>.805</td>
</tr>
<tr>
<td>SSP forward</td>
<td>8.00 (1.53)</td>
<td>6.73 (1.53)</td>
<td>6.50 (1.70)</td>
<td>7.70 (1.56)</td>
<td>&lt;.001</td>
<td>A+, N&lt;A, C</td>
<td>.112</td>
<td>.977</td>
</tr>
<tr>
<td>SSP backward</td>
<td>7.28 (2.01)</td>
<td>6.17 (1.18)</td>
<td>6.12 (1.68)</td>
<td>7.22 (1.50)</td>
<td>.001</td>
<td>A+, N&lt;A, C</td>
<td>.099</td>
<td>.956</td>
</tr>
<tr>
<td>DSC</td>
<td>66.84 (14.58)</td>
<td>60.33 (12.20)</td>
<td>53.85 (13.41)</td>
<td>63.00 (14.97)</td>
<td>.008</td>
<td>A+, A&lt;C</td>
<td>.069</td>
<td>.837</td>
</tr>
<tr>
<td>Trails A</td>
<td>26.69 (8.55)</td>
<td>31.68 (7.27)</td>
<td>34.91 (11.42)</td>
<td>29.49 (9.80)</td>
<td>.014</td>
<td>A+&gt; A</td>
<td>.063</td>
<td>.794</td>
</tr>
<tr>
<td>Trails B</td>
<td>71.97 (37.88)</td>
<td>78.40 (25.37)</td>
<td>90.90 (30.64)</td>
<td>67.85 (22.55)</td>
<td>.002</td>
<td>A+&gt; Cont</td>
<td>.086</td>
<td>.922</td>
</tr>
<tr>
<td>Measure</td>
<td>a-MCI Mean (SD)</td>
<td>na-MCI Mean (SD)</td>
<td>a-MCI+ Mean (SD)</td>
<td>Control Mean (SD)</td>
<td>p.</td>
<td>Post-hoc</td>
<td>Effect size (η²)</td>
<td>Power</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>----</td>
<td>----------</td>
<td>------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Stroop A</td>
<td>12.87 (3.53)</td>
<td>14.18 (3.11)</td>
<td>14.14 (2.72)</td>
<td>13.62 (3.14)</td>
<td>.392</td>
<td></td>
<td>.018</td>
<td>.270</td>
</tr>
<tr>
<td>Stroop C</td>
<td>28.92 (7.91)</td>
<td>38.34 (10.75)</td>
<td>39.82 (7.74)</td>
<td>29.61 (7.56)</td>
<td>&lt;.001</td>
<td>A+,N&gt;A,C</td>
<td>.234</td>
<td>1.00</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>50.44 (9.01)</td>
<td>51.77 (7.26)</td>
<td>44.04 (10.54)</td>
<td>52.60 (9.02)</td>
<td>.001</td>
<td>A+&lt;N, C</td>
<td>.101</td>
<td>.961</td>
</tr>
<tr>
<td>RCFT (copy)</td>
<td>33.44 (2.06)</td>
<td>34.07 (5.45)</td>
<td>30.85 (5.45)</td>
<td>34.47 (1.89)</td>
<td>&lt;.001</td>
<td>A+&lt;N, C</td>
<td>.174</td>
<td>.999</td>
</tr>
<tr>
<td>LNS total</td>
<td>11.36 (2.41)</td>
<td>9.43 (2.32)</td>
<td>8.54 (2.27)</td>
<td>10.63 (2.39)</td>
<td>&lt;.001</td>
<td>A+&lt;A, C; N&lt;A</td>
<td>.131</td>
<td>.992</td>
</tr>
</tbody>
</table>
Discussion

The aim of the screening phase was to assess individual performances across a range a
cognitive domains and subsequently classify those individuals according to published MCI
criteria (Winblad et al., 2004). A total of 25 individuals met the criteria for a-MCI; 30 for na-
MCI; 26 for a-MCI+; and 88 individuals met the criteria for the control group. The
examination of group differences at the screening stage was used to provide confirmation of
MCI subtype classification. The results indicated clear group differences that align with MCI
subtype profiles.

a-MCI

The a-MCI group displayed poor performance across all tests of visual and verbal episodic
memory and immediate verbal recall. As a group, a-MCI participants showed no impairment
on non-amnestic tasks that involved attentional or executive processing and working memory
capacity. Overall, the performance of the a-MCI group aligns with Petersen’s original
conception of MCI as a profile with pure amnestic deficits (Petersen et al., 1999). It is
important to note that membership to this subtype required one impaired performance (<10th
percentile) on a single test of episodic memory. Given evidence that healthy controls
frequently perform at an impaired level on at least one test within a neuropsychological test
battery (Brooks, Iverson, & White, 2007), it is possible that this subtype contains individuals
who may not meet the criteria of a-MCI at follow up. The use of a single impaired test
performance to classify MCI remains a contentious issue with some claiming that it reduces
the sensitivity and specificity of the diagnosis (Jak et al., 2009). Despite this, current criteria
(Winblad et al., 2004) and recent recommendations to adapt those criteria (Albert et al., 2011) do not require evidence of multiple test failures for an MCI diagnosis. By using the criterion of impaired performance on a single objective measure the present study has maintained alignment with published protocols.

**na-MCI**

The na-MCI group displayed no deficits to visual or verbal episodic memory which aligns with the profile of this variant. The na-MCI group did display poorer performance on tests of spatial short term memory (SSP forward and backward) when compared to the a-MCI and control groups. The na-MCI group also displayed significantly poorer performance on Stroop C, indicating poorer response inhibition. In terms of immediate verbal span, the na-MCI group performance was comparable to the control group. However, the na-MCI group was found to have a significantly shorter letter-number span (LNS) compared to the a-MCI group. This lowered performance is consistent with the na-MCI classification that encompasses impairments to working memory functioning. There is a genuine paucity of research examining the profile of the na-MCI variant. From the limited evidence tracking the trajectory of this subtype, it has been suggested that na-MCI individuals are more heterogeneous and may revert to normal levels of functioning at a higher rate than other MCI variants (Artero et al., 2006; Busse et al., 2006; Forlenza et al., 2009; Jungwirth et al., 2012). Further follow up of this variant will indicate whether this is the case.
**a-MCI+**

Consistent with their classification, the a-MCI+ group displayed a consistently poorer performance across multiple cognitive domains. Similar to the a-MCI subtype, the a-MCI+ group demonstrated poor performance on tests of visual and verbal episodic memory. The a-MCI+ group also displayed lowered performances across all non-amnestic domains (attention, visual and verbal working memory, language function, and executive processing). As a group, their performance confirms their classification within the multiple domain amnestic variant.

**Control**

Members of the control group were classified on the basis of test performances within the 10th percentile for age appropriate norms on all neuropsychological tasks. Overall, performance from this group reflected an unimpaired profile with the majority of measures revealing significantly greater performance compared to other MCI variants. For example, on tests of visual and verbal episodic memory, the control group had significantly higher recall compared to a-MCI and a-MCI+. This is not surprising given that the control group is defined by the absence of episodic memory impairment whereas the two amnestic variants are defined on the presence of episodic memory impairment. Similarly, the control group performed significantly better than the na-MCI group and a-MCI+ group on the majority of non-amnesic tasks confirming their classification as unimpaired controls.
Chapter 4

The self-fulfilling prophecy of episodic memory impairment in mild cognitive impairment (MCI): Do episodic memory deficits identified at classification remain evident when later examined with different memory tests?
Abstract

Previous studies of Mild Cognitive Impairment (MCI) have been criticised for using the same battery of neuropsychological tests during classification and longitudinal follow up. The key concern is that there is a potential circularity when the same tests are used to identify MCI and then subsequently monitor change in function over time. The aim of the present study was to examine the evidence of this potential circularity problem. The present study assessed the memory function of 72 MCI participants and 50 healthy controls using an alternate battery of visual and verbal episodic memory tests 9 months following initial comprehensive screening assessment and MCI classification. Individuals who were classified as multiple-domain amnestic MCI (a-MCI+) at screening show a significantly reduced performance in visual and verbal memory function at follow up using a completely different battery of valid and reliable tests. Consistent with their initial classification, those identified as non-amnestic MCI (na-MCI) or control at screening demonstrated the highest performance across the memory tasks. The results of the present study indicate that persistent memory deficits remain evident in amnestic MCI subgroups using alternate memory tests, suggesting that the concerns regarding potential circularity of logic may be overstated in MCI research.
Introduction

The concept of Mild Cognitive Impairment (MCI) emerged from a series of MAYO clinic epidemiological studies attempting to identify predictive risk factors for Alzheimer’s dementia (AD) (Petersen et al., 1997; Petersen et al., 1999; Petersen, Stevens, et al., 2001). The utility of MCI was perceived to be its ability to identify individuals most at risk of future cognitive decline, particularly those likely to transition to AD (Petersen et al., 1999). Subsequently, the clinical features used to classify MCI have gradually been replaced by MCI diagnostic criteria (Albert et al., 2011; Winblad et al., 2004), although a number of researchers question whether these criteria lack appropriate sensitivity and specificity to be considered diagnostic (Ganguli, Dodge, Shen, & DeKosky, 2004; Gauthier & Touchon, 2005; Larrieu et al., 2002; K. Palmer, Wang, Backman, Winblad, & Fratiglioni, 2002; K. Ritchie, Artero, & Touchon, 2001; Summers & Saunders, 2012; Tabert et al., 2006). Current MCI classification criteria include concern regarding a change in cognitive functioning; evidence of objective dysfunction (usually from neuropsychological assessment); relatively intact daily functioning; and an absence of dementia (Albert et al., 2011). According to the diagnostic criteria outlined by Winblad et al. (2004), amnestic subtypes are defined by the presence of an episodic memory deficit, whereas non-amnestic subtypes are defined by the presence of a non-memory deficit (e.g. attention, language, working memory). Both of these broad variants may be further classified as single domain (deficits are limited to one cognitive domain, e.g. episodic memory) or multiple domain (deficits are present in more than one domain, e.g. memory and attention) (Winblad et al., 2004).
The aim of many MCI studies is to follow an MCI cohort over time to identify the most sensitive predictors of future cognitive decline. Some of these studies classify patients with MCI and monitor cognitive function over time using the same battery of neuropsychological tests (e.g. Forlenza et al., 2009; Loewenstein, Acevedo, Agron, & Duara, 2007; N. L. J. Saunders & Summers, 2011). This has introduced a concern regarding the independence of the assessment of cognitive function over time from the initial diagnosis/classification of MCI. Specifically, it raises the question as to whether individuals who maintain a specific MCI classification at follow up do so because of genuine neuropsychological impairment or because of the self-fulfilling prophecy created from using the same psychometric instruments to identify MCI (Pearman & Storandt, 2004; K. Ritchie & Touchon, 2000). Other studies attempt to avoid this issue by using broad screening measures (e.g. MMSE, CERAD) to classify MCI and then track the progression of MCI cohorts using tests of discreet neuropsychological functions (e.g. Alexopoulos et al., 2006; Rabin et al., 2009). However, this introduces an alternative issue regarding the accuracy of the initial MCI classification. Research has revealed that broad screening measures lack the sensitivity to detect non-memory deficits (e.g. attention, language, working memory) in MCI, based on evidence that a majority of MCI cases demonstrate such deficits when assessed using reliable and valid neuropsychological measures (Alladi et al., 2006; Belleville et al., 1996; K. Ritchie & Touchon, 2000; N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). In attempt to avoid circularity, previous studies utilising restricted screening protocols may have misclassified MCI and/or missed classifying genuine cases of MCI.

The extent to which circular reasoning is an issue for the assessment of MCI remains unclear. One way of reducing its potential effects is by using a separate test battery to classify MCI, and an alternative test battery to assess cognitive function over time (Price et al., 2010). The
The present study represents an exploration into the potential issue of circular logic by examining memory function in an MCI cohort. We attempted to investigate whether amnestic dysfunction remained evident when groups were assessed using alternate tests of visual and verbal memory at screening and follow-up. It was hypothesised that if circularity of logic affects MCI classification, then MCI subtypes would display a change in their memory performance across two independent neuropsychological batteries.

Method

Study population

Community-residing older adults from Tasmania (Australia) were recruited using consecutive sampling from advertisements placed in local media (TV and radio) and local general medical practices. Participants were recruited to participate in a larger longitudinal study tracking the neuropsychological profile of MCI subtypes. Control participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment. Each participant provided fully informed consent prior to the commencement of the study, in accordance with the Human Research Ethics Committee (Tasmania) Network and National Health and Medical
Research Council (NHMRC) of Australia Human Research Guidelines, in accordance with the Declaration of Helsinki (1964).

Each participant underwent pre-screening via telephone to ensure that there were no medical, neurological, or psychological conditions that would impact their participation. In addition, each participant who passed pre-screening was assessed on a clinical neuropsychological battery spanning multiple memory and non-memory domains (see Table 6). This was important to avoid previous criticisms of erroneous classification of MCI cases due to inadequate classification protocols. The aim of the screening stage was to identify those who met the criteria for MCI (Winblad et al., 2004). Performances were classified as subclinically impaired where the performance was more than $1.28SD (\textless 10^{th} \text{ percentile})$ below age- and/or education based norms in accordance with previously established protocols (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Classification of MCI subtype as single domain amnestic MCI (a-MCI), single domain non-amnestic MCI (na-MCI), multiple domain amnestic MCI (a-MCI+), or multiple domain non-amnestic MCI (na-MCI+) was based on the presence of one or more subclinical impairments to one or more cognitive domains (Albert et al., 2011; N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). A total of 130 participants successfully complete pre-screening and classification screening. These participants composed the following groups: a-MCI ($n=24$); na-MCI ($n=23$); a-MCI+ ($n=27$); na-MCI+ ($n=6$); and healthy control ($n=50$). Due to the statistical issues associated with analysing small samples, the na-MCI+ group were collapsed to form a larger na-MCI group. Prior to the reassessment of episodic memory, eight participants withdrew, four for personal reasons and four due to emerging chronic health issues. The final sample of 122 participants ($\text{male} = 48$) formed the following groups: a-MCI ($n=23$); na-MCI ($n=25$); a-MCI+ ($n=24$); and healthy control ($n=50$).
### Table 6.

*Screening Test Battery Used for MCI Classification*

<table>
<thead>
<tr>
<th>Test</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler Test of Adult Reading (WTAR; WTAR, 2001)</td>
<td>Estimated premorbid IQ</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale (HADS; Snaith &amp; Zigmond, 1994)</td>
<td>Clinical anxiety and depression symptoms</td>
</tr>
<tr>
<td>Mattis Dementia Rating Scale, 2nd edition (DRS-2; Jurica et al., 2001)</td>
<td>Global cognitive functioning</td>
</tr>
<tr>
<td>Rey Complex Figure test (RCFT; Lezak et al., 2004)</td>
<td>Visual episodic memory</td>
</tr>
<tr>
<td>Logical Memory I &amp; II (LM; Wechsler, 1997b)</td>
<td>Verbal episodic memory</td>
</tr>
<tr>
<td>Digit Span (DSP; Wechsler, 1997a)</td>
<td>Immediate verbal memory span</td>
</tr>
<tr>
<td>Spatial Span (SSP; Wechsler, 1997b)</td>
<td>Immediate visual memory span</td>
</tr>
<tr>
<td>Letter-Number Sequencing (LNS; Wechsler, 1997a)</td>
<td>Working memory capacity</td>
</tr>
<tr>
<td>Stroop-Victoria version (Stroop; Lezak et al., 2004)</td>
<td>Executive functioning</td>
</tr>
<tr>
<td>Vocabulary (Vocab; Wechsler, 1997a)</td>
<td>Language function</td>
</tr>
<tr>
<td>Trail making test (TMT; Lezak et al., 2004)</td>
<td>Divided attention</td>
</tr>
<tr>
<td>Digit symbol coding (DSC; Wechsler, 1997a)</td>
<td>Sustained attention</td>
</tr>
</tbody>
</table>
Materials

Participants were screened on a test battery (see Table 6) comprised of tests selected on the basis of excellent reliability and validity in clinical and subclinical populations. Follow-up episodic memory assessment (experimental) involved alternate tests of episodic memory to those used at screening assessment. Tests assessing both verbal and visual memory were included at screening and the experimental stages as research has shown that episodic memory deficits may manifest both verbally and/or visually in MCI (Alladi et al., 2006). The experimental protocol included the Paired Associates Learning test (PAL; Cambridge Cognition Ltd, 2011) and the Rey Auditory Verbal Learning Test (RAVLT; Strauss et al., 2006). The PAL is a subtest of the Cambridge Neuropsychological Test Automated Battery (CANTAB). It is a visual measure of episodic memory and learning and is sensitive to medial temporal lobe function (Cambridge Cognition Ltd, 2011). The PAL has a demonstrated ability to accurately discriminate between individuals with AD and healthy controls as well as the capacity to predict future cognitive decline (Swainson et al., 2001). During the PAL, participants are presented with six white boxes that open up one at a time in random order. At trial one, the computer reveals two different patterns hidden in two separate boxes. The participant is required to recall the location of each pattern at the end of the presentation sequence. Correct detection of each pattern within the allocated ten attempts allows the participant to move on to the next phases where three, six, and eight patterns are hidden respectively. Failure to recall the correct location of each pattern after 10 trials results in termination of the test. The selected outcome measures for the PAL were total errors at 6 and 8 shapes (adjusted), which report the number of errors made at each of the stages respectively. These outcome measures were selected as they adjust the total score for those participants who fail to meet criterion on an earlier trial and do not complete the entire PAL
sequence (Cambridge Cognition Ltd, 2011). The RAVLT is a verbal assessment of episodic memory and learning. The RAVLT consists of five consecutive learning trials of an auditory presentation of 15 item word list. Following each learning trial, participants recall as many of the 15 words in any order. After the fifth learning trial, a distracter list of 15 new words is presented followed by a recall trial. Following this, the participant is required to recall as many words possible from the initial list. Outcome measures used in the following analysis were RAVLT trial 5; RAVLT total (trials 1-5); and RAVLT delayed. The RAVLT has been found to be reliable in distinguishing between healthy controls and individuals with AD, as well as differentiating between various neurodegenerative conditions (Tierney et al., 1994).

**Procedure**

Individual assessment sessions were conducted in a well-lit, well ventilated room and took approximately 90-120 minutes, including mandated rest breaks, to complete. Tasks assessing visual and verbal episodic memory were administered as part of a larger test battery examining the neuropsychological profile of MCI subtypes. Only results pertaining to episodic memory function were analysed for the present study. The CANTAB was administered on a laptop connected to an external 17-inch LCD touch screen monitor and response pad according to standard instructions. Participants sat approximately 50cm from the touch screen with the response pad positioned 15cm from the touch screen.
Results

Results were analysed using SPSS for Windows (version 19.0). MANOVA was used to control for potential inflation of Type 1 error due to analysing data from multiple tests within the same domain (episodic memory). Significant multivariate results were followed with one-way ANOVAs and post hoc analyses. Games-Howell was considered the appropriate post hoc analysis due to unequal sample sizes and breaches of homogeneity of variance (Howell, 2002).

Demographic variables were assessed to examine any potential group differences that may act as potential confounds (Strauss et al., 2006) (see Table 7). No group differences were detected in terms of age, education level, or HADS depression score. Group differences were detected on the WTAR with the a-MCI+ group having a significantly lower estimate of premorbid IQ compared to all groups. Group differences were also detected on HADS anxiety score however, due to insufficient power for the medium effect size evident a post hoc analysis failed to identify significant group differences, with a trend towards significance between the a-MCI+ and Control group ($p = .068$). Group differences in global cognitive function (DRS-2 score) were significant but in expected directions with the a-MCI+ having significantly lower scores than the control and na-MCI groups; and the a-MCI having significantly lower scores than the control group. While significant differences were found, no group had a mean DRS-2 score of clinical significance (all AEMSS $\geq 9$). There was no significant difference in gender ratio across the four groups ($\chi^2_{(3)} = 3.45$, $p_* = .327$).
Table 7.

*Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Measure (SD)</th>
<th>a-MCI</th>
<th>na-MCI</th>
<th>a-MCI+</th>
<th>Control (SD)</th>
<th>p</th>
<th>Post-hoc (at p &lt; .05)</th>
<th>Effect size (n_p^2)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean (SD)</td>
<td>70.61 (7.99)</td>
<td>70.60 (5.97)</td>
<td>69.29 (6.42)</td>
<td>72.66 (6.52)</td>
<td>.201</td>
<td></td>
<td>.038</td>
<td>.404</td>
</tr>
<tr>
<td>Education</td>
<td>Mean (SD)</td>
<td>14.43 (3.16)</td>
<td>15.04 (3.54)</td>
<td>12.46 (3.45)</td>
<td>14.20 (3.74)</td>
<td>.072</td>
<td></td>
<td>.057</td>
<td>.585</td>
</tr>
<tr>
<td>WTAR (est. FSIQ)</td>
<td>Mean (SD)</td>
<td>110.22 (5.42)</td>
<td>110.24 (4.97)</td>
<td>103.33 (7.99)</td>
<td>110.38 (5.76)</td>
<td>&lt;.001</td>
<td>a-MCI+ &lt; na-MCI, a-MCI, C</td>
<td>.178</td>
<td>.993</td>
</tr>
<tr>
<td>DRS-2 (AEMSS)</td>
<td>Mean (SD)</td>
<td>10.91 (2.17)</td>
<td>11.56 (1.89)</td>
<td>10.04 (1.97)</td>
<td>12.54 (2.14)</td>
<td>&lt;.001</td>
<td>a-MCI+ &lt; na-MCI, C; a-MCI &lt; C</td>
<td>.183</td>
<td>.994</td>
</tr>
<tr>
<td>HADS A</td>
<td>Mean (SD)</td>
<td>5.39 (2.79)</td>
<td>5.00 (2.68)</td>
<td>6.88 (3.71)</td>
<td>4.72 (2.56)</td>
<td>.028</td>
<td>Insufficient power</td>
<td>.074</td>
<td>.719</td>
</tr>
<tr>
<td>HADS D</td>
<td>Mean (SD)</td>
<td>3.04 (2.34)</td>
<td>2.52 (2.02)</td>
<td>3.38 (2.46)</td>
<td>2.72 (2.33)</td>
<td>.558</td>
<td></td>
<td>.017</td>
<td>.193</td>
</tr>
</tbody>
</table>

WTAR = Wechsler Test of Adult Reading; est FSIQ = estimated Full Scale Intelligence Quotient; DRS-2 = Dementia Rating Scale-2 (Age and Education corrected); HADS = Hospital Anxiety and Depression Scale (A=Anxiety score; D = Depression score); C = control
A MANOVA identified significant group differences in episodic memory (PAL 6 shapes adjusted; PAL 8 shapes adjusted; RAVLT trial 5; RAVLT total; RAVLT delayed) (Pillai’s trace = .260, $F_{(15, 348)} = 2.20$, $p = .006$, power = .975, $n^2_p = .087$). Group differences within each dependent variable were subsequently analysed by one-way ANOVA with post-hoc Games-Howell analysis.

Significant group differences were detected on PAL 6 shapes adjusted ($F_{(3,118)} = 6.69$, $p < .001$, power = .971, $n^2_p = .145$) and PAL 8 shapes adjusted ($F_{(3,118)} = 5.73$, $p = .001$, power = .943, $n^2_p = .127$). Post hoc analyses revealed that the a-MCI+ group made significantly more errors in attempting to recall the spatial location of six patterns compared to the na-MCI and control groups (see Figure 9). At eight patterns, the a-MCI+ group made significantly more errors than the control group (see Figure 9).
Figure 9. Group differences visual episodic memory (mean ± SEM)

Significant group differences were detected on RAVLT trial 5 \( (F_{(3,118)} = 6.61, p < .001, \) power = .969, \( n_p^2 = .144) \); RAVLT total \( (F_{(3,118)} = 5.16, p = .002, \) power = .917, \( n_p^2 = .116) \); and RAVLT delay \( (F_{(3,118)} = 7.17, p < .001, \) power = .980, \( n_p^2 = .154) \). Post hoc analyses revealed that the a-MCI+ group recalled significantly less words on average than the na-MCI and control groups at trial 5 (see Figure 10); across all RAVLT trials in total (see Figure 10); and at the delayed recall stage (see Figure 10).
The results of the present study indicate that individuals identified as a-MCI+ from a comprehensive screening assessment display significantly lower performances on different measures of verbal and visual episodic memory compared to control participants or those classified as na-MCI. Specifically, the a-MCI+ group made significantly more errors when
attempting to recall the spatial location of patterns (PAL 6 & 8 shapes adjusted). The a-MCI+ group also recorded the poorest performance on the final trial of a verbal learning task (RAVLT trial 5); lowest delayed verbal episodic memory recall (RAVLT delay); and poorest cumulative verbal learning across trials (RAVLT total). These results may seem unsurprising given that individuals within this group, by definition of their initial classification, scored at subclinical levels (< 10th percentile) on at least one memory and one non-memory test at screening. That this group performed poorly on a different set of memory measures compared to those used at screening strongly suggests that circular reasoning in MCI research may be less problematic than previously suggested.

While the a-MCI group appear to perform at an intermediate level between the a-MCI+ group and the control and na-MCI groups, these differences do not reach statistical significance. It could be argued that this is due to circular reasoning given that a new battery of memory tests was unable to identify significant group differences. However, a better explanation of these findings relates to stability. Previous research tracking MCI subtypes longitudinally suggests that the a-MCI profile is not only rare but highly unstable (Alladi et al., 2006; Kramer et al., 2006; Lopez et al., 2003; N. L. J. Saunders & Summers, 2010). That the a-MCI group performed at an intermediate level between the a-MCI+ group and the control and na-MCI groups may be a result of recovery of function of some individuals within this subtype. Therefore, it may be erroneous to conclude with certainty about circular reasoning in this group as performance differences could be confounded by false positive cases. Further, membership to the a-MCI subtype only requires a single impaired performance on a single test of episodic memory. As a result, individuals in the present study may have obtained a classification of a-MCI whilst experiencing a transient cause of cognitive dysfunction (e.g. anxiety, fatigue). Previous research indicates that it is not uncommon for healthy older
individuals to perform below clinical thresholds on a single neuropsychological measure (Brooks, Iverson, Holdnack, & Feldman, 2008; 2007). Despite this, current published criteria (Albert et al., 2011) state that a single aberrant test score is sufficient for a diagnosis of MCI. Future studies may circumvent this issue by requiring multiple test failures within a domain prior to diagnosis.

The present study attempted to address the issue of circular reasoning in MCI research. The above data suggests that circular reasoning may be less of an issue given that the a-MCI+ subtype displays evidence of depressed verbal and visual episodic memory function on alternate tests conducted 9 months after initial assessment. However, it may be argued that the notion of circular reasoning within this context is flawed as it relies on the premise that MCI is stable. Research demonstrate that MCI is far from stable with consistent evidence that of recovery of function is common (Summers & Saunders, 2012). As a theoretical construct, if MCI is a precursor stage to dementia if cannot be a stable entity. As a precursor to a neurodegenerative disease one would expect that MCI should display a pattern of deterioration cognitive function(s) over time until the clinical stage of dementia is reached. As such, those identified as MCI should continue to display evidence of cognitive difficulties that have either remained stable or deteriorated over time. However, there should not be evidence recovery of function in genuine MCI cases as this would indicate erroneous classification within the MCI spectrum.

Several factors warrant caution when interpreting the above data. First, the small sample size is likely to limit the generalisability of the present findings. Second, it could be argued that the issue of circular logic may have been better assessed by including a comparison group of individuals who were assessed with the same tests at screening and follow up. However, it is
not possible to obtain two identical clinical groups for comparison. Further, by adopting this approach it would be impossible to differentiate circularity effects from group differences and therefore confound the results. Third, it could be argued that circularity is inevitable unless there is complete independence between predictors and outcome measures (Tuokko & Frerichs, 2000). The use of different tests tapping the same domains is likely to result in some degree of circularity as performance is likely to be highly correlated. However, this study represents one of the first attempts to formally investigate circular reasoning in MCI and has several strengths compared to previous research. All MCI cases were assessed using a comprehensive test battery rather than the conventional approach of using screening tests to classify MCI. In addition, both visual and verbal episodic memory was assessed as part of the screening classification and the follow up memory assessment. Previous research that has only examined verbal memory may have inadvertently missed classifying or misclassified cases where the memory impairment was visual in nature (Alladi et al., 2006). This study also represents one of the few that has not compromised the comprehensiveness of the screening protocol by using global measures in attempt to avoid circularity.

Results of the present study show that when using different follow up tests, memory function remains compromised in individuals initially classified as a-MCI+. This suggests that circular reasoning in MCI research may be less of an issue than previously thought. Further it implies that researchers are not justified in using broad global measures at screening to avoid the issue of circularity. Potential MCI cases should always be assessed with comprehensive test protocols that enhance diagnostic accuracy. However, future studies wanting to minimize the influence of circularity should adopt different classification and follow up protocols. More research is required as to how this procedure may impact on the sensitivity and specificity of the MCI classification.
Chapter 5

Lowered performance in working memory and attentional sub-processes are most prominent in multi-domain amnestic mild cognitive impairment subtypes
Abstract

**Background**  Research suggests that working memory and attention deficits may be present in Mild Cognitive Impairment (MCI). However, the functional status of these domains within revised MCI subtypes remains unclear, particularly when previous studies examine these cognitive domains using the same tests as were used to classify MCI subtypes. The aim of this study was to examine working memory and attention function in MCI subtypes on a battery of neuropsychological tests that were distinct from those used to classify MCI subtypes.

**Methods**  A total of 122 adults aged 60-90 years were classified at baseline as amnestic (a-MCI), non-amnestic (na-MCI), and multi-domain amnestic (a-MCI+). The attentional and working memory capacity of participants was examined using a battery of tests distinct from those used to classify MCI at screening.

**Results**  The a-MCI+ group demonstrated the poorest performance on all working memory tasks and specific sub processes of attention. The na-MCI group had lowered performance on visual span and complex sustained attention only. There was no evidence of either attentional or working memory impairment in a-MCI participants.

**Conclusion**  When MCI cohorts are assessed on measures distinct from those used at classification, a-MCI+ demonstrates the greatest compromise to working memory and attention function. These results support previous findings that suggest a-MCI+ more closely resembles early stage AD and may be at increased risk of future cognitive decline compared to other MCI subtypes.
**Introduction**

Examination of the cognitive profile of early Alzheimer’s Disease (AD) suggest that deficits to episodic memory do not occur in isolation, but in conjunction with subtle deficits to executive functions, e.g. attention and working memory (Baddeley, Logie, Bressi, Dellasala, & Spinnler, 1986; Belleville et al., 2007; R. G. Morris & Baddeley, 1988; Perry & Hodges, 1999). More recently, there is evidence to suggest that the presence of attention and working memory deficits may predate episodic memory dysfunction in AD (Storandt, 2008) or indicate the imminent decline to dementia in individuals with pre-clinical AD who present with episodic memory dysfunction (Grober et al., 2008; Rapp & Reischies, 2005). Evidence of working memory and attention compromise has also been demonstrated in individuals classified with Mild Cognitive Impairment (MCI) (Belleville et al., 2007; N. L. J. Saunders & Summers, 2011). This suggests working memory and attention function may play an important role in understanding the trajectory of cognitively impaired but not yet demented older adults.

In a recent study, N. L. J. Saunders and Summers (2011) reported longitudinal deterioration to multiple sub-processes of attention in non-amnestic MCI (combined na-MCI and na-MCI+) and amnestic MCI (combined a-MCI and a-MCI+) participants, with the a-MCI group displaying prominent deficits to divided attention. In addition, na-MCI and a-MCI groups were found to have stable impairments to visual working memory. Belleville et al. (2007) also reported deficits to working memory and attention in an MCI sample (combined a-MCI and a-MCI+). This lead Belleville and colleagues to suggest that divided attention may show early compromise in MCI which is followed by difficulty with manipulating information in
short term memory. While these studies suggest that attention and working memory deficits are present in MCI, it is less clear if this profile is maintained when cohorts are: a) comprehensively assessed on valid and reliable neuropsychological tests; and, b) assessed on a battery that is distinct from that which is used when MCI classifications are established.

The importance of comprehensively assessing memory and non-memory function in potential MCI cases is now acknowledged (Kramer et al., 2006; N. L. J. Saunders & Summers, 2011). Previous research has demonstrated that MCI cases can be missed or misclassified if broad global measures are used to infer the integrity of non-memory domains (Alladi et al., 2006; de Jager & Budge, 2005). As a result, comprehensive neuropsychological assessment is emerging as the gold standard procedure to maximise the sensitivity and specificity of MCI classifications. Further, some researchers have classified and examined cognitive profiles in MCI using the same battery of neuropsychological tests. It has been argued that this creates a self-fulfilling prophecy in MCI classification (Pearman & Storandt, 2004; K. Ritchie & Touchon, 2000). If individuals presenting with performance decrements on specific neuropsychological tests are classified as MCI then it is hardly surprising that cross sectional analysis of their profiles reveals lowered performances within these domains. It has been suggested that MCI research should employ a distinct test battery to classify individuals with MCI to that which is used to examine their cognitive profile in order to reduce the potential influence of circular reasoning (Price et al., 2010).

The aim of the present study was to clarify the working memory and attention profile of MCI subtypes using a comprehensive test battery distinct from that which was used to classify individuals at screening. The findings reported here refer to a cohort classified at screening as amnestic, non-amnestic, multiple-domain amnestic, or control (Winblad et al., 2004). This
approach will provide greater understanding of the working memory and attention function of MCI subtypes whilst reducing the potential influence of circular reasoning. In line with the current conceptualisation of MCI subtypes, it was hypothesised that a-MCI and, na-MCI participants will display poorer performance on attention and working memory tasks compared to the control and a-MCI groups.

**Method**

**Study population**

The participants in this study were a sample of community-residing adults recruited in 2010 as part of a longitudinal study tracking the neuropsychological profile of MCI. Participants were recruited on the basis of the following criteria: (1) Presence of cognitive complaints (e.g. memory, attention); (2) preserved general cognition (as assessed using the DRS-2); (3) self-reported capacity to maintain independent daily functioning (confirmed by an informant); (4) no history of major medical, neurological, or psychiatric illness (medical screening questionnaire); (5) no history of major risk factors for vascular disease (medical screening questionnaire); (6) no history of sensory impairment or impairment to hand mobility (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Control participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment. Each participant provided informed consent prior to the commencement of the study, in accordance with the Human Research Ethics Committee (Tasmania) Network and
From initial recruitment and pre-screening ($n=286$), a total of 200 participants were screened via a comprehensive neuropsychological test battery (see Table 8) to classify participants according to existing MCI criteria (Winblad et al., 2004) and previous published research (N. L. J. Saunders & Summers, 2011; Summers & Saunders, 2012). Thirty one were excluded from data analysis and further assessment due to various psychological and medical reasons not disclosed at telephone screening. A total of 169 participant results were included in the screening cohort (66 males, 103 females). Participants were all of Anglo-Saxon or European decent.

Subclinical impairment was defined as a performance $1.28 \text{ SD}$ or greater ($<10^{th}$ percentile) below age appropriate normative references; a level of subclinical impairment that is consistent with previous research (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012) and recent recommendations outlining MCI assessment procedures (Albert et al., 2011). Classification of MCI subtype as amnestic MCI was based on the presence of subclinical impairments on one or more visual or verbal memory tasks. Likewise, a non-amnestic classification was based on the presence of subclinical impairment on one or more non-memory tasks (e.g. attention, working memory, language, visuo-spatial). Participants were further classified as single or multiple domain within their given subtype based on the presence of impaired memory and non-memory performance (amnestic multiple domain) or multiple non-memory impairments (non-amnestic multiple domain). Whilst all participants were recruited on the initial basis of a subjective cognitive complaint, only those with both subjective and objective impairments were included in the MCI sample. Controls were
recruited from a similar age cohort in the wider community with no history of significant medical history or subjective cognitive complaints. Control participants underwent the same screening procedure as MCI participants to ensure that they were suitable to act as controls.

After screening, the 169 participants met the classification criteria for the following groups: amnestic MCI (a-MCI; \( n = 25 \)); multiple domain amnestic MCI (a-MCI+; \( n = 26 \)); non-amnestic MCI (na-MCI; \( n = 24 \)); multiple domain non-amnestic MCI (na-MCI+; \( n = 6 \)); and control \( (n = 88) \). Due to a small number of participants meeting the criteria for na-MCI+ it was not possible to separately analyse this subgroup so they were collapsed into the na-MCI group. Of the 88 healthy controls recruited, 50 were selected to participate in the working memory and attention assessment on the basis of matching to the age and gender of MCI participants. A total of 10 participants withdrew from the study prior to the working memory and attention assessment (one geographic relocation; one transferred to an alternate study with competing methodology; four withdrew for personal reasons; and four withdrew due to developing health issues). A total of 122 participants aged 60 – 90 years (48 male, 74 female) completed the working memory and attention assessment: a-MCI \( (n = 23) \); na-MCI \( (n = 26) \); a-MCI+ \( (n = 23) \); control \( (n = 50) \).

**Materials**

Separate test batteries were utilised for MCI classification (screening) and the working memory and attention assessment (see Table 8). Tests were selected on the basis of excellent reliability and validity in clinical and subclinical populations. A number of tests were selected from the Cambridge Neuropsychological Test Automated Battery (CANTAB) (Cambridge
Cognition Ltd, 2011). The clinical validity and utility of the CANTAB has been well established, not only in detecting deficits early in the AD disease process but also in distinguishing between different neurodegenerative conditions (Robbins et al., 1998; Sahakian & Owen, 1992).

Table 8.

Test Batteries for Screening and Working Memory and Attention Assessment

<table>
<thead>
<tr>
<th>Screening assessment</th>
<th>Working memory and attention assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTAR</td>
<td>DRS-2</td>
</tr>
<tr>
<td>HADS</td>
<td>Simple Reaction Time (SRTI) 3</td>
</tr>
<tr>
<td>DRS-2</td>
<td>Choice Reaction Time (CRTI) 3</td>
</tr>
<tr>
<td>Digit Symbol Coding (DSC) 2</td>
<td>Digit Span (DSP) 2</td>
</tr>
<tr>
<td>Logical Memory (LMI) I 1</td>
<td>Spatial Span (SSP) 3</td>
</tr>
<tr>
<td>RCFT-copy</td>
<td>Letter Number Sequencing (LNS) 2</td>
</tr>
<tr>
<td>Digit Span (DSP) 2</td>
<td>HADS</td>
</tr>
<tr>
<td>Spatial Span (SSP) 1</td>
<td>Rapid Visual Processing (RVP) 3</td>
</tr>
<tr>
<td>RCFT-delay recall</td>
<td>Spatial Working Memory (SWM) 3</td>
</tr>
<tr>
<td>Letter Number Sequencing 2</td>
<td>Matching to Sample (MTS) 3</td>
</tr>
<tr>
<td>Trail Making Test A &amp; B (TMT)</td>
<td></td>
</tr>
<tr>
<td>Logical Memory II (LMII) 1</td>
<td></td>
</tr>
<tr>
<td>Stroop-Victoria version</td>
<td></td>
</tr>
<tr>
<td>Vocabulary (VOC) 2</td>
<td></td>
</tr>
</tbody>
</table>

1 subtest from the Wechsler Memory Scale, 3rd edition (WMS-III)
2 subtest from the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III)
3 subtest from the Cambridge Automated Neuropsychological Test Assessment Battery (CANTAB)

HADS = Hospital Anxiety and Depression Scale
WTAR = Wechsler Test of Adult Reading
DRS-2 = Dementia Rating Scale, 2nd edition
RCFT = Rey Complex Figure Test
RAVLT = Rey Auditory Verbal Learning Task
Procedure

Individual assessment sessions were conducted in a well-lit, well ventilated room and took approximately 90-120 minutes including rest breaks. Tasks assessing working memory and attention function were administered in the order presented in Table 8 as part of a larger test battery examining the neuropsychological profile of MCI subtypes. Only results pertaining to working memory and attention were analysed for the present study. In order to limit negative practice effects, participants were given a 10 minute break during the middle of the testing sequence. The CANTAB was administered on a laptop connected to an external 17-inch LCD touch screen monitor and response pad according to standard instructions. Participants sat approximately 50cm from the touch screen with the response pad positioned 15cm from the touch screen.

Results

Results were analysed using SPSS for Windows (version 19.0). MANOVA was used to control for potential inflation of Type 1 error due to analysing data from multiple tests of related cognitive domains. Significant multivariate results were followed with one-way ANOVAs and Games-Howell post hoc due to unequal sample sizes and some breaches of homogeneity of variance (Howell, 2002).

The neuropsychological tests used in the present study are sensitive to between group differences in demographic factors, estimated premorbid IQ, and global cognitive functioning (Strauss et al., 2006) and were analysed as potential confounders (see Table 9). There were
no significant differences between the groups on age, education level, or HADS depression scores. Group differences were detected on the WTAR with the a-MCI+ group having a significantly lower estimate of premorbid IQ compared to all groups. Group differences were also detected on HADS anxiety score however, due to insufficient power for the medium effect size evident a post hoc analysis failed to identify significant group differences. Group differences in global cognitive function (DRS-2 score) were significant but in expected directions with the a-MCI+ having significantly lower scores than the control and na-MCI groups; and the a-MCI having significantly lower scores than the control group. While significant differences were found, no group had a mean DRS-2 score of clinical significance (all AEMSS ≥9). There was no significant difference in gender ratio across the four groups ($X^2_{(3)} = 2.61, p = .456$).
Table 9.

*Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores*

<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI</th>
<th>na-MCI</th>
<th>a-MCI+</th>
<th>control</th>
<th>p</th>
<th>Post-hoc (at p &lt; .05)</th>
<th>Effect size (n²)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>23</td>
<td>26</td>
<td>23</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>70.61 (7.99)</td>
<td>70.58 (5.97)</td>
<td>69.26 (6.56)</td>
<td>72.66 (6.52)</td>
<td>.201</td>
<td></td>
<td>.038</td>
<td>.404</td>
</tr>
<tr>
<td>Education</td>
<td>14.43 (3.16)</td>
<td>14.92 (3.52)</td>
<td>12.48 (3.53)</td>
<td>14.20 (3.74)</td>
<td>.098</td>
<td></td>
<td>.052</td>
<td>.535</td>
</tr>
<tr>
<td>WTAR (est. FSIQ)</td>
<td>110.22 (5.42)</td>
<td>109.88 (5.20)</td>
<td>103.43 (8.15)</td>
<td>110.38 (5.76)</td>
<td>&lt;.001</td>
<td>a-MCI+ &lt; all groups</td>
<td>.163</td>
<td>.986</td>
</tr>
<tr>
<td>DRS-2 (AEMSS)</td>
<td>10.91 (2.17)</td>
<td>11.58 (1.86)</td>
<td>9.96 (1.97)</td>
<td>12.54 (2.14)</td>
<td>&lt;.001</td>
<td>a-MCI+ &lt; na-MCI, C; a-MCI &lt; C</td>
<td>.189</td>
<td>.996</td>
</tr>
<tr>
<td>HADS A</td>
<td>5.39 (2.79)</td>
<td>5.08 (2.65)</td>
<td>6.87 (3.79)</td>
<td>4.72 (2.56)</td>
<td>.034</td>
<td>Insufficient power</td>
<td>.071</td>
<td>.693</td>
</tr>
<tr>
<td>HADS D</td>
<td>3.04 (2.35)</td>
<td>2.46 (2.01)</td>
<td>3.48 (2.47)</td>
<td>2.72 (2.33)</td>
<td>.429</td>
<td></td>
<td>.023</td>
<td>.249</td>
</tr>
</tbody>
</table>

WTAR = Wechsler Test of Adult Reading; est FSIQ = estimated Full Scale Intelligence Quotient; DRS-2 AEMSS = Dementia Rating Scale-2 (Age and Education corrected MOANS scaled score); HADS = Hospital Anxiety and Depression Scale (A=Anxiety score; D = Depression score); a-MCI = single-domain amnestic MCI; na-MCI = nonamnestic MCI; a-MCI+ = multiple-domain amnestic MCI; C = control
MANOVA revealed significant group differences in immediate memory processes (DSP forward, DSP backward; SSP total) (Pillai’s trace = .244, $F_{(9, 354)} = 3.54, p < .001$, power = .988, $n_p^2 = .081$). Significant group differences were detected on forward visual span (SSP total) ($F_{(3, 118)} = 5.62, p = .001$, power = .939, $n_p^2 = .125$), with post-hoc analyses indicating that both the a-MCI+ and na-MCI groups displayed shorter visual spans compared to the a-MCI group (see Figure 11).

**Figure 11.** Mean spatial span length for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to a-MCI group at $p < .05$)
Group differences were also detected on forward ($F_{(3,118)} = 6.28, p = .001$, power $= .961$, $n_{p^2} = .138$) and backward ($F_{(3,118)} = 4.63, p = .004$, power $= .883$, $n_{p^2} = .105$) digit span. Post-hoc analysis showed that forward and backward verbal spans were shorter in the a-MCI+ group compared to the a-MCI and control groups (see Figure 12).

**Figure 12.** Mean forward Digit Span length forward and backward for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to control and a-MCI group at $p < .05$)

MANOVA revealed significant group differences in working memory capacity (LNS total; SWM total errors) (Pillai’s trace $= .138$, $F_{(6, 236)} = 2.91, p = .009$, power $= .891$, $n_{p^2} = .069$).
Post hoc analysis of verbal working memory (LNS total) \( (F_{(3,118)} = 4.09, p = .008, \text{ power } = .836, n_p^2 = .094) \) indicated that the a-MCI+ group displayed a shorter letter-number span than the a-MCI and control groups (see Figure 13). Significant group differences were also detected on visual working memory (SWM total errors) \( (F_{(3,118)} = 3.30, p = .023, \text{ power } = .740, n_p^2 = .077) \), with the a-MCI+ group making more errors than the a-MCI group (see Figure 14).

Figure 13. Mean letter-number sequencing for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to a-MCI and control groups at \( p < .05 \))
MANOVA revealed significant group differences in attentional control (SRTI; CRTI; MTS mean correct RT; RVP mean correct latency; RVP’A) (Pillai’s trace = .273, $F_{(15, 348)} = 2.32$, $p = .004$, power = .982, $n_p^2 = .091$). There were no significant group differences on measures of simple sustained attention (SRTI) ($F_{(3,118)} = 1.02$, $p = .386$, power = .271, $n_p^2 = .025$), or divided attention (CRTI) ($F_{(3,118)} = .020$, $p = .996$, power = .053, $n_p^2 = .000$) (see Table 10).
Significant group differences were detected on the measure of selective attention (MTS) \(F_{(3,118)} = 2.74, p = .046, \text{power} = .651, n_p^2 = .065\), however, due to insufficient power for the medium effect size evident a post hoc analysis failed to identify significant group differences (see Table 10). Group differences on complex sustained attention (RVP mean latency) were significant \(F_{(3,118)} = 3.53, p = .017, \text{power} = .772, n_p^2 = .082\), with post hoc analysis indicating that the control group were superior to the na-MCI group at maintaining attention control over time (see Figure 15). Significant group differences were evident on a measure target detection threshold (RVP A’) \(F_{(3,118)} = 6.26, p = .001, \text{power} = .961, n_p^2 = .137\), with post-hoc analysis indicating that the a-MCI+ group was more sensitive to errors than both the a-MCI and control groups (see Figure 16).

Table 10.

Non-significant Group Differences in Attention Performance

<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI</th>
<th>na-MCI</th>
<th>a-MCI+</th>
<th>control</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>23</td>
<td>26</td>
<td>23</td>
<td>50</td>
</tr>
<tr>
<td>STRI (msec)</td>
<td>306.81 (42.94)</td>
<td>321.34 (53.50)</td>
<td>305.60 (43.58)</td>
<td>324.41 (60.29)</td>
</tr>
<tr>
<td>CRTI (msec)</td>
<td>358.28 (54.03)</td>
<td>359.19 (43.24)</td>
<td>361.02 (46.85)</td>
<td>358.86 (42.91)</td>
</tr>
<tr>
<td>MTS mean correct latency (msec)</td>
<td>2276.18 (784.83)</td>
<td>2712.53 (857.87)</td>
<td>2726.86 (612.61)</td>
<td>2396.75 (603.39)</td>
</tr>
</tbody>
</table>

SRTI = simple reaction time; CRTI = choice reaction time; MTS = matching to sample
Figure 15. Mean RVP mean latency for control, a-MCI, na-MCI, and a-MCI+ groups (± SEM; * = significantly different to control group at $p < .05$)
Discussion

The results of the present study demonstrate that when participants are classified via a comprehensive test battery and then assessed using reliable and valid measures of working memory and attention function distinct from those used at screening, several group differences emerge. Most strikingly, the a-MCI+ group displayed poorer performances on
tasks tapping visual and verbal working memory as well as elements of attentional processing. Those identified as na-MCI at baseline (combined na-MCI and na-MCI+) demonstrated slower reaction time on complex sustained attention when compared to the control group, and a lowered performance in terms of visual short-term storage capacity of working memory relative to the a-MCI group. In keeping with their initial classification, the a-MCI group displayed a working memory and attention profile that did not differ significantly from the control group.

Key findings of the present study relate to the poor performance from the a-MCI+ group on tests that challenged the visual and verbal short-term storage components of working memory, as well as visual and verbal central executive processing. These findings are consistent with recent studies demonstrating working memory and attention compromise in amnestic variants of MCI (Belleville et al., 2007; N. L. J. Saunders & Summers, 2010, 2011) and early AD (Belleville et al., 1996; R. G. Morris, 1994; R. G. Morris & Baddeley, 1988). Further, research has indicated that individuals classified with a-MCI+ show persistent neuropsychological impairments over time and higher rates of conversion to AD (Alexopoulos et al., 2006; Bozoki et al., 2001; Summers & Saunders, 2012; Tabert et al., 2006). Therefore, compromised working memory and attention deficits may suggest a poorer prognosis for those classified as a-MCI+. Future studies following a-MCI+ groups over time could confirm whether reduced performance on working memory tasks represents a potential prognostic marker.

The a-MCI+ group also displayed additional difficulties with attentional regulation, as evident by significantly reduced target detection threshold on a complex visual attention task (RVP A`). The finding that lowered performance in attentional processing is limited to
specific MCI subtypes contradicts previous research. N. L. J. Saunders and Summers (2011) identified a slowing of reaction time on a simple sustained attention task in both na-MCI (combined na-MCI and na-MCI+) and a-MCI (combined a-MCI and a-MCI+) groups in a 20 month longitudinal study. The a-MCI group in the same study also showed marked decline in divided attention over a 20 month period. This led Saunders and Summers to suggest that the decline in simple sustained attention and divided attention may predate the emergence of decline in other cognitive areas. The results of the present study fail to identify significant group differences on measure of simple sustained attention and divided attention. However, the present study only assessed function at a single time point whereas Saunders and Summers detected groups differences at a 20 month follow up. It is possible that a decline in these functions may be detectable upon longitudinal follow up. Also, as Saunders and Summers concede, their use of Petersen’s criteria to classify MCI resulted in heterogeneous a-MCI and na-MCI groups, in which differences in functional status or severity between single and multiple domain subtypes could not be examined. Similarly, it was necessary in the present study to collapse the single and multiple-domain non-amnestic groups into a single heterogeneous group. As a result, the severity and breadth of deficits evident in na-MCI+ may be underestimated.

Evidence that the a-MCI+ and na-MCI groups demonstrated poorer functioning in working memory and attention tasks using alternative tests suggests that circular reasoning may be less of an issue than previously suggested. If the present study had failed to reveal significant group differences, such a result could have been partly attributed to circular reasoning. However, it may be argued that any claim of circular reasoning in MCI classification is logically flawed given the unstable nature of the MCI profile. Further, whilst we adopted different tests at the classification and follow up stages as a way of reducing the potential
impact of circular reasoning, performance on these measures is likely to correlate and may
reduce the extent to which circular logic can be avoided (L. J. Ritchie & Tuokko, 2010;
Tuokko & Frerichs, 2000). Despite this, the present study represents one of few MCI studies
that have attempted to avoid circular reasoning. Further, this was done using comprehensive
protocols at both screening and assessment stages. Many studies attempt to avoid circular
reasoning by using broad screening measures at screening, despite their lack of sensitivity at
the sub clinical level, with alternative test of neuropsychological function at follow up.

It is necessary to consider a major limitation of the present study that restricts the capacity to
extrapolate findings to the wider MCI population. Existing research indicates that there is a
high degree of instability in MCI classification (Han et al., 2012; Summers & Saunders,
2012), thus it is possible that individuals in the present study may have shifted MCI subtype
or may have recovered to age appropriate levels of functioning since baseline screening
testing. This latter point may be particularly relevant given the frequency at which healthy
older individuals perform within a subclinically impaired range (Brooks et al., 2008; Brooks
et al., 2007; Summers & Saunders, 2012) and the high proportion of MCI individuals seen to
revert to age appropriate levels of functioning at follow up testing (de Rotrou et al., 2005;
Summers & Saunders, 2012). Therefore, genuine neuropsychological change could be
masked by the inherent instability in MCI classification. As the present study did not examine
the stability of classification it is not possible to confirm whether individuals have maintained
MCI status. Further follow up of this cohort is underway and will help established stability
of classification over time.

Findings from the present study suggest that when an MCI cohort is classified according to a
comprehensive battery and examined on a separate test battery in terms of working memory
and attention function, the a-MCI+ subtype demonstrated the poorest performance. These results are consistent with earlier studies despite differences in classification method (Belleville et al., 2007; N. L. J. Saunders & Summers, 2010, 2011). The cognitive profile of a-MCI+ more closely resembles the presentation of clinical AD (Backman et al., 2004; Summers & Saunders, 2012), adding weight to the claim that this subtype may represent a clinically recognisable predementia syndrome. Follow up of this cohort will enable more direct comparison with longitudinal research implicating the a-MCI+ variant as the most at risk of future cognitive decline and a higher likelihood of conversion to AD (Ahmed et al., 2008; Alexopoulos et al., 2006). Future research should prioritise comprehensive testing at screening and follow up, and greater observation of those who present with an a-MCI+ profile as these individuals may be at the greatest risk of developing AD.
Chapter 6

Exploring the validity of Mild Cognitive Impairment (MCI) subtypes: Multiple-domain amnestic MCI is the only identifiable subtype at longitudinal follow up
Abstract

Research exploring risk factors for Alzheimer’s dementia has resulted in the identification of the Mild Cognitive Impairment (MCI) profile. However, the validity of MCI as a diagnostic entity remains uncertain. The aim of the present study was to examine the longitudinal neuropsychological profiles of MCI subtypes in 118 adults aged 60-90 years classified as amnestic (a-MCI), non-amnestic (na-MCI), and multiple-domain amnestic (a-MCI+). The a-MCI+ group displayed the poorest performance compared to all other groups in terms of episodic memory, working memory, attention, and executive functioning. These findings suggest the a-MCI+ subtype is the only variant that is recognisable via neuropsychological testing.
Introduction

Petersen and colleagues (1997) undertook a large study of elderly community members to identify potential risk factors for Alzheimer’s dementia (AD). From this study a subset of older adults were identified as suffering from Mild Cognitive Impairment (MCI), a condition characterized by isolated subclinical memory dysfunction and associated with an increased risk of 10-15% of converting to AD compared to a 1-2% risk in the general population. Subsequently, diagnostic criteria were developed (Petersen et al., 1999) and MCI has become the most widely adopted framework with which to examine subclinical impairment in older adults. The MCI concept was later revised to incorporate subtypes (amnestic and non-amnestic, single and multiple domains impaired) to account for the variation that began to emerge across MCI studies (Winblad et al., 2004). Despite ongoing efforts to elucidate the trajectory of these subtypes, research evidence continues to indicate that MCI is a heterogeneous condition (Han et al., 2012).

The heterogeneity of the trajectory of MCI raises concerns regarding the validity of the diagnosis of MCI and MCI subtypes (Summers & Saunders, 2012) which lacked supporting empirical data at the time of inception (Lopez et al., 2006). Instead, MCI subtypes appear to have assumed diagnostic status by virtue of their official criteria and extensive use in clinical and research settings (see Kendell & Jablensky, 2003). While there have been many attempts to validate these subtypes, multiple factors complicate comparison of studies. Some use test protocols that do not comprehensively assess non-amnestic functions leading to speculation about the adequacy of subtype classification (Artero et al., 2006; Busse et al., 2006; Fischer et al., 2007); some focus primarily on amnestic MCI (Kramer et al., 2006); and others have translated statistically derived profiles of neuropsychological functioning in broadly defined MCI as supporting evidence for the validity of these subtypes (Delano-Wood et al., 2009;
Libon et al., 2010). Further, many studies are epidemiological investigations of risk factors, prevalence and rates of conversion to dementia; rather than clinical studies designed to establish and validate diagnostic criteria. There is a dearth of longitudinal clinical studies of subpopulations utilizing comprehensive neuropsychological assessment to establish the validity of the MCI subtypes as discreet diagnostic entities.

In a cross sectional comparison of amnestic MCI and multiple-domain MCI (amnestic and non-amnestic) groups, it was found that the amnestic group showed greater impairment to memory function (visual and verbal) compared to controls and multi domain variants (Lopez et al., 2006). Further, multiple-domain MCI was associated with the poorest performance on tests of language, visuoconstructual ability; and fine motor control. Only measures of executive function failed to differentiate between MCI subtypes. However, as Lopez et al. classified MCI utilising broad screening measures (e.g. MMSE), it is possible that the sensitivity and specificity of the initial MCI classifications may be compromised. In a recent study, N. L. J. Saunders and Summers (2011) tracked MCI subtypes over a 20 month period using a comprehensive test battery spanning multiple memory and non-memory domains. It was reported that the amnestic MCI cohort (combined single and multiple-domain) had poorer visual and verbal memory compared to the non-amnestic group (combined single and multiple-domain). However, there were no differences detected between amnestic and non-amnestic groups on a majority of measures including complex sustained attention; target detection; strategy use; and visual working memory. Interestingly, the profiles from both MCI groups remained relatively stable over the 20 month duration of the study. The only area to show longitudinal decline in both groups was simple sustained attention, with additional decline in divided attention for the amnestic group.
While the above studies are only a small portion of MCI research, they represent two of the few studies focused on clarifying the neuropsychological profiles of MCI subtypes. If MCI does represent a genuine diagnostic entity, then the validity of existing MCI diagnostic criteria needs to be established. This can only be achieved through comprehensive research examining the status of discreet neuropsychological functions that were initially proposed as features of the subtype profiles. The present paper reports the findings of a prospective 20 month study, which utilised one comprehensive screening assessment used for MCI subtype classification and two comprehensive reassessments for neuropsychological profiling. The aim was to provide detail about the symptom features of MCI subtypes over time. Evidence of the symptom profile of these subtypes will provide important information to revise and strengthen current criteria. Further, it will provide evidence based on the clarification of symptom profiles rather than epidemiological risk factors, which has the potential to reduce the heterogeneity in MCI research.

Method

Participants

Participants were recruited during 2011 from a regional urban community of Tasmania via advertisements placed in local media (television and radio) and local general medical practices seeking older adults with subjective cognitive complaints. Figure 17 shows the progression of participants from recruitment to the final assessment phase (time 2). Control
participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment. The participant pool in the present study is a distinct cohort from those reported in previous publications from our group (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Participants were all of Anglo-Saxon or European decent. This project secured approval from the Human Research Ethics Committee (Tasmania) Network. Participants provided written fully informed consent in accordance with the HREC (Tasmania) Network approval and National Health and Medical Research Council (NHMRC) of Australia Human Research Guidelines in accordance with the ethical rules for human experimentation as stated in the Declaration of Helsinki.
At initial recruitment \( (n = 286) \) participants competed a telephone pre-screening to ensure that there were no medical, neurological, or psychological conditions that would impact their participation. A total of 200 participants passed pre-screening and were assessed on a comprehensive clinical neuropsychological test battery (see Table 11). Performance at screening assessment was used to classify participants according to MCI subtype criteria adapted from Winblad et al. (2004) and in accordance with previous published research (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). All participants demonstrated no evidence of dementia (DRS-2 AEMSS score ≥9); and had preserved
activities of daily living as assessed by a verbal interview with participant and informant. Subclinical impairment was defined as a performance 1.28 \( SD \) or greater (<10\(^{th}\) percentile) below age appropriate normative references; a level of subclinical impairment that is consistent with previous research (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012) and avoids the potential of cut-off scores of 1.5 \( SD \) being too severe (A. J. Mitchell & Shiri-Feshki, 2009). Classification of MCI subtype as single domain amnestic MCI (a-MCI), single domain non-amnestic MCI (na-MCI), multiple domain amnestic MCI (a-MCI+), or multiple domain non-amnestic MCI (na-MCI+) was based on the presence of one or more subclinical impairments to one or more cognitive domains (Albert et al., 2011; N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Control participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment.
Table 11.

*Screening and Experimental (Time 1 and Time 2) Test Batteries*

<table>
<thead>
<tr>
<th>Screening assessment</th>
<th>Time 1 assessment</th>
<th>Time 2 assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTAR</td>
<td>DRS-2</td>
<td>DRS-2</td>
</tr>
<tr>
<td>HADS</td>
<td>Boston Naming Test (BNT)</td>
<td>Paired Associates Learning (PAL)</td>
</tr>
<tr>
<td>DRS-2</td>
<td>Simple Reaction Time (SRTI)</td>
<td>RAVLT (List C)</td>
</tr>
<tr>
<td>Digit Symbol Coding (DSC)</td>
<td>Choice Reaction Time (CRTI)</td>
<td>Simple Reaction Time (SRTI)</td>
</tr>
<tr>
<td>Logical Memory (LMI) I</td>
<td>Digit Span (DSP)</td>
<td>Choice Reaction Time (CRTI)</td>
</tr>
<tr>
<td>RCFT-copy</td>
<td>Spatial Span (SSP)</td>
<td>Digit Span (DSP)</td>
</tr>
<tr>
<td>Digit Span (DSP)</td>
<td>Letter Number Sequencing (LNS)</td>
<td>Rapid Visual Processing (RVP)</td>
</tr>
<tr>
<td>Spatial Span (SSP)</td>
<td>Paired Associates Learning (PAL)</td>
<td>Letter Number Sequencing (LNS)</td>
</tr>
<tr>
<td>RCFT-delay recall</td>
<td>HADS</td>
<td>Matching to Sample (MTS)</td>
</tr>
<tr>
<td>Letter Number Sequencing</td>
<td>RAVLT (List A)</td>
<td>Spatial Span (SSP)</td>
</tr>
<tr>
<td>Trail Making Test A &amp; B (TMT)</td>
<td>Rapid Visual Processing (RVP)</td>
<td>HADS</td>
</tr>
<tr>
<td>Logical Memory II (LMII)</td>
<td>Spatial Working Memory (SWM)</td>
<td>Spatial Working Memory (SWM)</td>
</tr>
<tr>
<td>Stroop-Victoria version</td>
<td>Matching to Sample (MTS)</td>
<td>Boston Naming Test (BNT)</td>
</tr>
<tr>
<td>Vocabulary (VOC)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. subtest from the Wechsler Memory Scale, 3rd edition (WMS-III)
2. subtest from the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III)
3. subtest from the Cambridge Automated Neuropsychological Test Assessment Battery (CANTAB)

HADS = Hospital Anxiety and Depression Scale; WTAR = Wechsler Test of Adult Reading; DRS-2 = Dementia Rating Scale, 2nd edition; RCFT = Rey Complex Figure Test; RAVLT = Rey Auditory Verbal Learning Task
Those who remained in the study after screening were examined at two time points (time 1 and time 2) where they were assessed on a comprehensive battery utilizing a majority of measures that were distinct from those used in the screening clinical battery but assessing the same cognitive domains (see Table 11) to minimize potential circularity of logic when using the same test battery to classify and subsequently assess cognitive change in MCI (Price et al., 2010; K. Ritchie & Touchon, 2000; L. J. Ritchie & Tuokko, 2010). A total of 118 participants aged 60 – 90 years (46 male) completed all three assessments (screening, time 1, time 2) and were included in the analysis of the present study: a-MCI ($n = 22$); na-MCI ($n = 19$); a-MCI+ ($n = 23$); na-MCI+ ($n = 5$); Control ($n = 49$). Due to the small number of individuals who met the criteria for na-MCI+ ($n = 5$) we were unable to analyse this group separately so they were collapsed into the na-MCI group. At time 2 assessment, one participant displayed a clinical significant deterioration in cognitive test scores and was referred to a psychogeriatrician for an independent evaluation.

**Materials**

Two tests were administered at each time point (screening, time 1, and time 2) to verify the absence of dementia (DRS-2) (Jurica et al., 2001) and clinically significant anxiety and/or depression (HADS) (Snaith & Zigmond, 1994). An estimate of premorbid IQ (WTAR, 2001) was undertaken to ensure all groups were relatively homogenous with regard to premorbid intelligence (screening only). Separate neuropsychological test batteries were utilised for
MCI classification and longitudinal assessment (see Table 11). Tests were chosen on the basis of excellent reliability and validity in subclinical populations.

Procedure

Participants classified into one of the MCI groups at screening were tracked for 20 months during which they received two follow up neuropsychological assessments (time 1 and time 2). Average follow up times between screening - time 1; and time 1 - time 2 were nine (SD=3) and eleven (SD= 1) months respectively. Individual assessment sessions were conducted in a well-lit, well ventilated room and took approximately 90-120 minutes including rest breaks. Each assessment was administered in the order presented in Table 11. To limit fatigue effects, participants were given a 10 minute break during the middle of the testing sequence. Additional rest breaks were provided as needed for each individual participant. The CANTAB was administered on a laptop connected to an external 17-inch LCD touch screen monitor and response pad according to standard instructions. Participants sat approximately 50cm from the touch screen with the response pad positioned 15cm from the touch screen.

Results

Data analysis

Demographic data was analysed using repeated measures ANOVA and chi square for categorical variables. Repeated measures MANOVA was used to assess neuropsychological
data due to the inflated type I error rate when assessing related cognitive domains. Significant
results were followed up by factorial ANOVA (group x time) with ANOVA and Games-
Howell procedure due to unequal sample sizes and some breaches of homogeneity of
variance (Howell, 2002).

Demographics

Results were analysed using SPSS for Windows (version 19.0). There were no significant
differences between the four groups in terms of age, level of educational attainment, or
gender balance (see Table 12). The a-MCI+ group was found to have a significantly lower
estimated premorbid FSIQ compared to all other groups, however, all groups had above
average estimated premorbid FSIQ. A significant main effect of global cognitive functioning
(DRS-2) was detected (see Table 12); consistent with the classification of these groups within
the amnestic subtype of MCI. A significant main effect for anxiety was revealed although due
to insufficient power for the medium effect size evident, post hoc analysis was unable to
identify group differences. There were no differences detected in self-reported levels of
depression (HADS D; see Table 12).
<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI Mean (SD)</th>
<th>na-MCI Mean (SD)</th>
<th>a-MCI+ Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>F</th>
<th>p.</th>
<th>Post-hoc Effect size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>22</td>
<td>25</td>
<td>22</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years; screening)</td>
<td>69.82 (8.09)</td>
<td>69.36 (5.50)</td>
<td>68.82 (6.75)</td>
<td>71.82 (6.38)</td>
<td>1.424</td>
<td>.240</td>
<td></td>
<td>.036</td>
</tr>
<tr>
<td>Education (years)</td>
<td>14.64 (3.08)</td>
<td>14.72 (3.43)</td>
<td>12.32 (3.52)</td>
<td>14.06 (3.64)</td>
<td>2.320</td>
<td>.079</td>
<td></td>
<td>.058</td>
</tr>
<tr>
<td>WTAR (est. FSIQ)</td>
<td>110.27 (5.54)</td>
<td>109.72 (5.23)</td>
<td>103.55 (8.33)</td>
<td>110.67 (5.43)</td>
<td>7.65</td>
<td>&lt;.001</td>
<td>A+&lt;A,NA,C</td>
<td>.168</td>
</tr>
<tr>
<td>Gender (M:F)</td>
<td>9:13</td>
<td>6:19</td>
<td>9:13</td>
<td>22:27</td>
<td></td>
<td></td>
<td>V2 3.15</td>
<td>.369</td>
</tr>
</tbody>
</table>

| Measure               | Group 10.47 | <.001 | A+<A,NA,C; A+<C | .216 | .998 |
| Time 1                | .004        | .967  | .000             | .050 |
| Time 2                | .100        | .994  | .003             | .068 |

| Measure               | Group 3.05 | .031 | None p.<.05     | .074 | .703 |
| Time 1                | .264        | .608  | .002             | .080 |
| Time 2                | 1.19        | .317  | .030             | .312 |

| Measure               | Group 1.59 | .195 | .040             | .410 |
| Time 1                | .075        | .784  | .001             | .059 |
| Time 2                | .802        | .495  | .021             | .219 |

| Measure               | Group 2.18 | .095 | .054             | .541 |
| Time 1                | .383        | .537  | .003             | .094 |
| Time 2                | 1.09        | .356  | .028             | .288 |

Note. WTAR=Wechsler Test of Adult Reading; est. FSIQ = estimated Full Scale Intelligence Quotient; DRS-2 (AEMMS) = Dementia Rating Scale - 2 (Age and Education corrected MOANS [Mayo Older American Normative Studies] Scaled Scores; HADS = Hospital Anxiety [A] and Depression [D] Scale C = control, A = a-MCI, A+ = a-MCI=, NA = na-MCI
Verbal and Visual Episodic Memory

A repeated measures MANOVA was used to examine group differences on visual and verbal episodic memory (PAL 6 shapes adj; PAL 8 shapes adj; RAVLT trial 5; RAVLT total recall; RAVLT delayed recall) at time 1 and time 2. Significant main effects for group (Pillai’s trace = .258, $F_{(15,336)} = 2.106$, $p = .009$, power = .968, $n_p^2 = .086$) and time (Pillai’s trace = .206, $F_{(5,110)} = 5.708$, $p < .001$, power = .991, $n_p^2 = .206$) were identified. The interaction between group and time was non-significant (Pillai’s trace = .083, $F_{(15,336)} = .638$, $p = .843$, power = .422, $n_p^2 = .028$). Follow up analyses (see Table 13) indicate that a-MCI+ group made significantly more errors at the 6 shape and 8 shape stages of the PAL compared to the na-MCI and control groups. Further, the a-MCI+ group displayed lower word recall at trial 5, total recall (trials 1-5), and delayed recall of the RAVLT compared to the na-MCI and control groups. Additionally, word recall declined from time 1 to time 2 on trial 5 and at the delayed recall trial of the RAVLT (see Table 13).
Table 13.

*Episodic Memory Performance in MCI subtypes at Time 1 and Time 2*

<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI Mean (SD)</th>
<th>na-MCI Mean (SD)</th>
<th>a-MCI+ Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>F</th>
<th>p.</th>
<th>Post-hoc (at p. &lt; .05)</th>
<th>Effect size (η²)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAL errors 6 shapes (adjusted)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>7.95 (10.21)</td>
<td>4.28 (3.18)</td>
<td>12.59 (10.28)</td>
<td>5.10 (5.11)</td>
<td>.569</td>
<td>.452</td>
<td>A+&gt;NA, C</td>
<td>.167</td>
<td>.985</td>
</tr>
<tr>
<td>Time 2</td>
<td>9.23 (9.45)</td>
<td>4.48 (4.43)</td>
<td>11.77 (10.68)</td>
<td>6.04 (5.76)</td>
<td>.450</td>
<td>.718</td>
<td></td>
<td>.012</td>
<td>.138</td>
</tr>
<tr>
<td>PAL errors 8 shapes (adjusted)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>18.09 (17.39)</td>
<td>10.20 (7.63)</td>
<td>25.41 (20.72)</td>
<td>10.53 (8.41)</td>
<td>.009</td>
<td>.924</td>
<td>A+&lt;NA, C</td>
<td>.173</td>
<td>.989</td>
</tr>
<tr>
<td>Time 2</td>
<td>19.77 (18.52)</td>
<td>9.68 (6.97)</td>
<td>23.32 (20.92)</td>
<td>11.88 (9.24)</td>
<td>.621</td>
<td>.603</td>
<td></td>
<td>.016</td>
<td>.176</td>
</tr>
<tr>
<td>RAVLT trial 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>11.18 (2.92)</td>
<td>12.52 (1.98)</td>
<td>10.41 (2.48)</td>
<td>12.80 (2.16)</td>
<td>.496</td>
<td>.028</td>
<td>T2&lt;T1</td>
<td>.042</td>
<td>.598</td>
</tr>
<tr>
<td>Time 2</td>
<td>11.45 (2.81)</td>
<td>11.96 (2.05)</td>
<td>9.91 (2.25)</td>
<td>11.96 (2.39)</td>
<td>1.78</td>
<td>.155</td>
<td></td>
<td>.045</td>
<td>.453</td>
</tr>
<tr>
<td>RAVLT total recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>46.14 (11.98)</td>
<td>49.84 (10.21)</td>
<td>41.55 (10.76)</td>
<td>51.20 (9.15)</td>
<td>.563</td>
<td>&lt;.001</td>
<td>A+&lt;NA, C</td>
<td>.129</td>
<td>.939</td>
</tr>
<tr>
<td>Time 2</td>
<td>46.23 (11.22)</td>
<td>49.18 (9.13)</td>
<td>40.27 (7.97)</td>
<td>48.78 (9.00)</td>
<td>.735</td>
<td>.533</td>
<td></td>
<td>.019</td>
<td>.203</td>
</tr>
<tr>
<td>RAVLT delayed recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>8.77 (3.87)</td>
<td>11.04 (3.40)</td>
<td>8.00 (3.22)</td>
<td>11.18 (2.72)</td>
<td>.289</td>
<td>.833</td>
<td>T2&lt;T1</td>
<td>.008</td>
<td>.104</td>
</tr>
<tr>
<td>Time 2</td>
<td>7.95 (3.66)</td>
<td>9.72 (3.04)</td>
<td>6.68 (3.00)</td>
<td>9.80 (2.96)</td>
<td>.558</td>
<td>.523</td>
<td></td>
<td>.019</td>
<td>.203</td>
</tr>
</tbody>
</table>

RAVLT = Rey Auditory Verbal Learning Test; PAL = Paired Associated Learning (CANTAB); C = control, A = a-MCI, A+ = a-MCI=, NA = na-MCI
Attention Measures

A repeated measures MANOVA was used to examine group differences on attention function (SRTI; CRTI; MTS mean correct latency; RVP mean correct latency; RVP A’) at time 1 and time 2. Analysis revealed significant group (Pillai’s trace = .265, $F_{(15,336)} = 2.171, p = .007$, power = .973, $n_{p^2} = .088$) and time (Pillai’s trace = .180, $F_{(5,110)} = 4.818, p = .001$, power = .975, $n_{p^2} = .180$) main effects. The time by group interaction was non-significant (Pillai’s trace = .132, $F_{(15,336)} = 1.030, p = .424$, power = .673, $n_{p^2} = .044$). Subsequent analyses (see Table 14) revealed that the a-MCI+ group displayed a significantly reduced signal detection threshold (RVP A’) than the a-MCI and control groups. The na-MCI group displayed a slower reaction time on a measure of complex sustained attention (RVP latency), which approached statistical significance. At time 1, the a-MCI+ group demonstrated poorer performance than the control group on selective attention (MTS correct RT); however, this difference was not maintained post hoc. Improved performance over time was evident on three attention measures (SRTI, RVP latency, and RVP A’). No significant effects were detected using the CRTI measure of divided attention.
<table>
<thead>
<tr>
<th>Measure</th>
<th>a-MCI Mean (SD)</th>
<th>na-MCI Mean (SD)</th>
<th>a-MCI+ Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>F</th>
<th>p</th>
<th>Post-hoc (p&lt;.05)</th>
<th>Effect size (η²)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SRTI (msec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>307.38 (43.87)</td>
<td>314.23 (40.10)</td>
<td>306.07 (44.54)</td>
<td>323.61 (60.65)</td>
<td>4.79</td>
<td>.031</td>
<td>T1&gt;T2</td>
<td>.583</td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>298.59 (40.72)</td>
<td>310.85 (40.35)</td>
<td>296.59 (50.28)</td>
<td>305.02 (48.27)</td>
<td>.648</td>
<td>.585</td>
<td></td>
<td>.183</td>
<td></td>
</tr>
<tr>
<td><strong>CRTI (msec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>358.32 (55.30)</td>
<td>354.77 (40.34)</td>
<td>361.70 (47.83)</td>
<td>358.13 (43.04)</td>
<td>.235</td>
<td>.629</td>
<td></td>
<td>.077</td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>350.88 (45.22)</td>
<td>359.86 (46.97)</td>
<td>357.53 (75.99)</td>
<td>355.96 (57.81)</td>
<td>.315</td>
<td>.814</td>
<td></td>
<td>.109</td>
<td></td>
</tr>
<tr>
<td><strong>MTS correct RT (msec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>2281.53 (802.87)</td>
<td>2635.18 (777.54)</td>
<td>2764.21 (599.63)</td>
<td>2378.72 (595.88)</td>
<td>1.39</td>
<td>.251</td>
<td></td>
<td>.360</td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>2451.68 (1046.40)</td>
<td>2411.50 (649.57)</td>
<td>2558.59 (611.93)</td>
<td>2314.83 (612.23)</td>
<td>3.20</td>
<td>.026</td>
<td>A+C (p&lt;.05)</td>
<td>.725</td>
<td></td>
</tr>
<tr>
<td><strong>RVP latency (msec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>407.08 (90.50)</td>
<td>542.38 (109.68)</td>
<td>522.95 (155.29)</td>
<td>469.34 (91.30)</td>
<td>13.92</td>
<td>&lt;.001</td>
<td>T1&lt;T2</td>
<td>.959</td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>418.72 (67.16)</td>
<td>501.74 (110.66)</td>
<td>476.33 (117.01)</td>
<td>449.26 (94.06)</td>
<td>.573</td>
<td>.634</td>
<td></td>
<td>.165</td>
<td></td>
</tr>
<tr>
<td><strong>RVP A’</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>.90 (.05)</td>
<td>.88 (.06)</td>
<td>.85 (.05)</td>
<td>.90 (.05)</td>
<td>8.04</td>
<td>.005</td>
<td>T1&gt;T2</td>
<td>.803</td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>.91 (.05)</td>
<td>.90 (.06)</td>
<td>.87 (.07)</td>
<td>.91 (.06)</td>
<td>2.230</td>
<td>.876</td>
<td></td>
<td>.092</td>
<td></td>
</tr>
</tbody>
</table>

SRTI = Simple Reaction Time (CANTAB); CRTI = Choice Reaction Time (CANTAB); MTS = Match to Samples (CANTAB); RVP = Rapid Visual Processing (CANTAB); RVP A’ = Rapid Visual Processing target detection threshold (CANTAB); C = control, A = a-MCI, A+ = a-MCI+, NA = na-MCI
Immediate Memory Measures

A repeated measures MANOVA was used to examine group differences in immediate memory (DPS fwd; DSP backwards; SSP length) at time 1 and time 2. Analysis revealed a significant group main effect (Pillai’s trace = .242, $F_{(9,342)} = 3.340$, $p = .001$, power = .984, $r^2_p = .081$). The time main effect (Pillai’s trace = .033, $F_{(3,112)} = 1.274$, $p = .287$, power = .333, $r^2_p = .033$) and the interaction between time and group (Pillai’s trace = .054, $F_{(9,342)} = .690$, $p = .718$, power = .342, $r^2_p = .018$) were non-significant. Follow-up analyses (see Table 15) revealed that the a-MCI+ group had a shorter forward verbal span (DSP forward) and significantly shorter backward verbal span (DSP backward) compared to the a-MCI and control groups. On forward visual span (SSP), the na-MCI group displayed a significantly shorter span than the a-MCI and control groups, with the a-MCI+ group displaying a significantly shorter visual span than the a-MCI group.
Table 15.

Performance on Immediate Memory and Working Memory Measures in MCI Subtypes at Time 1 and Time 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>n</th>
<th>a-MCI Mean (SD)</th>
<th>na-MCI Mean (SD)</th>
<th>a-MCI+ Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>F</th>
<th>p</th>
<th>Post-hoc (at p.&lt;.05)</th>
<th>Effect size (η²)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSP forward</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>22</td>
<td>10.77 (2.22)</td>
<td>9.36 (2.18)</td>
<td>8.50 (1.74)</td>
<td>10.41 (2.29)</td>
<td>6.59</td>
<td>&lt;.001</td>
<td>A+ &lt; A,C</td>
<td>.148</td>
<td>.969</td>
</tr>
<tr>
<td>Time 2</td>
<td>25</td>
<td>10.82 (2.54)</td>
<td>9.28 (1.72)</td>
<td>8.86 (1.83)</td>
<td>10.69 (2.28)</td>
<td>.937</td>
<td>.335</td>
<td></td>
<td>.008</td>
<td>.160</td>
</tr>
<tr>
<td>DSP backward</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.422</td>
<td>.737</td>
<td></td>
<td>.011</td>
<td>.132</td>
</tr>
<tr>
<td>Time 1</td>
<td>22</td>
<td>7.50 (2.24)</td>
<td>6.72 (2.15)</td>
<td>5.41 (1.50)</td>
<td>7.14 (2.35)</td>
<td>4.71</td>
<td>.004</td>
<td>A+ &lt; A,C</td>
<td>.110</td>
<td>.888</td>
</tr>
<tr>
<td>Time 2</td>
<td>25</td>
<td>7.41 (2.40)</td>
<td>6.76 (1.94)</td>
<td>5.86 (1.61)</td>
<td>7.51 (2.13)</td>
<td>1.46</td>
<td>.230</td>
<td></td>
<td>.013</td>
<td>.223</td>
</tr>
<tr>
<td>SSP length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.46</td>
<td>.119</td>
<td>A+ &lt; A</td>
<td>.021</td>
<td>.343</td>
</tr>
<tr>
<td>Time 1</td>
<td></td>
<td>5.55 (.91)</td>
<td>4.72 (.74)</td>
<td>4.68 (.72)</td>
<td>5.20 (.84)</td>
<td>2.46</td>
<td>.119</td>
<td></td>
<td>.021</td>
<td>.343</td>
</tr>
<tr>
<td>Time 2</td>
<td></td>
<td>5.36 (.91)</td>
<td>4.96 (.54)</td>
<td>5.00 (.93)</td>
<td>5.39 (1.02)</td>
<td>1.31</td>
<td>.275</td>
<td></td>
<td>.033</td>
<td>.341</td>
</tr>
<tr>
<td>LNS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.93</td>
<td>&lt;.001</td>
<td>NA &lt; A,C; T1&lt;T2</td>
<td>.135</td>
<td>.950</td>
</tr>
<tr>
<td>Time 1</td>
<td></td>
<td>10.59 (2.59)</td>
<td>10.20 (2.08)</td>
<td>8.91 (1.66)</td>
<td>10.69 (2.18)</td>
<td>4.82</td>
<td>.030</td>
<td></td>
<td>.041</td>
<td>.586</td>
</tr>
<tr>
<td>Time 2</td>
<td></td>
<td>11.45 (2.67)</td>
<td>10.60 (2.65)</td>
<td>8.82 (1.40)</td>
<td>11.02 (2.00)</td>
<td>1.10</td>
<td>.353</td>
<td></td>
<td>.028</td>
<td>.290</td>
</tr>
<tr>
<td>SWM errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.98</td>
<td>.034</td>
<td>A+ &gt; A</td>
<td>.073</td>
<td>.692</td>
</tr>
<tr>
<td>Time 1</td>
<td></td>
<td>20.73 (17.78)</td>
<td>30.52 (17.31)</td>
<td>38.18 (17.08)</td>
<td>29.16 (18.37)</td>
<td>1.22</td>
<td>.272</td>
<td></td>
<td>.011</td>
<td>.194</td>
</tr>
<tr>
<td>Time 2</td>
<td></td>
<td>22.36 (19.72)</td>
<td>29.64 (15.85)</td>
<td>34.45 (21.72)</td>
<td>25.76 (18.64)</td>
<td>.732</td>
<td>.535</td>
<td></td>
<td>.019</td>
<td>.202</td>
</tr>
<tr>
<td>Measure</td>
<td>n</td>
<td>a-MCI Mean (SD)</td>
<td>na-MCI Mean (SD)</td>
<td>a-MCI+ Mean (SD)</td>
<td>Control Mean (SD)</td>
<td>F</td>
<td>p</td>
<td>Post-hoc (at p&lt;.05)</td>
<td>Effect size (ηp²)</td>
<td>Power</td>
</tr>
<tr>
<td>------------------</td>
<td>----</td>
<td>-----------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>------</td>
<td>---------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>-------</td>
</tr>
<tr>
<td>SWM (strategy)</td>
<td>22</td>
<td>29.64 (6.87)</td>
<td>33.52 (5.85)</td>
<td>33.23 (6.02)</td>
<td>30.41 (6.88)</td>
<td>2.58</td>
<td>.057</td>
<td>.064</td>
<td>.621</td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>25</td>
<td>33.52 (6.87)</td>
<td>33.23 (5.85)</td>
<td>33.23 (6.02)</td>
<td>30.41 (6.88)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>22</td>
<td>32.95 (7.86)</td>
<td>32.28 (5.57)</td>
<td>32.18 (7.26)</td>
<td>29.29 (6.47)</td>
<td>3.76</td>
<td>.055</td>
<td>.032</td>
<td>.486</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>interact</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| DSP = Digit Span (WAIS-III); SSP = Spatial Span (CANTAB); LNS = Letter-Number Sequencing (WAIS-III); SWM = Spatial Working Memory (CANTAB) C = control, A = a-MCI, A+ = a-MCI+, NA = na-MCI
**Working Memory Measures**

A repeated measures MANOVA was used to examine group differences on the working memory measures (LNS; SWM errors; SWM Strategy) at time 1 and time 2. Analysis revealed a significant main effect for group (Pillai’s trace = .217, $F_{(9,342)} = 2.958$, $p = .002$, power = .969, $n_p^2 = .072$). The main effect for time (Pillai’s trace = .064, $F_{(3,112)} = 2.556$, $p = .059$, power = .617, $n_p^2 = .064$) and the time by group interaction (Pillai’s trace = .055, $F_{(9,342)} = .703$, $p = .706$, power = .349, $n_p^2 = .018$) were non-significant. Follow-up analyses (see Table 15) revealed that the a-MCI+ group had a significantly shorter letter number span (LNS) than all other groups and made more errors during the spatial working memory task (SWM errors) than the a-MCI group. In addition, letter-number spans (LNS) significantly improved between time 1 and time 2.

**Language**

There were no significant group differences in language function or change in language function over time detected (see Table 12).

**Discussion**

The results of the present study indicate that at 20 month follow up, only the a-MCI+ group demonstrated a profile consistent with their initial MCI classification. The a-MCI+ group
demonstrated consistently poorer performance across the majority of neuropsychological measures. The a-MCI+ group displayed significantly poorer visual (PAL) and verbal (RAVLT) episodic memory, compared to the na-MCI and control groups, but not compared to the a-MCI group. These results contradict previous findings of significantly impaired visual and verbal episodic memory in a-MCI compared to other subtypes including a-MCI+ (Lopez et al., 2006). However, as the MCI classification criteria of Lopez et al. (2006) required evidence of at least two impaired test scores in a single cognitive domain, the a-MCI sample in their study may have presented with greater memory impairments than those in the present study.

The results of the present study indicate that in addition to episodic memory deficits, the a-MCI+ group displayed significantly shorter verbal spans (DSP) and were more sensitive to errors on a complex sustained attention task (RVP A’) compared to the a-MCI and control groups. The a-MCI+ group also demonstrated shorter verbal working memory span (LNS) compared to all other groups. Taken together, these results indicate reduced episodic memory and executive functioning are stable features of the a-MCI+ subtype. This finding aligns with evidence from studies of early stage AD which are associated with prominent deficits to episodic memory and accompanying deficits to central executive processing (Baddeley et al., 1986; Belleville et al., 2007). That the results of the present study indicate that a-MCI+ subtypes are readily differentiated from other subtypes and controls suggests that the a-MCI+ subtype is most likely to represent a discreet clinical diagnostic entity.

In contrast to the a-MCI+ subtype, the a-MCI and na-MCI subtypes frequently performed at a level equivalent to the healthy controls. The a-MCI subtype did not display impaired visual and verbal episodic memory performance in spite of participants in this group being
classified at the screening assessment on the basis of subclinical memory impairment. This suggests that single domain subtypes may lack stable clinical features to be considered separate diagnostic entities. Previous research indicates that approximately 30% of healthy individuals perform within an impaired range on a single neuropsychological test performance within a battery of tests (Brooks et al., 2008; Brooks et al., 2007). Therefore, using a single test impairment to define a-MCI or na-MCI may increase the rate of false positive MCI diagnosis. The lack of differentiation between the single domain and multiple-domain subtypes suggests that single domain MCI variants may not be reliably identifiable as diagnostic entities using clinical neuropsychological measures. This does not discredit the theoretical existence of single domain MCI subtypes, rather it highlights the inadequacies of currently neuropsychological measures to identify single domain subclinical impairments with sufficient sensitivity and specificity to result in minimal false positive identifications (Summers & Saunders, 2012).

Improvement in function over time was observed on measures of simple sustained attention (SRTI), complex sustained attention task (RVP), accuracy of target detection (RVP A’), and verbal working memory capacity (LNS). These results contradict previous findings of longitudinal decline in attentional processes, particularly simple sustained attention, in both a-MCI and na-MCI subtypes (N. L. J. Saunders & Summers, 2011). Improved performance within these domains may reflect practice effects that can occur from familiarity with the test format, or remembering specific aspects of the test itself (Duff, 2012). However, given the length of the retest interval (11 months), it is unlikely that participants were able to recall specific aspects of the test materials. It is more likely that improved performance may be an artefact of the degree of measurement error inherent in clinical neuropsychological tests capacity to detect impairment at a subclinical level with sensitivity and specificity. As
standard clinical neuropsychological tests are developed to differentiate clinically significant impairment from non-clinically significant impairment, it is likely that these tests lack the required degree of task difficulty to accurately differentiate subclinical from unimpaired levels of function. The increasing body of research highlighting the instability of single-domain MCI variants (Forlenza et al., 2009; Han et al., 2012; Hughes, Snitz, & Ganguli, 2011) may reflect initial diagnostic errors arising from the use of neuropsychological tests with insufficient sensitivity and specificity in the subclinically impaired range. In accordance with the results of the present study, previous studies indicate that multiple-domain amnestic MCI variants (a-MCI+) may be the only reliably identifiable MCI subtype demonstrating continued impairment over time (Brodaty et al., 2012; Forlenza et al., 2009; N. L. J. Saunders & Summers, 2011; Summers & Saunders, 2012) when examined using comprehensive neuropsychological assessment. While some suggest that the finding of improved diagnostic accuracy for a-MCI+ subtypes reflects the proximity of the diagnosis to the transition to dementia (Brainerd et al., 2013), such an explanation overlooks the psychometric limitations inherent in identifying a single domain subclinical impairment using standardized clinical neuropsychological measures.

The results of the present study should be considered in the context of its limitations. The approach of the present study differed from previous longitudinal studies of MCI subtypes. In attempting to circumvent the potential issue of circularity of logic by using a screening (classification) battery of tests that were predominantly different to the battery used to track cognitive function change over time we may have also introduced a potential confound. Those individuals meeting the criteria for MCI on the basis of subclinical impairments identified on screening tests may not subsequently fulfil MCI diagnostic criteria due to differences in the psychometric characteristics of the tests used at screening from those used
at follow-up assessment. The assumption that any test of a specific cognitive function has the same reliability, validity, sensitivity and specificity as another test of that same cognitive function is erroneous.

Findings from the present study suggest that after comprehensive longitudinal neuropsychological assessment, the a-MCI+ profile is the only variant consistent with its initial classification of impaired memory and non-memory functioning. This profile was largely stable over time and showed evidence of decline in the domain of verbal episodic memory. However, we were unable to clarify the symptom profiles of a-MCI and na-MCI suggesting that these variants may not represent discreet diagnostic entities or that we were unable to identify genuine cases with this condition due to our design parameters. Future studies may need to adopt more challenging test protocols to isolate those with genuine a-MCI or na-MCI or require evidence of more than one test impairment prior to MCI classification. Alternatively, it may be necessary to wait until single domain variants transition to multiple domain variants to identify genuine cases, particularly when recruiting prospective subjects for intervention or treatment studies. Future research should ensure that MCI is classified on the basis of comprehensive neuropsychological assessment and prioritise follow up examinations of individuals who fulfil the criteria for a-MCI+. 
Chapter 7

The learning profile of persistent Mild Cognitive Impairment (MCI): A potential diagnostic marker of persistent amnestic MCI
Abstract

**Background** Previous research examining Mild Cognitive Impairment (MCI) has highlighted heterogeneity of outcome in MCI sufferers. MCI is associated with greater risk of progression to dementia, however a substantial proportion of those identified with MCI have alternate outcomes including recovery to unimpaired status. This heterogeneity may in part reflect insufficient sensitivity and specificity in identifying subclinical memory impairment.

**Method** The present study examined learning in a sample of 109 adults aged 61-91 years with persistent amnestic MCI, persistent non-amnestic MCI, recovered MCI, and healthy controls. At the final assessment point, learning for words recalled across each trial of the RAVLT was examined for each group.

**Results** It was found that persistent amnestic MCI participants displayed significantly lower learning compared to recovered MCI, and healthy control groups.

**Discussion** The results of this study indicated that poor learning across trials may be a defining feature of persistent amnestic MCI. Further research is required to establish the predictive utility of within trial list learning performance to identify individuals with persistent and progressive variants of MCI.
Introduction

Mild Cognitive Impairment (MCI) is a subclinical condition associated with an elevated risk of progression to dementia, particularly AD (Petersen, 2004). Despite evidence of higher rates of conversion from MCI to dementia (A. J. Mitchell & Shiri-Feshki, 2009), recent research indicates that MCI is longitudinally unstable (Han et al., 2012; Summers & Saunders, 2012). Various studies report that 15-41% of MCI participants no longer meet MCI criteria at longitudinal follow up, having reverted to age appropriate levels of cognitive function (Brodaty et al., 2012; Busse et al., 2006; de Jager & Budge, 2005; de Rotrou et al., 2005; Fisk, Merry, & Rockwood, 2003; Ganguli et al., 2004; Gauthier et al., 2006; Larrieu et al., 2002; J. Mitchell et al., 2009; K. Ritchie et al., 2001; Summers & Saunders, 2012).

The heterogeneity of outcome in MCI diagnosis may reflect a combination of factors. Some suggest that the variance in rates of progression to dementia reflect differences in the source of sample recruitment (e.g., community sample versus memory clinic sample) (Petersen & Bennett, 2005; Petersen & Morris, 2005). The implication that the reliability of diagnostic criteria is dependent on sampling factors points to an inadequacy of the diagnostic criteria rather than variation in the sampling method. Rather, it is likely that MCI outcome heterogeneity arises from insufficient sensitivity and specificity of standardized memory tests to detect subclinical impairment, especially when a single memory test performance taken at a single time point is the criterion for memory impairment in amnestic MCI diagnosis (Brooks et al., 2008; Brooks et al., 2007; B. W. Palmer, Boone, Lesser, & Wohl, 1998). Such heterogeneity of outcome poses a significant challenge for research attempting to identify diagnostic biomarkers or examining the efficacy of treatment interventions in preventing transition to dementia from MCI. There is a pressing need to develop accurate and reliable
diagnostic criteria for MCI that exclude those who will recover from being diagnosed with MCI. By enriching intervention programs with confirmed MCI cases, researchers will increase the probability of identifying a diagnostic biomarker or a genuine therapeutic agent rather than misinterpreting recovered MCI cases as evidence of an effective treatment (J. Mitchell et al., 2009). It has been repeatedly established that increased sensitivity and specificity of identifying persistent neuropsychological impairment occurs when repeat testing is used, or where impairment is evident across multiple measures of related cognitive functions are used at a single time point (Brooks et al., 2008; Brooks et al., 2007; B. W. Palmer et al., 1998). However, few studies have diagnosed MCI samples based on a repeat testing protocol or a comprehensive neuropsychological examination possibly reflecting the prohibitive costs associated with such processes. The aim of the present study was to examine the potential for single time-point trial based learning performance (RAVLT) to differentiate between persistent MCI cases from those with unstable or recovered forms of MCI as well as healthy controls.

Previous studies have used measures of learning from list learning tasks to predict those who are likely to transition from MCI to AD (Arnaiz & Almkvist, 2003; Rabin et al., 2009; Tremont, Miele, Smith, & Westervelt, 2010). However, a majority of these studies infer the integrity of learning based on measures of delayed recall with little attention paid to trial performance. This is despite evidence that learning deficits are present in early AD and may occur prior to clinical onset (Bondi et al., 1999; Germano & Kinsella, 2005; Greene, Baddeley, & Hodges, 1996; Grober & Kawas, 1997), with lowered rates of learning being found in MCI samples (Ribeiro et al., 2007) and being associated with increased risk of developing AD within 2 years (Chang et al., 2010). Such findings suggest that examination of
learning performance may assist in isolating those MCI cases most at risk of developing dementia, in particular AD.

To the best of our knowledge no studies have specifically examined the trial performance of groups with longitudinally established MCI. To describe the learning curve of each subtype, the present study utilised cross sectional trial data from the RAVLT, an established test of verbal episodic memory. Whilst a majority of studies utilise delayed recall (retention) measures in potentially unstable MCI cohorts, the present study explores potential differences in learning curves (performance over trials) between persistent forms of amnestic MCI, persistent non-amnestic MCI, recovered-MCI, and healthy controls. If performance on specific trials differentiates between longitudinally persistent forms of MCI, there may be utility in adopting trial based cut off scores in addition to the standard practice of delayed recall measures. Further, this may help reduce the variability often reported in cross sectional designs and subsequently enhance the sensitivity and specificity of the MCI classification at a single time point. If amnestic MCI represents prodromal AD and impaired learning is a feature of AD, it is logical to suggest that similar deficits would be present in a persistent amnestic MCI cohort.
Method

Participants

The participants in this study were a sample of community-residing adults recruited in 2010 as part of a longitudinal study tracking the neuropsychological profile of MCI. Participants were recruited on the basis of the following criteria: (1) Presence of cognitive complaints (e.g. memory, attention) as reported by the participant and/or informant; (2) preserved general cognition; (3) independent daily functioning (confirmed by an informant); (4) no history of major medical, neurological, or psychiatric illness (DSM-IV Axis I or II); (5) no history of major risk factors for vascular disease; (6) no history of sensory impairment or impairment to hand mobility (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Figure 18 shows the progression of participants from recruitment to the final assessment phase (time 2). Control participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment. The participant pool in the present study is a distinct cohort from cohorts reported in previous publications from our group (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Each participant provided informed consent prior to the commencement of the study, in accordance with the Human Research Ethics Committee (Tasmania) Network and National Health and Medical Research Council (NHMRC) of Australia Human Research Guidelines and the Helsinki Declaration.
From initial recruitment and pre-screening ($n = 286$), a total of 200 participants were screened via a comprehensive neuropsychological test battery (see Table 16) to classify participants according to existing MCI criteria (Winblad et al., 2004). Participants were then assessed at two further time points (time 1 and 2) on a comprehensive battery utilizing measures distinct from those used in the screening clinical battery but assessing the same
cognitive domains (see Table 16). This was done to reduce the potential influence of circular logic into the classification process, an issue that has been previously raised by some MCI researchers (Price et al., 2010; K. Ritchie & Touchon, 2000). Following each assessment point, the performance of each individual participant was examined by an experienced clinical neuropsychologist (MS) to determine if the performance met the criteria for MCI (Winblad et al., 2004). The assessor was blinded to participant classification from previous assessment points. Performances were classified as subclinically impaired where the performance was more than 1.28SD (<10th percentile) below age- and/or education based norms in accordance with previously established protocols (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012) and recent guidelines outlining evaluation procedures for MCI samples (Albert et al., 2011). Classification of MCI subtype as amnestic MCI was based on the presence of subclinical impairments on one or more visual or verbal memory task (Albert et al., 2011; N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Likewise, a non-amnestic classification was based on the presence of subclinical impairment on one or more non-memory task (e.g. attention, working memory). Due to the small sample size, it was not possible to separately analyse single versus multiple domains impaired within each subtype. Mean follow up time from screening assessment to time 1 was nine months (SD=3 months). Mean follow up time between time 1 and time 2 assessments was 11 months (SD=1 month). Because the current study was interested in the MCI profile in terms of stability of classification, only the 118 participants (overall retention of 90.8%) who were assessed at all time points were included in the current analysis. A numerical breakdown of the classification groups by test phase is presented in Table 17 with the classification of individuals at each assessment phase depicted in Figure 19.
Table 16.

Test Batteries for Screening and Follow-up Testing

<table>
<thead>
<tr>
<th>Screening assessment</th>
<th>Time 1 assessment</th>
<th>Time 2 assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTAR</td>
<td>DRS-2</td>
<td>DRS-2</td>
</tr>
<tr>
<td>HADS</td>
<td>Boston Naming Test (BNT)</td>
<td>Paired Associates Learning (PAL)</td>
</tr>
<tr>
<td>DRS-2</td>
<td>Simple Reaction Time (SRTI)</td>
<td>RAVLT (List C)</td>
</tr>
<tr>
<td>Digit Symbol Coding (DSC)</td>
<td>Choice Reaction Time (CRTI)</td>
<td>Simple Reaction Time (SRTI)</td>
</tr>
<tr>
<td>Logical Memory (LMI)</td>
<td>Digit Span (DSP)</td>
<td>Choice Reaction Time (CRTI)</td>
</tr>
<tr>
<td>RCFT-copy</td>
<td>Spatial Span (SSP)</td>
<td>Digit Span (DSP)</td>
</tr>
<tr>
<td>Digit Span (DSP)</td>
<td>Letter Number Sequencing (LNS)</td>
<td>Rapid Visual Processing (RVP)</td>
</tr>
<tr>
<td>Spatial Span (SSP)</td>
<td>Paired Associates Learning (PAL)</td>
<td>Letter Number Sequencing (LNS)</td>
</tr>
<tr>
<td>RCFT-delay recall</td>
<td>HADS</td>
<td>Matching to Sample (MTS)</td>
</tr>
<tr>
<td>Letter Number Sequencing</td>
<td>RAVLT (List A)</td>
<td>Spatial Span (SSP)</td>
</tr>
<tr>
<td>Trail Making Test A &amp; B (TMT)</td>
<td>Rapid Visual Processing (RVP)</td>
<td>HADS</td>
</tr>
<tr>
<td>Logical Memory II (LMII)</td>
<td>Spatial Working Memory (SWM)</td>
<td>Spatial Working Memory (SWM)</td>
</tr>
<tr>
<td>Stroop-Victoria version</td>
<td>Matching to Sample (MTS)</td>
<td>Boston Naming Test (BNT)</td>
</tr>
<tr>
<td>Vocabulary (VOC)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 subtest from the Wechsler Memory Scale, 3rd edition (WMS-III)
2 subtest from the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III)
3 subtest from the Cambridge Automated Neuropsychological Test Assessment Battery (CANTAB)
HADS = Hospital Anxiety and Depression Scale; WTAR = Wechsler Test of Adult Reading; DRS-2 = Dementia Rating Scale, 2nd edition; RCFT = Rey Complex Figure Test; RAVLT = Rey Auditory Verbal Learning Task
Figure 19. Flow chart of classification of participants at each assessment phase

Table 17.

Classification of Participants at Screening, Time 1, and Time 2

<table>
<thead>
<tr>
<th>Classification group</th>
<th>a-MCI</th>
<th>na-MCI</th>
<th>Healthy</th>
<th>unstable</th>
<th>recovered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening</td>
<td>44</td>
<td>25</td>
<td>49</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Time 1</td>
<td>14</td>
<td>55</td>
<td>49</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Time 2</td>
<td>17</td>
<td>44</td>
<td>57</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Stability classification</td>
<td>15</td>
<td>36</td>
<td>29</td>
<td>9</td>
<td>29</td>
</tr>
</tbody>
</table>

a-MCI = amnestic MCI; na-MCI = non amnestic MCI; Stability classification = classification based on the stability of classification profile over three testing sessions
At the conclusion of Time 2, all participants were reclassified into groups according to stability of MCI classification across the three time points. Participants classified as persistent amnestic MCI (a-MCI) demonstrated an impaired performance (≤10th percentile) on visual and/or verbal episodic tasks at Time 1 and Time 2, or across all three time points. Participants classified as persistent non-amnestic MCI (na-MCI) demonstrated impaired performances (≤10th percentile) on non-amnestic tasks at Time 1 and Time 2, or across all three time points. Recovered-MCI refers to those participants who met the criteria for a MCI subtype at screening but consistently displayed unimpaired performances at times 1 and time 2. Healthy controls were those participants performing consistently at unimpaired levels across all measures at all three assessments. Nine participants fluctuated between MCI and healthy control status across the three time points and were classified as unstable. These cases were removed from the analysis due to the small sample size. The 109 participants comprised the following groups: persistent a-MCI (n = 15); persistent na-MCI (n = 36); recovered MCI (n = 29); healthy (n = 29). The rate of reversion to unimpaired status from a MCI diagnosis (recovery) is highly variable, ranging from 0-90% across multiple studies (Brooks et al., 2008). In the present study, 42% of the initial MCI sample displayed recovery of function over time, a rate of recovery that falls within the range reported by recent studies (Brodaty et al., 2012; J. Mitchell et al., 2009; Sachdev et al., 2013).

**Materials**

Participants were assessed on a comprehensive neuropsychological test battery at three time points (see Table 16). These assessments comprise a detailed longitudinal study examining
the profile of MCI over time. Performance on the Rey Auditory Verbal Learning test (Strauss et al., 2006) was used to assess learning over trials in MCI subtypes. The RAVLT is a 15 word list learning task with five trials. Participants are read the list aloud at a rate of one word per second. After each trial, participants attempt to recall as many of the 15 words as possible. The RAVLT has a long clinical and research history particularly in terms of differentiating between various neurodegenerative conditions (Tierney et al., 1996). RAVLT learning data was not used to classify MCI participants at any stage, with identification of verbal episodic memory impairment at time 1 and 2 being based on the delayed recall trial of the RAVLT.

Procedure

Individual testing sessions were conducted in a well-lit, well ventilated room at the School of Psychology Newnham campus. Tests were administered in the order indicated in Table 16. Only data collected from the RAVLT at time 2 (20 months) assessment was used for analysis of learning profiles because it was of primary interest to characterise learning after a stability classification had been established. Each assessment took approximately 2.5 hours including a mandated 10 minute rest break at the middle of the test sequence to minimise fatigue effects. Additional rest breaks were provided as needed for each individual participant.
Results

Data analysis

Demographic data was assessed using ANOVA and Chi Squared for categorical variables. Repeated measures ANCOVA with age entered as a covariate was used to assess RAVLT trial performance. Post hoc analysis used pairwise comparisons as well as ANOVA with Ryan Einot Gabriel Welsch Range (R-E-G-W) procedure (Howell, 2002). All results were analysed using SPSS for Windows (version 19.0).

The neuropsychological tests used in the present study are sensitive to between group differences in demographic factors, anxiety and depression, estimated premorbid IQ, and global cognitive functioning (Strauss et al., 2006) and were therefore analysed for any potential confounding effect (see Table 18). There were no significant group differences in years of formal education; global cognitive functioning (DRS-2); anxiety (HADS A); or depression (HADS D) (all $p. > .05$). The persistent na-MCI group had significantly lower estimated premorbid IQ (WTAR) than the healthy control group ($p. = .045$) although differences were not detectable post hoc. Despite this difference, all groups had above average premorbid IQ. Gender ratio was equally distributed between groups ($\chi^2 (3) = 2.97, p. = .396$). The persistent a-MCI and persistent na-MCI groups were significantly older on average than the recovered, and healthy control groups ($p. < .001$). Exploratory analysis revealed a significant linear relationship between age and all trials of the RAVLT, therefore age was entered as a covariate for further analyses.
Table 18.

*Group Differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS Scores*

<table>
<thead>
<tr>
<th>Measure</th>
<th>persistent a-MCI</th>
<th>persistent na-MCI</th>
<th>recovered</th>
<th>healthy control</th>
<th>p.</th>
<th>Post-hoc (at p.&lt;.05)</th>
<th>Effect size ($n_p^2$)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>15</td>
<td>36</td>
<td>29</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>77.80 (7.46)</td>
<td>74.58 (5.38)</td>
<td>68.07 (5.26)</td>
<td>70.93 (6.30)</td>
<td>&lt;.001</td>
<td>A,N&gt;R,H</td>
<td>.248</td>
<td>.999</td>
</tr>
<tr>
<td>Education</td>
<td>13.20 (3.86)</td>
<td>13.33 (3.28)</td>
<td>14.90 (3.26)</td>
<td>14.72 (3.88)</td>
<td>.173</td>
<td></td>
<td>.046</td>
<td>.432</td>
</tr>
<tr>
<td>WTAR (est. FSIQ)</td>
<td>108.67 (7.44)</td>
<td>106.67 (7.41)</td>
<td>110.03 (5.34)</td>
<td>111.07 (5.82)</td>
<td>.045</td>
<td>Insufficient power</td>
<td>.073</td>
<td>.656</td>
</tr>
<tr>
<td>DRS-2 (AEMSS)</td>
<td>10.67 (2.80)</td>
<td>11.56 (2.65)</td>
<td>11.59 (2.03)</td>
<td>12.38 (1.86)</td>
<td>.138</td>
<td></td>
<td>.051</td>
<td>.474</td>
</tr>
<tr>
<td>DRS-2 (raw)</td>
<td>135.47 (7.64)</td>
<td>139.22 (3.87)</td>
<td>140.48 (2.50)</td>
<td>141.17 (2.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS A</td>
<td>5.33 (2.38)</td>
<td>5.53 (3.08)</td>
<td>5.10 (3.31)</td>
<td>5.07 (2.78)</td>
<td>.919</td>
<td></td>
<td>.005</td>
<td>.080</td>
</tr>
<tr>
<td>HADS D</td>
<td>3.47 (2.26)</td>
<td>3.36 (2.61)</td>
<td>2.72 (2.40)</td>
<td>2.59 (2.06)</td>
<td>.446</td>
<td></td>
<td>.025</td>
<td>.241</td>
</tr>
</tbody>
</table>

WTAR = Wechsler Test of Adult Reading; est FSIQ = estimated Full Scale Intelligence Quotient; DRS-2 = Dementia Rating Scale-2 (Age and Education corrected); HADS = Hospital Anxiety and Depression Scale (A=Anxiety score; D = Depression score). A = persistent amnestic; N = persistent non-amnestic; R = recovered; H = healthy control
A repeated measures ANCOVA with age as a covariate revealed a significant main effect for group ($F_{(3,104)} = 11.57, p < .001, \text{power} = 999, \eta^2 = .250$) and a significant RAVLT trial x group interaction (Pillai’s trace =.303, $F_{(12,309)} = 2.89, p < .001, \text{power} = .989, \eta^2 = .101$). The main effect for RAVLT trial was non-significant (Pillai’s trace =.075, $F_{(4,101)} = 2.06, p = .092, \text{power} = .596, \eta^2 = .075$). Analysis identified a significant main effect for age ($F_{(1,104)} = 15.34, p < .001, \text{power} = .999, \eta^2 = .129$) but a non-significant group x age interaction (Pillai’s trace =.022, $F_{(4,101)} = .570, p = .685, \text{power} = .184, \eta^2 = .022$).

Post hoc comparisons of the main effect for group showed that the persistent a-MCI group performed the poorest compared to all other groups (all $p < .001$). The na-MCI group also performed more poorly than the healthy control group ($p < .001$) (see Figure 20).
Post hoc ANCOVAs with R-E-G-W procedure were used to examine the significant interaction. At trial 1, the persistent a-MCI group recalled significantly fewer words than the persistent na-MCI, healthy control and recovered-MCI groups; the persistent na-MCI recalled fewer words than the healthy control and recovered groups; and the healthy control recalled significantly fewer words than the recovered-MCI group ($F_{(3,105)} = 75.69, p < .001$, power = 1.00, $n_p^2 = .684$). The pattern of group differences on remaining trials was identical, with the
the persistent a-MCI group recalling significantly fewer words than all other groups; and the persistent na-MCI group recalling significantly fewer words than the recovered-MCI and healthy control groups trial 2 ($F_{(3,105)} = 136.22, p. < .001$, power = 1.00, $n_p^2 = .796$); trial 3 ($F_{(3,105)} = 117.28, p. < .001$, power = 1.00, $n_p^2 = .770$); trial 4 ($F_{(3,105)} = 206.51, p. < .001$, power = 1.00, $n_p^2 = .855$); and trial 5 ($F_{(3,105)} = 415.35, p. < .001$, power = 1.00, $n_p^2 = .922$).

Overall, there was a reduction of the learning curve of the persistent a-MCI group compared to all other groups across trials (see Figure 20). In contrast, the persistent na-MCI, recovered-MCI, and healthy control groups demonstrated a learning curve that increased steadily across each trial. As persistent a-MCI and persistent na-MCI groups were significantly different to recovered and healthy control groups, effect sizes were calculated to interpret the clinical significance of these differences (see Table 19).

Table 19.

<table>
<thead>
<tr>
<th>RAVLT trial</th>
<th>A - H</th>
<th>N - H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.43</td>
<td>.23</td>
</tr>
<tr>
<td>2</td>
<td>.58</td>
<td>.36</td>
</tr>
<tr>
<td>3</td>
<td>.65</td>
<td>.25</td>
</tr>
<tr>
<td>4</td>
<td>.70</td>
<td>.33</td>
</tr>
<tr>
<td>5</td>
<td>.75</td>
<td>.33</td>
</tr>
</tbody>
</table>

A = Persistent amnestic-MCI; N = Persistent non-amnestic MCI; H = Healthy Control
Discussion

Findings from present study indicate that participants identified as persistent a-MCI demonstrated a significantly lower learning curve on a verbal episodic memory task compared to participants with persistent na-MCI, healthy controls and those who displayed a recovered form of MCI. The effect sizes of these differences were large and increased in magnitude across the five trials (all $f > .4$). Findings also indicated that the learning curve of the persistent na-MCI group was lower compared to the recovered and healthy control groups. However, the persistent na-MCI group still demonstrated a consistent increase across recall trials. Further, the magnitude of these differences did not accumulate over trials and did not reach the large effects that were evident from the persistent a-MCI group.

Overall, reduced learning was most prominent in the persistent a-MCI group. Performance at trial 1 suggests that immediate memory span of the persistent a-MCI group was significantly lower than the persistent na-MCI, recovered-MCI, and healthy control groups. However, larger effects emerged as the trials progressed, particularly in terms of trials 2-5. The persistent a-MCI group learning profile remained significantly lower than all other groups indicating that they were unable to benefit from the repeated presentations of the word list to the extent as was evident in other groups. The finding that trial learning was poorest within the persistent a-MCI subtype is consistent with the limited research detailing learning profiles in MCI (Price et al., 2010; Ribeiro et al., 2007).

Research suggests that poor performance during learning tasks could be due to ineffective strategy use at the encoding stage (Germano & Kinsella, 2005; Price et al., 2010; Ribeiro et al., 2007). In a comparison of MCI, AD, and healthy older adults, Ribeiro et al. (2007) reported that the MCI group had greater acquisition over trials of a learning task compared to
the AD group, but significantly poorer acquisition than the healthy older adults. Although the MCI group displayed evidence of semantic clustering, Ribeiro et al. (2007) found that they were much less effective at applying this strategy to enhance their short term storage. In a recent study, Price et al. (2010) attempted to examine the mechanisms that underpin learning difficulties in a-MCI. Price et al. (2010) found that the a-MCI group were less effective at utilising semantic clustering to enhance learning when compared to healthy older adults. While the present study was not designed to assess strategy use in MCI, the above studies suggest that difficulty implementing strategies to enhance short term memory may partially explain the flattened learning curve of the a-MCI group in the present study. Further, as neither study contained longitudinally confirmed cases of MCI, the role of poor strategy use in the results of the present study remain speculative.

Ultimately, understanding the underlying cause of learning difficulties in MCI has less clinical utility than understanding how learning difficulties may assist in the diagnosis of MCI. At this preliminary stage, the results of the present study indicate that depressed verbal learning curves are evident in longitudinally persistent amnestic MCI, but not in longitudinally persistent non-amnestic MCI or in those individuals who display recovery from MCI. The notion of recovery from MCI is at odds with a concept of MCI as a prodrome to dementia. As such, recovery from MCI most likely represents cases of diagnostic error (false positive diagnosis) rather than genuine cases of MCI. That it is not yet possible to delineate cases of MCI that will recover from those that will not highlights the importance of research examining potential cognitive markers that differentiate these groups and may lead to a reduction in false positive diagnosis of MCI. Future research may extend these findings to identify optimal cut off scores for RAVLT learning trials with adequate sensitivity and specificity for differentiating between those individuals with apparent MCI who will recover
from those who will remain MCI who are presumably at a higher risk of progression. Such investigations could also include an analysis of a delayed recall measure, particularly given that a previous study has highlighted the benefit of assessing MCI cohorts on both trial learning and retention (Chang et al., 2010).

A major strength of the present study was that it minimised the heterogeneity characteristic of single time point classification by establishing stability of MCI classification across multiple time points. Within the present study, group differences between the persistent a-MCI and others revealed consistently large effects, which provide further support for future investigations examining trial data of longitudinally confirmed MCI cases. However, several limitations must be considered in light of the present findings. Firstly, the persistent a-MCI sample was small; meaning that extrapolation from this group must be approached with caution. Secondly, we were not able to differentiate between those with single or multiple domains impaired due to the size of the groups. This may mean that severity of deficits could be underestimated in some groups and overestimated in others. In our attempt to circumvent the potential issue of circularity, we may have confounded classification by using a separate test battery at screening to that at times 1 and 2. While care was taken to use psychometric measures with established reliability and validity, variation between the psychometric properties (e.g. reliability, normative samples) of individual tests may have influenced classification at single time points. However, as the stability classification was based on consistency of classification across all time points, we are confident that we have captured those with genuine impairments that would be evident on any reliable neuropsychological measure. In addition, although RAVLT trial data was not used to classify MCI at any stage, performance across the learning trials correlates with the measures used for classification (delayed recall) and therefore represents a potential confound. While future studies could
attempt to avoid this by using different measures for classification, tests of related domains are likely to correlate regardless. Finally, without full medical assessment, participants with undiagnosed metabolic or other causes of cognitive impairment could not be excluded.

The above results suggest that when MCI classification is established through comprehensive longitudinal neuropsychological examination, persistent a-MCI is associated with a reduced learning curve on a verbal learning task. Identifying an indicator of memory performance that differentiates persistent MCI from other MCI variants is of significant importance to both clinicians and researchers. Research attempting to isolate biomarkers predictive of developing dementia in MCI populations may be limited by the heterogeneity associated with current classification criteria. Likewise, clinical studies exploring possible pharmacological and non-pharmacological interventions for MCI groups are likely to be affected by the heterogeneity of MCI outcomes. Whilst the use of a test-retest protocol for classifying MCI is the preferred approach (Albert et al., 2011), such approaches are cost ineffective and time-consuming for larger scale trials. Ideally, establishing a single point measure with cut off scores that diagnose persistent MCI with a high degree of sensitivity and specificity is preferable. Serial list learning tasks, such as the RAVLT and CVLT, enable multiple measures of memory function, such as learning, immediate recall and delayed recall, from a single test delivered at a single time point. Evidence of impairments across both learning and recall may differentiate persistent a-MCI from other forms of MCI at a single time point assessment, obviating the need for repeat testing to establish classification stability. Further research is required to investigate the predictive utility of trial data in identifying those who are likely to maintain MCI status versus those who will recover, and those likely to convert to AD.
Chapter 8

Reducing false positive MCI diagnoses with neuropsychological assessment
Abstract

Background  Longitudinal studies of MCI report that a sizeable proportion of MCI cases revert to normal levels of functioning over time. The rate of recovery from MCI indicates that existing MCI diagnostic criteria result in an unacceptably high rate of false positive diagnosis and lack adequate sensitivity and specificity.

Method  The aim of the present study was to identify a set of neuropsychological measures able to differentiate between true positive cases of MCI from those who were unimpaired at 11 month follow up.

Results  A discriminant function analysis identified that a combination of measures of complex sustained attention, semantic memory, working memory, episodic memory, and selective attention, correctly classified outcome in more than 80% of cases. The rate of false positive diagnosis (5.93%) was considerably lower than evident in previously published MCI studies.

Discussion  The results of the present study indicate that the rate of false positive MCI diagnosis can be significantly reduced through the use of sensitive and specific neuropsychological measures of memory and non-memory functions.
Introduction

Petersen and colleagues (Petersen et al., 1995; Petersen et al., 1997; 1999) first introduced the notion of diagnostic criteria for Mild Cognitive Impairment (MCI) after longitudinal observations revealed that memory impaired older adults converted to Alzheimer’s dementia (AD) at an elevated rate. Since the initial description of MCI diagnostic criteria, revisions and recommendations have been proposed (Albert et al., 2011; Winblad et al., 2004). However, a critical flaw of the existing MCI diagnostic criteria is the absence of clinical evidence of diagnostic sensitivity and specificity. Rather, existing diagnostic criteria were devised from epidemiological studies examining risk factors for dementia in apparently cognitively intact older adults (Petersen et al., 1995; Petersen et al., 1997; Petersen et al., 1999). These epidemiological studies have been critical in understanding the presence of subclinical cognitive changes that represent risk factors for later dementia development. Some studies indicate that episodic memory dysfunction is a prominent feature (e.g. Albert et al., 2007; Lange et al., 2002; Mickes et al., 2007), whereas others suggests that non-memory deficits (e.g. executive processes) are more indicative of those likely to convert (e.g. Rapp & Reischies, 2005). Alternatively, other research implies that combined memory and non-memory dysfunction (e.g. attention, working memory, semantic memory) is better at predicting future cognitive decline (e.g. Amieva et al., 2004; Bozoki et al., 2001; Chapman et al., 2011; Summers & Saunders, 2012; Tabert et al., 2006).

Consistently across these studies is the finding that not all cases of MCI convert to dementia, with many reporting that a sizeable proportion of individuals return to age appropriate levels of functioning over time (e.g. Brodaty et al., 2012; de Jager & Budge, 2005; Han et al., 2012; Ravaglia et al., 2006; Summers & Saunders, 2012). Rates of recovery vary across clinical and
community studies, with some reporting rates as low as 7% (Loewenstein, Acevedo, Agron, Martinez, & Duara, 2007) or as high as 48% (de Rotrou et al., 2005). If “...MCI is a clinical diagnosis which is the same as are the diagnoses of dementia or AD” (Petersen, 2004, p. 193), recovery is not consistent with an MCI diagnosis. Rather, cases of recovery represent false positive diagnoses and create a significant source of variability for research in this area (Brooks et al., 2008). The high rate of false positive diagnoses of MCI arising from current diagnostic criteria highlights the errors that emerge when risk factors for dementia are erroneously utilised as clinical diagnostic criteria. In order to improve the resolution of current diagnostic criteria, clinical studies of adults displaying MCI are required. Sensitive and specific diagnostic criteria for MCI are essential to facilitate research examining factors that predict conversion to dementia and the potential ways in which this process may be slowed or prevented.

With the majority of MCI studies focusing on predicting the trajectory to dementia, fewer studies have examined the factors that reliably identify individuals with MCI (i.e. true positive cases). Summers and Saunders (2012) identified several neuropsychological measures that were useful in predicting stable MCI, recovery, conversion (to AD), and those who remained unimpaired. These measures spanned multiple domains including visual and verbal episodic memory, working memory, and attentional processing; and successfully classified 88% of MCI cases; 96% of controls; 65% of recovered cases and 100% of converters (AD). These results suggest that predicting outcome is enhanced via a comprehensive protocol that spans multiple cognitive domains. J. Mitchell et al. (2009) found that the combination of a visual memory measure (PAL) and a dementia screening tool (ACE) was the best at predicting whether individuals would convert, recover, or maintain MCI status at follow up. Mitchell and colleagues argue that this combination was successful
because of the challenging nature of the PAL task compared to traditional verbal memory measures; and because the ACE taps a range of cognitive domains including episodic memory, semantic memory, visuospatial skills and executive functioning, all of which can be compromised in early AD. The study reported 87.5% of recovered cases correctly classified and 66.7% of converters correctly classified, although it was not clear how many cases remained within the bounds of MCI. The pattern emerging from these studies is that accuracy in identifying MCI can be improved by using measures that tap memory and non-memory domains.

The aim of the present study was to identify a set of neuropsychological measures from a comprehensive neuropsychological battery that accurately differentiate between individuals with stable MCI (true positive) from those who remain unimpaired (true negative) over time. By increasing the accuracy with which true positive cases of MCI are identified, the rate of false positive MCI diagnoses evident in current research will be reduced. We examined data from a longitudinal study tracking the cognitive profile of an MCI cohort to identify specific neuropsychological measures that predict outcome (MCI versus unimpaired). Each participant in the study was classified as MCI or unimpaired (non-MCI) according to the stability of their neuropsychological profile across three assessment points. As findings from conversion studies suggest that combined memory and non-memory protocols enhance predictive accuracy, it was predicted that a similar spectrum of measures would be useful in predicting MCI and unimpaired status in the present study.
Method

Participants

The participants in this study comprised a sample of community-residing adults recruited in 2010 as part of a longitudinal project tracking the neuropsychological profile of MCI. Figure 21 shows the progression of participants from recruitment to the final assessment phase (time 2). Control participants were recruited from a similar age cohort in the wider community with no significant medical history or subjective cognitive complaint and displayed no evidence of subclinical impairment at screening assessment. The participant pool in the present study is a distinct cohort from cohorts reported in previous publications from our group (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Each participant provided informed consent prior to the commencement of the study, in accordance with the Human Research Ethics Committee (Tasmania) Network and National Health and Medical Research Council (NHMRC) of Australia Human Research Guidelines and the Helsinki Declaration.

At initial recruitment \((n = 286)\), each participant underwent telephone pre-screening to ensure that there were no medical, neurological, or psychological conditions that would impact their participation. A total of 200 participants successfully passed pre-screening and were subsequently assessed on a comprehensive neuropsychological screening battery (see Table 20) to classify participants according to existing MCI criteria (Winblad et al., 2004). Participants were then assessed at two further time points (time 1 and time 2) on a comprehensive battery utilizing measures distinct from those used in the screening battery but assessing the same cognitive domains (see Table 20). This was done to reduce the potential
influence of circular logic into the classification process, an issue that has been previously raised by some MCI researchers (Price et al., 2010; L. J. Ritchie & Tuokko, 2010). Following each assessment point, the performance of each individual participant was examined by an experienced clinical neuropsychologist (MS) to determine if the performance met the criteria for MCI. Participants were classified as subclinically impaired where the performance was more than 1.28SD (<10th percentile) below age- and/or education based norms in accordance with previously established protocols (N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Classification of MCI was based on the presence of subclinical impairments on one or more memory or non-memory task (Albert et al., 2011; N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). Mean follow up time from screening assessment to time 1 was nine months ($SD=3$ months). Mean follow up time from time 1 and time 2 assessments was 11 months ($SD=1$ month). Because the current study was interested in predicting stability outcome, only the 118 participants (overall retention of 90.8%) who were assessed at all three time points were included in the current analysis.
Figure 21. Participant progression chart
Table 20.

Test Batteries for Screening and Follow-up Testing

<table>
<thead>
<tr>
<th>Screening assessment</th>
<th>Time 1 assessment</th>
<th>Time 2 assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTAR</td>
<td>DRS-2</td>
<td>DRS-2</td>
</tr>
<tr>
<td>HADS</td>
<td>Boston Naming Test (BNT)</td>
<td>Paired Associates Learning (PAL)</td>
</tr>
<tr>
<td>DRS-2</td>
<td>Simple Reaction Time (SRTI)</td>
<td>RAVLT (List C)</td>
</tr>
<tr>
<td>Digit Symbol Coding (DSC)</td>
<td>Choice Reaction Time (CRTI)</td>
<td>Simple Reaction Time (SRTI)</td>
</tr>
<tr>
<td>Logical Memory (LMI)</td>
<td>Digit Span (DSP)</td>
<td>Choice Reaction Time (CRTI)</td>
</tr>
<tr>
<td>RCFT-copy</td>
<td>Spatial Span (SSP)</td>
<td>Digit Span (DSP)</td>
</tr>
<tr>
<td>Digit Span (DSP)</td>
<td>Letter Number Sequencing (LNS)</td>
<td>Rapid Visual Processing (RVP)</td>
</tr>
<tr>
<td>Spatial Span (SSP)</td>
<td>Paired Associates Learning (PAL)</td>
<td>Letter Visual Processing (LNS)</td>
</tr>
<tr>
<td>RCFT-delay recall</td>
<td>HADS</td>
<td>Matching to Sample (MTS)</td>
</tr>
<tr>
<td>Letter Number Sequencing</td>
<td>RAVLT (List A)</td>
<td>Spatial Span (SSP)</td>
</tr>
<tr>
<td>Trail Making Test A &amp; B (TMT)</td>
<td>Rapid Visual Processing (RVP)</td>
<td>HADS</td>
</tr>
<tr>
<td>Logical Memory II (LMII)</td>
<td>Spatial Working Memory (SWM)</td>
<td>Spatial Working Memory (SWM)</td>
</tr>
<tr>
<td>Stroop-Victoria version</td>
<td>Matching to Sample (MTS)</td>
<td>Boston Naming Test (BNT)</td>
</tr>
<tr>
<td>Vocabulary (VOC)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 subtest from the Wechsler Memory Scale, 3rd edition (WMS-III)
2 subtest from the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III)
3 subtest from the Cambridge Automated Neuropsychological Test Assessment Battery (CANTAB)

HADS = Hospital Anxiety and Depression Scale; WTAR = Wechsler Test of Adult Reading; DRS-2 = Dementia Rating Scale, 2nd edition; RCFT = Rey Complex Figure Test; RAVLT = Rey Auditory Verbal Learning Task
At the conclusion of Time 2, all participants received a stability classification based on the consistency of their classifications across screening, time 1, and time 2 (see Figure 22). Participants who met the criteria for MCI (≤10\textsuperscript{th} percentile on one or more memory and/or non-memory test) at all three time points were given the stability classification of \textit{MCI}. In addition, those who met the criteria for unimpaired at screening but went on to meet MCI criteria at time 1 or 2 were given the stability classification of \textit{MCI}. Participants who met the criteria for healthy control (≥10\textsuperscript{th} percentile on all tests) were given the stability classification of \textit{unimpaired}. In addition, those who met the criteria for MCI at screening but continued to show intact performance at time 1 and 2 were placed in the \textit{unimpaired} group. This decision was made on the basis that recovery of function is inconsistent with the definition of MCI as a clinical diagnosis. The 118 participants who completed all three time points formed the following groups: MCI (n = 60; 33 females) and unimpaired (n = 58; 39 female).
Participants were assessed on a comprehensive neuropsychological test battery at three time points (see Table 20). These assessments were part of a detailed longitudinal study examining the profile of MCI over time. Measures from the screening protocol were utilised solely for the purpose of classifying individuals into the study as well as contributing to the final stability classification.

**Figure 22.** Stability classification procedure

**Materials**

Participants were assessed on a comprehensive neuropsychological test battery at three time points (see Table 20). These assessments were part of a detailed longitudinal study examining the profile of MCI over time. Measures from the screening protocol were utilised solely for the purpose of classifying individuals into the study as well as contributing to the final stability classification.
DNA extraction

Participants provided a sputum sample and DNA extraction was performed using ORAGENE PrepIT extraction kit to manufactures instructions (DNAdenotek cat. no. OG-500).

PCR

The multiplex amplification refractory mutation system PCR (ARMS-PCR) was used to genotype DNA samples for the ApoE gene. ARMS primers were designed according to Donohoe, Salomäki, Lehtimäki, Pulkki, and Kairisto (1999). This PCR required 2 reaction mixes to allow genotype determination. The ARMS primers (designed to bind to Cys 158/Arg158 to produce a 588bp amplicon, and bind to Cys 112/Arg 112 to produce a 451bp amplicon) had matched Tm values, low self-commentary and matched GC content (Donohoe et al., 1999). Each PCR reaction mix contained the following (see below for sequences): saliva DNA (1μL), 1x REDExtract-N-Amp™ PCR ReadyMix™ (Sigma, catalogue no. R4775-1.2ML), ARMS common primer (2μM), and forward and reverse α antitrypsin positive control primers (0.05μM). Reaction mixture A contained, in addition to the above, Cys 112 (1μM) and Cys 158 (2μM) primers. Similarly, reaction mixture B contained Arg 112 (1μM) and Arg 158 (2μM) primers. Enough DEPC water was added for a final volume of 12μL. For each reaction (A and B) a no template control (NTC) containing 1μL DEPC water was included as a negative control. PCR amplification of DNA was performed with the initial denaturation stage at 95°C for 4 minutes, 35 cycles with denaturation at 96°C for 45 seconds, and annealing at 65°C for 45 seconds. The extension stage was carried out at 72°C for 45 seconds, followed by a final cycle of extension at 72°C for 5 minutes. Samples were held at 11°C until removed.
Custom DNA primer sequences from GeneWorks:

Arg112 rev 5’-CGCGGACATGGAGGACGTTC-3’
A158 rev 5’-ATGCCGATGACCTGCAGAATC-3’
ARMS primer 5’-GTTCAGTGATTGTGCCTGGCA-3’
Cys 158 5’-ATGCCGATGACCTGCAGAATT-3’
Cys 112 5’-CGCGGACATGGAGGACGTTT-3’
Alpha-AT for CCCACCTTCCCCCTCTCTCCAGGAAATGGG
Alpha-AT rev GGGCCTCAGTCCCAACATGGCTAAGAGGTG

Gel electrophoresis

Amplified multiplex products were run on a 2% agarose gel and visualized using SYBR® SAFE DNA Gel Stain (Life Technologies, catalogue no. S33102). The amplicons were sized against Quick-Load® 100 bp DNA Ladder (New England Biolabs, catalogue no. N04675). Gels images were captured using the Chemi Doc XRS scanner, and visualised using QuantityOne software. In order to determine genotype of samples using the ARMS PCR, the reference gel from Donohoe et al. (1999) was utilised. Analysis of genotype was performed with the observer blinded to participant MCI status.

Procedure

Individual testing sessions were conducted in a well-lit, well ventilated room at the School of Psychology Newnham campus. Tests were administered in the order indicated in Table 20. Each assessment took approximately 2.5 hours including a mandated 10 minute rest break in the middle of the test sequence to minimise fatigue effects. Additional rest breaks were provided as needed for each individual participant.
Results

All results were analysed using SPSS for Windows (version 19.0). To predict group membership (MCI vs. unimpaired), discriminant function analysis (DFA) was deemed most appropriate due to the small sample size (Hair, Anderson, Tatham, & Black, 1995). As the prediction of stability status (MCI vs. unimpaired) based on longitudinal classification profiles had not been previously examined, it was deemed appropriate to identify predictor variables statistically. Firstly, all time 1 outcome measures were entered into a MANOVA to control for the inflated Type 1 error rate that occurs when assessing cognitive functions from related domains. The overall MANOVA was significant (Pillai’s trace = .474, $F_{(20, 97)} = 4.38$, $p < .001$, power = 1.00, $r^2_p = .474$). Follow up one way ANOVAs indicated a number of measures significantly discriminated between the two groups. Partial correlations; inflation factor (VIF) and tolerance values were assessed to examine any potential multicollinearity issues for each of the potential predictor variables (Field, 2009; Hair et al., 1995). The following time 1 variables were identified as potential predictors due to significant group differentiation and absence of multicollinearity: RAVLT (delayed recall); RVP A’; RVP latency; LNS; BNT; SSP length; DSP backwards; SWM total errors; SWM strategy; MTS mean correct RT; Age (Table 21); and WTAR est. FSIQ (Table 21). Of the 118 participants, 113 had consented to APOE genotyping from salivary samples. The proportion of APOE e4 carriers was not significantly different between MCI and unimpaired groups ($X^2_{(1)} = .378$, $p = .539$), consequently APOE e4 status was not included as a predictor in the present model. Therefore, it was deemed appropriate to analyse the larger sample of 118 participants. Predictors were entered into a DFA which revealed a significant discriminant function ($\Lambda = .580$, $\chi^2_{(13, n=118)} = 59.67$, $p < .001$), which classified 100% of the cases and accounted for 52% of the variance in outcome group. Analysis of the structure matrix revealed ten significant predictors, RVP A’ (0.657); BNT (0.593); RAVLT-delayed recall (0.535); SSP
length (0.476); Age (-0.472); MTS mean RT (-0.416); RVP latency (-0.497); LNS (0.532); SWM errors (-0.471); PAL 6 (-0.439); and SWM strategy (-0.349) with WTAR estimated FSIQ as a poor predictor (0.241). Results indicated that 80.0% of MCI cases and 87.9% of unimpaired cases were correctly identified (see Table 22). Results also revealed that 7 (5.93%) unimpaired cases were incorrectly identified as MCI using the present model. In order to determine if classification of MCI using a DFA model improved diagnostic accuracy of MCI relative to standard diagnostic procedures (Winblad et al., 2004), we calculated the sensitivity and specificity for MCI diagnosis using existing Winblad et al. (2004) criteria for the entire sample at screening. Each participant’s diagnosis at screening was then compared to their diagnosis across time 1 and time 2. Using the Winblad et al. (2004) criteria it was found that 29 (24.58%) unimpaired cases were incorrectly identified as MCI, confirming that the classification of MCI cases using the DFA model resulted in a markedly lower false positive diagnostic rate of 5.93% (see Table 23).
Table 21.

Demographic Information and APOE E4 Status

<table>
<thead>
<tr>
<th></th>
<th>MCI</th>
<th>Unimpaired</th>
<th>p value</th>
<th>Post hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>56</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>73.55</td>
<td>68.57</td>
<td>&gt;.001</td>
<td>MCI &gt; U</td>
</tr>
<tr>
<td>Education</td>
<td>13.18</td>
<td>14.81</td>
<td>.012</td>
<td>MCI &lt; U</td>
</tr>
<tr>
<td>WTAR est FSIQ</td>
<td>107.63</td>
<td>110.55</td>
<td>.015</td>
<td>MCI &lt; U</td>
</tr>
<tr>
<td>DRS-2</td>
<td>11.07</td>
<td>12.07</td>
<td>.016</td>
<td>MCI &lt; U</td>
</tr>
<tr>
<td>HADS A</td>
<td>5.68</td>
<td>5.14</td>
<td>.318</td>
<td></td>
</tr>
<tr>
<td>HADS D</td>
<td>3.22</td>
<td>2.55</td>
<td>.118</td>
<td></td>
</tr>
<tr>
<td>APOE e4 carrier</td>
<td>55.6%</td>
<td>44.4%</td>
<td>.539</td>
<td></td>
</tr>
<tr>
<td>Gender (Male)</td>
<td>45%</td>
<td>33%</td>
<td>.173</td>
<td></td>
</tr>
</tbody>
</table>

HADS = Hospital Anxiety and Depression Scale; WTAR = Wechsler Test of Adult Reading; DRS-2 = Dementia Rating Scale, 2nd edition; APOE e4 genotyping n=102

Table 22.

Outcome Classification Predicted from Time 1 Neuropsychological Scores

<table>
<thead>
<tr>
<th>Predicted outcome (from time 1 scores)</th>
<th>MCI (% n)</th>
<th>Unimpaired (% n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCI</td>
<td>48 (80%)</td>
<td>12 (20%)</td>
</tr>
<tr>
<td>Unimpaired</td>
<td>7 (12.1%)</td>
<td>51 (87.9%)</td>
</tr>
</tbody>
</table>

99 out of 118 correctly classified (83.90%)
Table 23.

Sensitivity and Specificity of MCI Diagnosis Using Standard Diagnostic Procedures (Winblad et al., 2004) and DFA analysis

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>Winblad et al. (2004)</td>
<td>n = 40 (33.90%)</td>
</tr>
<tr>
<td></td>
<td>DFA</td>
<td>n = 48 (40.68%)</td>
</tr>
<tr>
<td>False</td>
<td>Winblad et al. (2004)</td>
<td>n = 28 (23.73%)</td>
</tr>
<tr>
<td></td>
<td>DFA</td>
<td>n = 7 (5.93%)</td>
</tr>
</tbody>
</table>

69 out of 118 correctly classified (58.47%) using the Winblad et al. criteria; 99 out of 118 correctly classified (83.90%)

Discussion

The aim of the present study was to examine the ability of neuropsychological measures from a single assessment to predict group classification (stable MCI versus unimpaired) over time. Consistent with our prediction, the discriminant function analysis revealed that several measures tapping multiple cognitive domains accurately classified group outcome, accounting for 83.9% of all cases. A measure of target detection (RVP A’) was the largest predictor variable in the model; followed by a measure of semantic memory retrieval (BNT); and verbal delayed recall (RAVLT delayed). This suggests that episodic memory dysfunction is predictive of stable MCI but only in conjunction with subclinically impaired performances in other domains such as attentional processing and semantic memory. The model also resulted in a false positive rate of 5.93% for MCI, which is considerably lower than the rate of false positive diagnosis identified using existing diagnostic criteria (23.73%). The finding that only 5.93% of unimpaired cases were classified as MCI using the present model is also
comparatively lower than many previous studies with similar samples (see Brooks et al., 2008).

Many MCI studies adopt test protocols that include discreet measures of episodic memory function (e.g. list learning tasks, paragraph recall) but utilise broad screening measures to assess non-memory function (e.g. MMSE; ACE) (e.g. Alexopoulos et al., 2006; Brodaty et al., 2012; Gavett et al., 2009; Jungwirth et al., 2012; Loewenstein, Acevedo, Agron, & Duara, 2007; Lonie et al., 2008). The present findings argue against an amnestic-centered approach to identifying individuals on the MCI spectrum. Rather, the present results in conjunction with those of other recent studies (J. Mitchell et al., 2009; Summers & Saunders, 2012) indicate that it is essential to employ reliable and valid measures of discrete memory and non-memory functions to enhance the identification of true positive MCI cases. Previous studies have highlighted the importance of using comprehensive test batteries to enhance sensitivity and specificity in MCI classification (Arnaiz & Almkvist, 2003; Bozoki et al., 2001; Summers & Saunders, 2012). Recent recommendations state that MCI studies should include an assessment of memory and non-memory functions (Albert et al., 2011), although little detail is provided regarding the assessment of non-amnestic function. If MCI is degenerative in prognosis, then recovery of function in individuals identified with MCI is diagnostically erroneous. The results of the present study indicate that comprehensive assessment of memory and non-memory function is essential to minimizing false positive diagnosis of MCI and consequently maximizing true positive diagnosis of MCI.

The discriminant function model from the present study identified RVP A’ as the strongest predictive variable of diagnostic stability. Previous research by our group found that RVP A’ was the largest predictor variable in predicting development of AD in an MCI sample.
These results suggest that RVP A’ may be a particularly sensitive measure to MCI. The RVP task from the Cambridge Automated Neuropsychological Test Assessment Battery (CANTAB; Cambridge Cognition Ltd, 2011) requires participants to correctly identify target sequences of numbers (e.g. 3-5-7) that appear within a stream of pseudo random single numbers at a rate of 100 digits per minute. The ratio of hits to false alarms, derived from signal detection theory (see Sahgal, 1987), constitutes RVP A’. Successful RVP task performance requires a participant to maintain sustained attention, utilize working memory to hold digits in short term storage, and use selective attention to identify target sequences (Coull, Frith, Frackowiak, & Grasby, 1996). Due to the demand placed on multiple executive processes, RVP is a considered particularly challenging (Sarter & Bruno, 2002). The task complexity of the RVP task may be the basis of the good discriminatory power of RVP A’ in distinguishing subclinical impairment (MCI) from normal aging (unimpaired). It is well established that broad screening measures such as the MMSE lack sensitivity to detect sub clinical conditions like MCI because of ceiling effects (e.g. Diniz, Yassuda, Nunes, Radanovic, & Forlenza, 2007). Complex cognitive tasks, such as RVP, may have greater discriminatory power because they are more cognitively taxing and have a higher ceiling. Further, recent research indicates that deficits to executive control are evident in early AD (Baddeley et al., 2001; Perry & Hodges, 1999), and MCI (Belleville et al., 2007; N. L. J. Saunders & Summers, 2010, 2011), and may be predictive of future cognitive decline (Rapp & Reischies, 2005). Sensitive measures of executive function, such as RVP, may therefore be assessing an underlying common diagnostic marker of MCI.

Unlike previous studies, individuals in the present study were defined as MCI or unimpaired according to their neuropsychological performance across three assessment time points. By using this method, we clinically identified a sample of participants with longitudinally stable
MCI. We were then able to identify neuropsychological markers that differentiated participants with stable MCI from those who were unimpaired; markers that represent potential diagnostic features for temporally stable MCI. Despite these strengths, some limitations need to be considered. Firstly, it could be argued that the present study is limited by circularity given that the tests used to predict group outcome were also used to classify MCI at time 1 and time 2. However, stability classification was based on the consistency of performance across screening, time 1 and time 2 and included a number of measures that were not included in the final model. In addition, the use of different tests to classify and track cohorts of cognitively impaired individuals does not completely avoid the issue of circularity given that test performance on tests in the same domain are likely to be correlated (L. J. Ritchie & Tuokko, 2010; Tuokko & Frerichs, 2000).

The results of the present study indicate that measures of multiple cognitive domains predicted stable MCI or unimpaired status in over 80% of all cases. The rate of false positive diagnoses was considerably lower using a predictive model based on performance on a range of cognitive measures than using existing diagnostic criteria (Winblad et al., 2004). Reducing the rate of false positive diagnoses is imperative for MCI research, particularly for intervention trials with individuals thought to be at highest risk of future cognitive decline (i.e. true positives). The high false positive rate associated with existing MCI diagnostic criteria may mask the true efficacy of treatment or intervention for MCI. Results from the present study indicate that an MCI diagnosis should only be made on the basis of impairment to multiple domains rather than isolated memory deficits. Further, these results suggest that current MCI diagnostic criteria should be modified to include sensitive and specific measures of complex sustained attention, semantic memory, working memory, episodic memory, and selective attention.
Chapter 9

Discussion
### Table 24.

**Summary of the Main Results of the Present Thesis**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Aim</th>
<th>Key Findings</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>To explore performance of screening data by classification group in a bid to ensure that each group's performance conceptually aligned with published MCI subtypes.</td>
<td>Amnestic subtypes showed impaired memory function compared to non-amnestic subtypes. Similarly, the na-MCI group showed evidence of impairment to non-memory domains. The control group maintained their status as unimpaired.</td>
<td>Results confirmed that individuals were appropriately classified within MCI subtypes according to their performance on the neuropsychological screening assessment.</td>
</tr>
<tr>
<td>3</td>
<td>To examine whether the use of a different set of memory measures at follow up tests yields similar patterns of amnestic performance in MCI subtypes.</td>
<td>The a-MCI+ group demonstrated the poorest performance on tests of visual and verbal episodic memory. These results suggest that a-MCI+ individuals demonstrate genuine difficulty with amnestic function that is unlikely to be related to a self fulfilling prophecy created by using the same test at classification and follow up.</td>
<td>Evidence of impaired amnestic functioning via the use of different episodic memory measures indicates that circular logic in MCI assessment may be less of an issue than previously indicated. However, potential circularity issues may be reduced using different test batteries to classify and track MCI over time.</td>
</tr>
<tr>
<td>4</td>
<td>To examine working memory and attention function of MCI subtypes post screening using a separate set of neuropsychological measures.</td>
<td>Compared to other MCI subtypes, the a-MCI+ group performed poorly on measures of complex sustained attention, visual and verbal short term stores, and verbal working memory. The na-MCI group showed some reduction to complex sustained attention and visual working memory.</td>
<td>When MCI is assessed using a comprehensive test battery that is distinct from that which is used at classification, working memory and selective attention processes are compromised in individuals classified as a-MCI+. This may indicate a less favourable outcome for those identified as a-MCI+.</td>
</tr>
<tr>
<td>Chapter</td>
<td>Aim</td>
<td>Key Findings</td>
<td>Conclusion</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>5</td>
<td>To examine the longitudinal performance of MCI subtypes using comprehensive neuropsychological assessment across a range of cognitive domains.</td>
<td>The a-MCI+ group showed stable impairments to episodic memory, working memory, immediate memory span, and complex sustained attention. Conversely, a-MCI and na-MCI groups performed at a level consistent with controls.</td>
<td>The a-MCI+ subtype demonstrated consistently poorer performance across multiple domains. The a-MCI+ group were the only group to demonstrate consistency with their initial classification. The a-MCI+ group may be the only variant that represents a distinct clinical entity that is identifiable using neuropsychological measures.</td>
</tr>
<tr>
<td>6</td>
<td>To explore the ability of learning trial performance to differentiate between longitudinally established variants of MCI</td>
<td>The persistent amnestic MCI group demonstrated significantly depressed learning over trials compared to persistent na-MCI, recovered MCI and healthy controls.</td>
<td>Poor learning across trials could be a diagnostic feature of persistent amnestic MCI. It may be beneficial for future investigations to examine learning trial performance in conjunction with delayed recall performance to increase sensitivity and specificity of MCI classification.</td>
</tr>
<tr>
<td>7</td>
<td>To examine which measures from a range of neuropsychological domains would be able to predict longitudinally established outcome (MCI vs unimpaired) in an MCI cohort.</td>
<td>Measures of sustained attention, working memory, semantic memory, episodic memory, and selective attention accurately predicted outcome in over 80% of all cases. The false positive rate for this model was 5.93% compared to 23.73% which was obtained using existing diagnostic criteria (Winblad et al., 2004).</td>
<td>Stable MCI can be predicted by poor performance across multiple cognitive domains. A diagnosis of MCI should not be made on the basis of isolated memory impairment. Studies should ensure that cohorts are assessed using comprehensive protocols that cover a wide range of executive and memory functions.</td>
</tr>
</tbody>
</table>
Chapter 4 Discussion

It has been argued that by using the same neuropsychological measures to classify MCI and then track MCI over time, individuals maintain MCI status due to a self-fulfilling prophecy rather than due to genuine neuropsychological impairment (K. Ritchie & Touchon, 2000; L. J. Ritchie & Tuokko, 2010). Many MCI studies attempt to avoid this issue by using broad screening measures for assessing initial classification and distinct batteries for neuropsychological follow up. Chapter 4 represents an examination of the issue of circular logic in MCI assessment. This concept was explored by examining amnestic functioning of MCI subtypes after a screening assessment classified them according to published MCI criteria. Different tests of visual and verbal memory with equivalent reliability and validity were then used to identify whether amnestic deficits remained evident. Evidence of poor episodic memory function using a different set of neuropsychological tests would suggest that performance is a result of genuine amnestic difficulty rather than circular logic.

Chapter 4 Key findings

Findings revealed that the a-MCI+ group continued to show poor performance on tasks of visual and verbal episodic memory compared to na-MCI and control groups. There were no significant group differences between the a-MCI variant and other MCI variants, although as a group, the a-MCI individuals performed at an intermediate level between the a-MCI+ and the na-MCI/control groups. Taken together, these results show that the amnestic variants, in particular a-MCI+, continue to demonstrate poorer episodic memory performance at follow
up when different measures are used. As discussed in chapter 4, the argument of circularity of logic in MCI assessment is partly undermined by the fact that MCI is a non-directionally unstable entity (e.g. Han et al., 2012). It may also be impossible to avoid circularity entirely given that test measures used for classification and prediction will be assessing the same domains and therefore highly correlated (L. J. Ritchie & Tuokko, 2010). Tuokko and Frerichs (2000) argue that the only way to completely avoid circularity would be to adopt measures that are completely independent e.g. neuropsychological measures versus neuropathological markers. It seems unlikely that this will be possible at the MCI stage given that there is still no gold standard for ante-mortem definitive diagnosis of clinical dementia (McKhann et al., 2011). These findings suggest that studies attempting to avoid circular logic by using broad screening measures at the classification stage and follow up protocols may be misguided. Researchers wanting to reduce the potential influence of circularity can utilise distinct test batteries of discreet neuropsychological functions. However, the use of broad screening measures as a way of avoiding circularity remains questionable as such measures lack sensitivity at the sub clinical level and may lead to erroneous classification of MCI individuals.

Chapter 5 Discussion

Individuals who go on to develop AD frequently experience deficits to working memory and attention in addition to impaired amnestic function (Rapp & Reischies, 2005; Storandt, 2008). Recent research has suggested that impaired working memory and attention is also evident at the MCI stage (Belleville et al., 2007; N. L. J. Saunders & Summers, 2008, 2010; Summers & Saunders, 2012). It remains to be seen whether such deficits are evident when
comprehensive neuropsychological testing is applied; or when circularity is reduced by using working memory and attention tests distinct from those used at classification. If working memory and attention dysfunction is evident at the MCI and AD stage, this would provide further evidence to support the claim that MCI represents a genuine precursor stage to AD. The objective of chapter 5 was to examine the performance of MCI subtypes on a range of tasks assessing working memory and attention function using a battery of tests distinct from those that were used to make initial MCI classifications.

Chapter 5 Key findings

Results indicated that when MCI variants were followed up using an alternative neuropsychological test battery, the a-MCI+ group showed the poorest performance in attention regulation as well as multiple components of working memory. Specifically, examination of performance on a complex sustained attention task (RVP A’) indicated that the a-MCI+ group were more prone to making errors when attempting to identify target sequences of numbers. Additionally, the a-MCI+ group also showed evidence of reduced visual and verbal span length indicating that short term storage systems of working memory may be compromised. The a-MCI+ group also demonstrated difficulty with visual and verbal tasks tapping central executive processes. The general function of working memory is to provide temporary storage for short term information, and coordinate the allocation and prioritisation of cognitive resources (Germano & Kinsella, 2005). While it is this latter function that is consistently impaired in early AD (Baddeley, Bressi, Della Sala, Logie, & Spinnler, 1991; Belleville, Peretz, & Malenfant, 1996; Morris, 1994), the results presented in Chapter 5 suggest that individuals at the a-MCI+ stage also experience a break down in these
processes. This suggests that the a-MCI+ may be closer to the clinical stage of AD (Backman et al., 2004; Hughes et al., 2011; Summers & Saunders, 2012). Further, if MCI represent a genuine pre-clinical phase of a neurodegenerative illness, and a breakdown of multiple cognitive processes evident in the a-MCI+ group, then this may be the variant that best represents MCI as it is currently conceptualised.

**Chapter 6 Discussion**

The inception of MCI subtypes was a response to increasing reports of instability within the overall condition of MCI. It was thought that different cognitive profiles at classification (e.g. amnestic vs non-amnestic) could indicate different underlying disease processes (AD vs non-AD) (Winblad et al., 2004). However, the revision of the MCI concept into various subtypes was conceptual in basis, rather than a product of extensive clinical research (Lopez et al., 2006). As a result, there are questions regarding the validity of the MCI subtypes, highlighted by the paucity of research clarifying the neuropsychological profile of these subtypes over time. Only by longitudinally examining MCI subtypes can researchers ensure that these subtypes capture the discreet pathological entities they were initially conceptualised to represent. The aim of chapter 6 was to examine the longitudinal cognitive profile of MCI subtypes using a comprehensive neuropsychological battery to assess the domains of episodic memory, working memory, attention/executive function, and semantic memory.
Chapter 6 Key findings

Results revealed that the a-MCI+ subtype was the only variant to demonstrate a cognitive profile consistent with their initial classification. At longitudinal follow up, the a-MCI+ group demonstrated poor performance on tasks of visual and verbal episodic memory, immediate memory span, target detection, and also verbal working memory. This is consistent with previous research indicating the a-MCI+ variant is more stable and more likely to be neurodegenerative in aetiology (K. Palmer, Backman, Winblad, & Fratiglioni, 2008; Summers & Saunders, 2012). Conversely, the a-MCI/na-MCI groups demonstrated performances that aligned with the control group. The lack of differentiation between control and a-MCI/na-MCI groups may suggest that a proportion of these participants no longer meet the criteria for MCI. This aligns with research showing that single domain subtypes are more likely to revert to age appropriate levels of functioning than their multiple domain counterparts (Brodaty et al., 2012). Alternately, the deficits of the single domain profiles may be too subtle to detect using neuropsychological measures. Even at the clinical level, neuropsychological measures can miss or misdiagnose conditions due to inadequate sensitivity and specificity. While the measures of cognitive function in this study were selected on the basis of good reliability and validity as well as clinical utility, the precision of these measures at the subclinical level remains questionable. The key finding is that the a-MCI+ group performed at a level that was consistent with their initial classification. This suggests the a-MCI+ profile may be the only subtype that represents a valid diagnostic entity capable of identification via neuropsychological assessment.
Chapter 7 Discussion

There is a growing body of evidence suggesting that MCI is an inherently unstable condition (Brodaty et al., 2012; Han et al., 2012; K. Ritchie et al., 2001; Summers & Saunders, 2012), particularly when classification is made on the basis of a single time point assessment (Tian, Bucks, Haworth, & Wilcock, 2003). A way of increasing the sensitivity and specificity of MCI classification is via repeat testing to confirm classification; or via multiple assessments at a single time point across related domains (Brooks et al., 2008; Brooks et al., 2007). The aim of chapter 7 was to confirm MCI subtype by examining the consistency of performance classification over three time points. From there, groups were examined on their performances on a list learning task that provides multiple measures of memory performance. Isolating features of longitudinally established MCI may negate the use of a test retest protocol to confirm classification and help reduce the heterogeneity within MCI cohorts.

Chapter 7 Key findings

Results revealed that those classified as persistent amnestic MCI had a significantly lower learning curve compared to those with persistent non-amnestic MCI, recovered MCI and healthy controls. As the trials progressed, the effect sizes of group differences became larger indicating that the persistent amnestic MCI group were unable to benefit from the repeated nature of the list learning task compared to other groups. These findings are consistent with the limited research examining learning trial performance in MCI (Chang et al., 2010; Price et al., 2010; Ribeiro et al., 2007). The finding that individuals classified with persistent a-MCI had the poorest learning may be partly explained by a difficulty in applying effective
strategies at the encoding stage (Price et al., 2010), although current data do not permit further comment on this. Overall, findings suggest that reduced trial based learning could be a feature of those who remain on the amnestic MCI spectrum. Further research may establish an optimal cut-off for trial-based learning measures such as the RAVLT in terms of identifying those who are likely to continue on an MCI trajectory versus those who are not. While a majority of studies currently use delayed recall measures to establish amnestic functioning, the additional examination of trial based learning may help reduce the variability that continues to be evident in MCI cohort studies.

Chapter 8 Discussion

MCI clinical criteria were initially developed from epidemiological studies highlighting risk factors for developing AD (Petersen et al., 1997; Petersen et al., 1999). Subsequently, there is a lack of clinical evidence demonstrating the ability of these criteria to identify genuine cases of MCI. This may partly explain the variability in outcome of those identified as MCI, including those demonstrating reversion to age appropriate levels of functioning (Summers & Saunders, 2012). There is a pressing need to improve the sensitivity and specificity of diagnostic criteria so as to adequately differentiate between those who will remain on the MCI spectrum versus those who will not. This is particularly important for intervention studies that rely on the correct identification MCI cases (true positives) to assess treatment efficacy, as well as conversion studies attempting to highlight predictive markers of future cognitive decline (J. Mitchell et al., 2009). The aim of chapter 8 was to identify neuropsychological predictors at time 1 that would be able to differentiate between individuals identified as MCI
(true positives) or unimpaired (true negatives) based on the consistency of their performance over time.

Chapter 8 Key findings

The results of this chapter indicated that multiple measures spanning various neuropsychological domains were able to classify individuals as either MCI or unimpaired in 83.9% of all cases. These included measures of sustained attention, semantic memory, episodic memory, visual and verbal working memory, and selective attention. This is consistent with previous findings indicating that the combination of memory and non-memory measures are better at predicting MCI status than memory measures alone (J. Mitchell et al., 2009; Summers & Saunders, 2012). Importantly, using this battery of measures resulted in a false positive rate of only 5.93% which is considerably lower than the false positive rate that would arise from using the Winblad et al. (2004) criteria (23.73%). This emphasizes that the assessment of performance on a range of memory and non-memory tasks is crucial to identifying individuals who are likely to remain on the MCI spectrum versus those who are not. Studies that primarily rely on amnestic functioning to make an MCI classification may misclassify individuals and contribute to the high false positive rates reported across MCI studies. Overall, these results suggest that sensitive and specific measures of complex sustained attention, semantic memory, working memory, episodic memory, and selective attention should be incorporated into current diagnostic criteria.
General discussion

The present study followed the longitudinal trajectory of a group of individuals identified as MCI using current clinical criteria (Winblad et al., 2004). An attempt was made to reduce the potential impact of circular reasoning by screening and classifying individuals on a battery of tests (chapter 2) that was different to the battery used to assess neuropsychological functioning at two follow up points (chapters 4-6). An emphasis was placed on utilising comprehensive test batteries to ensure that the functions implicated in MCI were adequately assessed and that classifications based on test performance were accurate.

MCI is an unstable condition and should not be considered a diagnostic entity

Findings from the present series of studies demonstrate that MCI criteria do not reach the level of sensitivity and specificity required for a diagnostic entity. This thesis provides several lines of evidence to support this proposition. As outlined in chapter 8, the use of current published criteria (Winblad et al., 2004) resulted in a false positive MCI diagnosis of 23.73% of the sample. That almost a quarter of healthy individuals were incorrectly diagnosed with MCI highlights the inadequacy of current MCI criteria as clinical diagnostic criteria. Similar rates of false positive diagnoses have been revealed in studies of a similar duration (Brodaty et al., 2012; Forlenza et al., 2009; J. Mitchell et al., 2009; Sachdev et al., 2013; Summers & Saunders, 2012) which raises similar concerns regarding the diagnostic status of MCI. Findings from chapters 4-6 also suggest that the diagnostic status of MCI is questionable. At each stage of neuropsychological follow up (chapters 4-6) single domain
variants (amnestic and non-amnestic) demonstrated profiles that were more consistent with
the control group than classification within their respective subtype. This suggests that these
groups demonstrate high levels of diagnostic instability when classified using current
published criteria. Similar findings have been reported in previous studies (Brodaty et al.,
2012; Forlenza et al., 2009; Han et al., 2012; Koepsell & Monsell, 2012; J. Mitchell et al.,
2009) suggesting that MCI is a temporally unstable and heterogeneous condition.

The lack of sensitivity and specificity of MCI criteria can be traced back to their evolution.
MCI criteria were developed from a series of epidemiological studies that were successful in
identifying risk factors associated with future development of AD (Petersen, 2000; Petersen
et al., 1997; Petersen et al., 1999). However, the term “risk factor” (Petersen et al., 1999) was
subsequently substituted with “diagnostic criteria” (Petersen, 2000, 2004; Petersen, Doody, et
al., 2001; Petersen & Morris, 2005) without evidence of the sensitivity and specificity of
epidemiologically-derived risk factors as diagnostic markers of a prodromal
neurodegenerative disorder. Consequently, while these criteria were able to capture a sub-set
of older adults who would experience future cognitive decline (typically AD), they also
captured a subset of individuals who would return to age appropriate levels of functioning. If,
by definition, MCI as a diagnostic entity is a precursor phase to a neurodegenerative illness,
then individuals diagnosed with MCI must either stay stable or show evidence of
deterioration. Evidence of spontaneous recovery of function is inconsistent with the
neurodegenerative profile of AD or any other dementia This suggests that many clinicians
and researchers have used the term ‘diagnosis’ erroneously (Milwain, 2000; Summers &
Saunders, 2012). Overall, the development of clinical criteria from epidemiologically derived
risk factors has led to an inability to adequately differentiate between individuals with
stable/deteriorating deficits and those without. For this reason, MCI as it is currently defined should not be considered a clinical diagnostic entity.

The a-MCI+ profile is a better representation of the MCI condition

Findings from the present thesis suggest that the concept of MCI as a preclinical phase of a neurodegenerative illness is more adequately captured by the multiple-domain amnestic (a-MCI+) profile. This is emphasized in chapters 4-6 which show that, consistent with their profile, the a-MCI+ group perform consistently lower than other MCI variants across a range of domains. This aligns with previous research which suggests that the a-MCI+ profile is more stable than other variants (Alexopoulos et al., 2006) and is associated with the poorest prognosis (Ahmed et al., 2008; Rasquin, Lodder, Visser, Lousberg, & Verhey, 2005; Summers & Saunders, 2012; Tabert et al., 2006). Further, results from chapter 8 indicate that the identification of persistent MCI state is best predicted by a group of tests spanning multiple cognitive domains e.g. attention control, semantic memory, episodic memory, and working memory. This implies that a stable state of MCI is best characterised by the neuropsychological profile of the multiple domain amnestic variant. However, this does not negate the existence of single domain variants of MCI. As evidence indicates that AD related pathological changes occur prior to clinically recognizable symptoms (Chao et al., 2010), it is plausible that focal deficits (e.g. memory) would be evident at some stage. It may be that single domain presentations are difficult to detect because current neuropsychological measures lack the sensitivity and specificity to reliably detect single domain subclinical impairment. In other words, that the a-MCI+ profile was found to be a better representation
of stable MCI may reflect the underlying constraints of neuropsychological assessment rather than the concept itself.

The psychometric properties of clinical assessment tools limit the capacity to sensitively and specifically detect subclinical levels of impairment (Brooks et al., 2007). This problem is exacerbated when a deficit is defined as a single test performance aberration (Brooks et al., 2008; Brooks et al., 2007). The sensitivity and specificity of MCI classification increases when impairment is defined across more than one measure, and across multiple cognitive domains (N. L. J. Saunders & Summers, 2011; Summers & Saunders, 2012). For this reason, it should not be surprising that individuals classified as a-MCI+ showed consistently poorer performance during the present study (chapters 4-6). It may be that such individuals are temporally closer to the point of developing clinical symptoms of dementia. However, measures used to identify MCI are clinical tools designed to measure performance at a clinical level, typically 2 SD below the mean. Even at this level, these measures are associated with a degree of error that impacts the reliability of detecting genuine neuropsychological impairment. These issues become magnified when they are used within subclinical ranges. While these properties are well understood by neuropsychologists in the field, a majority of MCI research fails to acknowledge the psychometric limitations of neuropsychological assessment measures and how this may impact research findings.
Limitations

Follow up/study duration

Follow up periods and the overall duration of the study were comparatively short. As a result, capturing the progression of each individual may be limited, particularly those experiencing slower rates of cognitive decline. Aretouli et al. (2013) suggest that between 3-6 years may be a suitable time frame to map the progression of preclinical neurodegenerative states. A longer follow up time would be required to confirm whether a-MCI+ individuals are more likely to progress to a clinical state (e.g. AD). It would also help clarify the fate of single domain variants that appear highly unstable according to this short follow up time.

Circular reasoning and classification stability

The attempt in this thesis to avoid circular reasoning by using separate screening and experimental protocols may have introduced classification instability. With different neuropsychological measures having different rates of reliability and measurement error, variation in performances across different batteries of tests may have resulted in individuals being classified as MCI at the screening assessment when they may not have if they had been assessed on the battery of tests used at follow up. Such performance variation may reflect differences in the reliability, validity and standard error of measurement between different tests of the same cognitive domain. It is also possible that while the tests used have similarity in terms of assessing a cognitive domain, subtle differences in how that domain is assessed may have impacted on an individual’s performance. If some individuals have been
misclassified, this could influence the conclusions made about group performances (chapters 4-6). However, there was an effort made to maximize the equivalence of measures at screening and follow up. Further, as the later chapters (7 & 8) examined classification stability across all three time points, any issues with single time point classification would be considered in their final stability classification. For example, cases that met the criteria for MCI at screening but not at follow up would be labeled as recovered (chapter 7) or unimpaired (chapter 8) and would not be included in the analysis of an MCI group. This would have helped to ensure that each group displayed greater homogeneity and as a result, less performance variability. Despite the potential impact on classification stability, using a separate test battery for classification allowed for an investigation into circular reasoning in MCI research. Although these findings suggest that circular reasoning may be less of an issue than previously thought, the approach taken here represents one of the first attempts to explore this aspect in MCI research.

**MCI subtypes**

This thesis was originally intended to assess the validity of MCI subtypes. However, due to the low prevalence of individuals meeting the criteria for na-MCI+, we were unable to analyse this group separately. By collapsing the non-amnestic subtypes for analysis, the severity of deficits within this profile may be underestimated. There remains a paucity of literature examining non-amnestic variants of MCI. This may reflect a lower prevalence of non-amnestic variants in the general population or the fact that many individuals with non-amnestic deficits concurrently display amnestic difficulties and are classified as a-MCI+. Because of the small sample, it was also not possible to analyse single versus multiple
domain differences in learning performance (chapter 7). Hughes et al. (2011) have suggested that it is more important to be able to differentiate those with MCI from those without MCI. When further empirical validation of MCI variants is obtained, it is hoped that sub-typing individuals will provide useful information about underlying aetiology and prognosis (Hughes et al., 2011).

**Conversion to Dementia**

Although the present study was primarily designed to track the MCI profile over time, the rate of conversion to dementia is lower than expected. At the final stage of testing, only one case was diagnosed with possible AD. While conversion is considered to be the eventual outcome for genuine MCI cases, the rate of conversion reported in the literature is variable. Typically higher annual conversion rates are reported by memory clinic studies than with community based samples (A. J. Mitchell & Shiri-Feshki, 2009) such as is utilised in the present study. Given the 20 month study duration, it is not inconceivable that the rate of conversion would be small. Age is a prominent risk factor for conversion and the participants in the present study may be considered young-old (mean age at baseline = 70 years, $SD = 6.5$) which may also contribute to the low conversion rate observed.

**Community versus clinic samples**

Issues with sampling variation continue to be raised as potential factors for variation in MCI research. As previously mentioned, conversion rates for clinic cohorts tend to yield higher rates of conversion to AD due to greater homogeneity of symptoms and higher levels of
impairment (Panza et al., 2007). In contrast, community based studies frequently exhibit higher rates of variance and subsequently lower rates of conversion to AD (Petersen & Morris, 2005). Because the participant sample of the present study was drawn from the general community, it is possible that this has resulted in higher levels of variation. Although the present study was designed to provide careful screening of psychological and neuropsychological functioning, it may not have been as rigorous as those reported in specialist clinics which are equipped to screen out additional causes of MCI (Shankle et al., 2005). Despite this limitation, community based samples remain important because they provide information about the prevalence and trajectory of MCI in the general population (Huey et al., 2013). Further, it is common for clinic based studies to enrich their sample with particular characteristics, e.g. a-MCI only, which limits their ability to extrapolate to the broader MCI profile within the community.

**Single test failure for MCI classification**

As with many previous studies, MCI classification was made on the basis of a subclinical impairment on a single test at a single time point. Therefore, to obtain a classification of MCI, a participant need only perform poorly on one measure at screening. Issues relating to this procedure have been examined by Brooks and colleagues (2008; 2007) who discuss the frequency with which healthy older adults perform below clinical thresholds. Evidence from these studies indicates that when healthy older adults are assessed on a neuropsychological test battery, approximately 30% will register a performance that reaches clinical impairment thresholds. As discussed in chapter 6, the use of a single aberrant test performance to classify an individual with MCI may have captured individuals who do not have MCI or who are suffering from transient cause of cognitive dysfunction (e.g. anxiety, fatigue). Therefore, the
present sample may have been over-represented with individuals unlikely to progress given the short duration of the study. Interestingly, MCI criteria make no indication as to whether individuals should demonstrate more than one test aberration to receive an MCI classification. Recent recommendations (Albert et al., 2011) are also silent on the issue suggesting that it is appropriate for researchers and clinicians to make classifications on the basis of single test failures despite evidence that this may increase rates of false positive diagnoses.

**Cut off scores to define subclinical impairment**

Related to the above issue is the choice to apply the cut off score of $1.28SD (<10^{\text{th}}$ percentile) as indicative of subclinical impairment when assessing neuropsychological performance. It could be argued that applying this cut off resulted in capturing a higher proportion of healthy adults and subsequently reducing the rate of conversion. The selection of a cut off of $<10^{\text{th}}$ percentile ($<1.28 SD$) was made on the basis of previous published work (Alladi et al., 2006; J. Mitchell et al., 2009; N. L. J. Saunders & Summers, 2010, 2011; Summers & Saunders, 2012). In addition, there is no accepted cut off for neuropsychological test scores in classifying MCI. Recent recommendations for MCI criteria (Albert et al., 2011) state that a cut off ranging from 1.0-1.5 $SD$ is acceptable, indicating that the choice to use 1.28 $SD$ cut off is within the suggested range.
Group based assessment versus individual assessment

The present set of studies examined group performances on a range of neuropsychological measures over time. Some researchers suggest that using this approach for examining MCI cohorts may be less appropriate due to potential variation in cognitive profiles (Belleville et al., 2007). Specifically, it has been argued that the comparison of average group performances may underemphasize or overemphasize ability within a given domain and potentially lead to erroneous conclusions about MCI profiles. This is supported by recent findings that indicate MCI is unstable and likely to contain individuals who represent a cognitively heterogeneous group (Han et al., 2012; Summers & Saunders, 2012). However, group based approaches are effective at investigating commonalities within a syndrome. As the resolution of MCI symptoms remains unclear, a clinical group design is useful to identify common symptoms that indicate whether an individual has the disorder. While there is likely to be individual variability using this approach, individual factors are of greater interest once common symptoms have been identified. Future studies may benefit from combining group and individual approaches in order to fully understand the similarities and differences evident at the MCI stage.

Implications

The finding that stable MCI can be predicted in approximately 80% of cases using a range of memory and non-memory domains has major implications for both clinical research and practice. By classifying MCI on a comprehensive set of memory and non-memory measures (e.g. complex sustained attention; episodic memory; semantic memory; working memory and


selective attention), researchers will increase the accuracy with which they can identify individuals who will remain on the MCI spectrum (true positive cases). In turn, this will reduce the number false positive cases and subsequent variability within samples. Treatment regimes assessing the efficacy of an intervention will be less likely to make erroneous statements about treatment effectiveness if cohorts are enriched with greater proportions of genuine MCI cases. Further, the likelihood of identifying a biomarker/s predictive of future cognitive decline will increase if cohorts have greater homogeneity in terms of a behavioural profile. Despite the sizeable research effort examining prognostic biomarkers in MCI, a biomarker with sufficient sensitivity and specificity to accurately identify predementia cases has yet to be established (e.g. Grimmer et al., 2013). As biomarker studies examine individuals with MCI identified using current criteria (Albert et al., 2011; Winblad et al., 2004), the inability to identify biomarkers that accurately identify older adults who will develop dementia likely reflects the inherent lack of diagnostic sensitivity and specificity in existing MCI criteria as identified in this thesis. However, identification of an accurate prognostic biomarker may be more likely if diagnostic criteria for MCI include sensitive measures of the domains implicated in this thesis. Additional implications revealed by the present set of studies will be discussed below.

**Comprehensive test protocols**

For research protocols, comprehensive testing of memory and non-memory domains is essential for classifying and tracking MCI over time. The use of broad screening measures (e.g. measures of dementia severity,) should not be used to infer the integrity of non-memory domains. Rather, sensitive and specific neuropsychological tests of discreet cognitive functions should be used. A growing body of evidence indicates that there is increased
sensitivity and specificity of MCI classification when comprehensive neuropsychological assessment is adopted compared to briefer approaches (Artero et al., 2006; Johns et al., 2012; Ribeiro, de Mendonca, & Guerreiro, 2006; N. L. J. Saunders & Summers, 2011; Summers & Saunders, 2012; Traykov et al., 2007). This evidence suggests that the heterogeneity of outcome evident across studies of various MCI subtypes may reflect variance in accuracy of MCI classification arising from inadequate cognitive testing, rather than actual heterogeneity of outcome for sufferers of MCI. Further, it would be beneficial to investigate the issue of task complexity in measures used to examine MCI cohorts. Performance on complex tasks may have greater sensitivity and specificity in delineating subclinical impairment from unimpaired levels of function. It is possible that the use of tests designed to detect clinical impairment lack sufficient sensitivity in detecting subclinical impairment to be valid for classifying or diagnosing MCI.

Use of the term ‘mild’

There is an emerging trend for the term MCI to be used to describe a range of episodes of mild cognitive decline that may be unrelated to a prodromal phase of dementia e.g. stroke (Rasquin et al., 2004). If MCI leads to dementia, then using the same term that describes a range of conditions that may or may not lead to a clinical condition is problematic. As de Jager and Budge (2005, p. 78) suggest, it may be necessary to redefine MCI that leads to dementia by using an alternative term such as “multiple-domain progressive cognitive decline”. It could be argued that this is simply a semantic issue; however, the term MCI has become ambiguous with regards to what it describes and the likely prognosis of someone affected by MCI. Ultimately, this is unhelpful for clinicians, researchers, and those who may be classified/diagnosed with MCI.
Conclusion

This thesis has demonstrated that the concept of MCI has progressed from an epidemiologically based risk profile to a diagnostic entity (Albert et al., 2011; Petersen, 2004; Petersen & Morris, 2005; Petersen & Negash, 2008) without sufficient clinical trial evidence that it meets the appropriate level of sensitivity and specificity required for a clinical diagnosis. As a result, researchers and clinicians should avoid using the term ‘diagnosis’ when describing MCI. Importantly, this thesis has identified a syndrome of subclinical cognitive impairments that predict those who display longitudinally stable MCI. This syndrome closely aligns with the conceptualisation of the a-MCI+ subtype, a variant associated with poor performance across a range of domains including attentional control, episodic memory, semantic memory and working memory. Unfortunately, current research is not at a stage where MCI can be identified with diagnostic accuracy. This may in part reflect the psychometric limitations of neuropsychological measures. While researchers are optimistic that biomarkers will play major role in future MCI detection, sensitive and specific biomarkers will not be identifiable until the behavioural syndrome can be isolated. Research attempting to identify biomarkers for MCI is subject to the same heterogeneity experienced by neuropsychologically defined samples, limiting the likelihood of an accurate diagnostic biomarker being identified.

In conclusion, some researchers treat MCI as a prodromal dementia diagnosis (Loewenstein et al., 2012; J. C. Morris et al., 2001) whereas others deal with MCI as it was initially conceived, a research classification characterised by an increased risk for future cognitive decline (Bennett et al., 2002; Gauthier et al., 2006; N. L. J. Saunders & Summers, 2011). Alternatively, there are those who claim it has no diagnostic utility at all (Gauthier & Touchon, 2005). Results from the present thesis demonstrate that epidemiological risk factors
have poor resolution as diagnostic criteria and that the status of MCI as a diagnostic entity is questionable. This does not discount the utility of defining such a condition nor does it discredit the conceptualization of a dementia prodromal phase such as MCI. However, if the benefit of early identification is maximising the effectiveness of intervention strategies, then the emphasis should be on how those individuals are best conceptualised and the way in which we can identify them with the most accuracy. As a result, future MCI studies should prioritize the identification and follow up of individuals presenting with the a-MCI+ profile as these individuals appear more likely to demonstrate continued impairment over time. Further, the a-MCI+ profile more closely aligns with the initial conceptualization of MCI as a prodromal phase of AD and is likely to have greater diagnostic utility as a result.
References


Appendix A - Recruitment poster and pamphlet

Refer to enclosed CD-ROM disc
Appendix B - Information Sheets and Consent forms

Refer to enclosed CD-ROM disc
Appendix C - Record forms

Refer to enclosed CD-ROM disc
Appendix D - Data analyses for Chapters 3, 4, 5, 6, 7 & 8

Refer to enclosed CD-ROM disc
Appendix E - Klekociuk & Summers (2013) reprint
Clinical Study

The Self-Fulfilling Prophecy of Episodic Memory Impairment in Mild Cognitive Impairment: Do Episodic Memory Deficits Identified at Classification Remain Evident When Later Examined with Different Memory Tests?

Shannon Zofia Klekociuk¹ and Mathew James Summers¹²

¹ School of Psychology, University of Tasmania, Launceston, TAS, 7250, Australia
² Wicking Dementia Research & Education Centre, School of Medicine, University of Tasmania, TAS, Australia

Correspondence should be addressed to Mathew James Summers; mathew.summers@utas.edu.au

Received 27 April 2013; Revised 17 July 2013; Accepted 25 July 2013

Academic Editor: Changiz Geula

Copyright © 2013 S. Z. Klekociuk and M. J. Summers. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Previous studies of mild cognitive impairment (MCI) have been criticised for using the same battery of neuropsychological tests during classification and longitudinal follow-up. The key concern is that there is a potential circularity when the same tests are used to identify MCI and then subsequently monitor change in function over time. The aim of the present study was to examine the evidence of this potential circularity problem. The present study assessed the memory function of 72 MCI participants and 50 healthy controls using an alternate battery of visual and verbal episodic memory tests 9 months following initial comprehensive screening assessment and MCI classification. Individuals who were classified as multiple-domain amnestic MCI (a-MCI+) at screening show a significantly reduced performance in visual and verbal memory function at follow-up using a completely different battery of valid and reliable tests. Consistent with their initial classification, those identified as non-amnestic MCI (na-MCI) or control at screening demonstrated the highest performance across the memory tasks. The results of the present study indicate that persistent memory deficits remain evident in amnestic MCI subgroups using alternate memory tests, suggesting that the concerns regarding potential circularity of logic may be overstated in MCI research.

1. Introduction

The concept of Mild Cognitive Impairment (MCI) emerged from a series of MAYO clinic epidemiological studies attempting to identify predictive risk factors for Alzheimer’s dementia (AD) [1–3]. The utility of MCI was perceived to be its ability to identify individuals most at risk of future cognitive decline, particularly those likely to transition to AD [2]. Subsequently, the clinical features used to classify MCI have gradually been replaced by MCI diagnostic criteria [4, 5], although a number of researchers question whether these criteria lack appropriate sensitivity and specificity to be considered diagnostic [6–12]. Current MCI classification criteria include concern regarding a change in cognitive functioning; evidence of objective dysfunction (usually from neuropsychological assessment); relatively intact daily functioning; and an absence of dementia [4]. According to the diagnostic criteria outlined by Winblad et al. [5], amnestic subtypes are defined by the presence of an episodic memory deficit, whereas non-amnestic subtypes are defined by the presence of a non-memory deficit (e.g., attention, language, working memory). Both of these broad variants may be further classified as single domain (deficits are limited to one cognitive domain, e.g., episodic memory) or multiple domain (deficits are present in more than one domain, e.g., memory and attention) [5].

The aim of many MCI studies is to follow an MCI cohort over time to identify the most sensitive predictors of future cognitive decline. Some of these studies classify patients with MCI and monitor cognitive function over time using
the same battery of neuropsychological tests (e.g., [13–15]). This has introduced a concern regarding the independence of the assessment of cognitive function over time from the initial diagnosis/classification of MCI. Specifically, it raises the question as to whether individuals who maintain a specific MCI classification at follow up do so because of genuine neuropsychological impairment or because of the self-fulfilling prophecy created from using the same psychometric instruments to identify MCI [16, 17]. Other studies attempt to avoid this issue by using broad screening measures (e.g., MMSE, CERAD) to classify MCI and then track the progression of MCI cohorts using tests of discreet neuropsychological functions (e.g., [18, 19]). However, this introduces an alternative issue regarding the accuracy of the initial MCI classification. Research has revealed that broad screening measures lack the sensitivity to detect non-memory deficits (e.g., attention, language, working memory) in MCI, based on evidence that a majority of MCI cases demonstrate such deficits when assessed using reliable and valid neuropsychological measures [6, 13, 17, 20–22]. In attempt to avoid circularity, previous studies utilising restricted screening protocols may have misclassified MCI and/or missed classifying genuine cases of MCI.

The extent to which circular reasoning is an issue for the assessment of MCI remains unclear. One way of reducing its potential effects is by using a separate test battery to classify MCI, and an alternative test battery to assess cognitive function over time [23]. The present study represents an exploration into the potential issue of circular logic by examining memory function in an MCI cohort. We attempted to investigate whether amnestic dysfunction remained evident when groups were assessed using alternate tests of visual and verbal memory at screening and follow up. It was hypothesised that if circularity of logic affects MCI classification, then MCI subtypes would display a change in their memory performance across two independent neuropsychological batteries.

2. Method

2.1. Study Population. Community-residing older adults from Tasmania (Australia) were recruited using consecutive sampling from advertisements placed in local media (TV and radio) and local general medical practices. Participants were recruited to participate in a larger longitudinal study tracking the neuropsychological profile of MCI subtypes. Each participant provided fully informed consent prior to the commencement of the study, in accordance with the Human Research Ethics Committee (Tasmania) Network and National Health and Medical Research Council (NHMRC) of Australia Human Research Guidelines, in accordance with the Declaration of Helsinki (1964).

Each participant underwent pre-screening via telephone to ensure that there were no medical, neurological, or psychological conditions that would impact their participation. In addition, each participant who passed pre-screening was assessed on a clinical neuropsychological battery spanning multiple memory and non-memory domains (see Table 1). This was important to avoid previous criticisms of erroneous classification of MCI cases due to inadequate classification protocols. The aim of the screening stage was to identify those who met the criteria for MCI [5]. Performances were classified as subclinically impaired where the performance was more than 1.28SD (<10th percentile) below age- and/or education based norms in accordance with previously established protocols [6, 13, 21]. Classification of MCI subtype as single domain amnestic MCI (a-MCI), single domain non-amnestic MCI (na-MCI), multiple domain amnestic MCI (a-MCI+), or multiple domain non-amnestic MCI (na-MCI+) was based on the presence of one or more subclinical impairments to one or more cognitive domains [4, 6, 13, 21]. A total of 130 participants successfully complete pre-screening and classification screening. These participants composed the following groups: a-MCI (n = 24); na-MCI (n = 25); a-MCI+ (n = 27); na-MCI+ (n = 6); and healthy control (n = 50). Due to the statistical issues associated with analysing small samples, the na-MCI+ group were collapsed to form a larger na-MCI group. Prior to the reassessment of episodic memory, eight participants withdrew, four for personal reasons and four due to emerging chronic health issues. The final sample of 122 participants (male = 48) formed the following groups: a-MCI (n = 23); na-MCI (n = 25); a-MCI+ (n = 24); and healthy control (n = 50).

3. Materials

Participants were screened on a test battery (see Table 1) comprised of tests selected on the basis of excellent reliability and validity in clinical and subclinical populations. Follow-up episodic memory assessment (experimental) involved alternate tests of episodic memory to those used at screening assessment. Tests assessing both verbal and visual memory were included at screening and the experimental stages as research has shown that episodic memory deficits may manifest both verbally and/or visually in MCI [22]. The experimental protocol included the Paired Associates Learning Test (PAL; [30]) and the Rey Auditory Verbal Learning Test (RAVLT; [31]). The PAL is a subtest of the Cambridge Neuropsychological Test Automated Battery (CANTAB). It is a visual measure of episodic memory and learning and is sensitive to medial temporal lobe function [30]. The PAL has a demonstrated ability to accurately discriminate between individuals with AD and healthy controls as well as the capacity to predict future cognitive decline [32]. During the PAL, participants are presented with six white boxes that open up one at a time in random order. At trial one, the computer reveals two different patterns hidden in two separate boxes. The participant is required to recall the location of each pattern at the end of the presentation sequence. Correct detection of each pattern within the allocated ten attempts allows the participant to move on to the next phases where three, six, and eight patterns are hidden, respectively. Failure to recall the correct location of each pattern after 10 trials results in termination of the test. The selected outcome measures for the PAL were total errors at 6 and 8 shapes (adjusted), which report the number of errors made at each
of the stages, respectively. These outcome measures were selected as they adjust the total score for those participants who fail to meet criterion on an earlier trial and do not complete the entire PAL sequence [30]. The RAVLT is a verbal assessment of episodic memory and learning. The RAVLT consists of five consecutive learning trials of an auditory presentation of 15 item word list. Following each learning trial, participants recall as many of the 15 words in any order. After the fifth learning trial, a distracter list of 15 new words is presented followed by a recall trial. Following this, the participant is required to recall as many words possible from the initial list. Outcome measures used in the following analysis were RAVLT trial 5; RAVLT total (trials 1–5); and RAVLT delayed. The RAVLT has been found to be reliable in distinguishing between healthy controls and individuals with AD, as well as differentiating between various neurodegenerative conditions [33].

3.1. Procedure. Individual assessment sessions were conducted in a well-lit, well ventilated room and took approximately 90–120 minutes, including mandated rest breaks, to complete. Tasks assessing visual and verbal episodic memory were administered as part of a larger test battery examining the neuropsychological profile of MCI subtypes. Only results pertaining to episodic memory function were analysed for the present study. The CANTAB was administered on a laptop connected to an external 17 inch LCD touch screen monitor and response pad according to standard instructions. Participants sat approximately 50 cm from the touch screen with the response pad positioned 15 cm from the touch screen.

4. Results

Results were analysed using SPSS for Windows (version 19.0). MANOVA was used to control for potential inflation of Type 1 error due to analysing data from multiple tests within the same domain (episodic memory). Significant multivariate results were followed with one-way ANOVAs and post hoc analyses. Games-Howell was considered the appropriate post hoc analysis due to unequal sample sizes and breaches of homogeneity of variance [34].

Demographic variables were assessed to examine any potential group differences that may act as potential confounds [31] (see Table 2). No group differences were detected in terms of age, education level, or HADS depression score. Group differences were detected on the WTAR with the a-MCI+ group having a significantly lower estimate of premorbid IQ compared to all groups. Group differences were also detected on HADS anxiety score however, due to insufficient power for the medium effect size evident a post hoc analysis failed to identify significant group differences, with a trend towards significance between the a-MCI+ and Control group (\( P = 0.068 \)). Group differences in global cognitive function (DRS-2 score) were significant but in expected directions with the a-MCI+ having significantly lower scores than the control and na-MCI groups; and the a-MCI having significantly lower scores than the control group. While significant differences were found, no group had a mean DRS-2 score of clinical significance (all AEMSS ≥ 9).

There was no significant difference in gender ratio across the four groups (\( \chi^2 = 3.45, P = 0.327 \)).

A MANOVA identified significant group differences in episodic memory (PAL 6 shapes adjusted; PAL 8 shapes adjusted; RAVLT trial 5; RAVLT total; RAVLT delayed) (Pillai's trace = 0.260, \( F_{(15,348)} = 2.20, P = 0.006, \) power = 0.975, \( n_p^2 = 0.087 \)). Group differences within each dependent variable were subsequently analysed by one-way ANOVA with post-hoc Games-Howell analysis.

Significant group differences were detected on PAL 6 shapes adjusted (\( F_{(3,118)} = 6.69, P < 0.001, \) power = 0.971, \( n_p^2 = 0.145 \)) and PAL 8 shapes adjusted (\( F_{(3,118)} = 5.73, P = 0.001, \) power = 0.943, \( n_p^2 = 0.127 \)). Post hoc analyses revealed that the a-MCI+ group made significantly more errors in attempting to recall the spatial location of six patterns compared to the na-MCI and control groups (Figure 1). At eight patterns, the a-MCI+ group made significantly more errors than the control group (Figure 1).

Significant group differences were detected on RAVLT trial 5 (\( F_{(3,118)} = 6.61, P < 0.001, \) power = .969, \( n_p^2 = .144 \));
Table 2: Group differences in Age, Education, Estimated Premorbid FSIQ, DRS-2, and HADS scores.

<table>
<thead>
<tr>
<th>Measure</th>
<th>n</th>
<th>a-MCI Mean (SD)</th>
<th>na-MCI Mean (SD)</th>
<th>a-MCI+ Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>P</th>
<th>Post-hoc (at P &lt; 0.05)</th>
<th>Effect size (n²)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>23</td>
<td>70.61 (7.99)</td>
<td>70.60 (5.97)</td>
<td>69.29 (6.42)</td>
<td>72.66 (6.52)</td>
<td>0.201</td>
<td></td>
<td>0.038</td>
<td>0.404</td>
</tr>
<tr>
<td>Education</td>
<td>25</td>
<td>14.43 (3.16)</td>
<td>15.04 (3.54)</td>
<td>12.46 (3.45)</td>
<td>14.20 (3.74)</td>
<td>0.072</td>
<td></td>
<td>0.057</td>
<td>0.585</td>
</tr>
<tr>
<td>WTAR (est. FSIQ)</td>
<td></td>
<td>110.22 (5.42)</td>
<td>110.24 (4.97)</td>
<td>103.33 (7.99)</td>
<td>110.38 (5.76)</td>
<td>&lt;0.001</td>
<td>a-MCI+ &lt; na-MCI, a-MCI, C</td>
<td>0.178</td>
<td>0.993</td>
</tr>
<tr>
<td>DRS-2 (AEMSS)</td>
<td></td>
<td>10.91 (2.17)</td>
<td>11.56 (1.89)</td>
<td>10.04 (1.97)</td>
<td>12.54 (2.14)</td>
<td>&lt;0.001</td>
<td>a-MCI+ &lt; na-MCI, C; a-MCI &lt; C</td>
<td>0.183</td>
<td>0.994</td>
</tr>
<tr>
<td>HADS A</td>
<td></td>
<td>5.39 (2.79)</td>
<td>5.00 (2.68)</td>
<td>6.88 (3.71)</td>
<td>4.72 (2.56)</td>
<td>0.028</td>
<td>Insufficient power</td>
<td>0.074</td>
<td>0.719</td>
</tr>
<tr>
<td>HADS D</td>
<td></td>
<td>3.04 (2.34)</td>
<td>2.52 (2.02)</td>
<td>3.38 (2.46)</td>
<td>2.72 (2.33)</td>
<td>0.558</td>
<td></td>
<td>0.017</td>
<td>0.193</td>
</tr>
</tbody>
</table>

WTAR: Wechsler Test of Adult Reading; est FSIQ: estimated Full Scale Intelligence Quotient; DRS-2: Dementia Rating Scale-2 (Age and Education corrected); HADS: Hospital Anxiety and Depression Scale (A: Anxiety score; D: Depression score); C: control.

5. Discussion

The results of the present study indicate that individuals identified as a-MCI+ from a comprehensive screening assessment display significantly lower performances on different measures of verbal and visual episodic memory compared to control participants or those classified as na-MCI. Specifically, the a-MCI+ group made significantly more errors when attempting to recall the spatial location of patterns (PAL 6 & 8 shapes adjusted). The a-MCI+ group also recorded the poorest performance on the final trial of a verbal learning task (RAVLT trial 5); lowest delayed verbal episodic memory recall (RAVLT delay); and poorest cumulative verbal learning across trials (RAVLT total). These results may seem unsurprising given that individuals within this group, by definition of their initial classification, scored at subclinical levels (<10th percentile) on at least one memory and one non-memory test at screening. That this group performed poorly on a different set of memory measures compared to those used at screening strongly suggests that circular reasoning in MCI research may be less problematic than previously suggested.

While the a-MCI group appear to perform at an intermediate level between the a-MCI+ group and the control and na-MCI groups, these differences do not reach statistical significance. It could be argued that this is due to circular
reasoning given that a new battery of memory tests was unable to identify significant group differences. However, a better explanation of these findings relates to stability. By definition, membership to the a-MCI subtype requires a single impaired performance on a single test of episodic memory. Previous research tracking MCI subtypes longitudinally suggests that the a-MCI profile is not only rare but highly unstable [21, 22, 35, 36]. That the a-MCI group performed at an intermediate level between the a-MCI+ group and the control and na-MCI groups may be a result of recovery of function of some individuals within this subtype. Therefore, it may be erroneous to conclude with certainty about circular reasoning in this group as performance differences could be confounded by false positive cases.

The present study attempted to address the issue of circular reasoning in MCI research. The above data suggests that circular reasoning may be less of an issue given that the a-MCI+ subtype displays evidence of depressed verbal and visual episodic memory function on alternate tests conducted 9 months after initial assessment. However, it may be argued that the notion of circular reasoning within this context is flawed as it relies on the premise that MCI is stable. Research demonstrates that MCI is far from stable with consistent evidence that of recovery of function is common [6]. As a theoretical construct, if MCI is a precursor stage to dementia, it cannot be a stable entity. As a precursor to a neurodegenerative disease, one would expect that MCI should display a pattern of deterioration cognitive function(s) over time until the clinical stage of dementia is reached. As such, those identified as MCI should continue to display evidence of cognitive difficulties that have either remained stable or deteriorated over time. However, there should not be evidence recovery of function in genuine MCI cases as this would indicate erroneous classification within the MCI spectrum.

Several factors warrant caution when interpreting the above data. First, the small sample size is likely to limit the generalisability of the present findings. Second, it could be argued that the issue of circular logic may have been better assessed by including a comparison group of individuals who were assessed with the same tests at screening and follow up. However, it is not possible to obtain two identical clinical groups for comparison. Further, by adopting this approach, it would be impossible to differentiate circularity effects from group differences and therefore confound the results. Third, it could be argued that circularity is inevitable unless there is complete independence between predictors and outcome measures [37]. The use of different tests tapping the same domains is likely to result in some degree of circularity as performance is likely to be highly correlated. However, this study represents one of the first attempts to formally investigate circular reasoning in MCI and has several strengths compared to previous research. All MCI cases were assessed using a comprehensive test battery rather than the conventional approach of using screening tests to classify MCI. In addition, both visual memory and verbal episodic memory were assessed as part of the screening classification and the follow up memory assessment. Previous research that has only examined verbal memory may have inadvertently missed classifying or misclassified cases where the memory impairment was visual in nature [22]. This study also represents one of the few that have not compromised the comprehensiveness of the screening protocol by using global measures in attempt to avoid circularity.

Results of the present study show that when using different follow up tests, memory function remains compromised in individuals initially classified as a-MCI+. This suggests that circular reasoning in MCI research may be less of an issue than previously thought. Further, it implies that researchers are not justified in using broad global measures at screening to avoid the issue of circularity. Potential MCI cases should always be assessed with comprehensive test protocols that enhance diagnostic accuracy. However, future studies wanting to minimize the influence of circularity should adopt different classification and follow up protocols. More research is required as to how this procedure may impact the sensitivity and specificity of the MCI classification.

Acknowledgments

This research received no specific grant from any funding agency and commercial or not-for-profit sectors. The researchers thank the participants and their families for generously supporting this research as well as the medical practitioners from Northern Tasmania for supporting this study. Shannon Zofia Klekociuk received an Australian Postgraduate Award Scholarship and postgraduate scholarship from the Wicking Dementia Research and Education Centre (WDREC) for the duration of this study.

References


Submit your manuscripts at http://www.hindawi.com
Appendices F, G & H have been removed for copyright or proprietary reasons.
Appendix I - Manuscript submission for chapter 8
Reducing false positive diagnoses in MCI: The importance of comprehensive neuropsychological assessment

<table>
<thead>
<tr>
<th>Journal:</th>
<th>European Journal of Neurology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>EJoN-14-0158</td>
</tr>
<tr>
<td>Wiley - Manuscript type:</td>
<td>Original Papers</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>05-Feb-2014</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Klekociuk, Shannon; University of Tasmania, School of Medicine Summers, Jeffery; University of Tasmania, School of Medicine; Liverpool John Moores University, Research Institute for Sport and Exercise Sciences Vickers, James; University of Tasmania, Wicking Dementia Research and Education Centre; University of Tasmania, School of Medicine Summers, Mathew; University of Tasmania, Wicking Dementia Research &amp; Education Centre; University of Tasmania, School of Medicine</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Neuropsychology &lt; NEUROLOGICAL DISORDERS, Mild cognitive impairment &lt; Cognitive disorders and dementia &lt; NEUROLOGICAL DISORDERS, Alzheimer's disease &lt; Cognitive disorders and dementia &lt; NEUROLOGICAL DISORDERS, Geriatric &lt; PATIENT CATEGORIES, Executive function &lt; Neuropsychology &lt; NEUROLOGICAL DISORDERS, Memory &lt; Neuropsychology &lt; NEUROLOGICAL DISORDERS, Attention disorders &lt; Neuropsychology &lt; NEUROLOGICAL DISORDERS</td>
</tr>
</tbody>
</table>
Reducing false positive diagnoses in MCI: The importance of comprehensive neuropsychological assessment

Shannon Zofia KLEKOCIUK¹, Jeffery Joseph SUMMERS¹,², James C VICKERS¹,³
& Mathew James SUMMERS¹,³ *

¹ School of Medicine, University of Tasmania, Australia
² Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom.
³ Wicking Dementia Research & Education Centre, University of Tasmania, Australia

* Corresponding author: Dr M.J. Summers, School of Psychology, Locked Bag 1342, Launceston, Tasmania, Australia, 7250; tel +61 3 6324 3266; facsimile +61 3 6324 3168; email: Mathew.Summers@utas.edu.au

Abstract - 163 words
Manuscript – 3,063 words
Tables - 4
Figures – 2
References - 45
Abstract

Background  Longitudinal studies of MCI report that a sizeable proportion of MCI cases revert to normal levels of functioning over time. The rate of recovery from MCI indicates that existing MCI diagnostic criteria result in an unacceptably high rate of false positive diagnosis and lack adequate sensitivity and specificity.

Method  The aim of the present study was to identify a set of neuropsychological measures able to differentiate between true positive cases of MCI from those who were unimpaired at 11 month follow up.

Results  A discriminant function analysis identified that a combination of measures of complex sustained attention, semantic memory, working memory, episodic memory, and selective attention, correctly classified outcome in more than 80% of cases. The rate of false positive diagnosis (5.93%) was considerably lower than evident in previously published MCI studies.

Discussion  The results of the present study indicate that the rate of false positive MCI diagnosis can be significantly reduced through the use of sensitive and specific neuropsychological measures of memory and non-memory functions.

Key words: neuropsychology; stability; diagnosis; dementia; mild cognitive impairment; recovery
Introduction

Petersen and colleagues [1-3] first introduced the notion of diagnostic criteria for Mild Cognitive Impairment (MCI) after longitudinal observations revealed that memory impaired older adults converted to Alzheimer’s dementia (AD) at an elevated rate. Since the initial description of MCI diagnostic criteria, revisions and recommendations have been proposed [4, 5]. However, a critical flaw of the existing MCI diagnostic criteria is the absence of clinical evidence of diagnostic sensitivity and specificity. Rather, existing diagnostic criteria were devised from epidemiological studies examining risk factors for dementia in apparently cognitively intact older adults [1-3]. These epidemiological studies have been critical in understanding the presence of subclinical cognitive changes that represent risk factors for later dementia development. Some studies indicate that episodic memory dysfunction is a prominent feature [e.g. 6, 7, 8], whereas others suggests that non-memory deficits (e.g. executive processes) are more indicative of those likely to convert [e.g. 9]. Alternatively, other research implies that combined memory and non-memory dysfunction (e.g. attention, working memory, semantic memory) is better at predicting future cognitive decline [e.g. 10, 11-14].

Consistently across these studies is the finding that not all cases of MCI convert to dementia, with many reporting that a sizeable proportion of individuals return to age appropriate levels of functioning over time [e.g. 14, 15, 16-18]. Rates of recovery vary across clinical and community studies, with some reporting rates as low as 7% [19] or as high as 48% [20]. If “...MCI is a clinical diagnosis which is the same as are the diagnoses of dementia or AD” [21], recovery is not consistent with an MCI diagnosis. Rather, cases of recovery represent false positive diagnoses and create a significant source of variability for research in this area.
The high rate of false positive diagnoses of MCI arising from current diagnostic criteria highlights the errors that emerge when risk factors for dementia are erroneously utilised as clinical diagnostic criteria. In order to improve the resolution of current diagnostic criteria, clinical studies of adults displaying MCI are required. Sensitive and specific diagnostic criteria for MCI are essential to facilitate research examining factors that predict conversion to dementia and the potential ways in which this process may be slowed or prevented.

With the majority of MCI studies focusing on predicting the trajectory to dementia, fewer studies have examined the factors that reliably identify individuals with MCI (i.e. true positive cases). Summers and Saunders [14] identified several neuropsychological measures that were useful in predicting stable MCI, recovery, conversion (to AD), and those who remained unimpaired. These measures spanned multiple domains including visual and verbal episodic memory, working memory, and attentional processing; and successfully classified 88% of MCI cases; 96% of controls; 65% of recovered cases and 100% of converters (AD). These results suggest that predicting outcome is enhanced via a comprehensive protocol that spans multiple cognitive domains. Mitchell, et al. [23] found that the combination of a visual memory measure (PAL) and a dementia screening tool (ACE) was the best at predicting whether individuals would convert, recover, or maintain MCI status at follow up. Mitchell and colleagues argue that this combination was successful because of the challenging nature of the PAL task compared to traditional verbal memory measures; and because the ACE taps a range of cognitive domains including episodic memory, semantic memory, visuospatial skills and executive functioning, all of which can be compromised in early AD. The study reported 87.5% of recovered cases correctly classified and 66.7% of converters correctly classified, although it was not clear how many cases remained within the bounds of MCI.
pattern emerging from these studies is that accuracy in identifying MCI can be improved by using measures that tap memory and non-memory domains.

The aim of the present study was to identify a set of neuropsychological measures from a comprehensive neuropsychological battery that accurately differentiate between individuals with stable MCI (true positive) from those who remain unimpaired (true negative) over time. By increasing the accuracy with which true positive cases of MCI are identified, the rate of false positive MCI diagnoses evident in current research will be reduced. We examined data from a longitudinal study tracking the cognitive profile of an MCI cohort to identify specific neuropsychological measures that predict outcome (MCI versus unimpaired). Each participant in the study was classified as MCI or unimpaired (non-MCI) according to the stability of their neuropsychological profile across three assessment points. As findings from conversion studies suggest that combined memory and non-memory protocols enhance predictive accuracy, it was predicted that a similar spectrum of measures would be useful in predicting MCI and unimpaired status in the present study.

**Method**

**Participants**

The participants in this study comprised a sample of community-residing adults recruited in 2010 as part of a longitudinal project tracking the neuropsychological profile of MCI. Figure 1 shows the progression of participants from recruitment to the final assessment phase (time 2). The participant pool in the present study is a distinct cohort from cohorts reported in previous publications from our group [14, 24, 25]. Each participant provided
informed consent prior to the commencement of the study, in accordance with the
Human Research Ethics Committee (Tasmania) Network and National Health and
Medical Research Council (NHMRC) of Australia Human Research Guidelines and the
Helsinki Declaration.

[INSERT FIGURE 1 HERE]

At initial recruitment ($n = 286$), each participant underwent telephone pre-screening to
ensure that there were no medical, neurological, or psychological conditions that would
impact their participation. A total of 200 participants successfully passed pre-screening and
were subsequently assessed on a comprehensive neuropsychological screening battery (Table
1) to classify participants according to existing MCI criteria [4]. Participants were then
assessed at two further time points (time 1 and time 2) on a comprehensive battery utilizing
measures distinct from those used in the screening battery but assessing the same cognitive
domains (Table 1). This was done to reduce the potential influence of circular logic into the
classification process, an issue that has been previously raised by some MCI researchers [26,
27]. Following each assessment point, the performance of each individual participant was
examined by an experienced clinical neuropsychologist (MS) to determine if the performance
met the criteria for MCI. Participants were classified as subclinically impaired where the
performance was more than $1.28SD$ (<10th percentile) below age- and/or education based
norms in accordance with previously established protocols [14, 24, 25]. Classification of MCI
was based on the presence of subclinical impairments on one or more memory or non-
memory task [5, 14, 24, 25]. Mean follow up time from screening assessment to time 1 was
nine months ($SD=3$ months). Mean follow up time from time 1 and time 2 assessments was
11 months ($SD=1$ month). Because the current study was interested in predicting stability
outcome, only the 118 participants (overall retention of 90.8%) who were assessed at all three time points were included in the current analysis.

[INSERT TABLE 1 HERE]

At the conclusion of Time 2, all participants received a stability classification based on the consistency of their classifications across screening, time 1, and time 2 (see Figure 2). Participants who met the criteria for MCI (≤10th percentile on one or more memory and/or non-memory test) at all three time points were given the stability classification of MCI. In addition, those who met the criteria for unimpaired at screening but went on to meet MCI criteria at time 1 or 2 were given the stability classification of MCI. Participants who met the criteria for healthy control (≥10th percentile on all tests) were given the stability classification of unimpaired. In addition, those who met the criteria for MCI at screening but continued to show intact performance at time 1 and 2 were placed in the unimpaired group. This decision was made on the basis that recovery of function is inconsistent with the definition of MCI as a clinical diagnosis. The 118 participants who completed all three time points formed the following groups: MCI (n=60, 33 females) and unimpaired (n=58, 39 female).

[INSERT FIGURE 2 HERE]

Materials

Participants were assessed on a comprehensive neuropsychological test battery at three time points (Table 1). These assessments were part of a detailed longitudinal study examining the profile of MCI over time. Measures from the screening protocol were utilised solely for the
purpose of classifying individuals into the study as well as contributing to the final stability
classification.

**Genetic analysis**

Participants provided a sputum sample and DNA extraction was performed using ORAGENE
PrepIT extraction kit to manufactures instructions (DNAdenotek cat. no. OG-500). The
multiplex amplification refractory mutation system PCR (ARMS-PCR) was used to genotype
DNA samples for the *ApoE* gene. ARMS primers and PCR amplification of DNA were
conducted according to Donohoe *et al.* [28]. Amplified multiplex products were run on a 2%
agarose gel and visualized using SYBR® SAFE DNA Gel Stain (Life Technologies,
catalogue no. S33102). The amplicons were sized against Quick-Load® 100 bp DNA Ladder
(New England Biolabs, catalogue no. N04675). Gels images were captured using the Chemi
Doc XRS scanner, and visualised using QuantityOne software. In order to determine
-genotype of samples using the ARMS PCR, the reference gel from Donohoe *et al* was utilised
[28]. Analysis of genotype was performed with the observer blinded to participant MCI
status.

**Procedure**

Individual testing sessions were conducted in a well-lit, well ventilated room at the School of
Psychology Newnham campus. Tests were administered in the order indicated in Table 1.

Each assessment took approximately 2.5 hours including a mandated 10 minute rest break in
the middle of the test sequence to minimise fatigue effects. Additional rest breaks were
provided as needed for each individual participant.
Results

All results were analysed using SPSS for Windows (version 19.0). To predict group membership (MCI vs. unimpaired), discriminant function analysis (DFA) was deemed most appropriate due to the small sample size [29]. As the prediction of stability status (MCI vs. unimpaired) based on longitudinal classification profiles had not been previously examined, it was deemed appropriate to identify predictor variables statistically. Firstly, all time 1 outcome measures were entered into a MANOVA to control for the inflated Type 1 error rate that occurs when assessing cognitive functions from related domains. The overall MANOVA was significant (Pillai’s trace = .474, $F_{(20, 97)} = 4.38, p < .001$, power = 1.00, $n_p^2 = .474$). Follow up one way ANOVAs indicated a number of measures significantly discriminated between the two groups. Partial correlations; inflation factor (VIF) and tolerance values were assessed to examine any potential multicollinearity issues for each of the potential predictor variables [29, 30]. The following time 1 variables were identified as potential predictors due to significant group differentiation and absence of multicollinearity: RAVLT (delayed recall); RVP A’; RVP latency; LNS; BNT; SSP length; DSP backwards; SWM total errors; SWM strategy; MTS mean correct RT; Age (Table 2); and WTAR est. FSIQ (Table 2). Of the 118 participants, 113 had consented to APOE genotyping from salivary samples. The proportion of APOE e4 carriers was not significantly different between MCI and unimpaired groups ($\chi^2_{(1)} = .378, p = .539$), consequently APOE e4 status was not included as a predictor in the present model. Therefore, it was deemed appropriate to analyse the larger sample of 118 participants. Predictors were entered into a DFA which revealed a significant discriminant function ($\Lambda = .580, \chi^2_{(13, n=118)} = 59.67, p < .001$), which classified 100% of the cases and accounted for 52% of the variance in outcome group. Analysis of the structure matrix revealed ten significant predictors, RVP A’ (0.657); BNT (0.593); RAVLT-delayed recall
For Peer Review

1.0; SSP length (0.476); MTS mean RT (-0.416); RVP latency (-0.497); LNS (0.532); SWM errors (-0.471); PAL 6 (-0.439); and SWM strategy (-0.349) with WTAR estimated FSIQ as a poor predictor (0.241). Results indicated that 80.0% of MCI cases and 87.9% of unimpaired cases were correctly identified (see Table 3). Results also revealed that 7 (5.93%) unimpaired cases were incorrectly identified as MCI using the present model. In order to determine if classification of MCI using a DFA model improved diagnostic accuracy of MCI relative to standard diagnostic procedures [4], we calculated the sensitivity and specificity for MCI diagnosis using existing Winblad, et al. [4] criteria for the entire sample at screening. Each participant’s diagnosis at screening was then compared to their diagnosis across time 1 and time 2. Using the Winblad, et al. [4] criteria it was found that 29 (24.58%) unimpaired cases were incorrectly identified as MCI, confirming that the classification of MCI cases using the DFA model resulted in a markedly lower false positive diagnostic rate of 5.93% (Table 4).

[INSERT TABLE 2 HERE]

[INSERT TABLE 3 HERE]

[INSERT TABLE 4 HERE]

Discussion

The aim of the present study was to examine the ability of neuropsychological measures from a single assessment to predict group classification (stable MCI versus unimpaired) over time. Consistent with our prediction, the discriminant function analysis revealed that several measures tapping multiple cognitive domains accurately classified group outcome, accounting for 83.9% of all cases. A measure of target detection (RVP A’) was the largest
predictor variable in the model; followed by a measure of semantic memory retrieval (BNT); and verbal delayed recall (RAVLT delayed). This suggests that episodic memory dysfunction is predictive of stable MCI but only in conjunction with subclinically impaired performances in other domains such as attentional processing and semantic memory. The model also resulted in a false positive rate of 5.93% for MCI, which is considerably lower than the rate of false positive diagnosis identified using existing diagnostic criteria (23.73%). The finding that only 5.93% of unimpaired cases were classified as MCI using the present model is also comparatively lower than many previous studies with similar samples [see 22].

Many MCI studies adopt test protocols that include discreet measures of episodic memory function (e.g. list learning tasks, paragraph recall) but utilise broad screening measures to assess non-memory function (e.g. MMSE; ACE) [e.g. 16, 31-35]. The present findings argue against an amnestic-centered approach to identifying individuals on the MCI spectrum. Rather, the present results in conjunction with those of other recent studies [14, 23] indicate that it is essential to employ reliable and valid measures of discrete memory and non-memory functions to enhance the identification of true positive MCI cases. Previous studies have highlighted the importance of using comprehensive test batteries to enhance sensitivity and specificity in MCI classification [10, 14, 36]. Recent recommendations state that MCI studies should include an assessment of memory and non-memory functions [5], although little detail is provided regarding the assessment of non-amnestic function. If MCI is degenerative in prognosis, then recovery of function in individuals identified with MCI is diagnostically erroneous. The results of the present study indicate that comprehensive assessment of memory and non-memory function is essential to minimizing false positive diagnosis of MCI and consequently maximizing true positive diagnosis of MCI.
The discriminant function model from the present study identified RVP A’ as the strongest predictive variable of diagnostic stability. Previous research by our group found that RVP A’ was the largest predictor variable in predicting development of AD in an MCI sample [14]. These results suggest that RVP A’ may be a particularly sensitive measure to MCI. The RVP task from the Cambridge Automated Neuropsychological Test Assessment Battery [CANTAB; 37] requires participants to correctly identify target sequences of numbers (e.g. 3-5-7) that appear within a stream of pseudo random single numbers at a rate of 100 digits per minute. The ratio of hits to false alarms, derived from signal detection theory [see 38], constitutes RVP A’. Successful RVP task performance requires a participant to maintain sustained attention, utilize working memory to hold digits in short term storage, and use selective attention to identify target sequences [39]. Due to the demand placed on multiple executive processes, RVP is a considered particularly challenging [40]. The task complexity of the RVP task may be the basis of the good discriminatory power of RVP A’ in distinguishing subclinical impairment (MCI) from normal aging (unimpaired). It is well established that broad screening measures such as the MMSE lack sensitivity to detect sub clinical conditions like MCI because of ceiling effects [e.g. 41]. Complex cognitive tasks, such as RVP, may have greater discriminatory power because they are more cognitively taxing and have a higher ceiling. Further, recent research indicates that deficits to executive control are evident in early AD [42, 43], and MCI [24, 25, 44], and may be predictive of future cognitive decline [9]. Sensitive measures of executive function, such as RVP, may therefore be assessing an underlying common diagnostic marker of MCI.

Unlike previous studies, individuals in the present study were defined as MCI or unimpaired according to their neuropsychological performance across three assessment time points. By using this method, we clinically identified a sample of participants with longitudinally stable
MCI. We were then able to identify neuropsychological markers that differentiated participants with stable MCI from those who were unimpaired; markers that represent potential diagnostic features for temporally stable MCI. Despite these strengths, some limitations need to be considered. Firstly, it could be argued that the present study is limited by circularity given that the tests used to predict group outcome were also used to classify MCI at time 1 and time 2. However, stability classification was based on the consistency of performance across screening, time 1 and time 2 and included a number of measures that were not included in the final model. In addition, the use of different tests to classify and track cohorts of cognitively impaired individuals does not completely avoid the issue of circularity given that test performance on tests in the same domain are likely to be correlated [27, 45].

The results of the present study indicate that measures of multiple cognitive domains predicted stable MCI or unimpaired status in over 80% of all cases. The rate of false positive diagnoses was considerably lower using a predictive model based on performance on a range of cognitive measures than using existing diagnostic criteria [4]. Reducing the rate of false positive diagnoses is imperative for MCI research, particularly for intervention trials with individuals thought to be at highest risk of future cognitive decline (i.e. true positives). The high false positive rate associated with existing MCI diagnostic criteria may mask the true efficacy of treatment or intervention for MCI. Results from the present study indicate that an MCI diagnosis should only be made on the basis of impairment to multiple domains rather than isolated memory deficits. Further, these results suggest that current MCI diagnostic criteria should be modified to include sensitive and specific measures of complex sustained attention, semantic memory, working memory, episodic memory, and selective attention.
CONFLICT OF INTEREST DECLARATION

Dr. M. Summers reports personal fees from Eli Lily (Australia) Pty Ltd, grants from Novotech Pty Ltd, outside the submitted work; Prof. Vickers reports grants from Bupa Foundation (Australia) Limited, grants from J.O. & J.R. Wicking Trust, outside the submitted work; Prof. J Summers and Ms. Klekociuk have nothing to disclose.

DESCRIPTION OF AUTHOR ROLES

SK collected and analysed the data as part of a doctoral project. SK prepared the manuscript. MS contributed with the study design and manuscript preparation. JS contributed with manuscript presentation. JV contributed with genetic analyses and manuscript preparation.

ACKNOWLEDGEMENTS

This research received no specific grant from any funding agency, commercial or not-for-profit sectors. S. Klekociuk is a recipient of an Australian Postgraduate Award Scholarship and postgraduate scholarship from the Wicking Dementia Research and Education Centre (WDREC). The authors acknowledge the assistance of Mr Graeme McCormack (WDREC) and Ms Rachel Bain (WDREC) in assaying saliva samples for APOE genotyping.
References


Figure 1: Participant progression chart.
Initial enrolment
Response from recruitment
\(n = 286\)

Telephone screening
Successfully completed
\(n = 200\)
Excluded
\(n = 86\)

Screening
(0 months)
Completed
\(n = 200\)
Used in analysis
\(n = 122\)
Excluded (\(n = 70\))
Medical 31
Healthy 39
Attrition \(N = 8\)
Medical 6
Relocated 1
Personal 1

Time 1
(9 months)
Completed assessment
\(n = 122\)
Attrition
\(n = 4\)
Deceased 1
Medical 1
Relocated 1
Withdrawn 1

Time 2
(20 months)
Completed assessment
\(n = 118\)
Figure 2: Stability classification procedure
Table 1

*Test Batteries for Screening and Follow-up Testing*
<table>
<thead>
<tr>
<th>Screening assessment</th>
<th>Time 1 assessment</th>
<th>Time 2 assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTAR</td>
<td>DRS-2</td>
<td>DRS-2</td>
</tr>
<tr>
<td>HADS</td>
<td>Boston Naming Test (BNT)</td>
<td>Paired Associates Learning (PAL)³</td>
</tr>
<tr>
<td>DRS-2</td>
<td>Simple Reaction Time (SRTI)³</td>
<td>RAVLT (List C)</td>
</tr>
<tr>
<td>Digit Symbol Coding (DSC)²</td>
<td>Choice Reaction Time (CRTI)³</td>
<td>Simple Reaction Time (SRTI)³</td>
</tr>
<tr>
<td>Logical Memory (LMI) I¹</td>
<td>Digit Span (DSP)²</td>
<td>Choice Reaction Time (CRTI)³</td>
</tr>
<tr>
<td>RCFT-copy</td>
<td>Spatial Span (SSP)³</td>
<td>Digit Span (DSP)²</td>
</tr>
<tr>
<td>Digit Span (DSP)²</td>
<td>Letter Number Sequencing (LNS)²</td>
<td>Rapid Visual Processing (RVP)³</td>
</tr>
<tr>
<td>Spatial Span (SSP)¹</td>
<td>Paired Associates Learning (PAL)³</td>
<td>Letter Number Sequencing (LNS)²</td>
</tr>
<tr>
<td>RCFT-delay recall</td>
<td>HADS</td>
<td>Matching to Sample (MTS)³</td>
</tr>
<tr>
<td>Letter Number Sequencing²</td>
<td>RAVLT (List A)</td>
<td>Spatial Span (SSP)³</td>
</tr>
<tr>
<td>Trail Making Test A &amp; B (TMT)</td>
<td>Rapid Visual Processing (RVP)³</td>
<td>HADS</td>
</tr>
<tr>
<td>Logical Memory II (LMII)¹</td>
<td>Spatial Working Memory (SWM)³</td>
<td>Spatial Working Memory (SWM)³</td>
</tr>
<tr>
<td>Stroop-Victoria version</td>
<td>Matching to Sample (MTS)³</td>
<td>Boston Naming Test (BNT)</td>
</tr>
<tr>
<td>Vocabulary (VOC)²</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ subtest from the Wechsler Memory Scale, 3rd edition (WMS-III)
² subtest from the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III)
³ subtest from the Cambridge Automated Neuropsychological Test Assessment Battery (CANTAB)

HADS = Hospital Anxiety and Depression Scale
WTAR = Wechsler Test of Adult Reading
DRS-2 = Dementia Rating Scale, 2nd edition
RCFT = Rey Complex Figure Test
RAVLT = Rey Auditory Verbal Learning Task
Table 2.

Demographic Information and APOE E4 Status
<table>
<thead>
<tr>
<th></th>
<th>MCI</th>
<th>Unimpaired</th>
<th>p value</th>
<th>Post hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>56</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>73.55</td>
<td>68.57</td>
<td>&gt;.001</td>
<td>MCI &gt; U</td>
</tr>
<tr>
<td>Education</td>
<td>13.18</td>
<td>14.81</td>
<td>.012</td>
<td>MCI &lt; U</td>
</tr>
<tr>
<td>WTAR est FSIQ</td>
<td>107.63</td>
<td>110.55</td>
<td>.015</td>
<td>MCI &lt; U</td>
</tr>
<tr>
<td>DRS-2</td>
<td>11.07</td>
<td>12.07</td>
<td>.016</td>
<td>MCI &lt; U</td>
</tr>
<tr>
<td>HADS A</td>
<td>5.68</td>
<td>5.14</td>
<td>.318</td>
<td></td>
</tr>
<tr>
<td>HADS D</td>
<td>3.22</td>
<td>2.55</td>
<td>.118</td>
<td></td>
</tr>
<tr>
<td>APOE e4 carrier</td>
<td>55.6%</td>
<td>44.4%</td>
<td>.539</td>
<td></td>
</tr>
<tr>
<td>Gender (Male)</td>
<td>45%</td>
<td>33%</td>
<td>.173</td>
<td></td>
</tr>
</tbody>
</table>

HADS = Hospital Anxiety and Depression Scale; WTAR = Wechsler Test of Adult Reading; DRS-2 = Dementia Rating Scale, 2nd edition; APOE e4 genotyping n=102
Table 3

*Outcome Classification Predicted from Time 1 Neuropsychological Scores*
<table>
<thead>
<tr>
<th></th>
<th>MCI</th>
<th>Unimpaired</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$   (%)</td>
<td>$n$   (%)</td>
</tr>
<tr>
<td>MCI</td>
<td>48 (80%)</td>
<td>12 (20%)</td>
</tr>
<tr>
<td>Unimpaired</td>
<td>7 (12.1%)</td>
<td>51 (87.9%)</td>
</tr>
</tbody>
</table>

99 out of 118 correctly classified (83.90%)
Table 4

Sensitivity and Specificity of MCI Diagnosis Using Standard Diagnostic Procedures [4] and DFA analysis
<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winblad et al. (2004)</td>
<td>$n = 40$ (33.90%)</td>
<td>$n = 29$ (24.57%)</td>
</tr>
<tr>
<td>DFA</td>
<td>$n = 48$ (40.68%)</td>
<td>$n = 51$ (43.22%)</td>
</tr>
<tr>
<td><strong>False</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winblad et al. (2004)</td>
<td>$n = 28$ (23.73%)</td>
<td>$n = 21$ (13.56%)</td>
</tr>
<tr>
<td>DFA</td>
<td>$n = 7$ (5.93%)</td>
<td>$n = 12$ (17.80%)</td>
</tr>
</tbody>
</table>

69 out of 118 correctly classified (58.47%) using the Winblad et al. criteria; 99 out of 118 correctly classified (83.90%)