Autobiographical Memory and Cognitive Theory of Mind in non-Mild Cognitive Impairment Parkinson's Patients

Meisha Naomi Nicolson

A thesis submitted in partial fulfilment of the requirements for the Degree of

Master of Science

Department of Psychology

University of Canterbury

2016

Table of Contents

Acknowledgements iii

Abbreviations iv

Introduction 2

1.1 Parkinson's disease 2

1.2 Mild cognitive impairment and pre-mild cognitive impairment 3

1.3 Autobiographical Memory in Parkinson's disease 5

1.3.1 Autobiographical memory interview 10

1.3.2 Theory of mind 12

1.4 A Link between Theory of Mind and Autobiographical Memory: Default Mode Network 25

1.5 The current study 27

Method 29

2.1 Participants 29

2.2 Initial Cognitive Screening Tests 32

2.3 Neuropsychological Assessment to establish non-MCI status 32

2.4 Cognitive Testing session for the current study 33

2.4.1 The Autobiographical Memory Interview (AMI) 33

2.4.2 Theory of Mind Task - card-sequencing task (Langdon et al., 1999) 34

2.4.3 Other Cognitive Tests 35

2.4.4 Psychiatric Tests 36

2.5 Procedure 36

2.6 Statistical Analyses 37

Results 38

3.1 Initial Cognitive Screening 38

3.2 The Brief Visuospatial Memory Test-Revised (BVMT) 40

3.3 The Symbol Digit Modalities Test 43

3.4 The Sydney Language Battery (SYDBAT) 44

3.5 Theory of Mind (ToM) and sequencing tasks 45

3.6 The Autobiographical Memory Interview (AMI) 46

3.6.1 Autobiographical (episodic) Memory 46

3.6.2 Personal Semantic Memory 48

3.7 Correlations between the cognitive decline risk score and other measures 50

3.8 Correlations between clinical Parkinson's disease measures and key test scores 51

3.9 Correlations between test scores across both PD and HC group 54

Discussion 56
4.1 Main findings 56
4.2 Comparison of main findings with prior research 57
4.3 Limitations 63
4.4 Future Directions 64
4.5 Conclusions 65
References 66
Appendices 77

Acknowledgements

Firstly I would like to extend my sincere thanks to my primary supervisor Professor John Dalrymple-Alford. Thank you for your time, your guidance and expertise throughout this project; it has been very much appreciated. Secondly to Leslie Livingston, thank you for spending the time to teach me the detailed neuropsychological testing and how to interact with Parkinson's participants with ease and empathy.

Thank you to the New Zealand Brain Research Institute for providing the space and resources for the duration of this project. I have enjoyed my time here immensely, surrounded by researchers and fellow students; it has been a great atmosphere to work in.

To members of the wider study group at the NZBRI that this thesis was part of, a very big thank you. To Megan Livingston, Beth Elias, Jeremy Goh and Marie Goulden, it has been wonderful to work with you all, thank you for your comradeship and support.

A very big thank you to the volunteer participants who took part in this research, whose willingness and enthusiasm was very much appreciated. I enjoyed listening to all of your stories and thank you for your time.

Finally, I would like to thank my family and dear friends for their support and encouragement over the past 12 months during this process.

Abbreviations

AD	Alzheimer's disease
ADAS-Cog	Alzheimer's Dementia Assessment Scale-Cognitive
AI	Autobiographical Interview
AM	autobiographical memory
AMI	Autobiographical Memory Interview
BDS	Blessed Dementia Scale
BVMT	Brief Visuospatial Memory Test
CVLT	California Verbal Learning Test
DMN	default mode network
DRS-2 (AESS)	Dementia Rating Scale (age and education scales score)
FPR	Faux Pas Recognition
fMRI	functional magnetic resonance imaging
GDS	Geriatric Depression Scale
HADS	Hospital Anxiety and Depression Scale
HC	Healthy Controls
Hx	History (medical)
MDRS	Mattis Dementia Rating Scale
MoCA	Montreal Cognitive Assessment
MMSE	Mini-Mental Status Exam
NP	Neuropsychological
PD	Parkinson's disease
PD-MCI	Parkinson's disease with mild cognitive impairment
PDD	Parkinson's disease with Dementia
PFC	Prefrontal cortex
TEA	Test of Everyday Attention
ToM	theory of mind
RME	Reading the Mind in the Eyes
SDMT	Symbol Digit Modality Test
SYDBAT	Sydney Language Battery
TMT	Trail Making Test
UPDRS	Unified Parkinson's Disease Rating Scale
VOSP	Visual Object and Space Perception

The current study examined autobiographical memory and theory of mind (ToM), both of which are associated with overlapping subsystems in the default mode network, in a group of Parkinson's disease (PD) patients who were characterised as not representing a stage of mild cognitive impairment, but at risk of future cognitive decline. The Autobiographical Memory Interview (AMI), which separately measures both personal episodic memory and personal semantic memory across the lifespan, was used for the first time in PD; a card sequencing task measured cognitive Theory of Mind (ToM) in these patients. Twenty non-MCI PD participants (18 above a threshold of 29% risk of future cognitive decline and 2 below this threshold) were compared with 15 healthy age and education matched controls (HC). PD participants showed significantly poorer personal episodic memory but unimpaired personal semantic memory, but neither measure was related to the cognitive risk score. Similarly their impaired ToM scores were unrelated to their risk scores. However, the Brief Visuospatial Memory Test showed a greater effect size than any other measures, and performance in the PD group was associated with the risk score, suggesting it may be a useful addition to improving a risk score in this patient group. Future research should examine these measures in larger sample sizes and in PD-MCI and PDD groups, and evaluate their MRI correlates. PD patients who do not meet criteria for PD-MCI nonetheless show a subtle range of cognitive changes, but only a subset may be useful predictors of significant decline in cognition.

Introduction

1.1 Parkinson's disease

Parkinson's disease (PD) is a progressive neurodegenerative condition causing significant motor and non-motor impairments, which eventuates in dementia (PDD) for over 80% of patients (Aarsland & Kurz, 2010a). The pre-requisite weakness of voluntary motor control in PD is associated with the loss of dopaminergic neurons projecting to the striatum (Jenkins et al., 1992) and functional disconnection in the striato-thalamo-frontal loop (Eimeren et al., 2009). The core motor symptoms include resting tremor, bradykinesia, rigidity and postural instability resulting from reduced functional integrity of the prefrontal cortex. However, it is now widely recognised that a range of non-motor symptoms also characterise PD. These symptoms include mood disorders, depression, hallucinations, sleep disorders, sensory and autonomic dysfunction, behavioural disorders and especially cognitive dysfunction (Poewe, 2008). In fact many non-motor symptoms may begin long before motor symptoms become evident, while new or worsening problems generally emerge throughout the disease course, significantly impacting quality of life. Dementia is associated with more rapid progression of the disease, reduced quality of life, increased care-giver burden and mortality (Aarsland & Kurz, 2010a).

Braak staging provides a general model for progression in which cognitive decline is obvious at the last stages, 5 and 6, but some decline may begin as early as stage 2 (Figure 1, Hawkes et al., 2010). Clinical (motor) onset does not begin until stages 3 to 4. Cognitive decline in PD often leads to mild cognitive impairment (PD-MCI) (Dalrymple-Alford et al., 2011; Litvan et al., 2011). PD-MCI patients are at high risk of PD with dementia (PDD) within 4 years (Wood et al., 2016; Janvin et al., 2006; Pedersen et al., 2013).

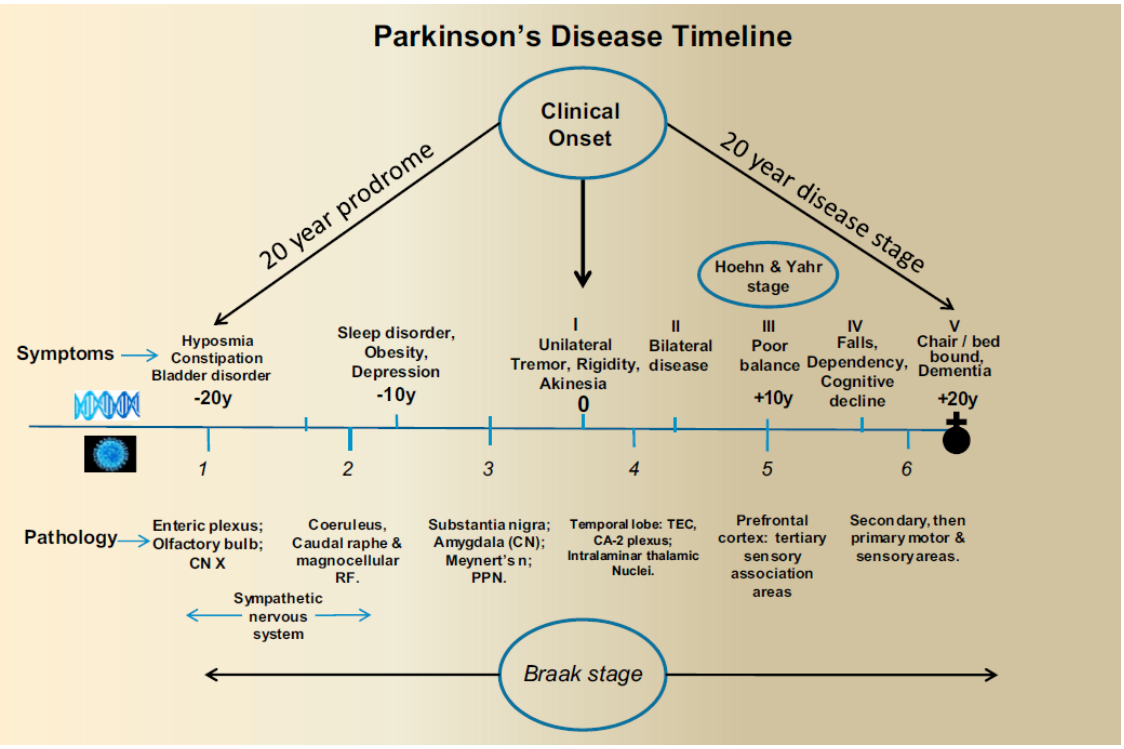


***Figure 1*. Braak staging from symptom onset to end of disease (Hawkes, Tredici & Braak, 2010).**

1.2 Mild cognitive impairment and pre-mild cognitive impairment

Mild cognitive impairment (MCI) is cognitive decline that has not yet interfered significantly with daily life, but is greater than expected for the individual's age and education level and often represents an intermediate stage prior to dementia (Gauthier et al., 2006). MCI is generally divided into two subgroups, amnesic MCI (aMCI) and non-amnesic MCI (nMCI). Amnesic MCI is the presence of memory impairment, whereas non-amnesic MCI is a non-memory cognitive impairment (Peterson, 2004). The idea of MCI has now been adapted in the context of PD and about 25-30% of non-demented Parkinson's patients have mild cognitive impairment (Litvan et al., 2011; Wood et al., 2016). Cognitive change associated with PD-MCI is significantly associated with older age at disease onset, longer disease duration, more severe motor symptoms, and advanced disease stage (Aarsland et al., 2010b). Early recognition of cognitive change is important if we are to attempt to intervene to improve outcomes and slow progression.

The novel concept of 'pre-MCI' has been introduced to define a risk of worsening cognition in the future. This term has mostly been used in Alzheimer's research, and marks a transitory stage between normal cognition and MCI. In non-PD, pre-MCI patients have been shown to decline less than MCI patients but most eventually progress to Alzheimer's (91% at autopsy within this cohort, Storandt et al., 2006). Although the term 'pre-MCI' has not been used in PD research thus far, researchers are starting to examine the changes that occur between PD with normal cognition and PD-MCI. At baseline, Parkinson's disease patients who had no cognitive impairment, but who later converted to MCI over 18 months, show bilateral temporal cortex thinning relative to the Parkinson's disease with no cognitive impairment stable patients that did not progress (Mak et al., 2015). This suggests that there are physical changes in the brain prior to any overt cognitive changes. The value of a pre-MCI status is that if an intervention can be administered at the earliest detection of cognitive change before progression to PD-MCI, then outcomes could be improved.

This thesis investigated the differences between a non-MCI PD group, most of whom were at risk for future progression to MCI or dementia (i.e. pre-MCI), and healthy age-matched controls. The current study focussed on autobiographical memory and theory of mind tasks (ToM). Performance on these tasks is influenced by frontal brain areas which are also implicated in Parkinson's disease. Early stage PD patients show impairments in frontal lobe functions and atrophy in the left and right prefrontal cortex (Bruck, 2004). If autobiographical memory and ToM tasks show deficits in the non-MCI PD group compared to controls, then these tasks may provide suitable measures of future progression in PD. It was hypothesised that the non-MCI PD group will perform more poorly on a range of cognitive tasks than the control group, but most especially tasks that depend on the integrity of the frontal brain regions. It was also hypothesised that there will be a correlation between performance on the theory of mind component of the card sequencing task and the

Autobiographical Memory Interview (AMI), as explained below. The primary novelty of this research was to examine the AMI (Kopelman, Wilson & Baddeley, 1989) for the first time with PD participants, in particular a sample of non-MCI PD participants many of whom were shown to be at risk for future progression, based on an initial "screening risk score" (Myall et al., 2015 AWCBR).

1.3 Autobiographical Memory in Parkinson's disease

Autobiographical memory concerns both personal episodic memories and personal semantic memories, from an individual's life. The personal episodic memories include specific details about experiences, events, people and times. Personal semantic memory includes more general factual information from these memories. The value of the AMI, designed by Kopelman, Wilson and Baddeley (1989) is that it is a standard instrument designed to assess 'personal episodic' and 'personal semantic' memory separately in distinct and detailed sections. The AMI is broken up into two components, a personal semantic schedule and an autobiographical incident schedule with participants producing memories for each of these components from different periods of their life. This allows a comparison of recall of facts and incidents in a way that has not been possible in previous tests (Kopelman et al., 1989).

Autobiographical memory is associated with both prefrontal cortex and medial temporal lobe function. Damage to the medial temporal lobe can impair ability to store new memories, and can cause loss of old memories (Squire, 1992). A correlational analysis of the amygdala, hippocampus and right inferior frontal gyrus found functional connectivity between these regions during episodic autobiographical memory retrieval but not semantic retrieval (Greenberg et al., 2005). These areas that show functional connectivity during autobiographical memory retrieval are often impaired in PD. This suggests that individuals with PD may have problems with autobiographical memory retrieval.

Only three studies have investigated autobiographical memory with PD groups (Table 1). These studies focussed on public semantic and personal episodic memory (Sagar et al., 1988), personal episodic and personal semantic memory (Smith, Souchay & Conway, 2010), and personal episodic memory only (Souchay & Smith, 2013). Each study found that PD participants experience some deficits in retrieving autobiographical memories (Table 1). However, these studies did used minimal cognitive testing to ascertain the cognitive status of their participants. The Mini-Mental Status Exam (MMSE) was the main cognitive measure for two of the aforementioned studies (Smith et al., 2010; Souchay and Smith, 2013), while the third used the Blessed Dementia Scale (BDS) (Sagar et al., 1988). One study added the Trail making and Stroop Tests (Souchay and Smith, 2013). A larger cognitive test battery would however be more accurate at determining the cognitive status of patients given their heterogeneous impairments (Hoops et al., 2009).

An early form of autobiographical memory test, used by Sagar et al. (1988), showed that both PD and AD groups are more impaired at recalling content of personal remote events (personal events from a long time ago) than recent events. This test was called the Personal Remote Memory Test (modified from Crovitz and Schiffman 1974). Participants had to generate personal memories from any lifetime period in response to ten high frequency noun cues (for example, tree). Participants were given up to 4 minutes to recall autobiographical memories with as much detail as possible in response to each cue. If participants could not respond within the first two minutes, they were prompted with nonspecific cues (e.g. ‘tell me more’, can you think of one instance?’). The following day without warning, participants were given a second recall test for each word cue. If they failed to recall as much as they had recalled the day beforehand they were prompted after two minutes with key words from their memories (e.g. Onyx’). Healthy controls recalled full and detailed memories on day 1, with or without cues, and were able to fully reproduce these memories on day two. Participants

with AD were impaired at autobiographical recall on day 1, whether cued or un-cued, and were impaired further on their recall on day 2. AD participants with higher BDS scores performed worse. PD participants were mildly impaired at recalling autobiographical memories on day 1, with or without cues, and showed a deficit at recalling these memories on day 2. Although when cued, the PD participants were able to recall as well as they had on day 1. The high BDS score group of PD (4 participants, 3 of which had PDD) performed worse than the low BDS group (normal BDS score) on both days. Impaired performance was characterised by failure to recall time-specific events. Recall of public events (episodic and semantic), assessed by the Public Scenes test, also showed impairment in the PD and AD groups. Participants were shown photographs of famous scenes from the 1940's to the 1980's and had to indicate if they recognised them, and recall event, content, scenario and date information for each photo. PD and AD participants showed a gradient deficit, recent events were recalled more poorly than remote events. Again the high BDS PD group performed more poorly than the low BDS PD group.

The other two studies used an autobiographical memory fluency task, where participants had 2 minutes to recall personal events (episodic) and 2 minutes to recall personal facts (semantic) in each of 5 lifetime periods, 0-18, 19-30, over 30 up until last 5 years, the last 5 years, excluding the last 12 months, and the last 12 months (Souchay and Smith, 2013; Smith, Souchay & Conway, 2010). In the first study, non-demented PD participants recalled fewer personal events (episodic and semantic) for recent time periods and had more difficulty recalling autobiographical events rather than autobiographical knowledge (Smith, Souchay & Conway, 2010). This difficulty was attributed to over-generality, as PD patients could not recall specific memories. The second study analysed different PD participants and their recall of autobiographical memories with the same fluency task, but added a delayed recall phase (Souchay & Smith, 2013). There were three parts to the

delayed recall phase. First, free recall: participants were required to recall the personal events they had already given in exactly the same way. Second, cued-recall: participants were given the lifetime periods as general cues to retrieve events. Third, self-generated cues: participants were given a memory title they had previously assigned to each memory as individual memory cues. Compared to controls, PD participants were impaired at recalling memories in free recall and cued recall but performed comparable to controls in response to self-generated cues. This difference was attributed to retrieval process. Specific self-generated autobiographical memory cues may provide a direct route to memory retrieval while general cues may involve a search elaboration process, which is more cognitively demanding (Souchay & Smith, 2013).

A potential limitation of the autobiographical memory fluency task is that the personal semantic knowledge section is limited. Participants were only required to recall names of people (excluding family or relatives) for each time period. The AMI, by contrast, provides an improved personal semantic memory measure as it takes into account not only names but dates, addresses and locations, which is a more realistic measure of semantic knowledge.

Table 1
PD studies researching Autobiographical Memory. The order is chronological.

Reference	**Study samples**	**Average Duration PD – since diagnosis**	**Average Age**	**Cognition**	**Used Paradigms**	**Main Findings**
Sagar et al., 1988	23 PD -3 of these PDD 32 AD 37 HC	PD = 7.9 (range 1-19)	PD = 64 AD = 64.8 HC = 62.2	BDS PD average = 3.48 Normal BDS group = scores 0-2 High BDS group scores 3-12	Tests for public events: The Famous Scenes Test, Verbal Multiple Choice Recognition Test Tests for personal events: Modified Crovitz Personal Remote Memory Test	Personal and public episodic and semantic memory impaired Participants with high BDS scores (n = 4, 3 of which had PDD) were more severely impaired
Smith, Souchay & Conway., 2010	16 PD - dementia excluded 16 HC	6 years (SD= 3.53) – only patients with mild to moderate rigid-akinetic form of PD included	PD = 72.13 HC = 72.5	MMSE scores above 26 NART	Autobiographical Memory Fluency Task - Participants recalled autobiographical memories (2 mins for personal events and 2mins for personal facts) from given time points (0-18, 19-30, 30-to last 5 years, last 5 years-last 12 months, and during the last 12 months).	PD impaired for recent time periods for episodic and semantic memory PD group made over-general responses
Souchay and Smith 2013	16 PD - dementia excluded (different patient group to above) 16 HC	8 years (SD = 16.28) – only patients with mild to moderate rigid-akinetic form of PD included	PD = 75.18 HC = 74.51	MMSE scores above 26 Trail making Task Stroop Task	Autobiographical Memory Fluency Task + cues for AM retrieval three recall periods included free phase recall, general lifetime period cued-recall and recall cued by self-generated cues	PD group impaired in free recall phase and to general lifetime period cues. PD group same as controls retrieving memories to self-generated cues.

Abbreviations: AD = Alzheimer's disease; AM = Autobiographical memory; BDS = Blessed Dementia Scale; HC = Healthy control; MMSE = Mini-Mental Status Exam; NART = The National Adult Reading Test; PD = Parkinson's disease; PDD = Parkinson's disease with dementia; SD = Standard deviation

1.3.1 Autobiographical memory interview

The AMI is a standard instrument designed to assess 'personal episodic' (autobiographical) and 'personal semantic' memory in distinct and detailed sections. The AMI addresses three broad time periods, childhood, early adult and recent life, each of which is further broken down into three more specific time periods. To assess autobiographical episodic memory, participants are required to recall specific memories from each of these 9 time periods. Detailed rather than general memories are encouraged, and prompts are used where the participants fail to recall on their own. To assess personal semantic memory participants are asked questions about their past that require knowledge of facts, such as dates of births and marriages, and names of friends. The broad range and specificity of the time periods covered in the AMI provides detailed information from the participants for both semantic and episodic memories. This test makes it possible to dissociate between semantic and episodic impairments and to compare performance across different time periods.

Most AMI studies have examined Alzheimer's disease patients. Autobiographical memory impairments are evident in the early stages of AD (Levy et al., 1998). In recall of autobiographical incidents, minimal and mild AD groups show a temporally graded loss, recalling fewer memories for recent time periods than remote time periods (Levy et al., 1998). Amnestic MCI patients, who are at high risk for developing Alzheimer's disease, also experience difficulty recollecting events from their entire lifespan, in particular, recently experienced events (Tramoni et al., 2012). Comparatively older adult controls show the opposite temporal gradient, as they can generally recall more autobiographical memories for recent time periods than earlier lifetime periods (Kopelman, 1989). This may reflect a general retrieval deficit for memories that occurred a long time ago, rather than specific memory impairment.

There is a similar impairment in recall of personal semantic memories and episodic memories in AD. Early stage AD patients also show a temporal gradient in semantic recall, recalling recent memories poorer than earlier memories (Levy et al., 1998). Kopelman et al.(1989) found a similar pattern of results for the AMI with their cohort of AD participants, while their control participants performed in the reverse temporal gradient, recalling recent semantic memories better than childhood memories.

A potential problem with using a recollection of personal events is checking accuracy. For example, confabulation is thought to occur in patients with both frontal lobe damage and memory deficits. Kopelman et al. (1989) assessed the accuracy of recall with Korsakoff's and Alzheimer's patients tested with the AMI by consulting their relatives. They found that the overall accuracy was high and that any inaccuracy and confabulation was minimal when compared to the magnitude of the deficit compared to controls. By correlating the patients initial scores with the scores after consulting with relatives, they found a high correlation (0.88) indicating that detailed checking of responses is probably not necessary for most patients.

The Autobiographical Interview (AI) of Levine is another measure of Autobiographical memory; however it is less sensitive to impairment than the AMI. The AI requires participants to recall a memory, and scores the semantic and episodic components from the same memory, compared to the distinctly separate sections in the AMI for each component. The AMI may sample fewer time periods (childhood, early adult life, and recent life) than the AI (early childhood, adolescence, early adulthood, middle age, and the previous year), but the AI requires fewer memories per time period. In a recent study, individuals with dementia of the Alzheimer's type (DAT) when tested with the AMI were again significantly impaired on both episodic and semantic memory with a significant temporal gradient sparing childhood memories (Barnabe, et al., 2012). By comparison DAT individuals tested with the

AI in the same study only showed impairment on recall of episodic memories with a slight temporal gradient (Barnabe, et al., 2012).

No studies have examined autobiographical memory in Parkinson's disease using the AMI. Kopelman's AMI is of value to PD research as it can be sensitive to temporal gradients in both episodic and semantic memory recall and separates these two components of lifetime memories. It appears that the current study is the first to use the AMI with PD patients.

1.3.2 Theory of mind

Theory of Mind (ToM) is the ability to understand and predict the behaviour of others. ToM is influenced by the integrity of the prefrontal cortex, which is often impaired in PD, so deficits in Theory of Mind tasks are anticipated. There is evidence to suggest that PD patients perform worse on ToM tasks than controls (Mengelberg and Siegert, 2003). However there is a large range of ToM tasks, not all of which show deficits with PD patients (Poletti et al., 2011). Table 2 shows the range of affective theory of mind tasks used in PD and Table 3 shows the range of cognitive theory of mind tasks used in PD. The tables are ordered by ToM 'tasks used' to group together the main findings for each task, so there is some repetition of researchers, participant, and cognitive measures information.

Table 2

Affective Theory of Mind tasks used in PD studies. Tasks are listed in alphabetical order.

Tasks	**Researchers**	**Test Details**	**Participants**	**Hoen and Yahr Stage**	**Cognitive Measures**	**Main Findings**
Faux pas recognition	Roca et al., 2010	Participants read a story that may contain a social faux pas. The affective component of the test is recognizing that the person committing the faux pas was unaware of saying something inappropriate, and recognizing that the person hearing the faux pas might have felt hurt or insulted.	36 PD (16 medicated 20 drug free), 35 HC	Medicated 1.42 (0.57) Drug free 1.33 (0.54)	MMSE (score: >24), categorical and phonologic verbal fluency, short version Boston Naming test, short version Token test, Rey auditory verbal learning test, delayed recall of the Complex Rey figure and attention, TMT A and B, digit forwards and backwards span, Wisconsin Card Sorting Test.	Medicated and drug free preserved
	Peron et al., 2009		17 early PD nondemented on and off meds (abstain from meds night before assessment, 26 HC	On 1.5 (0.7) Off 1.5 (0.7)	MDRS (score >130), plus the following for the advanced PD patients: MCST, TMT, categorical and literal fluency tasks, the Stroop Test.	Preserved (no difference between early on/off)
			27 advanced PD nondemented on and off, 26 HC	On 1.3 (0.9) Off 2.5. (1.0)		Preserved

	Kawamura & Koyama, 2007		11 PD nondemented, 20 HC	NA	MMSE	Preserved
Reading the mind in the eyes	Bodden et al., 2010	Photographs of eye regions presented, participants read the emotional expression and choose one of four words that best describes the mental state expressed by the eyes.	21 PD nondemented (20 medicated), 21 HC	2.5 (Range 1.0 to 3.0)	MMSE (cut-off: >26), PANDA, memo test (immediate and delayed recall) digit span forward and backward from WMS-R, verbal and semantic fluency tasks (FAS and animals), TMT, subtest 4 of the German intelligence test battery LPS.	Impaired
	Peron et al., 2010		13 PD pre- and post-STN DBS (medicated), 13 HC	Pre-DBS 1.9 (0.9) Post-DBS 1.2 (1.0)	MDRS (score >130), MCST, TMT, categorical and literal fluency tasks, the action (verb) fluency task, the Stroop Test.	Pre-DBS preserved, post-DBS Impaired

	Roca et al., 2010	36 PD (16 medicated 20 drug free), 35 HC	Medicated 1.42 (0.57) Drug-free 1.33 (0.54)	MMSE (score: >24), categorical and phonologic verbal fluency, short version Boston Naming test, short version Token test, Rey auditory verbal learning test, delayed recall of the Complex Rey figure and attention, TMT A and B, digit forwards and backwards span, Wisconsin Card Sorting Test.	Medicated and drug free preserved
	Euteneuer et al., 2009	21PD nondemented (medicated), 23 HC	2.5	MMSE, DemTect, Subtest 4 of German intelligence test battery LPS, MCST, verbal fluency (FAS and animals) .	Preserved
	Peron et al., 2009	17 early PD on nondemented, 26 HC 27 advanced PD nondemented, 26 HC	On 1.5 (0.7) Off 1.5 (0.7) On 1.3 (0.9) Off 2.5. (1.0)	MDRS (score >130), plus the following for the advanced PD patients: MCST, TMT, categorical and literal fluency tasks, the Stroop Test.	Preserved (no difference on/off) Preserved (no difference on/off)

	Mimura et al., 2006		18 PD nondemented (medicated), 40 HC	2-3	MMSE, WCST, maze-tracing subtest of WISC-R, The Stroop Color Word naming Test, Verbal fluency with Japanese phonemes [ka], [ki], or [ma].	Slightly impaired (PD scored lower than HC but both scored highly).
Yoni task	Bodden et al., 2010	60 items presented on a computer screen, a face named Yoni in the middle with 4 coloured pictures of faces or objects in the corners affective first order = Yoni's mental state inferred ('Yoni likes…') Affective second order = Yoni's ToM process inferred ('Yoni likes the fruit that…likes')	21 PD nondemented (20 medicated), 21 HC	2.5 (Range 1.0-3.0)	MMSE (cut-off: >26), PANDA, memo test (immediate and delayed recall) digit span forward and backward from WMS-R, verbal and semantic fluency tasks (FAS and animals), TMT, subtest 4 of the German intelligence test battery LPS.	First order ToM Preserved Second order ToM Impaired

Abbreviations: DBS = Deep Brain Stimulation; F, A, S = letters tested ; FPR = Faux Pas Recognition; HC = Healthy Controls; LPS = Leistungsprufsystem; MCST = Modified Wisconsin Card Sorting Test; MDRS = Mattis Dementia Rating Scale; MMSE = Mini-Mental Status Exam; PANDA = Parkinson Neuropsychometric Dementia Assessment; PD = Parkinson's Disease; RME = Reading the Mind in the Eyes; TMT = Trail Making Test; WISC-R = Wechsler Intelligence Scale for Children; WMS-R = Wechsler memory Scale-Revised

Table adapted from Poletti, Enrici, Bonuccelli & Adenzato, 2011.

Table 3

Cognitive Theory of Mind tasks used in PD studies. Tasks are listed in alphabetical order.

Tasks	**Researchers**	**Test Details**	**Participants**	**Hoen and Yahr Stage**	**Cognitive Measures**	**Main Findings**
Deception task	Saltzman et al., 2000	A clip was hidden under 4 egg cups, the 'examiner' and another person pointed to where they thought the clip was (examiner always correct). The participant had to indicate where they thought the clip was.	11 PD nondemented (medicated), 8 HC	2.5	MMSE (cut-off: >26), Vocab subtest of WAIS-R, CCST, verbal fluency task (FAS), five-point fluency task.	Preserved
False beliefs	Monetta et al., 2009	Participants listened to a story and were provided with a written copy. The first order false belief question required participants to describe a person's belief about the world. The second order false belief question required the participant to indicate what the protagonist believed about the second characters knowledge.	11 PD nondemented (medicated), 11 HC	2.5 (0.9)	MDRS, Auditory working memory test, color TMT, tower of London, Warrington Recognition Memory test for faces and words, Benton phoneme discrimination and face recognition subtests, verbal fluency test. MMSE (cut-off : >25), NART	First order false belief mostly preserved (some PD impaired) Second order false belief impaired

Mengelberg & Siegert, 2003	Visual false belief task: the card sequencing task: 18 story sequences presented on 4 cards using cartoon drawings, a false belief category involving a character acting on a false belief because he/she is unaware of a prior event + control conditions.	13PD nondemented (medicated), 11HC	4 II 6 III 2IV (average 2.83)	MMSE (cut-off : >25), NART	Impaired
Mengelberg & Siegert, 2003	Participants read a story about two characters involved in a routine activity, the first order false belief question tested the ability to make an inference about another person's mental state. A second story was read and participants had to answer a second order false belief story that required the understanding of "One character thinks that another character thinks…'	13PD nondemented (medicated), 11HC	4 II 6 III 2IV (average 2.83)	MMSE (cut-off: >26), Vocab subtest of WAIS-R, CCST, verbal fluency task (FAS), five-point fluency task.	First order impaired Second order preserved
Saltzman et al., 2000	False belief stories: participants were read a story and also provided with a written copy. Participants were asked to make a prediction regarding the behaviour of one of the characters in the story based on the information available. The answer required inferences based on mental states and actions of characters.	11 PD nondemented (medicated), 8 HC	2.5		Impaired

Faux pas recognition	Roca et al., 2010	Participants read a story that may or may not contain a social faux pas. 10 stories contain a faux pas and 10 do not. After reading each story the participant is asked if the character committing the faux pas was aware that something inappropriate had been said, and if so, why it was inappropriate.	36 PD (16 medicated 20 drug free), 35 HC	Medicated 1.42 (0.57) Drug-free 1.33 (0.54)	MMSE (score: >24), categorical/phonologic verbal fluency, short version Boston Naming, short version Token test, Rey auditory verbal learning test, delayed recall of the Complex Rey figure and attention, TMT A and B, digit forwards and backwards span, WCST.	Medicated and drug free impaired
	Peron et al., 2009	A shortened version of the faux pas test, using 5 stories with faux pas and 5 stories without.	17 early PD nondemented on and off meds (abstain from meds night before assessment, 26 HC 27 advanced PD nondemented, 26 HC	On 1.5 (0.7) Off 1.5 (0.7) On 1.3 (0.9) Off 2.5. (1.0)	MDRS (score >130), plus the following for the advanced PD patients: MCST, TMT, categorical and literal fluency tasks, the Stroop Test.	Preserved (no difference between on/off) Impaired

	Kawamura & Koyama, 2007		11 PD nondemented, 20 HC	NA	MMSE	Impaired
Perspective taking	Saltzman et al., 2000	Participants were shown a full picture and then half of it was hidden. They were then asked what other characters would think the picture was.	11 PD nondemented (medicated), 8 HC	2.5	MMSE (cut-off: >26), Vocab subtest of WAIS-R, CCST, verbal fluency task (FAS), five-point fluency task.	Preserved
Short passage task	Mengelberg & Siegert, 2003	Participants were required to read a story and answer questions about the text. The theory of mind questions required the participant to make inferences about the characters feelings, thoughts or intentions.	13PD nondemented (medicated), 11HC	4 II 6 III 2IV (average 2.83)	MMSE, NART	Impaired
Spy task	Saltzman et al., 2000	Participants were the spy on a game board and had to retrieve a document without getting caught. Participants were told they were dripping wet and had to demonstrate or explain a solution to this puzzle, (e.g wiping up wet tracks as they went). A correct solution demonstrated the understanding that they had to hide their path.	11 PD nondemented (medicated), 8 HC	2.5	MMSE (cut-off: >26), Vocab subtest of WAIS-R, CCST, verbal fluency task (FAS), five-point fluency task.	Impaired

Yoni task	Bodden et al., 2010	60 items presented on a computer screen, a face named Yoni in the middle with 4 coloured pictures of faces or objects in the corners Cognitive first order = Yoni's mental state inferred (Yoni likes the fruit that … likes') Second order = Yoni's ToM process inferred ('Yoni is thinking of the car that…wants').	21 PD nondemented (20 medicated), 21 HC	2.5 (Range 1.0-3.0)	MMSE (cut-off: >26), PANDA, memo test (immediate and delayed recall) digit span forward and backward from WMS-R, verbal and semantic fluency tasks (FAS and animals), TMT, subtest 4 of the German intelligence test battery LPS.	First order preserved Second order impaired

Abbreviations: CCST = California Card Sorting Task; DBS = Deep Brain Stimulation; F, A, S = letters tested; FPR = Faux Pas Recognition; HC = Healthy Controls; LPS = Leistungsprufsystem; MDRS = Mattis Dementia Rating Scale; MMSE = Mini-Mental Status Exam; NART = National Adult Reading Test; PANDA = Parkinson Neuropsychometric Dementia Assessment; PD = Parkinson's Disease; RME = Reading the Mind in the Eyes; TMT = Trail Making Test; WAIS-R = Wechsler Adult Intelligence Scale-Revised; WCST = Wisconsin Card Sorting Test; WMS-R = Wechsler memory Scale-Revised

Table adapted from Poletti, Enrici, Bonuccelli & Adenzato, 2011.

It is proposed that there are two separate systems involved in ToM tasks (Poletti, Enrici, Adenzato, 2012). Cognitive ToM tasks involve processing inferences about others beliefs and intentions and inferring their mental states, whereas the affective tasks involve processing inferences about other people's emotions and feelings. Each of these ToM categories may have different underlying neural correlates: the dorsolateral PFC influences cognitive ToM performance, whereas the ventromedial PFC influences affective ToM (Figure 2). Across studies, PD patients have greater difficulty performing in cognitive theory of mind tasks than affective theory mind tasks (Tables 2 & 3). That is, most of the affective ToM tasks shown in Table 2 are preserved, whereas most of the cognitive tasks in Table 3 show impairments.

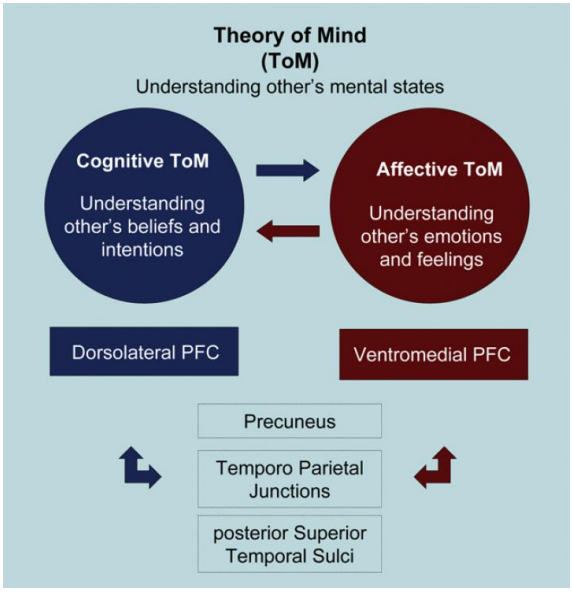


Figure 2. **A model describing the relationship between the two neural systems for cognitive and affective ToM processing (Poletti, M., Enrici, I., & Adenzato, M., 2012).**

Of the affective theory of mind tasks outlined in Table 2, one study reported impairments. Another study showed impairments in the PD group after DBS but not beforehand; the dysfunction may be due to STN micro lesions or stimulation (Peron et al., 2010). The two impairments in affective ToM were found in one study using the Yoni test and Reading the Mind in the Eyes test (Bodden et al., 2010). The fact that these tasks show impairment in this PD group but previous research has shown a lack of impairment may be due to the patient group being in a more progressed stage of the disease. The average Hoen and Yahr stage of the PD participants who showed impairment is 2.5 (Bodden et al., 2010), this is the highest of all the studies in Table 2. One theory proposes that affective ToM may not be influenced until later stages of the disease because of the spatio-temporal depletion of

dopamine (Poletti et al., 2011). In the later stages of PD, the frontostriatal circuits that connect the basal ganglia with medial regions of the prefrontal cortex are affected by dopamine depletion and these areas are thought to be important in affective ToM performance. However the one study in Table 2 that did assess performance of early and late stage PD and affective ToM found no significant differences between patients at either stages of the disease (Peron et al., 2009).

Of the cognitive ToM tests (Table 3), most showed impairments in PD groups, but some tasks did not show impairments in PD groups compared to controls. Two tasks that showed preserved ToM were the deception task and the perspective taking task from Saltzman et al. (2000). These tasks had variable performance levels by PD participants even though the task is usually completed by very young children. Peron et al., (2009) found that performance on the faux pas recognition test was unimpaired in early PD but impaired in advanced stage PD. They suggest that a deficit may not be evident until a degenerative process has spread beyond the dopaminergic pathways. But these results contradict two other studies that did find impairment in early PD with this task (Roca et al., 2010; Kawamura & Koyama, 2007). The results of Peron et al. (2009) may not be comparable to the results of Roca et al. (2010) and Kawamura and Koyama (2007) as they used a shortened version of the faux pas recognition task and had a much smaller sample size. PD participants were unimpaired on the first order component of the Yoni test but were impaired on the second order component (Bodden et al., 2010). First order tasks are easier than second order tasks; however, Mengelberg and Siegert (2003) showed the opposite effect. In the first order false belief stories, participants had impaired functioning compared to controls but there was no significant difference between PD and controls on the second order task. The equal performance of PD and controls could be due to a high working memory load, as both PD and controls performed poorly. This shows that some cognitive ToM tasks may instead

reflect executive functioning deficits and thus provide ambiguous measures of cognitive ToM.

One study suggested there may be a relationship between measures of cognitive ToM and executive functioning, but did not assess comprehension or reasoning ability in order to control for deficits in these areas (Saltzman et al., 2000). By contrast, Mengelberg and Siegert (2003) replicated the study of Saltzman et al. (2000) and added control questions along with different measures of ToM (visual and text based) (Table 2). PD participants scored significantly lower on three out of four ToM tasks (card sequencing task, short passages, and first order false belief story), but were not impaired on any of the non-ToM components which included inference, understanding and memory questions (Mengelberg & Siegert, 2003). Hence impairments in executive functioning may not be necessary for the PD participants to show poor ToM performance.

The card sequencing task designed by Langdon et al., (1997) is a useful ToM task as it dissociates between theory of mind impairment and executive functioning deficits. This task will be used to examine whether performance on Theory of Mind tasks differs between at-risk non-MCI PD patients and healthy controls. This card sequencing task is a good measure of theory of mind as it enables control conditions to address non theory of mind issues like poor sequencing, social understanding and complex reasoning. There are four card categories: false belief, social situations, mechanical and capture. Highly functioning autistic children and Schizophrenic adults can have selective difficulty sequencing false belief stories but can perform equally or better on social script and mechanical conditions (Langdon & Coltheart, 1999). The card sequencing task is also a useful measure as it is a visual task and does not require reading or memory retrieval to respond correctly as the false belief test based stories do.

1.4 A Link between Theory of Mind and Autobiographical Memory: Default Mode Network

Research has consistently identified a distributed group of brain regions that form the DMN. These are the medial posterior cortex (particularly the posterior cingulate cortex and the precuneus), the medial frontal cortex, and the bilateral inferior parietal and posterior temporal areas around the temporoparietal junction (Mars et al., 2012). Figure 3 shows the default

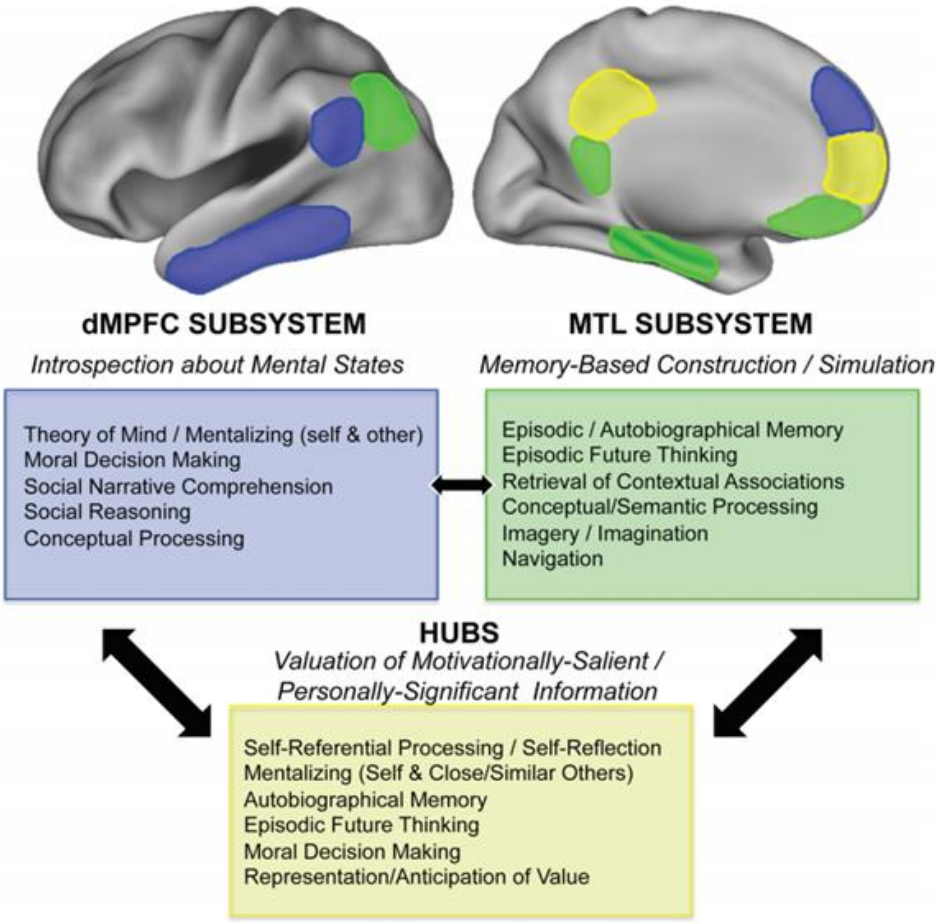


Figure 3. **Major default mode network components. The arrows represent approximate strength of connectivity between components. The DMN hubs are yellow (PCC , aMPFC), the dMPFC subsystem is blue and the MTL subsystem is green. (Andrews-Hanna, 2012).**

mode network (DMN) hubs, plus the two key subsystems and each component's hypothesised function (Andrews-Hanna, 2012). The DMN is a distribution of separate brain regions that show functional connectivity and high basal activity when an individual is at rest

and not engaged in an externally demanding task. The DMN appears to be active during external-stimulus independent thought. By contrast, there is generally decreased activity in the default mode network during externally directed task-related activity.

Relevant to the current study, the DMN has been shown to be activated by both autobiographical memory tasks and theory of mind tasks (Andrews-Hanna, 2012). The fact that the same network underlies functioning on both of these tasks suggests that performance on these tasks may be correlated. While autobiographical memory and theory of mind have traditionally been studied separately, recent fMRI studies have revealed extensive functional overlap in the core brain networks that underlie these processes (Spreng et al., 2008). That is, the default mode network increases in activity during both autobiographical memory and theory of mind tasks. This overlap is exhibited most reliably in the medial prefrontal, medial-temporal and medial lateral parietal cortices, that is, key regions of the DMN. Specifically, the hubs highlighted in Figure 3, namely the posterior cingulate cortex and the anterior medial PFC, underlie the functioning of both autobiographical memory and ToM (Andrews-Hanna, 2012). However, relative differences between the two phenomena are found in that autobiographical memory tasks predominantly engage the MTL subsystem and theory of mind tasks predominantly engage the dMPFC subsystem (Andrews-Hanna, 2012). The theory of mind tasks that were used to evaluate the dMPFC, involved inferring other people's mental states, although these were mainly affective ToM tasks. Research also suggests that midline regions like the medial prefrontal cortex help with coding emotions in mental state attribution (affective ToM) and that the lateral prefrontal cortex supports externally generated representations and might code cognitive aspects of mental state attribution (cognitive ToM) (Olsson & Ochsner, 2007). Nonetheless, the overlap at the level of the key hubs, suggest overlap in terms of memory-based construction (autobiographical episodic retrieval) and cognitive theory of mind tasks (mentalizing the state of others and social reasoning).

Individuals with PD may have reduced functional integrity of the DMN, which may contribute to the development of cognitive decline. Cognitively unimpaired PD patients (defined by cognitive domain z scores that are not below the population mean in any of 3 domains) and who show no significant structural differences compared to controls, nonetheless show functional disruption of the DMN (Tessitore et al., 2012). These patients were at stages 1 and 2 of the Hoehn and Yahr scale, which represent mild disease stage. PD patients, showed decreased functional connectivity in the right medial temporal lobe and bilateral inferior parietal cortex in the DMN. Given that the DMN shows early functional disruption, detecting cognitive changes in tasks related to DMN functioning could be a good predictive measure of future decline. That is, it is of value to identify functional neuropsychological tests that can identify cognitive change in early PD prior to significant changes to patient's cognitive status, such as PD non-MCI to PD-MCI. Autobiographical memory and theory of mind tests may be suitable to detect these early cognitive changes as performance on these tasks may be dependent on the functioning of the DMN.

1.5 The current study

The current study examined both autobiographical memory and theory of mind in individuals who are in early stages of PD but do not have MCI. Performance on these tasks was compared to that shown by age, education and sex matched healthy controls (HC). The AMI was used as a standard instrument to assess both autobiographical and personal semantic memory explicitly. The card sequencing task mentioned above was used to assess cognitive theory of mind. It was expected PD participants would show deficits in both autobiographical memory and theory of mind compared to the HC group, and that there would be an association between the two measures that may be related to overlapping changes to the default mode network. In addition, these measures were assessed in terms of their potential as early markers of future cognitive decline. For this latter aim, the association between the AMI

and ToM measures was assessed with an independent risk score derived from a short screen developed at the NZBRI (Myall et al., 2015 AWCBR). This risk score takes into account the scores of the MoCA, the Stroop-inference, Trails B, the Test of Everyday Attention (TEA) Map Search, as well as age of an individual, to discriminate between converters to PD-MCI or PDD within 4 years and non-converters.

Method

2.1 Participants

Patient volunteer participants were sought from current Christchurch Neurology and Movement disorders clinics. Of 238 participants contacted, 195 completed initial cognitive screening that lead to 20 participants eligible to participate in cognitive testing for the current study (see Figure 3). This group included 2 participants below 29% risk of conversion to dementia within 4 years and 19 participants above 29% risk of conversion to dementia within 4 years. This percentage risk of conversion was determined with an algorithm designed by Daniel Myall based on the scores of the initial screening assessment in an earlier cohort and based on an analysis of progression over 4 years of PD normal (PD-N) patients to either PD-MCI/PDD. The 29% was determined to be the optimal ROC cut-off (Youden Index) for progression to PD-MCI/PDD that gives the optimal sensitivity and specificity, to discriminate PD patients who do/do not show cognitive decline to PD-MCI /PDD across the next four years after screening. So the over 29% risk score group can be considered 'pre-MCI'. The patients eligible for the study as determined by the screening test, undertook two sessions of detailed neuropsychological cognitive status testing (covered by a HRC project) to establish PD-MCI or non-MCI status. Inclusion criteria were a diagnosis of PD without confounding atypical movement disorders. Exclusion criteria were involvement with studies that include pharmacological intervention, history of major medical or psychiatric illness in the past 12 months, current or history of other neurological or psychiatric conditions, medications impacting cognition, history of alcohol or substance abuse, history of learning disability, poor comprehension of the English language, patients under the age of 60 and over the age of 85, and patients with PD who meet criteria for PD with mild cognitive impairment (PD-MCI) or PD with dementia (PD-D).

Table 4
Clinical and Demographics of Participants (mean (SD))

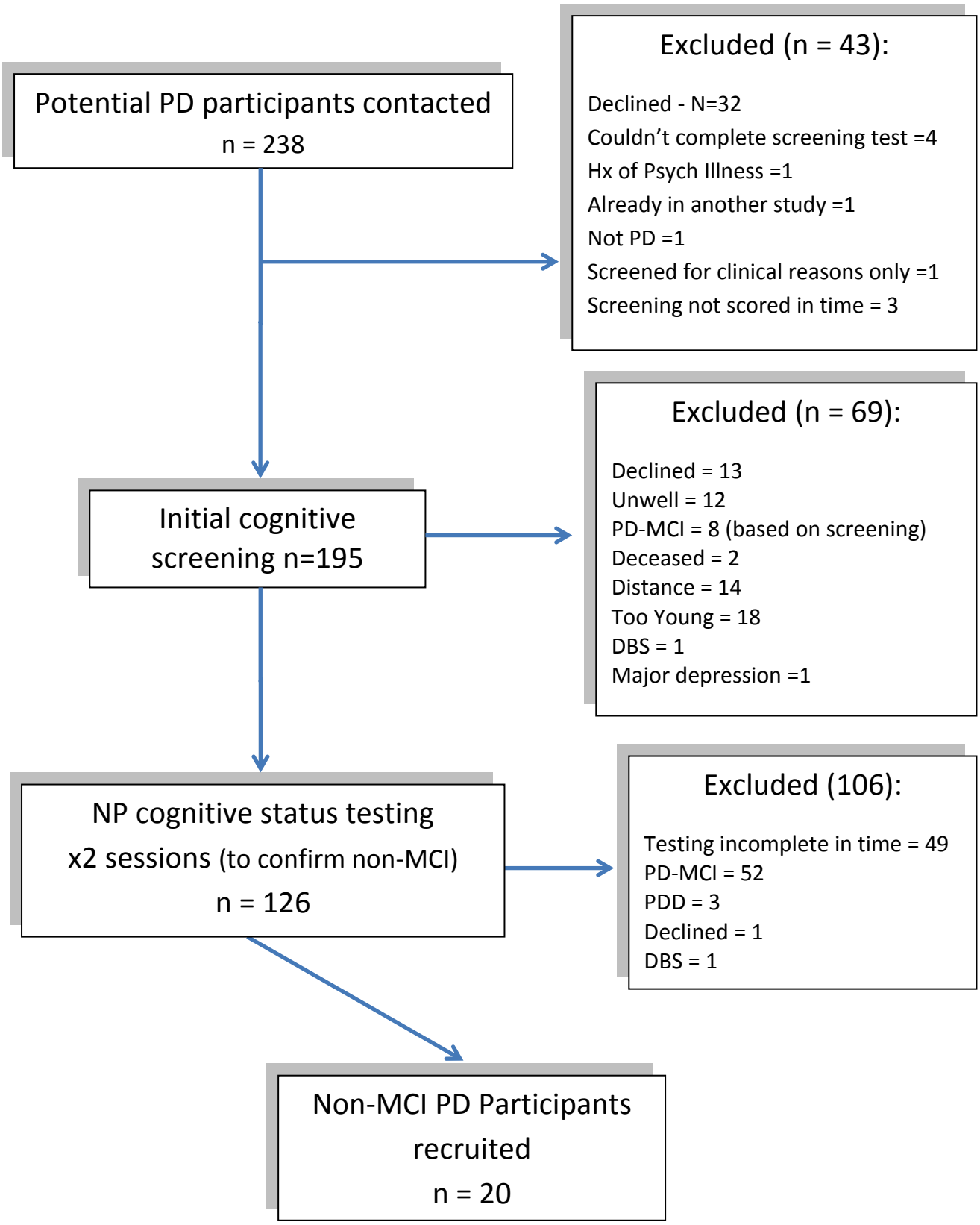
	PD (n = 20)	Healthy Controls (n = 15)	p
Sex, M:F	13:7	9:6	
Age	73.25 (6.53)	74.27 (6.23)	0.65
Education (y)	13.05 (2.21)	14.53 (2.31)	0.08
HADS Anxiety	4.55 (3.07)	3.07 (3.31)	0.18
HADS Depression	3.45 (2.31)	2.27 (2.15)	0.13
GDS	0.15 (0.67)	0.13 (0.52)	0.94
Symptom duration (y)	8.55 (5.24)	-	
Risk score	51.67 (0.22)	-	
Hoehn and Yahr stage	2.55 (0.58)	-	
UPDRS (motor)	34.75 (14.38)	-	
DRS-2 (AESS)	11.55 (3.15)	-	
ADAS-Cog	6.54 (2.07)	-	

Abbreviations: ADAS-Cog = Alzheimer's Dementia Assessment Scale-Cognitive; DRS-2 (AESS) = Dementia Rating Scale (age and education scales score); GDS = Geriatric Depression Scale; HADS = Hospital Anxiety and Depression Scale; UPDRS = Unified Parkinson's Disease Rating Scale

Fifteen healthy controls identified from the NZBRI database were contacted and matched to fifteen of the PD participants based on age, education (see table 4 for a summary of the participants).

Figure 3.

Process of PD participant selection and testing



2.2 Initial Cognitive Screening Tests

Parkinson's patients underwent an initial screening assessment to determine their cognitive status as part of a larger HRC study on PD-MCI at the NZBRI. The screening assessment took approximately 20-30 minutes and consists of the MoCA, the Stroop test, the Trails test version A and B, and the Test of Everyday Attention Map Search. The NZBRI has shown that these screening tests, when adjusted for age, provide a sensitive composite measure of probability of future progression to PDD (Myall et al., 2015 AWCBR). Healthy controls also completed this short screening test.

2.3 Neuropsychological Assessment to establish non-MCI status

Testing was completed across two sessions that took 2 to 3 hours each using procedures prescribed in the appropriate manuals. The tests used covered the five cognitive domains required by the Movement Disorders task force Level II criteria (Litvan et al., 2011). Attention, working memory and processing speed was assessed using the Digits Forwards and backwards Test, Digit Ordering (Wechsler, 2008), TEA (MAP Search) (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994), Trails A, and DKEFS Stroop colour and Stroop word (Delis, Kaplan & Kramer, 2001). Executive function was assessed by Trails B, D-KEFS letter fluency, category fluency, category switching and the colour word inference test (Stroop) (Delis, Kaplan & Kramer, 2001) and action fluency (Piatt, Fields, Paolo, & Troster, 2004). Visuospatial functioning was tested with the judgement of line orientation (JLO) (Benton, Hasher, Varney & Spreen, 1983), the fragmented letters subtest of the visual object and space perception (VOSP) (Warrington & James, 1991), the Rey-Osterrieth Complex Figure Copy (Meyers & Meyers, 1995), pentagons and the Wechsler Adult Intelligence Scale IV picture completion task (Wechsler, 2008). Learning and memory was assessed learning with the short from California Verbal Learning Test II (Delis, Kramer, Kaplan & Ober, 2000) and immediate and delayed recall of the Rey-Osterrieth Complex Figure (Meyers & Meyers,

1995). Language was assessed with the Boston naming test (Kaplan, Goodglass, & Weintraub, 1983), the similarities component in the Dementia Rating Scale-2 (Jurica, Leitten & Mattis, 2001), and the language component of the ADAS-Cog (Mohs et al., 1997).

2.4 Cognitive Testing session for the current study

2.4.1 The Autobiographical Memory Interview (AMI)

Using the AMI (Kopelman et al., 1989), participants were required to recall personal semantic details and autobiographical episodic incidents by answering questions from three lifetime periods: childhood, early adult life and recent life. Within each lifetime period there were three sections that focused on key events specific to that time period. Within the 'childhood' period, the subsections focused on: period before school, first school, secondary school. Within the 'early adult life' period, the subsections focused on: career, wedding and children. Within the 'recent life' period, the subsections focus on: present hospitalisation, previous hospitalisation, last Christmas, and a holiday.

Each subsection of the AMI contains personal semantic questions and an autobiographical incident (episodic) question. The personal semantic questions require participants to recall facts such as names of teachers or friends, date and place of wedding and the time and location of holidays. When the participant was unable to answer, prompts were provided to stimulate a response, for example if a participant is struggling to recall teachers or friends names they were asked, 'can you name your form teacher? A friend?' The autobiographical questions required the participants to provide specific and detailed memories for each subsection, each specific to a different time frame. Prompts were provided to stimulate a response when the participant could not think of an incident. For example, for an incident at college they were asked, 'can you think of an incident from your first day at college, or an incident with a friend?' Participant responses were recorded as accurately as

possible and verbal responses recorded on a Zoom H4N recorder, with their permission, to improve accuracy. AMI testing took 30 to 50 minutes on average.

2.4.2 Theory of Mind Task - card-sequencing task (Langdon et al., 1999)

This task used short story sequences presented on cards depicting black and white cartoon-like drawings. Each story had 4 cards (21cm x 15cm). The cards for any given story were presented face down in a line in front of the participant. The participant was asked to turn the cards over of any given story and rearrange the cards in the correct sequential order. There were two practice sequences (mechanical) followed by 4 stories of 4 cards in each for each of four conditions: (a) False belief: a character acts on a false belief because they are unaware of an event that occurred prior; (b) A social script: character(s) involved in a daily activity; (c) Mechanical: Objects interacting; (d) Capture: depicting routine activities but a misleading cue makes the sequence more complex. The last three conditions were used to control for sequencing errors caused by factors other than false belief ToM deficits. These control conditions tested participant's ability to reason logically (social script), to infer causal relations (mechanical), and to disengage from a cognitively salient misleading cue (capture). The semi-random order of the cards for any given story was determined by coloured dots on the back (orange, blue, yellow, green) so that the cards were presented to each participant in the same random order. The order that the stories and order of the conditions were presented in was also semi-random. Appendix A shows the scoring sheet which also shows the order the cards are presented in.

2.4.3 Other Cognitive Tests

2.4.3.1 The Brief Visuospatial Memory Test-Revised (BVMT) (Benedict, 1997)

Participants were presented with 6 figures (objects) on a single page in the recall stimulus booklet to study for 10 seconds and then asked to recall each figure by drawing them in the response form as they appeared and in their correct location on the page. This procedure was repeated for a total of 3 learning trials. After a delay of 25 mins of verbal tasks, a delayed recall trial was administered. Participants were given points for drawing the correct object shape, and for drawing it in the correct location, there was a maximum of 12 points for each trial. A recognition trial was administered immediately after the delayed recall trial. Participants were shown the 12 figures one at a time and had to indicate if they had or had not previously seen the figures on the display (6 figures were from the display and 6 figures were new).

2.4.3.2 The Symbol Digit Modalities Test (SDMT) (Smith, 1982)

Participants were administered the written response format of the SDMT. Participants were required to use a coded key to match 9 abstract symbols paired with numerical digits. Participants were given 10 practice items and corrected immediately if they made any errors. Participants were then required to match as many symbols to the correct numbers within 90 seconds.

2.4.3.3 The Sydney Language Battery (SYDBAT) (Savage et al., 2013)

Participants completed the four subtests of the SYDBAT: naming, repetition, word comprehension and semantic associations, there were 30 items in each subtest. For the naming task, participants were required to provide the name of an item shown in a colour photograph, presented one at a time for a total of 30 items. Mispronunciations of words due

to phonemic, phonological, or semantic substitutions were considered errors. For the repetition task, participants were required to listen and repeat each of 30 words after the tester. For the word comprehension task, participants were required to select a picture shown on screen that best matched the word spoken by the tester (30 word were used). The correct picture had to be selected out of one target item and six foils, which were semantically close to the target or visually similar. The semantic association task required participants to select one picture from a set of four options that was most closely associated with the target picture (30 pictures were used). The four options were semantically related to each other but only one was closely related to the target. Administration of all four subtests took approximately 20 minutes.

2.4.4 Psychiatric Tests

All participants were given the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) and the 15 item brief version of the Geriatric Depression Scale (GDS) (Yesavage et al., 1983).

2.5 Procedure

This study was approved by the Northern B Health and Disability Ethics Committee of New Zealand. Informed consent was given by all participants prior to the initial screening session. Participants completed the screening session in quiet test rooms at the NZBRI or at the Movement Disorders clinic in Christchurch. Participants were then invited to take part in the neuropsychological assessment (2 sessions) at the NZBRI to establish their PD-N or PD-MCI status. Based on this full neuropsychological assessment, participants who were eligible were invited to take part in the testing for the current study. For the current study, participants received the SDMT, the Autobiographical Memory Interview, the card sequencing task, the BVMT, and the SYDBAT. All tests were given in that order during one session which took 1

to 1.5 hours, except for the SYDBAT which was given to the PD participants as part of another session a few days later (this session was part of a larger study using the same participant group). The testing for the current study took place within 7 months of the full neuropsychological assessment. Matched healthy control participants completed the last list of neuropsychological tests listed above along with the initial screening tests (they did not complete the two sessions for the main neuropsychological assessment).

2.6 Statistical Analyses

Group comparisons using independent t-tests assessed the majority of measures, with non-parametric equivalents (Mann-Whitney) when appropriate. Between-within ANOVA was used when repeated measures were examined in the AMI and the BVMT. Additional factors were added as covariates in a series of ANCOVAs. The association between two scores were assessed using the Pearson r statistic or Spearman's non-parametric equivalent.

Results

3.1 Initial Cognitive Screening

Table 5 summarises the performance of the PD and HC participants in the screening tests and the specific cognitive tests included in the current study. Scores were age and education adjusted where possible using normative data supplied by manuals. In the case of the SDMT, normative data was taken from a study using a sample of 14,456 Australian participants (Kiely, Butterworth, Watson & Wooden, 2014). Normative data for the SYDBAT was taken from the average performance of the control group provided by Savage et al. (2013). Of the screening measures, the MoCA, Stroop-inference, and TEA Map Search showed large mean differences between the PD and HC groups, (see effect sizes (ES), Table 5). A similar large effect of impaired performance in the PD group was found for SDMT, but the BVMT total and delayed recall showed even larger impairments in the PD group. No impairment was found in the PD group for any of the four language measures. Appendix B shows the results of the more general cognitive status testing for the PD participants. As a group, Appendix B shows that the PD participants performed above average when compared to normative data on most tests (Digits Forward and backward, Stroop word and inference, Trails A and B, Letter fluency, Category fluency, VOSP, Picture Completion, CVLT free recall, short delay recall and long delay recall, Rey Immediate, Rey delay, and the Language ADAS-cog). Some tests do suggest a mild impairment in the PD group (digit ordering, TEA Map Search, Action Fluency and the Rey Copy). However, no individual PD participant was impaired (more than 1.5SD below the mean) at more than one test across different domains. Results of this more general neuropsychological testing show that none of the participants met PD-MCI criteria. Appendix C shows the scores of the two under 29% risk score participants for the screening tests and other key measures. The two under 29% risk score participants scored equal to or

higher than the average HC on the MoCA, and higher than HC on all other screening measures except the TEA Map Search.

Table 5

Means (SD) and effect sizes of the screening test and other cognitive measures.

	PD mean (SD) (n = 20)	**HC mean (SD) (n=15)**	**Effect size (d) (95% CI)**
Screening Tests			
MoCA** (raw score)	26 (2.22)	27.80 (1.61)	0.91* (0.2 to 1.6)
Stroop-colour (z-score)	-0.20 (0.57)	0.18 (1.01)	0.48 (-0.2 to 1.2)
Stroop-word (z-score)	0.05 (0.64)	0.40 (0.82)	0.48 (-0.2 to 1.2)
Stroop-inference** (z-score)	0.03 (0.99)	0.80 (0.78)	0.85* (0.2 to 1.5)
Trails A (z-score)	0.38 (0.72)	0.78 (0.55)	0.61* (-0.1 to 1.3)
Trails B** (z-score)	0.50 (0.90)	0.61 (0.89)	0.12* (-0.6 to 0.8)
TEA Map Search** (z-score)	-0.82 (0.62)	2.69 (5.52)	0.97* (0.3 to 1.7)
Other Cognitive Measures			
Visual Memory (z-scores)			
BVMT – Total Recall	-0.72 (1.11)	1.07 (0.95)	1.71* (1.0 to 2.4)
BVMT – Delayed Recall	-0.25 (0.94)	0.99 (0.65)	1.49* (0.8 to 2.2)
Attention/Recognition			
SDMT (raw score)	30.25 (8.96)	40.33 (9.85)	1.08* (0.4 to 1.8)
SDMT (z-score)	-1.02 (0.91)	0.00 (1.00)	
Language (raw scores)			
Sydney language battery			
Naming	26.25 (1.83)	26.47 (2.72)	0.10 (-0.6 to 0.8)
Repetition	30.00 (0.00)	30.00 (0.00)	0.00
Word Comprehension	29.45 (0.76)	29.00 (0.85)	-0.56 (-1.3 to 0.1)
Semantic Association	27.80 (1.44)	28.33 (1.18)	-0.40 (-0.3 to 1.1)

Note: * p<.05

** = tasks used in the model for the PD-MCI/PDD risk score

BVMT Total recall is the sum of recall from Trial 1, 2 and 3.

3.2 The Brief Visuospatial Memory Test-Revised (BVMT)

Figure 4 shows the mean scores for learning trials 1 to 3 and the delayed recall trial. There was a significant main effect of group, HC correctly recalled more figures than PD over trials 1, 2 and 3 ($F(1,33) = 24.17$, $p < 0.001$). This main effect remained when controlling for age, sex and education ($F(3,28) = 7.03$, $p < 0.002$) and anxiety and depression ($F(3,29) = 7.42$, $p < 0.001$). Both PD and HC groups correctly recalled a greater number of figures in their correct locations as the learning trials increased ($F(2,66) = 148.74$, $p <0.001$). The improvement in recall was greater between trial 1 to 2 ($F(2,32) = 12.00$, $p <0.001$), than between trial 2 to 3 for both groups ($F(2,32) = 10.64$, $p < 0.001$). However, there was no group by trial interaction ($F(2,66) = 1.14$, $p = 0.33$). Recall at trial 3 and at delayed recall 25 minutes later was similar, with no significant difference in scores between these two recall periods ($F(1,33) = 3.10$, $p = 0.09$), but the HC group again recalled significantly more than the PD group at delayed recall ($t(33) = 7.56$, $p < 0.001$).

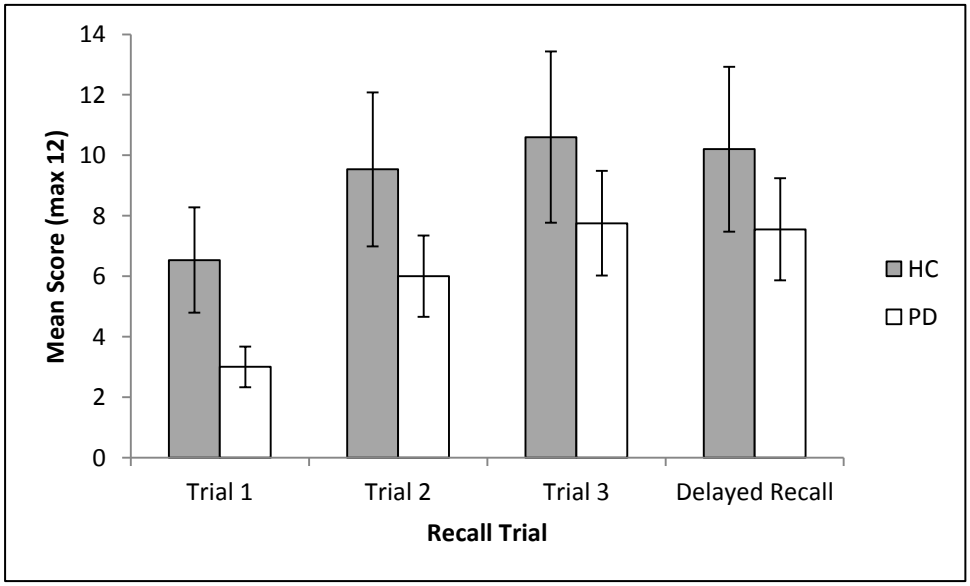


Figure 4. **Means (SEM) scores across recall trials in the BVMT.**

Figure 5 shows the distribution of individualised raw scores across trials for both groups, showing a broad range of scores in the PD group across all trials. Trial 3 had the most variation in raw scores between PD participants (SD = 2.45), whereas for HC participants, trial 3 had the smallest range of raw scores out of the different recall periods (SD = 1.40). When the analysis of scores for trials 1 to 3 were re-run without the 2 under 29% risk of conversion participants, there was still a main effect of group ($F(3,29) = 7.44$, $p < 0.001$. The two under 29% risk score participants scored 3 and 3 for trial 1, 8 and 6 for trial 2 and 10 and 10 for trial 3. These scores can be seen in Figure 5. These scores were on or above average compared to the mean of PD participant scores. The trial 3 recall scores of the two under 29% risk score participants were 10 and 10 and the delayed recall scores were 10 and 9, so there

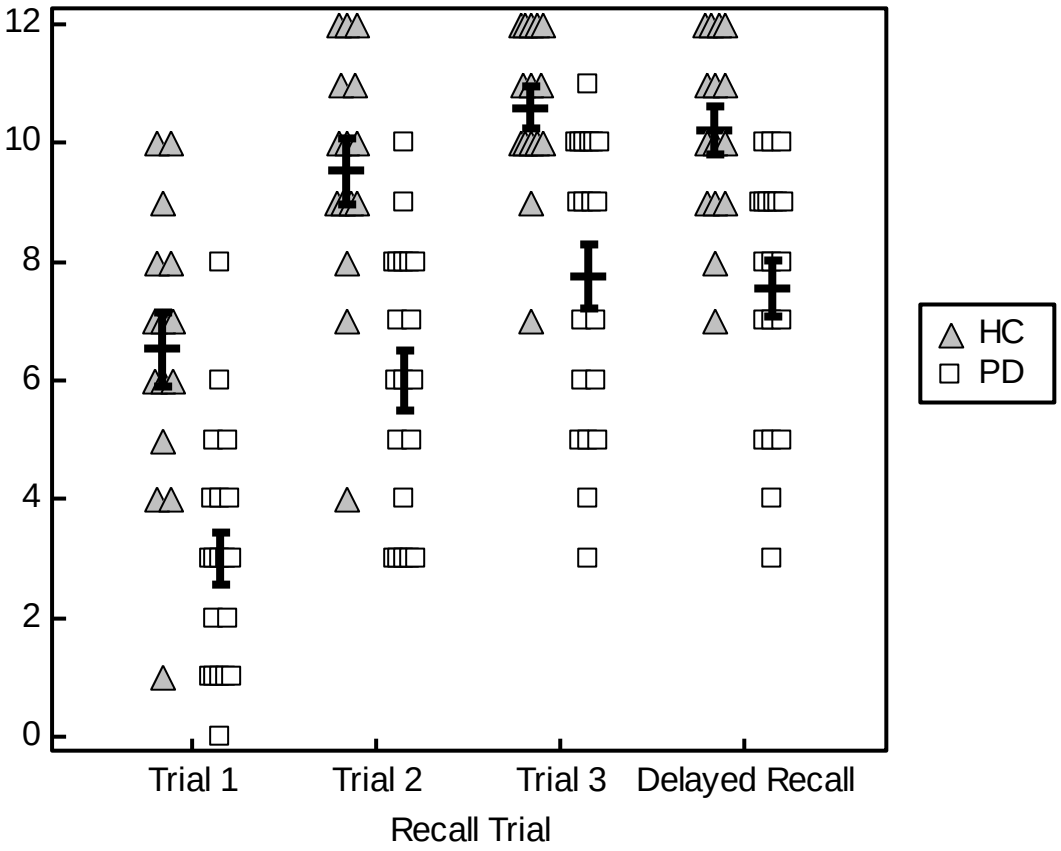


Figure 5. **Distribution of individual raw scores across BVMT trials for both groups. Horizontal lines represent means; vertical lines = SEM.**

was no difference between the overall recall of the PD group at these two recall periods when these were removed. The significant difference between PD and HC at delayed recall remained ($t(31) = 4.35$, $p < 0.001$). These delayed recall scores were 0.9 and 0.5 standard deviations above the mean.

Table 6 shows the mean recognition scores for both PD and HC groups. Recognition hits show the correct identification of figures already displayed during the BVMT recall trials (max 6), false alarms show the incorrect recognition of figures that were not displayed in the recall trials (max 6) and the recognition discrimination index reflects the ability to discriminate previously presented target stimuli from non-target stimuli, this is the best measure of recognition memory. As shown in Table 6, there is no difference between PD and HC groups on any recognition measure (Mann-Whitney U pairwise comparison $p > 0.2$)

Table 6

BVMT Recognition Trial: Hits, False Alarms and Discrimination Index (median and range)

	PD (n = 20)	**HC (n = 15)**
Recognition Hits	6 (4-6)	6 (4-6)
Recognition False Alarms	0 (0-2)	0 (0)
Recognition Discrimination Index	6 (3-6)	6 (4-6)

3.3 The Symbol Digit Modalities Test

Figure 6 shows the age and education weighted z scores of both PD and HC. PD participants scored significantly lower on the SDMT than the HC group ($t(33) = 3.16$, $p<.01$). There was greater variation in the z scores of the PD group; however, the majority of PD participants are performing below average (13 out of 20 PD, as opposed to 7 out of 15 HC). When the analysis of scores for the SDMT were re-run without the 2 under 29% risk of conversion participants, there was still a significant difference between PD and HC performance ($t(31) = 3.85$, $p <0.001$). The two under 29% risk score participants achieved high z scores of 1.32 and 1.66, which was well above average compared to the majority of PD and HC participants.

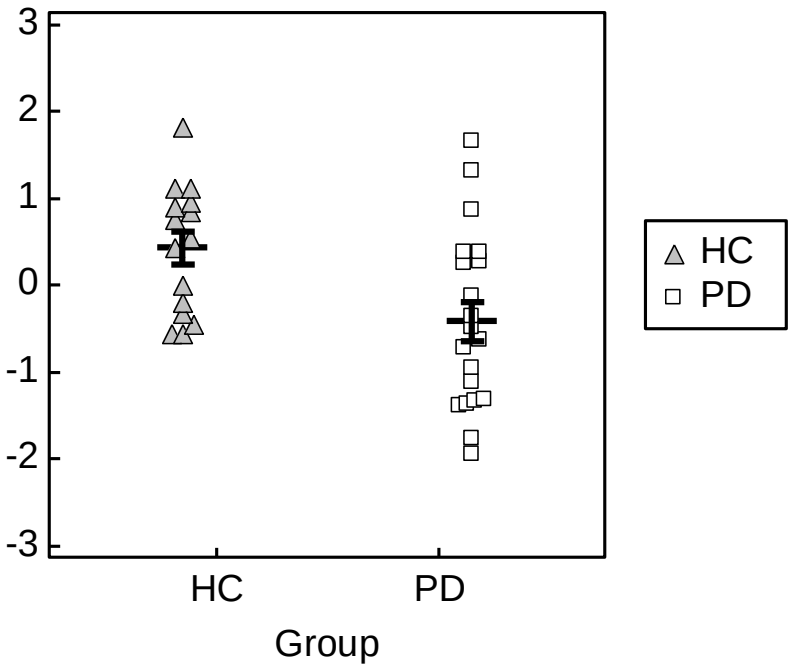


Figure 6. Z score distribution of SDMT scores for Parkinson's and healthy control participants

3.4 The Sydney Language Battery (SYDBAT)

Figure 7 shows the distribution of scores for PD and HC for each subtest of the SYDBAT. There were no significant differences between PD and HC on any of the four subtests (Mann-Whitney U, $p > 0.50$). Every participant from both the PD and HC group scored a perfect score (30) for the repetition subtest. Participants from both groups scored highly for the other subtests of naming, word comprehension and semantic association (means (SD) and effect sizes are shown in Table 5). The horizontal lines represent the mean of a sample of healthy controls from Savage et al. (2013). Both HC and PD groups perform slightly lower than the normative mean control group for the naming subtest but similar to that control group on the three other subtests.

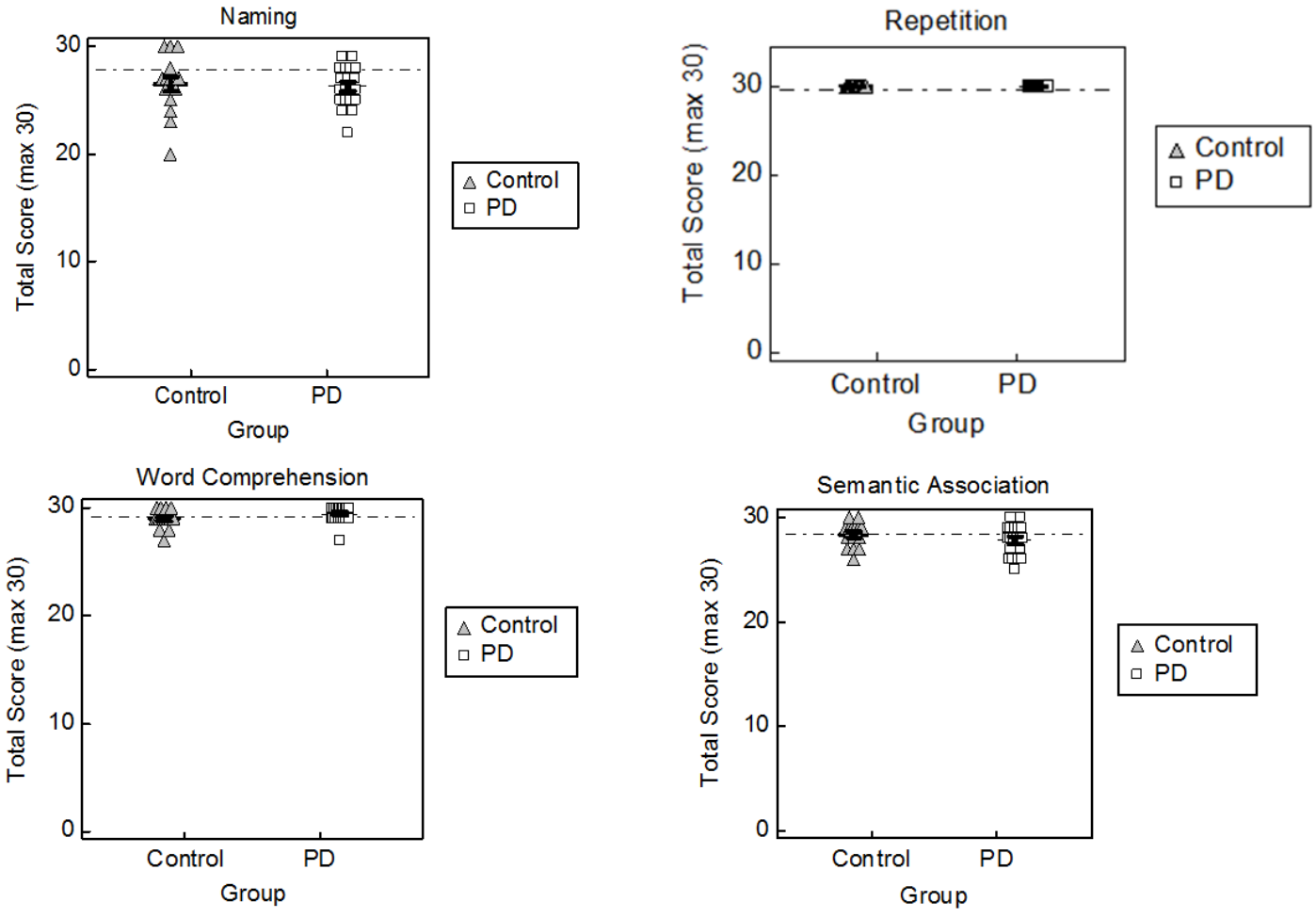


Figure 7. **Distribution of scores on the four subtests of the SDMT for Parkinson's disease and healthy control participants. The horizontal lines show the means for each subtest from the control sample from Savage et al. (2013).**

3.5 Theory of Mind (ToM) and sequencing tasks

Figure 8 shows the raw scores for PD and HC across the 4 card sequencing tasks. Only the false belief task (ToM condition) showed a difference between the two groups, with poorer accuracy in the PD group ($t(33) = 2.55$, $p = 0.02$; Newman-Keuls $p=0.01$; ES = 0.87 (0.2 to 1.6)). This difference between groups remained when controlling for age, sex and education ($F(1,30) = 4.48$, $p < 0.05$) and when controlling for anxiety and depression ($F(1,31) = 7.31$, $p < 0.02$). No group differences were evident for any of the other 3 card sequencing tasks ($t(33) < 1.17$, $p > 0.24$). Some PD participants performed above the average score of the control group for the false belief task (n = 5), but most scores were clustered in the low score region with 15 out of 20 PD participants scoring below the average score of controls. When the analysis between PD and HC performance on the false belief task was re-run, removing the 2 under 29% risk of conversion participants (one scored 11, the other 9), the result was similar, $t(31) = 2.34$, $p = 0.03$. The scores of the two under 29% risk score participants were close to

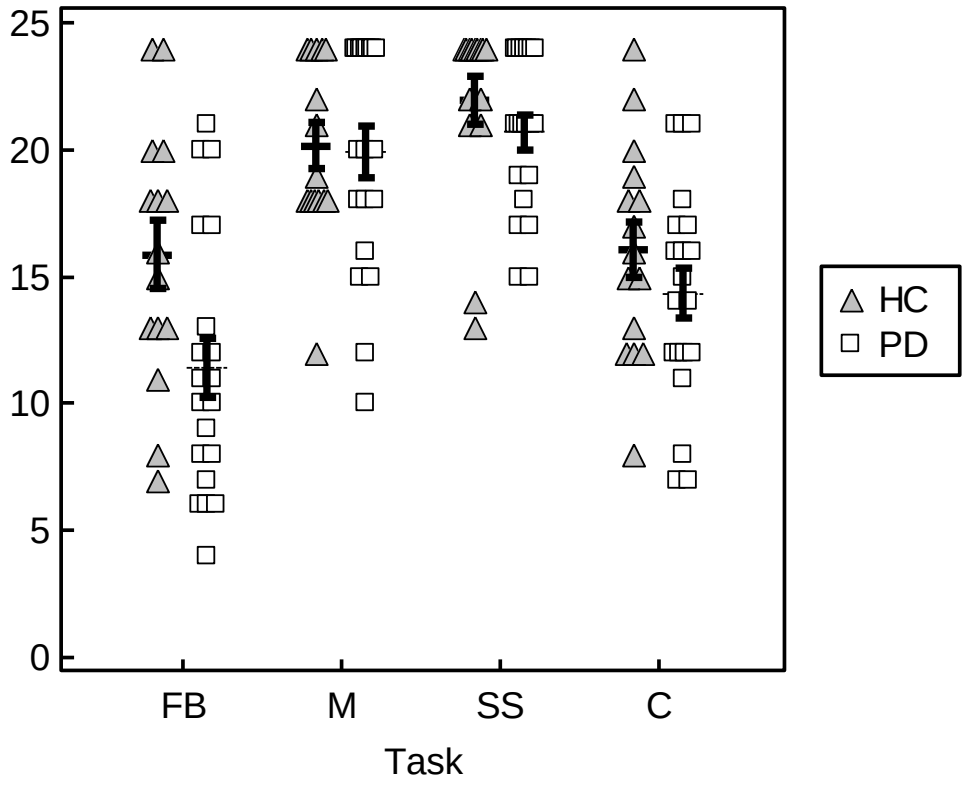


Figure 8. **Raw scores for Parkinson's disease and healthy control participants in each of the four card sequencing tasks. Error bars show standard error of the mean.**
FB = False Belief; SS = Social Script, M = Mechanical, C = Complex Reasoning
Error bars show standard error. Note: Maximum score for each card type is 24.

the average ToM score of the PD group, they scored 9 (-1.34 SD below the mean of HC) and 11 (-0.95 SD below mean).

3.6 The Autobiographical Memory Interview (AMI)

3.6.1 Autobiographical (episodic) Memory

Figure 9 shows the mean autobiographical incident (episodic) memory scores for the two groups across the 3 AMI time periods. The PD group scored lower than the HC group across the three time points (main effect, $F(1,33) = 9.04$, $p = 0.01$; ES (d) = 1.03 (0.3 to 1.7)). The main effect for group remained when controlling for age, sex and education ($F(1, 30) = 7.73$, $p = 0.01$) and when controlling for anxiety and depression ($F(3,29) = 3.40$, $p = 0.03$). Although the group difference was greater for childhood than for young adult and recent autobiographical memory, there was no significant group x time period interaction ($F(2,66) = 2.48$, $p = .09$). However Newman-Keuls post hoc tests suggested a significant difference between PD and HC on autobiographical memory recall only during the childhood period ($p<.01$).

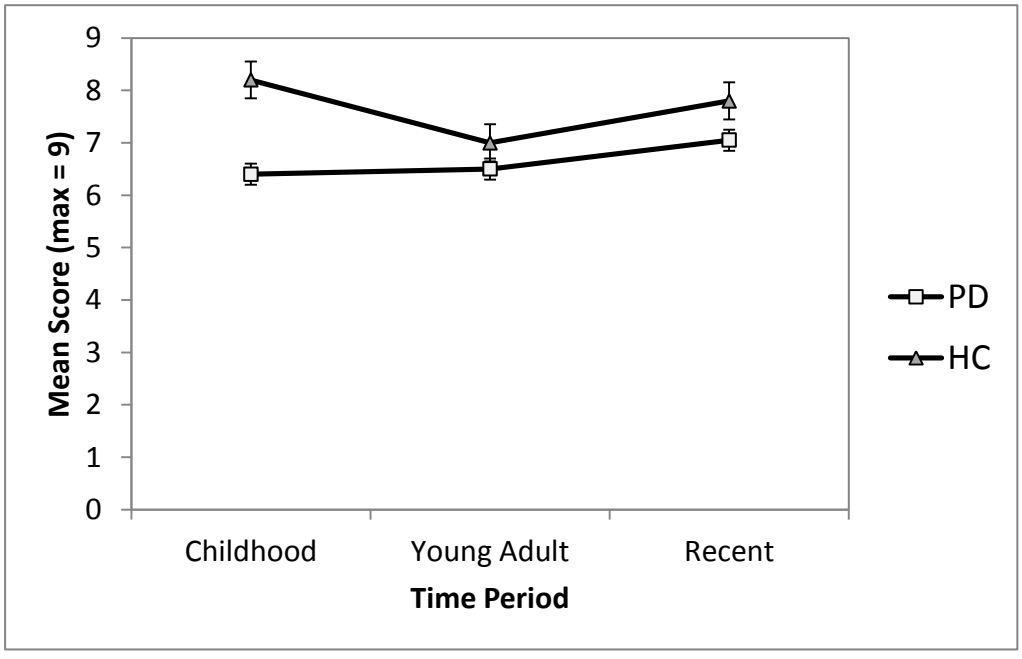


Figure 9. **Mean autobiographical memory scores for Parkinson's disease and healthy control participants at three time points during the Autobiographical Memory Interview.**

When the analysis between PD and HC for autobiographical memory recall was re-run without the 2 under 29% risk of conversion participants, the results were similar, $F(1,31) = 6.83$, $p = 0.01$. The two under 29% risk score participants scored 4 (-3 SD below mean of HC) and 5 (-3 SD below mean) for the childhood period, 6 (-0.8 SD below mean) and 5 (-1.6 SD below mean) for the young adult period and 5 (-2.98 SD below mean) and 8 (0.21 SD above mean) for the recent period. These scores can be seen in Figure 10 which shows the distribution of raw scores for incident (episodic) autobiographical memory. Apart from the score of 8, these scores are below average for the HC and PD participants' autobiographical memory recall. Overall PD scores were much more dispersed for the childhood period.

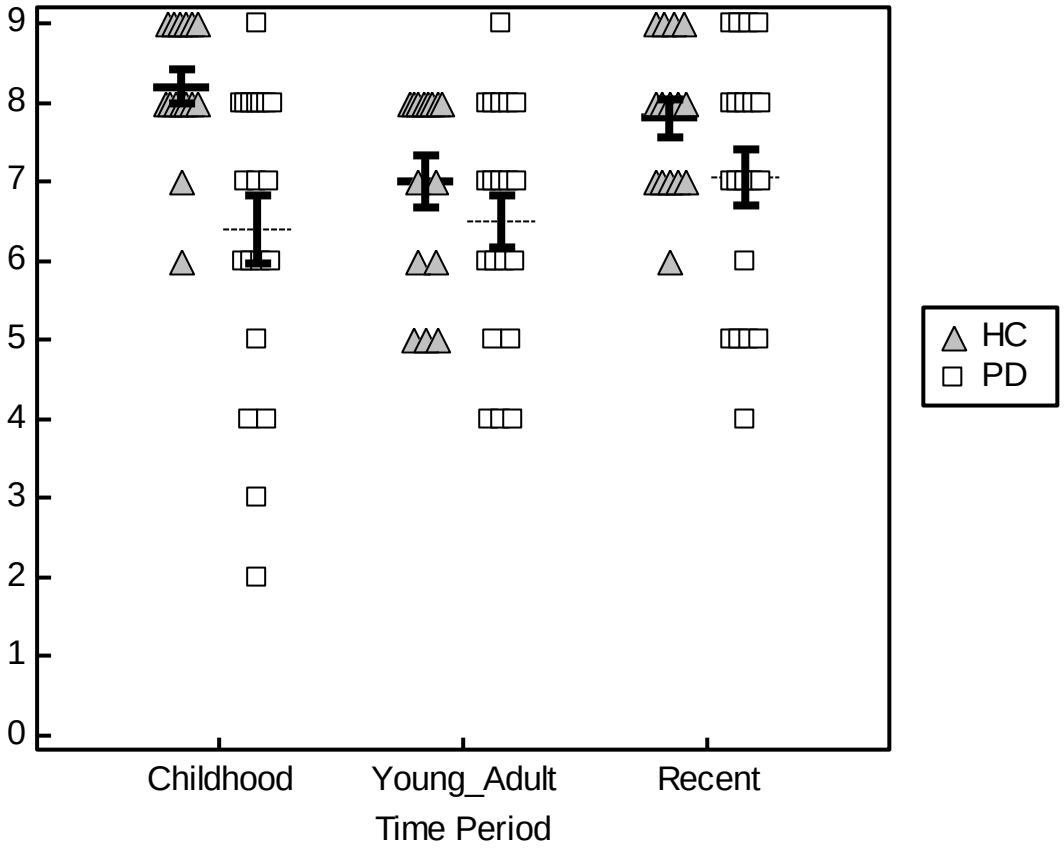


Figure 10. **Distribution of autobiographical incident (episodic) memory scores for PD and HC for each time period in the AMI.**

3.6.2 Personal Semantic Memory

Figure 11 shows the mean personal semantic memory scores for the two groups across the three AMI time periods. There was no significant difference between the personal semantic recall of PD or HC groups ($F(1,33) = 1.49$, $p=.23$; ES (d) = 0.42 (-0.3 to 1.1)). There was a significant main effect of time ($F(2,66) = 16.91$, $p<.001$). Both PD and HC groups had poorer recall for the childhood period but achieved an almost perfect score for the recent period. However when controlling for age, sex and education this difference across time points was no longer significant, $F(2,60) = 1.52$, $p =.22$, the same was found when controlling for anxiety and depression ($F(3,29) = 0.129$, $p = 0.90$).

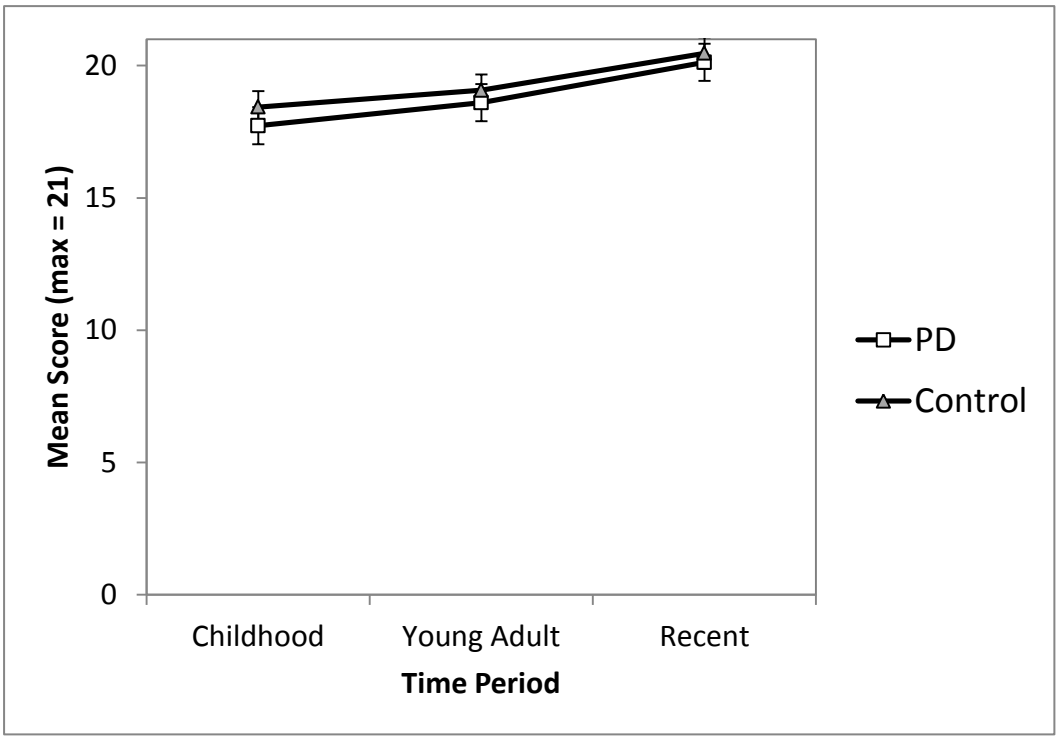


Figure 11. **Mean personal semantic memory scores of Parkinson's disease and healthy controls at each time point during the Autobiographical Memory Interview.**

When the analysis between PD and HC for semantic memory recall was re-run without the 2 under 29% risk of conversion participants, the main effect of group remained non-significant ($F(1,31) = 1.94$, $p=.17$) and the effect of time was similar ($F(2, 62) = 16.35$, $p < 0.001$). The two under 29% risk score participants scored 19 (0.23 SD above mean of HC) and 18 (-0.17 SD below mean) for the childhood period, 19.5 (0.29 SD above mean) and 21 (1.30 SD above mean) for the young adult period and 21 (0.79 SD above mean) and 19.5 (-1.45 SD below mean) for the recent period. These scores can be seen in Figure 12 which shows the distribution of raw scores for personal semantic memory. The scores of the two under 29% risk score participants are above the average of the PD participant group for all 3 time periods. Overall Figure 12 shows both PD and HC are clustered together apart from a few low scores in each group.

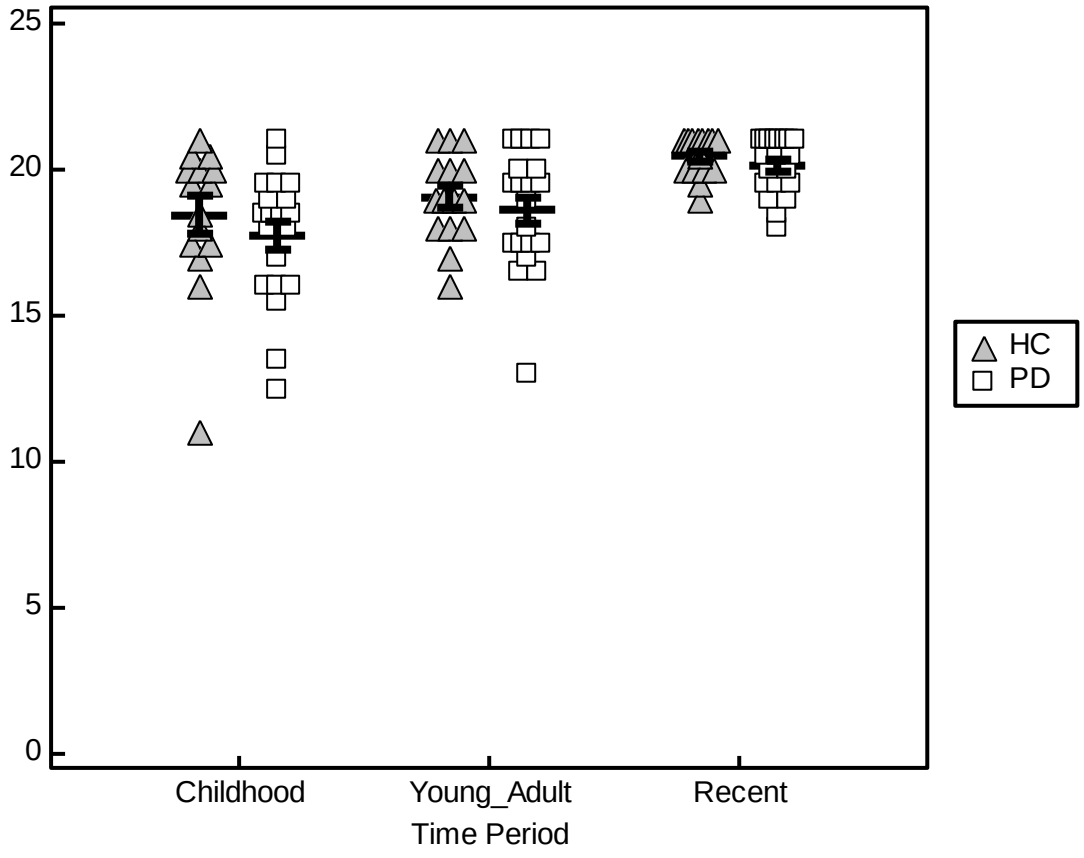


Figure 12. **Distribution of semantic memory scores for PD and HC for each time period of the AMI.**

3.7 Correlations between the cognitive decline risk score and other measures

Table 7 shows correlations between test scores of the PD participants only (n= 20) and the risk score (percentage likelihood of future progression to PD-MCI or PDD). Only two tests were significantly correlated with the risk score: the BVMT delayed recall and the SDMT. Table 7 shows a moderate negative correlation between the BVMT delayed recall and the progression score, r = -0.47. Generally as BVMT delayed recall scores decrease, the risk score increases. There was a moderate negative correlation between the SDMT and the risk score, r = -0.56. Generally as SDMT scores decrease, the risk score increases. Table 7 shows correlations between the scores of two other tests: BVMT delayed recall with BVMT total recall, and BVMT delayed recall with the SDMT. As expected there was a strong positive correlation between the BVMT total recall and the BVMT delayed recall, r = 0.80. There was also a moderate positive correlation between the SDMT and delayed recall, r = 0.60. There was also a modest but non-significant correlation between BVMT delayed recall and false belief (ToM) (r = 0.30) and between BVMT delayed recall and AMI incidents (-0.36).

Table 7

Correlations between test scores of the Parkinson's group and the percentage decline risk score

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Risk score	1.00										
2. False belief	-0.00	1.00									
3. AMI incidents	0.10	-0.09	1.00								
4. AMI semantic	-0.14	-0.14	0.32	1.00							
5. BVMT total recall	-0.11	0.37	-0.42	-0.18	1.00						
6. BVMT delayed recall	-0.47*	0.30	-0.36	-0.09	0.80*	1.00					
7. SDMT	-0.56*	0.08	0.04	0.28	0.36	0.60*	1.00				
8. SYDBAT Naming	-0.31	0.18	-0.01	0.07	-0.15	-0.21	-0.20	1.00			
9. SYDBAT Repetition**									1.00		
10. SYDBAT Word Comp	-0.11	-0.12	0.08	-0.27	-0.23	-0.31	-0.34	0.10		1.00	
11. SYDBAT Semantic Association	-0.12	0.08	0.25	-0.09	-0.28	-0.28	-0.32	0.26		0.42	1.00

Note *p<.05

**no correlation because max SYDBAT repetition scores

Abbreviations: AMI = Autobiographical Memory Interview; BVMT = Brief Visuospatial Memory Test; SYDBAT = The Sydney Language Battery

3.8 Correlations between clinical Parkinson's disease measures and key test scores

Table 8 shows the correlations between key Parkinson's disease stage measures and key testing measures form the current study (n=20). There was a moderate significant correlation between the UPDRS motor and Hoehn and Yahr stage (0.66). There was a moderate significant correlation between the DRS-2 and BVMT delayed recall (0.66). There was a moderate correlation between HADS Anxiety and HADS Depression (0.56) and between

HADS Anxiety and AMI semantic recall (-0.56). There was also a correlation between HADS Depression and AMI semantic recall (-0.49).

There were modest but non-significant correlations between the risk score and DRS-2 (-0.41), between the AMI incident recall and the DRS-2 (-0.39), between the ADAS-Cog and DRS-2 (-0.33), between false belief and the ADAS-Cog (-0.33), between AMI incident recall and the ADAS-Cog (0.34) and between AMI semantic recall and disease duration (0.33).

Table 8.

Correlations between Parkinson's disease participant motor, cognition and disease stage measures and key testing measures

	UPDRS motor	**DRS-2**	**H & Y**	**ADAS -Cog**	**ADL-IS**	**HADS Anxiety**	**HADS Depression**	**Disease Duration**
UPDRS motor	1.00							
DRS-2	-0.05	1.00						
H & Y	0.66	0.44	1.00					
ADAS-Cog	-0.10	-0.33	0.02	1.00				
ADL-IS**	0.25	-0.35	-0.02	-0.19	1.00			
HADS Anxiety	-0.09	-0.23	-0.08	-0.13	-0.15	1.00		
HADS Depression	-0.01	-0.14	0.02	-0.08	0.01	0.56	1.00	
Disease Duration	0.04	0.26	0.26	0.01	0.29	-0.35	0.19	1.00
AMI incidents	0.17	-0.39	-0.07	0.34	0.07	-0.29	-0.28	-0.24
AMI semantic	0.12	0.02	-0.09	0.05	0.23	-0.56	-0.49	0.33
False belief	0.17	-0.06	0.00	-0.33	0.17	0.01	0.19	-0.19
Risk score	0.02	-0.41	0.02	0.06	0.03	0.15	0.27	-0.18
BVMT total recall	0.13	0.37	0.28	-0.02	-0.12	0.00	0.18	0.01
BVMT delayed recall	0.12	0.66	0.37	-0.10	-0.26	-0.08	-0.08	0.02

Note *p<.05

** n = 17 for this measure (1 participants data was missing for an unknown reason and 2 participants' data could not be collected)

Abbreviations: ADAS-Cog = Alzheimer's Dementia Assessment Scale-Cognitive; ADL-IS = Activities of Daily Living-International Scale; AMI = Autobiographical Memory Interview; BVMT = Brief Visuospatial Memory Test; DRS-2 (AESS) = Dementia Rating Scale (age and education scales score); H & Y = Hoehn and Yahr stage; HADS = Hospital Anxiety and Depression Scale; UPDRS = Unified Parkinson's Disease Rating Scale

3.9 Correlations between test scores across both PD and HC group

Table 9 shows the correlations between different overall test scores using the scores of both PD and HC (n=35). Five significant correlations were found between test scores. A moderate positive relationship was found between false belief (ToM) and BVMT total recall, r = 0.46, and a moderate positive relationship between false belief (ToM) and BVMT delayed recall, r = 0.41. Generally, as BVMT total and delayed recall increase, performance on the ToM task also increases. However, when BVMT total recall and delayed recall were correlated with AMI incidents, this significance dropped out. As expected there was a strong positive relationship between BVMT delayed recall and BVMT total recall, r = 0.87. There was a moderate positive relationship between the SDMT and BVMT total recall (r = 0.51) and BVMT delayed recall (r = 0.65). Generally as SDMT performance increases, so too does BVMT total and delayed recall. Significant relationships were not found between autobiographical memory scores and semantic memory scores (r = .20) or between false belief scores (ToM) and autobiographical incident recall (r = 0.18).

Table 9.

Correlations between tests including scores of Parkinson's disease participants and healthy controls

	1.	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**
1. False belief	1.00									
2. AMI incidents	0.18	1.00								
3. AMI semantic	-0.20	0.20	1.00							
4. BVMT total recall	0.46*	0.08	0.10	1.00						
5. BVMT delayed recall	0.41*	0.08	0.10	0.87*	1.00					
6. SDMT	0.21	0.20	0.30	0.51*	0.65*	1.00				
7. SYDBAT Naming	0.03	0.02	0.07	0.14	0.01	0.04	1.00			
8. SYDBAT Repetition								1.00		
9. SYDBAT Word Comp	-0.15	-0.04	-0.20	-0.20	-0.23	-0.21	0.18		1.00	
10. SYDBAT Semantic Association	0.00	0.24	0.06	0.04	0.00	-0.01	0.30		0.18	1.00

Note *p<.05
**no correlation because max SYDBAT repetition scores

Abbreviations: AMI = Autobiographical Memory Interview; BVMT = Brief Visuospatial Memory Test; SYDBAT = The Sydney Language Battery

Discussion

4.1 Main findings

The current study was the first to examine both personal episodic and personal semantic lifetime memories in PD. Moreover all but two participants were identified as pre-MCI PD, a novel group not yet defined in the PD literature that is at risk of cognitive decline over the next four years. There was a mild but selective deficit in autobiographical (episodic) memories in this group of patients whereas personal semantic memory was unimpaired. There was a suggestion that the poorer autobiographical recall in the PD group was greatest for the childhood period but they were generally impaired across all three lifetime periods examined. Secondly, a cognitive ToM task also revealed a deficit in this group of PD participants, shown by poorer performance in the false belief card sequencing task. However, the PD and control groups performed equally on the other sequencing (non-ToM) tasks, suggesting a specific ToM deficit rather than a more general cognitive problem.

Contrary to the hypothesis, there was no association between autobiographical memory and the ToM in this group of PD participants. There was also no association between the cognitive decline risk score, which was based on age of an individual plus the four cognitive screening tests, and either autobiographical memory or ToM. However, these latter negative findings may be because this PD group has relatively normal cognition (pre-MCI) and an association may exist in a more cognitively impaired group of PD participants. Conversely, the BVMT delayed recall did show a moderate correlation with the cognitive decline risk score, suggesting that the BVMT may be a useful measure to test memory deficits in individuals with early stage PD and thus could be added to future models of risk of cognitive decline. The quality of the results found in the current study is superior to previous research as the cognitive status of the PD participant group was thoroughly characterised. The

relevant prior research often used the MMSE or a few neuropsychological tests and did not fully characterise their patients other than specifying them as PDD or PD without dementia. The novelty of the current research findings contributes to a growing body of literature that is helping researchers and practitioners to understand the nature of early PD.

4.2 Comparison of main findings with prior research

The PD group was significantly impaired at recalling autobiographical memories across the three lifetime periods compared to controls. This was consistent with previous research of autobiographical episodic memory in PD (Sagar et al., 1988; Smith et al., 2010; Souchay & Smith, 2013). However the pattern of temporal gradient of the episodic memory recall was contrary to what the prior PD literature on episodic autobiographical memory might suggest. PD participants on average recalled less autobiographical memories (incidents) for the childhood period than for the young adult or recent period. The three articles that investigated autobiographical memory in PD (discussed in the introduction) all found recent autobiographical events were recalled more poorly than remote autobiographical events (Sagar et al., 1988; Smith et al., 2010; Souchay & Smith, 2013). This pattern of episodic memory recall is also evident in AD, including mild AD (Kopelman et al., 1989; Levy, 1988). The pattern of episodic memory recall of the PD group in the current study is more similar to the controls in previous literature. For example, Kopelman et al. (1989) found that control participants recalled more detailed recent personal episodic memories on the AMI, compared to less detailed childhood episodic memories. However, the controls in the current study actually achieved better recall for their childhood memories compared to the other time periods. The raw data for autobiographical incident recall (Figure 10), reveals that a small group of 5 PD participants skewed the average with poorer recall for the childhood period than for the other time periods. Some participants commented that recalling details from their childhood was especially difficult.

The amygdala, hippocampus and right inferior frontal gyrus show functional connectivity during episodic memory recall (Greenberg et al., 2005). These brain regions often show atrophy of functional loss in PD. The non-MCI PD participants in the current study showed only a mild deficit in episodic memory recall, so the connectivity across these regions may be disrupted to a lesser degree than in later stage PD participants, such as PD-MCI, who show more significant episodic memory impairments. MRI images of the PD participants in this study were taken as part of the broader study at the NZBRI, so future research could analyse the connectivity in these areas and correlate it with autobiographical memory and ToM performance in these patients. In addition, PD-MCI participants could be tested, and included in analyses to determine comparisons across different levels of cognitive status.

Unlike personal episodic memories, both PD and HC groups performed equally well at personal semantic memory recall, with high scores across all time periods. Recall was generally more detailed for recent memories and least detailed for childhood memories in both groups. Examination of the raw scores for personal semantic memory (Figure 12) shows that both PD and HC participants are clustered near the max score of 21 for all three time periods, especially the recent period. The scores for the childhood period were the most varied. Only a couple of scores fell below the cluster in the childhood period, with one control scoring a low score of 11 and two PD scoring 12.5 and 13.5, and in the young adult period only one PD participant scored below the cluster with a score of 13. There was no correlation between the recall of personal episodic memory and personal semantic memory scores, providing additional evidence of the independence of these two measures of autobiographical memory recall. The similar performance of both PD and controls at personal semantic memory recall is similar to the results of Smith, Souchay & Conway (2010), who also found no group difference at this measure. However, Smith, Souchay and Conway

(2010) reported the opposite temporal gradient, with fewer names recalled in the recent time periods compared to early time periods for both PD and HC (Smith, Souchay & Conway, 2010). Their results, however, came from a different autobiographical memory test, the 'autobiographical fluency task', which could account for differences in the pattern of findings.

The same pattern of results was found for the different card sequencing tasks in the current study as was reported by Mengelberg and Siegert (2003). That is, PD participants scored significantly lower on the false belief category, with mean values almost identical to this previous report, but there were no significant difference between groups on the three control categories, suggesting a specific ToM deficit. The only difference between the results of the current study and that of Mengelberg and Siegert (2003) was performance on the capture (complex reasoning) card category. This was the most difficult of the control card tasks. On the capture task, Mengelberg and Seigert's controls scored on average about 12.5 out of 24, and their PD group scored about 11.5 out of 24. The controls in the current study scored on average 16.07, and the PD scored on average 14.35. This suggests that the PD participants in the current study had better complex reasoning, yet were still impaired to the same level at ToM.

One key hypothesis was that there would be a correlation between the episodic measure of the AMI and ToM because these tasks are both associated with blood oxygen level-dependent (BOLD) functional activity in DMN subsystems. The current study found, however, no significant relationship between personal episodic autobiographical memory and ToM. While there is a modest cognitive ToM impairment, there was a less pronounced deficit in episodic autobiographical memory. There was reasonable variance in both episodic memory and ToM scores, so the lack of correlation cannot be attributed to the cognitively

unimpaired status of the PD group. These results suggest that there may be dissociation between the regions that underlie these two tasks. The default mode network is frequently activated by both autobiographical memory and ToM tasks (Andrews-Hanna, 2010), but each task is correlated to a different subsystem within the DMN. Dissociation of these subsystems could explain this lack of correlation. The dorso medial prefrontal cortex (dMPFC) is associated with ToM functioning, whereas the medial temporal lobe subsystem is thought to underlie episodic/autobiographical memory. These two regions may degenerate or become dysfunctional at different times during the PD disease process. It is possible that the dMPFC is associated more specifically to affective ToM than to cognitive ToM. This could explain the lack of correlation between autobiographical memory and cognitive ToM. Previous research has suggested that affective ToM is only impaired in PD patients during the later stages of the disease (Poletti et al., 2011), and the current study shows that autobiographical memory may show more significant impairment in PD participants who are more cognitively impaired. It would of interest in future research to see if there is a correlation between autobiographical memory and affective ToM, both of which may be more significantly impaired in later stage PD.

It was hypothesised that there might be a relationship between the cognitive decline risk score of the PD participants and their AMI and ToM performance, but this was not found. Higher risk scores did not predict poorer performance on the AMI or ToM. There was a good range of scores for autobiographical recall and ToM performance across participants so lack of variation in scores does not explain the failure to observe this association. Some PD participants who performed poorly in the screening assessment (which formed the risk score) performed well on the ToM task, and some participants who performed poorly in the ToM task performed well in the screening assessment. One possibility to explain this lack of correlation could be that the risk score is predicting a decline in areas that are not associated

with autobiographical memory or ToM. If the subsystems that underlie autobiographical memory and ToM are distinct enough in their function from areas that underlie the screening tests, then these areas may degenerate at a different rate. Two of the cognitive screening tests that are part of the risk score measure executive function (Stroop-interference, Trails B), another measures attention, working memory and processing (TEA Map Search) and the MoCA provides a brief global function measure. As per Mengelberg and Siegert (2003), ToM performance was unrelated to executive function deficits. The ToM functions seem independent to those other cognitive domains. Given the discussion above about the dissociation between cognitive and affective ToM, future research should investigate the correlation between affective ToM and the risk score, as this measure could contribute to indicative risk of future decline.

Results of the initial screening tests showed that although there was a significant difference between the two groups on all key measures (MoCA, Stroop-interference, Trails B, and the TEA Map Search; also Trails A), performance of PD and HC groups on Stroop-colour and stroop-word was similar. Neither motor symptoms, based on the UPDRS, nor neuropsychiatric symptoms, based on the HADS, had any impact on the scores of the PD participants on the key measures in the current study (episodic AMI, ToM card sequencing task and BVMT). However anxiety and depression did show a significant correlation with AMI semantic recall scores, but not with episodic memory recall, which was the important measure in this study. There was also no significant difference on the four language measures across PD and HC, and both groups which scored highly on those measures, suggesting that basic language functions are not impaired in early PD.

An unexpected finding was that the BVMT delayed recall score showed a greater effect size between performance of the PD and HC group than any of the other measures used in the initial screening test. This was surprising because prior work at the NZBRI using other

memory tests (Rey-Osterrieth Complex Figure Copy and CVLT recall) found these measures relatively intact in non-MCI patients with PD. There was also a significant relationship between the cognitive decline risk score and poor BVMT delayed recall score in the patient group in the current study. Together these findings suggest that the BVMT delayed recall may be a good predictor of early cognitive change. This measure may be a more useful addition to screening measures used in this particular participant group of PD patients (pre-MCI). A moderate correlation between the BVMT delayed recall score and the risk score, rather than a high correlation, is a useful finding as it shows that the BVMT could add more value to the risk score; a high correlation would show that the risk score and the BVMT delayed recall essentially overlap and the BVMT does not add new information. The BVMT total and delayed recall was correlated with the ToM task, but not with the AMI (episodic) incidents. This dissociation between the two episodic memory tasks, one for visuospatial recent memory and one for past personal incidents, suggests that these two measures have a different neural basis.

A strength of this research is the better characterisation of the participant group compared to the 3 other studies that have investigated autobiographical memory in PD. Two of the previous studies (Smith et al., 2010; Souchay and Smith 2013) used the MMSE as their main cognitive measure which has been shown to be less sensitive than even the MoCA for PD (Dalrymple-Alford et al., 2010). The third study used only the Blessed Dementia Scale, another general cognitive measure (Sagar et al., 1988). A broader range of tests is needed to assess the cognition of PD patients (Litvan et al., 2012), a prerequisite which this study satisfies. It is difficult to compare results between studies that have poorly defined participant groups. Participants at different stages of PD can have different cognitive performance on tasks, so without well-defined participant groups the results are less meaningful. Research is increasingly focussing on determining cognitive change and following progression, thus

careful attention needs to be paid to participant characterisation at different stages in the disease course to build a reliable body of literature.

4.3 Limitations

Although the sample size in the current study is larger than most of the studies discussed in the introduction, it is still relatively small. Small sample sizes place limits on the power of the study and potential accuracy of the ensuing results. It is possible that the correlations between measures found in this study may not be a true reflection of the relationships due to this small sample size. A larger sample size would have also allowed for more detailed analyses of poorer performers within each task. In addition, having only 2 participants from the under 29% risk score category made it impossible to compare fully their results with that provided by patients in the over 29% risk score group. A larger sample of the under 29% participants would have been valuable to compare between PD participants with lower and higher risk scores, but this was not possible in the time available.

Results of this study would have been strengthened with a comparison to a PD-MCI group. This would have given a clearer understanding of how autobiographical memory and ToM are affected in more significant cognitive impairment. There was also lack of any imaging in the current study which meant correlations between task performance and associated brain regions could not be analysed. Brain imaging would have been useful especially to analyse the dissociation between areas associated with autobiographical memory and ToM, and to see if imaging was able to differentiate between lower and higher cognitive decline risk scores.

Although the current study did examine the relationship between motor severity and test scores, there was no analysis of the motor sub-types of PD, such as non-tremor dominant (rigid-akinetic) subtype and tremor subtype. Patients with the tremor dominant motor sub-

type tend to have shorter disease duration, less severe motor symptoms and less cognitive impairment (Burn et al., 2012). Future research should differentiate between these two subtypes to investigate a difference in AMI and ToM test performance and use imaging to investigate structure and functioning in the brain.

4.4 Future Directions

Future research should test the correlation between the cognitive decline risk score and the AMI and ToM with PD-MCI participants to see if a relationship between these measures eventuates in a more severely impaired participant group. The relationship between these measures should be tested with a larger participant group and should include not only PD-MCI but lower risk pre-MCI participants (under 29%) as well as higher risk pre-MCI participants (over 29%), together with PDD participants. This would provide information for a range of participants across disease stages, allowing for a comparison between these groups and analysis of significant changes at each stage of PD. As mentioned earlier, MRI images were taken for the participants in this study as part of a broader study at the NZBRI. These images should be analysed in future research to see if there is any correlation between performance on autobiographical memory and ToM with degeneration or poor function in associated brain areas. For example, analyses of MRI images may show a relationship between the risk score and cortical thinning, as previous research has shown that cortical thinning is evident prior to overt cognitive changes (Mak et al., 2015).

The current study also suggests that the BVMT should be considered as a screening measure for participants in future research and added to predictive progression models like the risk score used in this study. The BVMT is sensitive to mild impairment that other tests may not recognise, so it may be an especially useful measure for detecting early cognitive change.

4.5 Conclusions

The aim of the current study was to examine autobiographical memory and ToM in a well characterised PD patient group in whom cognition was relatively unimpaired (non-MCI) but who were showing risk of future decline. This study used the AMI for the first time with a PD participant group and found significant impairments in personal episodic memory compared to HC participants. There was also a significant impairment of PD participants on the ToM task compared to controls. Although it was hypothesised that there would be a correlation between autobiographical memory and ToM due to an overlapping of associated brain areas, this was not found. There was also no correlation between these measures and the cognitive decline risk score, suggesting the AMI and card sequencing task are not useful as predictors of future cognitive decline. However, the BVMT delayed was associated with the risk score and surprisingly had a larger effect size between groups than any other measure in the screening test. The BVMT may be useful to add to predictive models of future cognitive decline and may detect cognitive change in the early stages of cognitive decline in PD. The results of this study contribute to a growing body of literature on the non-MCI PD patients and may be of clinical relevance to detect early cognitive change.

References

Aarsland, D., & Kurz, M. W. (2010a). The epidemiology of dementia asscociated with Parkinson disease. *Journal of Neurological Sciences, 289*, 18-22.

Aarsland D., Bronnick, K., Williams-Gray, C., Weintraub D., Marder, K., Kulisevsky, J. … Emre, M. (2010b). Mild cognitive impairment in Parkinson disease: A multicentre pooled analysis. *Neurology 75*, 1062-1069.

Andrews-Hanna, J. R. (2012). The Brain's Default Network and its Adaptive Role in Internal Mentation. *The Neuroscientist, 18*(3), 251-270.

Barnabe, A., Whitehead, V., Pilon, R., Arsenault-Lapierre, G., Chertkow, H. (2012). Autobiographical memory in mild cognitive impairment and Alzheimer's disease: A comparison between the Levine and Kopelman interview methodologies. *Hippocampus, 22*(9), 1809-1825.

Benedict R. H. B. (1997). Brief visuospatial memory test - revised: Professional manual. Lutz, FL: *Psychological Assessment Resources.*

Benedict, R. H. B. (1988). *Brief Visuospatial Memory Test-Revised*. Lutz, Florida, USA:PAR.

Benton, A. L., Hasher, K., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment: A clinical manual.* New York: Oxford.

Bodden, M., E., Dodel, R., & Kalbe, E. (2009). Theory of Mind in Parkinson's disease and related basal ganglia disorders: a systematic review. *Movement Disorders, 25*(1), 13-27.

Bruck, A., Kurki, T., Vahlberg, J., &Rinne, J. O. (2004) Hippocampal and prefrontal atrophy in ptietns with early non-demented Parkinson's disease is related to cognitive impairment. *J Neurol Neurosurg Psychiatry, 75*, 1467-1469.

Burn, D., J., Landau, S., Hindle, J. V., Samuel, M. Wilson, K. C., Hurt, C. S., & Brown, R. G. (2012). Parkinson's Disease motor subtypes and mood. Movement Disorders, 27(3), 379-386.

Colcombe, S., & Kramer, A. F. (2003). Fitness Effects on the Cognitive Function of Older Adults: A Meta-Analytic Study. *Psychological Science, 14*(2), 125-130.

Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society, 4*, 517-518.

Dalrymple-Alford, J. C., MacAskill, M. R., Nakas, C. T., Livingston, L., Graham, C., Crucian, G. P. … & Anderson, T.J. (2010). The MoCA: Well-suited screen for cognitive impairment in Parkinson's disease. *Neurology*, *75*, 1717-1725.

Dalrymple-Alford, J. C., Livingston, L., MacAskill, M., Graham, C., Melzer, T. R., Porter, R., Watts, R., Anderson, T. J. (2011). Characterising mild cognitive impairment in Parkinson's' disease. *Movement Disorders, 26*(4), 629-636.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test* (2nd ed.). San Antonio, Texas: Pschological Corporation.

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis-Kaplan Executive Function Scale (D-KEFS).* San Antonio, Texas: The Psychological Corporation.

Dritschel, B. H., Williams, J. M. G., Baddeley, A. D., & Nimmi-Smith, I. (1992). Autobiographical fluency: A method for the study of personal memory. *Memory & Cognition, 20*(2), 133-140.

Eimeren, T., Monchi, O., Ballanger, B., & Strafella, A. P. (2009). Dysfunction of the Default Mode network in Parkinson disease: A functional magnetic resonance imaging study. *Arch Neurol, 66*(7), 877-883.

Euteneur, F., Schaefer, F., Stuermer, R., Boucsein, W., Timmermann, L., Barbe, M. T. ... & Kalbe, E. (2009). Dissociation of decision making under ambiguity and decision making under risk in patients with Parkinson's disease: a neuropsychological and psychophysiological study. *Neuropsychologia, 47*, 2882–90.

Gates, N., & Valenzuela, M. (2010). Cognitive Exercise and Its Role in Cognitive Function in Older Adults. *Curr Psychiatry Rep, 12*, 20-27.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Belleville, S., Brodaty, H., Bennett, D., Chertkow, H., Cummings, J. L., de leon, M., Feldman, H., Ganguli, M., Hampel, H., Scheltens, P., Tierney, M. C., Whitehouse, P., & Winblad, B. (2006). Mild cognitive impairment. *The Lancet, 367*, 1262-70.

Gilbert, B., Belleville, S., Bherer, L., Chouinard, S. (2005). Study of verbal working memory in patients with Parkinson's disease. *Neuropsyhcology, 19*, 106-14.

Graham, K. S., & Hodges, J. R. (1997). Differentiating the roles of the hippocampus complex and the neocortex in long-term memory storage: Evidence from the study of semantic dementia and Alzheimer's disease. *Neuropsychology, 11*(1), 77-89.

Greenberg, D. L., Rice, H. J., Cooper, J. J., Cabeza, R., Rubin, D. C., & LaBar, K. S. (2005). Co-activation of the amygdala, hippocampus and inferior frontal gyrus during autobiographical memory retrieval. Neuropsychologia, 659-674.

Greene, J. D. W., & Hodges, J. R. (1995). Autobiographical memory and executive function in early dementia of Alzheimer type. *Neuropsychologia, 33*(12), 1647-1670.

Hawkes, C. H., Tredici, K. D., & Braak, H. (2010) A timeline for Parkinson's disease. *Parkinsonism and Related Disorders, 16,* 79-84.

Hindle, J. V., Petrelli, A., Clare, L., & Kalbe, E. (2013). Nonpharmacological Enhancement of Cognitive Function in Parkinson's Disease: A systematic Review. *Movement Disorders*, 28(8), 1034-1049.

Hoops, S., Nazem, S., Siderowf, A. D., Duda, A. D. Xie, J. E., Sern, M. B., & Weintraub, D. (2009). Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology, 73,* 1738-1745.

Hou, C. E., Miller, B. L., & Kramer, J.H. (2005). Patterns of autobiographical memory loss in dementia. *International Journal of Geriatric Psychiatry, 20*(9), 809-815.

Janvin, C. C., Larsen, J. P., Aarsland, D. & Shugdahl, K. (2006). Subtypes of mild cognitive impairment in Parkinson's disease: progression to dementia. *Mov. Disord. 21,* 1343–1349.

Jenkins, I. H., Fernandez, W., Playford, E. D., Lees, A. J., Frackowiak, R. S. J., Passingham, R. E., & Brooks, D. J. (1992). Impaired activation of the supplementary motor area in Parkinson's disease is reversed when akinesia is treated with apomorphine. Annals of neurology, 32(6), 749-757.

Jurica, S. J., Leitten, C. L., & Mattis, S. (2001). *Dementia Rating Scale: Professional manual*. Odessa, Fl: Psychological Assessment Resources.

Kaplan, E., Goodglass, H., Weintraub, S. (1983). *Boston Naming Test.* Philadelphia: Lean & Febiger.

Kawamura, M. & Koyama, S. (2007). Social cognitive impairment in Parkinson's disease. *J Neurol, 254*, 49-53.

Kiely, K. M., Butterworth, P., Watson, N., & Wooden, M. (2014). The Symbol Digit Modalities test: Normative data from a large nationally representative sample of Australians. *Archives of Clinical Neuropsychology, 29*, 767-775.

Kopelman, M. D., Wilson, B.A. & Baddeley, A. D. (1989). The autobiographical memory interview: A new assessment of autobiographical and personal semantic memory in amnesic patients. *Journal of Clinical and Experimental Neuropsychology, 11*(5), 724-744.

Langdon, R., Michie, P. T., Ward, P. B., McConaghy, N., & Catts, S. V. (1997). Defective self and/or other mentalising in Schizophrenia: a cognitive neuropsychological approach. *Cognitive neuropsychiatry, 2*(3), 167-193.

Langdon, R., & Coltheart, M. (1999). Mentalising, schizotypy, and schizophrenia. *Cognition, 71*, 43-71.

Levy, J. P., (1998). Autobiographical memory in mild and moderate dementia of the Alzheimer's type. (Unpublished Doctoral thesis). Louisiana State University, United States.

Litvan, I., Aarsland, D., Adler, C. H., Goldman, J. G., Kulisevsky, J., Mollenhauer, B., Rodriguez-Oroz, M. C., Troster, A I., & Weintraub, D. (2011). *Movement Disorders, 26*(10), 1814-1824.

Mak, E., Su, L., Williams, G. B., Firbank, M. J., Lawsom, R. A., Lawson, R. A. … O'Brien, J. T. (2015) Baseline and longitudinal grey matter changes in newly diagnosed Parkinson's disease: ICICLE-PD study. *Brain, 138*(10), 2974-2986.

Mars, R. B., Neubert, F., Noonan, M. P., Sallet, J., Toni, I., &Rushworth, M. F. S.(2012). On the relationship between the "default mode network" and the 'social brain'. *Frontiers in Human Neuroscience, 6*(189).

Mengelberg, A., & Siegert, R. (2003). Is theory-of-mind impaired in Parkinson's disease? *Cognitive Neuropsychiatry, 8(*3), 191-209.

Meyers, J. E. & Meyers, K, R. (1995). Rey complex figure test under four different administration procedures. *The Clinical Neuropsychologist, 9*(1), 63-67.

Mimura, M., Oeda, R., & Kawamura, M. (2006). Impaired decision-making in Parkinson's disease. *Parkinsonism Related Disorders, 12,* 169–75.

Mohs, R. C., Knopman, D., Peterson, R. C., Ferris, S. H., Ernesto, C., Grundman, M., … Thal, L. (1997). Development of cognitive instruments for use in clinical trials of anti-dementia drugs: additions to the Alzheimer's disease assessment scale that broaden its scope. *Alzheimer Disease and Associated Disorders, 11*(2), S13-S21.

Monetta. L., Grindrod, C. M., & Pell, M. D. (2009). Irony comprehension and Theory of Mind deficits in patients with Parkinson's disease. *Cortex, 45*, 972–81.

Moreau, N., Viallet, F., & Champagne-Lavau, M. (2013). Using memories to understand others: the role of episodic memory in theory of mind in Alzheimer disease. *Ageing Research Reviews, 12*, 833-839.

Myall, D. J., Wood, K. L., Livingston, L., Pitcher, T. L., MacAskill, M. R., Anderson, T. J., & Dalrymple-Alford, J. C. (2015). Individualised medicine using an abbreviated cognitive screen for dementia prediction in Parkinson's disease. *Movement Disorders, 30*, 346.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment. *Journal if the American Geriatrics Society, 53*, 695-699.

Olsson, A. & Ochsner, K. N. (2007). The role of social cognition in emotion. *Trends in Cognitive Neuroscience, 12*(2), 65-71.

Pedersen, K. F., Larsen, J. P., Tysnes, O. & Alves, G. (2013). Prognosis of mild cognitive impairment in early Parkinson's disease. *JAMA Neurol. 70*, 580–586.

Peron, J., Vicente, S., Leray, E., Drapier, S., Drapier, D., Cohen, R. … & Verin, M. (2009). Are dopamin- ergic pathways involved in Theory of Mind? A study in Parkinson's disease. Neuropsychologia, 47, 406–14.

Peron, J., Le Jeune, F., Haegelen, C., Dondaine, T., Drapier, D., Sauleau, P. … & Verin, M. (2010). Sub- thalamic nucleus stimulation affects Theory of Mind network: a PET study in Parkinson's disease. *PLoS ONE, 5*(3), e9919.

Petersen, R. C., (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256*, 183-194.

Piatt, A. L., Fields, J. A., Paolo, A. M., & Troster, A. I. (2004). Action verbal fluency normative data for the elderly. *Brain Lang, 89*(3), 580-583.

Piolino, P., Belliard, S., Desgranges, B., Perron, M., & Eustache, F. (2003). Autobiographical memory and autonoetic consciousness in the case of semantic dementia. *Cognitive Neuropsychology, 20*(7), 619-639.

Poletti, M., Enrici, I., Bonuccelli, U., & Adenzato, M. (2011). Theory of Mind in Parkinson's disease. *Behavioural Brain Research, 219*, 342-350.

Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective theory of mind in neurodegenerative diseases: Neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience and Biobehavioural Reviews, 36*, 2147-2164.

Poewe, W. (2008). Non-motor symptoms in Parkinson's disease. *European Journal of Neurology, 15*(1), 14-20.

Roca, M., Torralva, T., Gleichgerrercht, E., Chade, A., Arevalo, G. G. , Gershanik, O. … & Facundo, M. (2010). Impairments in social cognition in early medicated and unmedicated Parkinson's disease. *Cogn Behav Neurol, 23*, 152–8.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). The Test of Everyday Atetntion (TEA). *The Application of Occupational Psychology to Employment and Disability, 4*(1), 51-55.

Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of Mind and executive functions in normal human aging and Parkinson's disease. *J Int Neuropsychol Soc, 6,* 781–8.

Sagar, H. J., Cohen, N. J., Sullivan, E. V., Corkin, S., Growdon, J. H. (1988). Remote memory function in Alzheimer's-disease and Parkinson's-disease. *Brain, 111*, 185-206.

Savage, S., Hsieh, S., Leslie, F., Foxe, D., Piguet, O., & Hodges, J.R. (2013). Distinguishing subtypes in Primary Progressive Aphasia: Application of the Sydney Language Battery. *Dementia and Geriatric Cognitive Disorders, 35*, 208-218.

Smith, A. (1982). Symbol Digit Modalities Test. Los Angeles: *Western Psychological Services*.

Smith, S. J., Souchay, C., & Conway, M. A. (2010). Overgeneral autobiographical memory in Parkinson's disease. *Cortex, 46*, 787-793.

Souchay, C., & Smith, S. J. (2013). Autobiographical memory in Parkinson's disease: A retrieval deficit. *Journal of Neuropsychology, 7*, 164-178.

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2008). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience, 21*(3), 489-510.

Storandt, M., Grant, E. A., Miller, J. P., & Morris, J. C. (2006). Longitudinal course and neuropathologic outcomes in original vs revised MCI and pre-MCI. *Neurology, 67*, 467-473.

Squire, L. R. (1992). Declarative and non-declarative memory: multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4*(3), 232-243.

Tessitore, A., Esposito, F., Vitale, C., Santangelo, G., Amboni, M., Russo, A., Corbo, D., Cirillo, G., Barone, P., & Tedeschi, G. (2012). Default-mode network connectivity in cognitively unimpaired patients with Parkinson disease. *American Academy of Neurology, 79*, 2226-2232.

Thomas-Anterion, C., Jacquin, K., & Laurent, B. (2000). Differential mechanisms of impairment of remote memory in Alzheimers's and frontotemporal dementia. *Dementia and Geriatric Cognitive Disorders, 11*(2), 100-106.

Tramoni, E., Felician, O., Koric, L., Balzamo, M., Joubert, S., & Ceccaldi, M. (2012). Alteration of autobiographical memory in amnestic mild cognitive impairment. *Cortex, 48,* 1310-1319.

Uc, E. Y., Doerschug, K. C., Magnotta, V., Dawson, J. D., Thomsen, T. R., Kline, J. N. … Darling, W. G. (2014). Phase I/II randomised trial of aerobic exercise in Parkinson disease in a community setting. *American Academy of Neurology, 83*, 413-425.

Warrington, E.K., & James, M. (1991). *The Visual Object and Space Perception Battery*. Bury St. Edmunds, UK: Thames Valley Test Company.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale-Fourth Edition* (4th ed.). San Antonio, Texas: Pearson.

Yesavage J. A., Brink T. L., Rose T. L., Lum O., Huang V., Adey M. B., Leirer V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research 17,* 37-49.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale-Fourth Edition (4th ed.). San Antonio, Texas: Pearson.

Wood, K., Myall, D.J., Livingston, L., Melzer, T.R., Pitcher, T.L., MacAskill, M.R., Geurtsen, G.J., Anderson, T.J., Dalrymple-Alford, J.C. (2016). Different PD-MCI criteria and risk of dementia in Parkinson's disease: Four year longitudinal study. *npj Parkinson's Disease*, *2*, 15027

Yesavage, J. A., & Brink, T. L. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17*(1), 37-49.

Zigmond, A. S. & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*,*67*(6), 361–370.

Appendices

Appendix A.

Story	Layout order	Time taken	Subject order (colour id → no. id)	Correct order	Score	Story
Prac 1						Prac 1
Prac 2						Prac 2
3	OGYB			BYGO		3
9	YGBO			BYGO		9
17	OBYG			GYOB		17
12	GYOB			GYBO		12
18	YBOG			BGOY		18
10	YOBG			GYBO		10
5	YGBO			YOBG		5
14	OYBG			YBOG		14
8	YBOG			GBYO		8
15	BGYO			YOGB		15
7	GOYB			OBYG		7
11	BGYO			YBGO		11
4	BGOY			OBYG		4
13	GBOY			OBYG or OBGY		13
16	GYOB			BOYG		16
6	GYBO			OYGB		6

Appendix B.

Neuropsychological Domain	**Neuropsychological Test**	**PD mean (SD) (n = 20)**	**Test of means against an expected population mean**
Attention, Working Memory and Processing	Digits Forward and Backward	0.48 (0.99)	p = 0.04
	Digit ordering	-0.60 (1.04)	p = 0.02
	TEA (Map Search)	-0.80 (0.62)	p = 0.00
	Stroop colour	-0.20 (0.57)	p = 0.13
	Stroop word	0.05 (0.64)	p = 0.73
	Trails A	0.38 (0.72)	p = 0.03
Executive Function	Letter Fluency	1.02 (1.09)	p = 0.00
	Action Fluency	-0.74 (0.90)	p = 0.00
	Category Fluency	0.75 (0.61)	p = 0.00
	Category Switching	-0.07 (0.93)	p = 0.75
	Trails B	0.50 (0.85)	p = 0.02
	Stroop interference	0.03 (0.99)	p = 0.88
Visuoperceptual/Visuospatial	Judgement of Line Orientation	-0.15 (0.84)	p = 0.43
	VOSP	0.33 (0.71)	p = 0.05
	Rey Copy	-0.62 (0.94)	p = 0.01
	Picture Completion*	0.65 (0.59)	p = 0.00
Learning and Memory	CVLT Free recall	0.20 (0.95)	p = 0.37
	CVLT Short delay	0.30 (1.24)	p = 0.29
	CVLT Long delay	0.10 (1.08)	p = 0.68
	Rey Immediate*	0.02 (1.31)	p = 0.94
	Rey Delay*	0.17 (1.32)	p = 0.57
Language	Boston Naming	0.47 (0.65)	p = 0.00
	Language DRS-2 similarities	-0.19 (0.77)	p = 0.30
	Language ADAS-Cog	0.05 (0.52)	p = 0.69

*n=19 for these tests as one participants' data was missing for an unknown reason

Appendix C.

Screening test results and results of key measures for the two under 29% risk score participants.

%	MoCA	Stroop-interference	Trails A	Trails B	TEA Map Search	BVMT total recall	BVMT delayed recall	SDMT	SYDBAT Naming	SYDBAT Repetition	SYDBAT Word	SYDBAT Semantic	ToM	AMI episodic	AMI semantic
15%	27	1.33	1.21	1.58	0.33	19	9	1.32	27	30	29	25	9	17	58.5
24%	30	1.33	1.38	1.61	0.00	21	10	1.66	24	30	29	28	11	16	59.5