The Clinical Utility of the Four Mountains Test in the diagnosis of

dementia: relationship to hippocampal and medial temporal

lobe atrophy.

Maura Scanlon

D.Clin.Psy. Thesis (Volume 1), 2017
University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has

been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Name: Maura Scanlon

Date: 02/12/17

OVERVIEW

This thesis explores aspects of dementia diagnosis by investigating the quality of instruments used in detecting anxiety in the dementia population and the clinical utility of a spatial memory test in early diagnosis of dementia.

Part 1 consists of a systematic literature review of the instruments which purport to measure self-reported anxiety in individuals with dementia. A total of ten studies reported on the methodological quality of nine instruments. The Rating Anxiety in Dementia (RAID) scale demonstrated the strongest weight of evidence in terms of the quality of measurement properties. This review highlighted a lack of high quality, high powered studies in this area and demonstrated the need to increase involvement of individuals with dementia in the validation of anxiety instruments.

Part 2 is an empirical paper examining the clinical utility of The Four Mountains Test (4MT) in the early diagnosis and differentiation of Alzheimer's (AD). In a secondary analysis of data, 15 structural MRI scans were analysed and compared with 4MT performance and other neuropsychological measures that are typically assessed as part of a diagnostic assessment for dementia. Contrary to prediction, there were no positive associations between the volume and thickness measurements of Regions of Interest (ROI's) and 4MT scores across the sample. Clinical and research implications and limitations are discussed.

Part 3 is a critical appraisal which reflects on the technical and methodological challenges of undertaking an analysis of MRI scans. It also reflects on the gaps between research and clinical practice in addition to conceptual and ethical considerations when working with individuals living with dementia.

TABLE OF CONTENTS

6

List of Tables and Figures

Literature Review

Empirical paper

ACKNOWLEDGEMENTS

PART 1: LITERATURE REVIEW

Title: A systematic review of self-report instruments that purport to measure anxiety in individuals with dementia

ABSTRACT

Aims: There is a high prevalence of anxiety disorders among individuals with dementia which affect quality of life, degree of care burden, and dementia progression. Screening for anxiety is important for long-term management and treatment. The objectives of the current review were to 1) Establish the quality of instruments that purport to measure self-reported anxiety in individuals with dementia, considering the methodological quality of studies that report on measurement properties. 2)Establish whether self-report or informant instruments provide the most reliable measurements. **Methods:** A range of databases were searched and articles were selected if their primary purpose was the development or assessment of measurement properties of a self-reported anxiety scale with a dementia sample. Methodological quality was assessed using the COnsesnus based Standard for the selection of health status Measurement INstruments checklist (COSMIN). **Results:** A total of ten papers reviewed nine instruments. The Rating Anxiety in Dementia scale (RAID) had the strongest weight of evidence for use in dementia. This was followed by evidence for The Brief Anxiety and Depression Scale (BADS). Self-report anxiety instruments demonstrated a degree of validity in mild-moderate dementia. However, few of these instruments have sufficient content validity or reliability and some lacked adequate factor analyses to determine structural validity. **Conclusions:** There have been improvements in the methodological quality of research in this field since a review by Seignourel, Kunik, Snow, Wilson, & Stanley (2008). However, further validation studies of existing instruments are needed to improve utility, detection, and treatment of anxiety in dementia.

INTRODUCTION

It is estimated that approximately 46.8 million people are living with dementia worldwide and this figure is expected to rise to 131.5 million in 2050, if age related prevalence remains as it is (World Alzheimer's Report, 2016). In the UK, it is estimated that there are 850,000 people living with dementia at a total cost of £26billion per year. This figure incorporates NHS, social care funding, and the costs covered by individuals with dementia and their families' and is expected to rise to £55 billion by 2040 (Prince et al., 2014). There is a large body of research examining the neuropsychiatric and behavioural problems associated with dementia, with a focus on observable behaviours such as agitation, wandering, and aggression (Lai, 2014, Reisberg et al., 2014).

Anxiety is a common feature of dementia with considerable variability in estimated prevalence rates. These have varied for different types of dementia, e.g. 38% in Alzheimer's disease (AD) and 72% in vascular dementia (Ballard et al., 2000). The prevalence rates for individuals with an anxiety disorder ranges from 5-21%, while anxiety-related symptoms range from 8-71% (Seignourel, Kunik, Snow, Wilson & Stanley, 2008). Prevalence rates also differ depending on place of residence (community-dwelling vs. care homes). Twenty-five to sixty percent of participants across outpatient samples are reported to experience anxiety or anxiety-related symptoms, which include, nervousness, fears, irritability, agitation, muscle tension, day/night disturbance, and motor restlessness (Hwang, Masterman, Ortiz, Fairbanks, & Cummings, 2004; Starkstein, Jorge, Petracca, & Robinson, 2007; Steinberg et al., 2008). Irrespective of specific focus, individuals with dementia are more likely to have anxiety than healthy older adults (McClive-Reed & Gellis, 2011). Anxiety in dementia is associated with worse quality of life (QoL), reduced

function in activities of daily living, increased risk of nursing home placement, and higher caregiver burden than dementia alone (Gibbons et al., 2002; Schultz, Hoth, & Buckwalter, 2004; Mc Clive-Reed et al., 2011).

The variability in prevalence rates reflects a wider lack of consensus in defining what is meant by anxiety in dementia. One of the reasons for this is the overlap of symptoms common to both mild-moderate anxiety and dementia, e.g. cognitive, physical, and functional deficits (Neville & Teri, 2011). Research into late-life anxiety and cognitive impairment is largely focused on the diagnosis of Generalised Anxiety Disorder (GAD; Beaudreau & O'Hara, 2008). The DSM–IV diagnostic criteria highlight excessive anxiety or worry as the primary symptom of GAD, which must be present for 6 months or longer, in addition to related somatic symptoms, e.g. irritability, concentration problems, restlessness, sleep difficulty, muscle tension, and fatigue (DSM-IV; American Psychiatric Association, 2000). Such symptoms can also be present in dementia and whether or how to determine if these are due to anxiety or dementia is a key issue for the assessment of anxiety. In a number of studies of people living with dementia (PLWD), standardised diagnostic tools are used while ignoring the aetiology of the anxiety (e.g., Diefenbach, Bragdon, & Blank, 2014; Mansbach, Mace, & Clark, 2014). This approach avoids speculation about the cause of the symptoms but increases the risk of inflating prevalence rates of anxiety disorders because of overlap with dementia symptoms. Alternatively, Starkstein et al. (2007) have outlined revised criteria for GAD diagnoses specific to AD which miss out potentially co-morbid symptoms. Fears, irritability, restlessness, muscle tension, and respiratory symptoms were significantly associated with excessive anxiety and worry, while difficulty concentrating, and sleep disturbance/ fatigue were not (Starkstein et al., 2007). This approach introduces rigor to the

process, however these adapted criteria do not appear to be widely used in the literature and require further validation studies (Seignourel et al., 2008).

A further issue is the difficulty in differentiating anxiety from depression which frequently co-occur in individuals with Mild Cognitive Impairment (MCI) and dementia (Beaudreau & O'Hara, 2008; Sinoff & Werner, 2003). Significant overlap between these constructs has been demonstrated in MCI and dementia groups (e.g., Gibbons et al., 2002; Diefenbach et al., 2014). This may be a measurement issue but may also potentially point to the possibility that anxiety and depression do not represent distinct clinical entities in dementia. There are mixed results as to this from studies that have conducted factor analyses and it is not possible to draw overall conclusions due to various confounding factors across studies (Seignourel et al., 2008), e.g. different types of dementia and use of different instruments. The differentiation of anxiety and depression symptoms are important for researchers and clinicians when developing and evaluating targeted treatments. This has implications for how anxiety is measured in dementia and whether instruments used in dementia are sensitive and specific enough to differentiate these constructs.

There is a question as to whether self-reported anxiety by PLWD or carer/clinician-ratings are more reliable. Recent evidence suggests that PLWD are able to participate in surveys and provide consistent and accurate responses to quantitative questions (e.g. Snow et al, 2005; Clark, Tucke, & Whitlatch, 2008). In addition, an increasing number of people are being diagnosed with dementia at an earlier stage of disease progression (Department of Health, 2013) and are subsequently more likely to be able to reliably report on their own mood states. People with amnesia have demonstrated the ability to provide valid data about their current emotions, long after the source of the emotion is forgotten (Feinstein, Duff,

& Tranel, 2010). In addition, limbic structures involved in processing emotions are relatively preserved in individuals with Alzheimer's Disease (Barnes et al., 2006). However, difficulties with memory and language are likely to have an effect on the accuracy of self-report. Anxiety in people with AD may be more reactive to the current environment when compared with controls (Kolanowski, Hoffman, & Hofer, 2007). This has implications for the reliability of reporting of anxiety symptoms which may be overly influenced by the current context or environment. The ability of PLWD to report their anxiety symptoms will also be dependent on their dementia severity. These problems with self-report have been navigated by using caregiver-rated measures in dementia, e.g. the Rating Anxiety in Dementia Scale (RAID: Shankar et al. 1999), however anxiety symptoms have been shown to be rated at a higher frequency in caregiver reports than in those with dementia themselves.(Burke et al., 1998). This has been attributed to lack of reliability of carers' ratings, particularly when it comes to reporting on the cognitive symptoms of anxiety, e.g. excessive worry or PLWD's subjective feelings and experiences (Dawson et al., 2012), where carer's reports may be invalid.

Due to these conceptual issues, it cannot be assumed that anxiety measures standardised with adults or older adults can be used with the dementia population and the validity and reliability of anxiety measures need to be examined in a dementia population. Various self-report measures of anxiety and depression have been developed or adapted for use with PLWD. Heidenblut & Zank (2014) suggest a number of adaptations to instruments for use with PLWD: short and simple questions; fixed response sets; suitable for verbal administration and easy to understand and respond to. Reliable measurement of anxiety in dementia is essential to identifying this issue and intervening to limit its impact on health outcomes for

PLWD. There are no existing reviews of self-report anxiety instruments in dementia that use a level of evidence approach where studies are systematically ranked based on the rigor of their methods (Park, Reilly-Spong & Gross, 2013).

The primary aim is to review the quality of the instruments, considering the methodological quality of the studies. This will be achieved by using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) guidelines (Mokkink et al., 2010). The review will address the following questions:

1) What is the quality of instruments that purport to measure self-reported anxiety in individuals with dementia, considering the methodological quality of studies that report on measurement properties

2) Do self-report or informant instruments provide the most reliable measurement of anxiety in dementia?

METHOD

Search strategy

The electronic databases Ovid Medline® (1946 through December 2016), PsycINFO® (1806 through December 2016 s), and CINAHL® (1979 through December 2016) were searched using the subject words *dementia* combined with *anxiety.* This was the search strategy adopted by a previous review of anxiety in dementia by Seignourel et al. (2008). A hand search of the references of the included papers and relevant reviews was carried out in addition to the electronic search. The search was limited to studies of human beings and articles published in English.

Selection criteria

Articles were selected if their primary purpose was to develop or evaluate the measurement properties of an anxiety instrument. The instrument had to quantify self-reported anxiety in PLWD. and could be either a self-report measure or a measure incorporating self and informant reports. Studies that included mixed samples, i.e. those with and without a confirmed dementia diagnosis were excluded. Measures which included and defined separate subscales for anxiety and depression were included. If an article was not full-text or original (e.g. dissertations, reviews, or commentaries), it was excluded. Articles were also excluded if the primary aim was to test the efficacy of an intervention for the treatment of anxiety in dementia. The rationale for excluding efficacy studies was outlined by De Vet, Terwee, Mokkink & Knol (2011), who concluded that these studies normally provide indirect evidence of the measurement properties of an instrument. 2014). Quality of Life measures in dementia were also excluded, in addition to measures of caregiver anxiety or rating scales used to assess the effectiveness of interventions for carers.

One reviewer (M.S.) carried out the screening of titles and abstracts retrieved in the search and selected the included articles. Two reviewers (M.S. and J.S.) assessed the full text of articles for inclusion and jointly made decisions regarding the final included articles. The steps involved in identifying and selecting the studies are illustrated in Figure 1. Results from the three databases were combined and duplicates removed, identifying a total of 1323 papers. The titles of all papers were screened and some were excluded based on the relevance of the titles. The abstracts of all the remaining papers were read to identify potentially eligible studies. The main reasons for exclusion at this stage included measures of caregiver anxiety, if

there was no dementia sample, if the instrument in the study was not a measure of anxiety, or if the study was a treatment efficacy study. Following this, seventy-two papers were retrieved, read in full, and compared against the inclusion and exclusion criteria. Grounds for exclusion at this stage included: absence of original psychometric properties; instruments which did not assess the construct of anxiety or worry; or there was no dementia sample. A total of twelve papers met the inclusion criteria and formed the set of papers for the current review.

Figure 1: Flow chart of search and selection process

1323 papers retrieved from databases (once duplicates removed)

Titles of all retrieved papers

↓

904 papers excluded

Main reasons for exclusion:

- Caregiver measure
- No dementia sample
- Not a measure of anxiety in dementia
- Not an original article
- Efficacy or treatment effect study

420 abstracts reviewed

↓

358 papers excluded

Main reasons for exclusion:

➤ Neuropsychiatric measures
➤ Not a dementia sample
➤ Not a measure of anxiety in dementia

62 full papers reviewed

↓

49 papers excluded

➤ No original psychometric properties.
➤ Mood measure not specific to anxiety
➤ No adequate assessment of dementia

13 papers included in the review

Measurement properties

The COSMIN taxonomy distinguishes three domains to assess measurement quality: reliability, validity, and responsiveness (Mokkink et al., 2010). Reliability is defined as the degree to which the instrument is free from measurement error and contains three subcategories: internal consistency (the degree of interrelatedness among the items), measurement error (the error of a patient's score that is not attributed to true changes in the construct to be measured), and reliability (the proportion of the total variance in the measurement which is because of ''true'' differences among patients). Validity is the degree to which the instrument measures the construct(s) it claims to measure and is broken down into content, construct, and criterion validity. Content validity is the degree to which the content of the instrument is an adequate reflection of the content being measured. It includes face validity; items of the instrument are in line with the construct being measured. Construct validity includes structural validity, hypothesis testing, and cross-cultural validity. Structural validity refers to the degree to which scores on the instrument measure the dimensionality of the construct. Hypothesis testing refers to the instrument's relationship with other measures that claim to measure the same construct and differ significantly from instruments that claim to measure different constructs. Criterion validity refers to the extent to which the instrument correlates with an accepted 'gold standard'. There is currently no accepted 'gold standard' instrument to measure anxiety in dementia (see Siegnoruel et al., 2008), therefore an evaluation of criterion validity will not be conducted.

The COSMIN checklist and assessment of measurement quality

One reviewer (M.S) extracted data from the selected articles and evaluated the measurement quality using the 4-point scale COSMIN checklist as a guide (Terwee et al., 2012). Each measurement property was rated as either: *excellent, good, fair,* or *poor.* The "worst score counts" algorithm was used in rating measurement properties. For example, if one element of a measurement property was defined as *poor*, the overall measurement property was considered *poor*. Where articles presented measurement properties for more than one instrument, these were presented, evaluated and rated as separate studies (see Table 4).

Best evidence synthesis

COSMIN also assesses whether studies provide *positive*, *negative* or *indeterminate* results for each measurement property based on study quality. Criteria for these ratings are outlined by Terwee et al. (2007) (see Table 1). Table 2 presents the criteria used when combining the results from the assessment of measurement properties and the quality of the studies reviewed

Table 1: Quality criteria for assessment of measurement properties

adapted from Terwee et al. (2007) and Park et al. (2013)

Property	Rating	Quality Criteria
Internal consistency		
	+	Subscale unidimensional and Cronbach/s alpha(s) ≥0.70
	?	Dimensionality not known or Cronbach's alpha not determined
	-	(Sub)scale not unidimensional OR Cronbach's alpha(s) <0.70
Reliability		
	+	ICC/weighted Kappa ≥ 0.70 or Pearson's r ≥ 0.80
	?	Neither ICC/weighted Kappa, nor Pearson's r ≥ 0.80
	-	ICC/weighted Kappa <0.70 OR Pearson's r< 0.80
Content Validity		
	+	The target population considers all items in the questionnaire to be relevant and considers the questionnaire to be complete
	?	No target population involvement
	-	The target population considers items on the questionnaire to be incomplete/no information found on target population
Structural		
	+	Factors should explain at least 50% of the variance or good or adequate fit (see goodness-of-fit criteria for CFA or EFA[a]
	?	Explained variance not mentioned OR equivocal fit by goodness-of-fit criteria for CFA or EFA[a]
	-	Factors explain <50% of the variance OR poor fit by goodness-of-fit criteria for a CFA or EFA[a]
Hypothesis Testing		
	+	Correlation with an instrument measuring the same construct r=≥0.50
	?	Solely correlations determined with unrelated constructs
	-	Correlation with an instrument measuring the same construct r=<0.50

[a] Good or adequate fit: comparative fit index (CFI) ≥0.90, root mean square of approximation (RMSEA) ≤0.08, standardized root means square residual (SRMR) <0.10. Inadequate fit: CFI ≤0.85, RMSEA ≥0.10, SRMR ≥0.10; Indeterminate fit: the values of the fit indexes ranged in between the adequate criteria and inadequate criteria.

Table 2: Levels of evidence for the overall quality of the measurement properties

Level	Rating	Criteria
Strong	+++ or ---	Consistent findings in multiple studies of *good* methodological quality OR in one study of *excellent* methodological quality.
Moderate	++ or --	Consistent findings in multiple studies of *fair* methodological quality OR in one study of *good* methodological quality.
Limited	+ or -	One study of *fair* methodological quality.
Conflicting	±	Conflicting findings from studies of comparable quality
Indeterminate	?	Findings from *excellent*, *good* or *fair* studies were not definitively positive or negative
None	na	Findings from *excellent, good* or *fair* studies were not available

Table from Park et al. (2013) was used. This was originally adapted from Van Tulder et al. (2003)

+positive result; -negative result; ±both positive and negative findings have been reported by studies of adequate quality; ? findings from studies of adequate quality were not definitively positive or negative; na findings from studies of adequate quality were not available.

RESULTS

Across ten studies, nine instruments were evaluated and a summary of the findings were described for each instrument. Table 3 presents the number of studies reviewed per anxiety instruments. Table 4 gives details and item examples of the instruments reviewed. Table 5 demonstrates the methodological quality for each study, rated as *excellent, good, fair, or poor* per measurement property. Table 6 shows the level of evidence synthesis based on the Terwee et al, 2007 criteria (see Table 2). This synthesis combines the *positive, negative or indeterminate* ratings for each measurement property. It also incorporates the methodological quality of the studies, and the consistency of their findings.

Table 3: Number of studies reviewed per anxiety instrument

Anxiety Measures	No. of studies
Rating Anxiety in Dementia Scale (RAID; Shankar, Walker, Frost, & Orwell, 1999)	4
Geriatric Anxiety Inventory (GAI; Pachana et al. 2007).	1
The STAI-S (Form Y 1) modified by Ward, Wadsworth, & Peterson (1994)	2
The Participant Anxiety Scale modified by Gibbons, Teri, ,Logsdon, & McCurry, 2006	1
The Penn State Worry Questionnaire Abbreviated (PSWQ-A; Hopko et al. 2003)	1
The Worry Scale (La Barge, 1993).	1
The Brief Anxiety and Depression Rating Scale (BADS; Mansbach et al., 2015)	1
The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983)	1
The Empirical Behavioral Pathology in Alzheimer's Disease (E-BEHAVE-AD) Rating Scale (Auer, Monteiro & Reisberg, 1996)	1

Table 4: Characteristics and item examples from the included instruments

Instrument	Construct assessed	Recall period	Dimensions (number of items)	Number of subscales	Response options (range)	Ease of scoring and administration (range of scores)	Sample items
RAID	Anxiety in dementia	Two weeks	Worry (10) Apprehension and vigilance (3) Motor tension (2) Autonomic hyperactivity (2) Phobias and panic attacks (1) Total (18)	4	Four point-scale (0= absent to 3=severe)	Moderate: final score determined by clinician based on interview with the participant and collateral report. (0-18)	"Worry about physical health" (Worry) "Sensitivity to noise, exaggerated startle response" (Apprehension and vigilance "Trembling" (Motor tension)
GAI	Anxiety	One week	Gastrointestinal (6) Hyperarousal (5) Excessive worry (8)	4	Two-point scale (agree/disagree)	Moderate: 20 items	"I get an upset stomach due to my worrying" (Gastrointestinal) "I often feel upset" (Hyperarousal)

			Difficulty-making decisions (1) Total (20)				"I worry a lot of the time" (Excessive worry)
STAI-S	State anxiety	Not specified	20 items	1	Two-point scale (Yes/No)	Moderate: 20 items	"I am tense" "I am worried" "I feel calm"
PAS	Anxiety	None-current symptoms	12 items, dimensions not reported	Not reported	Two-point scale: "yes" or "no" to presence/absence of symptoms	Easy	Not reported
PSWQ-A	Worry	No specific timeframe outlined	Worry (8)	1	Five-point scale: (0="not at all typical of me") to 5 ("very typical of me").	Moderate: ability to answer scaling questions may be impacted by cognitive impairment.	"I know I should not worry about things, but I just cannot help it" "When I am under pressure, I worry a lot"
The Worry Scale	Dementia worry	Unspecified	Worry (10)	1	Five-point scale: "always", "often", "sometimes", "rarely" and "never"	Moderate: response options require temporal judgement	"I feel worthwhile and satisfied with myself" "The changes in my memory make me feel I've lost a very important part of myself"

BADS-Anxiety Factor	Anxiety	No specific timeframe outlined	Anxiety (6)	1	Three-point scale: "no" "somewhat or "yes"	Easy: clinician reads statements, average of 5 minutes to administer for participants with dementia	"Recent increase in worrying" "Worrying about future" "Overwhelmed" "Nervousness" "Controlling worry"
HADS-Anxiety scale	Anxiety	One week	Anxiety (7)	1	Three-point scale rating frequency (0-3)	Easy: Administered by research assistant (0-21)	I feel tense or 'wound up' Worrying thoughts go through my mind'
E-BEHAVE AD	Anxieties and Phobias	None	General anxieties Fear of being left alone	12	0=absent 1=Mildly present 2=Moderately present 3=Severely present	Easy: Administered by trained clinician (0-36)	Patient is anxious regarding his/her memory Other anxieties, e.g., about health, money, the future, their children, etc. Stranger anxiety (gets anxious when confronted with examiner)

Rating for Anxiety in Dementia Scale

Description of the measure

The Rating Anxiety in Dementia Scale (RAID; Shankar et al., 1999) contains a total of 18 items rated on a four-point scale (see Table 4). The original study specified how final ratings were decided by a clinician following interviews with the individual with dementia and an informant. A score of 11 or more suggests clinically significant anxiety. Two studies in this review (Gibbons et al., 2006; Twelftree & Qazi, 2006) used only caregiver ratings of anxiety.

For the RAID, study quality ratings have been reported for the following measurement properties: internal consistency, reliability, content validity, structural validity, and hypothesis testing. Internal consistency for the RAID was *positive* assessed in three studies assessed to be of *poor* quality (Shankar et al. 1999; Snow et al. 2012; Twelftree & Qazi, 2006). Cronbach's alpha ranged from 0.75-0.83 for the RAID and 0.75 for the RAID-SI (Snow et al. 2012). Inter-rater reliability was rated as *negative* in a *fair* quality study (Shankar et al. 1999) and *positive* in a *poor*-quality study (Snow et al., 2012), thus overall can be considered *poor*. Content validity was rated as *excellent* (Shankar et al., 1999) and this original study was subsequently referenced by the other studies in this review. Structural validity assessment using a Principal Components Analysis (PCA) suggested a one factor structure (Shankar et al., 1999). However, sample size was small meaning that further Exploratory Factor Analysis (EFA) or Confirmatory Factor Analysis (CFA) studies should be carried out.

All four studies evaluated construct validity through hypothesis testing. The RAID showed *positive* moderate–high correlations with other observer-rated anxiety scales (Clinical Anxiety Scale; CAS, The Anxiety Status Inventory; ASI and GAI-collateral). There were moderate-high correlations with self-report scales (STAI-S and PSWQ-patient rating). Three of these studies were assessed to be of *fair* methodological quality (Shankar et al., 1999; Snow et al., 2012; Twelftree &Qazi; 2006), while Gibbons et al. (2006) was rated as *good* for methodological quality. The RAID shows weaknesses in its ability to distinguish anxiety from other related constructs, e.g. there was a moderate correlation with a measure of agitation; The Cohen Mansfield Agitation Inventory (CMAI; Cohen-Mansfield, Marx, & Rosenthal, 1989). In addition, correlations with a measure of depression the Cornell Scale for Depression in Dementia (CSSD; Alexopoulos, Abrams, Young, & Shamoian, 1988a) were greater than those with other measures of anxiety.

Two studies assessed predictive validity. Shankar et al. (1999) used modified DSM-IV criteria for Generalized Anxiety Disorder (GAD), where a score of 11 or more on the RAID had a 90% sensitivity and 78.5% specificity rating. Snow et al. (2012) used the The Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998) as a diagnostic tool for anxiety in their dementia sample. With a cut-off score of 10, the RAID-Structured Interview demonstrated a 90% sensitivity rating and a 67% specificity rating and an area under the curve (AUC) at 0.80 (SE=0.08; 95% CI= 0.64-0.96).

Table 6 shows the results of combining both quality criteria for measurement properties and the methodological quality of each of the studies in analysing these measurement properties. The RAID had *strong* content validity and *moderate* structural validity and hypotheses testing validity. The reliability of the RAID is *limited* and the *poor* quality of the studies means that internal consistency findings were given no weight in this final synthesis.

The Geriatric Anxiety Inventory

Description of the measure

The Geriatric Anxiety Inventory (GAI) is a twenty-item self-report instrument designed to measure anxiety in older people (Pachana et al., 2007). Psychometric properties of the measure have been reported for patients with cognitive impairment, demonstrating good reliability, construct, and predictive validity in diagnosing GAD (Byrne, Pachana, Arnold, Chalk, & Appadurai, 2008; Boddice, Pachana, & Byrne, 2013).

GAI study quality ratings have been reported for the following measurement properties; internal consistency, content validity and hypothesis testing. The reliability of the GAI will be addressed in a later section (see carers vs. self-report). Internal consistency was rated as *positive*. A high Cronbach's alpha value ($\alpha=0.92$) was reported in a study of *fair* methodological quality (Bradford et al. 2013). The GAI was developed to assess the symptoms of GAD, (e.g. fearfulness, worry, and physiological symptoms). However, there has been no assessment of the relevance of the items for people with cognitive impairment or a diagnosis of dementia. (Bradford

30

et al, 2013; Gerolimatos et al., 2013). Bradford et al. (2013) adopted *fair* statistical

methods to assess the predictive validity. The GAI performed above chance in

predicting an anxiety disorder with an AUC of 0.62 (SE=0.08). In terms of

hypothesis testing, there was a high correlation with a measure of depression yielding

a *positive* result for hypothesis testing as the constructs of depression and anxiety are

related (Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Sinoff, Ore,

Zlotogorsky, & Tamir, 2002).

Levels of evidence conclusions

Evidence for the internal consistency and hypotheses testing of the GAI with

dementia participants is *limited*. Reliability of the measure is *indeterminate*, while

adequate content validity analysis has not been conducted. Further studies of content

validity are required involving both PLWD and carers to judge the relevance and

comprehensiveness of the items for this population.

The State-Trait Anxiety Inventory-Short form (STAI-S)

Description of the measure

The STAI (Form Y-1; Speilberger, 1983) was adapted by Ward et al. (1994). This is

a 20-item scale and instead of choosing one of the 4 degrees of anxiety severity (as is

the procedure in the original measure), participants were asked to endorse each item

by answering "yes" or "no". Higher scores on this measure indicate greater anxiety.

STAI-S study quality ratings have been reported for the following

measurement properties: internal consistency, content validity, and hypothesis

testing. Cronbach's alpha for the measure was *positive* across both studies (α=0.88;

31

Twelftree &Qazi, 2006) and (α =0.91: Ward et al., 1994) however, the methodological quality of both were *poor* as there was no factor analyses referenced. In terms of content validity, ratings for quality were *indeterminate* and the methodological quality was *poor* in both studies. Ward et al. (1994) made adaptations to the original STAI to account for participants with dementia, e.g. reducing the response options and using it as a clinician-administered scale, however the relevance of the items was not assessed in the target population. Hypothesis testing yielded *positive* quality ratings in studies of *fair* methodological quality. The STAI-S correlated moderately with the worry subscale of the RAID (Twelftree & Qazi, 2006).

Levels of evidence conclusions

There is *moderate* evidence for hypothesis testing of the STAI-S. For internal consistency and content validity, there is insufficient evidence to draw conclusions about the quality of these measurement properties.

The Participant Anxiety Scale

Description of the measure

The Participant Anxiety Scale is an adapted version of the Clinical Anxiety Scale (CAS; Westhuis & Thyer, 1989). The items for the CAS were derived from the criteria for anxiety disorders in the DSM-III (3rd ed.; *DSM–III*; American Psychiatric Association, 1980). It was originally designed for use in measuring anxiety in clinical settings and not intended for use with a dementia population.

The study quality of Gibbons et al. (2006) is rated for the following measurement properties: internal consistency, content validity, and hypothesis

testing. Methodological quality for assessing internal consistency and content validity was *poor* and therefore these statistics are not reported. The content validity of the measure was taken into consideration and adaptations made to improve ease of use for dementia patients, e.g. adapting the response options to "yes" or "no". However, there were no details provided as to which items were removed nor are there reported statistical methods for how these decisions were made by the authors. Hypothesis testing yielded *negative* results and the quality of this study was *good*. Correlations between the PAS and various measures of anxiety were weak (RAID and NPI-Anxiety Scale). There were several possible explanations for this, including; different symptoms measured by each of the scales, different time periods assessed, e.g. past two weeks with the RAID vs. current symptoms with the PAS, and variable rating of symptoms, e.g. the PAS is scored based on presence or absence of an anxiety symptom vs. the RAID score based on severity of the symptom.

Levels of evidence conclusions

For the internal consistency and content validity, findings from studies of adequate quality were not available. One study of *good* methodological quality; Gibbons et al. (2006) demonstrated a *negative* result for hypothesis testing.

The Penn State Worry Questionnaire-Abbreviated

Description of the measure

The PSWQ-A is an abbreviated and simplified version of the PSWQ (Meyer, Miller, Metzger & Borkovec, 1990). It was validated with 160 older adults with GAD diagnoses and CFA found that a single factor model with an 8-item scale fit the data most optimally (Hopko et al., 2003).

Bradford et al. (2013) reported data on the following measurement properties: internal consistency, reliability, content validity, predictive validity. The methodological quality of Bradford et al. (2013) was rated as *fair* for each measurement property, except for the methods used to assess content validity which were rated as *poor*. Internal consistency was rated as *positive* (α=0.84) for PSWQ-A-Participant and 0.89 for PSWQ-A-Collaterals. There was a *negative* result for reliability. Intra-class correlations (ICCs) considering participants and collateral ratings fell below the quality criteria (ICC of 0.417) with collaterals' ratings exceeding participants' ratings by an average of 5.5 points. Methods used to assess the content validity of this measure were *poor*. In terms of predictive validity, the PSWQ-A-Participant had an AUC of 0.691 (SE = 0.08) using the Mini-International Neuropsychiatric Inventory (MINI: Sheehan et al., 1998) as a comparison instrument.

Levels of evidence conclusions

There is *limited* evidence for the internal consistency and reliability of the PSWQ-A in a dementia sample. Evidence for content validity was *indeterminate*, while structural validity and hypothesis testing were not assessed.

The Worry Scale

The Worry Scale (LaBarge, 1993) is a 10 item self-report instrument designed for use with individuals with mild dementia. The items are a series of statements about feelings experienced in the context of memory loss, (see Table 4), therefore this scale specifically assesses worry related to dementia symptoms rather than anxiety symptoms. Study quality ratings have been reported for the following measurement

properties: internal consistency, reliability, content validity, and hypothesis testing. Internal consistency of the measure (using 8 items) was *positive* (α=0.85) in a study of *fair* methodological quality. In terms of reliability analysis, split-half methods were used demonstrating Pearson's *r* values for the 'no- dementia' group (0.82), 'very mild' (0.77) and 'mild' (0.80). As per the Terwee et al. (2007) criteria (see Table 1), the 'no-dementia' group and the 'mild dementia' group met criteria for a *positive* rating (r≥0.80). Content validity was *indeterminate* and the methods used were *poor*. Items were generated by professionals with experience in working in an Alzheimer's research centre but there was no target population involvement. Structural validity was demonstrated using a PCA which indicated a unidimensional scale for the 'very mild' and 'mild' groups collectively. Two items were removed due to poor factor weightings ('I am able to express my feelings now' and 'I talk to someone who understands what is happening to me'). Factor weightings for the remaining 8 items ranged from 0.45 to 0.78 in a study of *fair* methodological quality. Hypothesis testing was rated as *negative* using *good* methods. There were moderate correlations with both the State and Trait Anxiety measures (r= 0.55 for both). As expected, there was also a moderate correlation with a measure of depression (r= 0.66) and weak correlations with state anger (r=0.32) and trait anger (r=0.31). However, we would not expect the correlation with depression to exceed the correlation with another measure of anxiety, suggesting The Worry Scale does not discriminate between these two constructs.

Levels of evidence conclusions

There is *limited* evidence of *positive* results for the internal consistency and reliability of The Worry Scale. There is no available evidence to draw conclusions on

the content validity of this measure. Finally, there is *moderate* evidence to suggest *negative* results for hypothesis testing, i.e. the ability of The Worry Scale to discriminate between anxiety and depression.

The Brief Anxiety and Depression Scale (BADS)

Description of the measure

The BADS is an 8-item scale assessing symptoms of anxiety and depression based on Major Depressive Disorder (MDD) and GAD. Additional items were added including behavioural and somatic symptoms (e.g. physical complaints and signs of agitation) as these are often reported by depressed older adults (Mansbach et al., 2015).

The quality of measurement properties was reported for the following: internal consistency, content validity, structural validity and hypothesis testing by Mansbach et al. (2015). Internal consistency was rated as *positive* ($\alpha=0.75$), however the quality of the methods used were *poor* as alpha values were not calculated for the anxiety and depression subscales separately. In terms of content validity, both the measurement property and study methods were rated as *excellent.* Mansbach et al. (2015) outlined the process of item selection, whereby various health care professionals reviewed the items at each stage of its development. The final 8 items were pilot tested in clinical settings and revised to improve clarity. Structural validity was rated as *positive* in a study of *excellent* methodological quality. Exploratory PCA demonstrated two separate components for anxiety and depression. Percentages of variance accounted for were 36.72% and 13.49% respectively. Only statistics

relevant to the assessment of the anxiety component of this measure will be reported. For the Anxiety Factor, a cut-score of > 4 (scores 4 and below indicating no GAD) yielded the optimal balance of sensitivity (73%) and specificity of (81%) for identifying GAD. A ROC curve was calculated and the AUC was 0.85 (SE=; 95%CI=0.80-0.90). These results suggest that the BADS Anxiety scale can detect clinical anxiety in the presence of dementia.

In terms of hypothesis testing, the study indicated *positive* results, however the quality of the methods used were deemed to be *fair.* The BADS anxiety factor is supported by a strong correlation with the Generalized Anxiety Disorder 7-item scale (GAD-7) (Spitzer, Kroenke, Williams, & Lowe, 2006).

Levels of evidence conclusions

There is *strong* evidence for the content validity of the measure and *moderate* evidence for its structural validity. Hypothesis testing shows *limited positive* evidence in the current study.

The Hospital Anxiety and Depression Scale (HADS)

Description of the measure

The Hospital Anxiety and Depression Scale (HADS) (Zignond and Snaith, 1983) is a fourteen-item self-report measure of anxiety and depression. Each item is rated from 0 to 3 and higher scores indicate greater severity of anxiety or depression.

The aim of the Stott et al. (2016) study was to establish the structural validity of the HADS in a dementia sample and this was the only element of validity reported on.

37

Structural validity was rated as *indeterminate* in a study of *excellent* methodological quality. The percentage of missing data was outlined and listwise deletion of cases was employed before conducting the CFA. Three different pre-determined models were tested; the two-factor model suggesting anxiety and depression as separate constructs (Zigmond and Snaith ,1983), the one factor model suggesting one overall distress factor (Razavi, Delvaux, Farvacques, & Robaye, 1990), and the three-factor non-hierarchical model which defines 'anxiety', 'depression' and 'negative affect' (Dunbar, Ford, Hunt & Der, 2000). Both the two factors and three factors models met criteria for good or adequate fit as described by Terwee et al. (2007). The analysis did not adequately distinguish between the two-factor and three-factor models which makes the interpretation of the HADS in dementia uncertain. Additional specification searching demonstrated that the fourth item on the anxiety scale "I can sit at ease and feel relaxed" did not relate to the underlying construct of anxiety. Stott et al. (2016) suggested that removal of this item should be considered for people with dementia and future research into developing adjusted cut-off scores is required.

Levels of evidence conclusions

This factor analytic study provides *indeterminate* evidence for the structural validity of the HADS in individuals with mild-moderate dementia.

The E-BEHAVE-AD

The E-BEHAVE-AD is a 12-item clinician-rated instrument developed to assess behavioural pathology in Alzheimer disease and related dementia. Clinician-ratings

are based on a 20-minute clinical interview with the PLWD without a care-giver present. The symptomatic model for this scale comes from the caregiver-rated Behaviour Pathology in Alzheimer disease questionnaire (BEHAVE-AD). It assesses six pathological domains: paranoid and delusional ideation; hallucinations and activity disturbance; aggressiveness; affective disturbance; anxieties and phobias. Twelve symptoms are assessed on a 4-point severity scale, where 0 represents absence of the symptom across the observation period and a score of 3 represents the symptom with severe magnitude (Reisberg et al., 2014).

For the E-BEHAVE-AD study quality ratings have been reported for reliability and hypothesis testing.

Reliability was rated as *positive* in a study of *poor* methodological quality. A rater independently interviewed each participant while a second rater observed. Raters did not communicate and independently rated the interview. The interviewing rater was altered randomly. For 'general anxieties' the intra class correlation for raters-fixed was 0.86 and 0.84 for raters-random. For 'fear of being left alone' the intra class correlations were 1.00 for both raters-fixed and raters-random. (Auer et al., 1996). The study quality is *poor* for the assessment of reliability due to the small sample size (N=20). Hypothesis testing was rated as *negative* in a study of *fair* methodological study. There was a significant correlation between the E-BEHAVE-AD and the carer-rated BEHAVE-AD for the 'anxieties and phobias' category was statistically significant but below the r=0.50 threshold (see Table 1).

Carers report vs self-report

The results from the reviewed studies presented a mixed picture and did not give a definitive answer as to whether self-report or carer-report measures are better in

measuring anxiety in dementia. The RAID scoring was informed by both caregivers and clinician ratings and it is not clear the degree to which the self-report of PLWD accounted for the final clinician ratings. It is likely as dementia severity increases; self-report becomes less reliable (Shankar et al., 1999). When the RAID was used as an informant only measure, 40% of the sample were rated as scoring above the clinical cut-off score of 11 for anxiety (Twelftree & Qazi, 2006). This result calls into question the reliability of carer-only ratings as typical prevalence rates for anxiety among individuals with dementia ranges from 5 to 21% (Seignourel et al., 2008).

Collateral ratings on the GAI yielded a higher area under the curve AUC (0.81, SE=0.08) than participant ratings (0.69, SE=0.08). This indicates that both are only modestly accurate in predicting a clinical diagnosis of anxiety (Bradford et al., 2013). The optimal cut-off score maximising sensitivity and specificity was 8 for participants and 10 for collaterals. It was demonstrated that other factors such as the type of relationship between the participant and collateral (e.g. spouse, adult child) or living arrangement did not account for the degree of discrepancy between the ratings, however if the collateral's gender was female, ratings were less concordant. These findings warrant further research into the characteristics of collaterals that may influence their anxiety ratings.

PSWQ-A collateral ratings yielded a higher area under the curve AUC (0.77, SE=0.08) than participant ratings (0.69, SE=0.08), however this difference was not

significant (Bradford et al., 2013). Similar to the GAI, optimal cut-off scores were lower for participants (17) than for collaterals (22).

The correlation between anxiety symptoms as rated by the E-BEHAVE AD and the carer-rated BEHAVE-AD was *negative* in a study of fair methodological quality.

Overall, there is a mixed picture of results across the different anxiety measures as to the reliability of carer vs. self-report measures. The evidence suggests that both self and informant reports of anxiety symptoms are warranted in the assessment of anxiety in dementia.

Table 5 Methodological quality of each study per measurement property and instrument

Measure	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Population (demographics, setting, diagnoses)
RAID						
Shankar et al. 1999	Poor	Fair	Excellent	Good	Fair	N=83, M age=79.1. %F=62 UK Inpatient/day hospitals. Dementia (DSM IV).
Twelftree & Qazi 2006	Poor		Excellent		Fair	N=40, M age=79.4 years. %F=70 UK day hospitals/community. Mild-moderate cognitive impairment (MMSE)
Gibbons et al. 2006			Excellent		Good	N=95, M age=79.9. % F=66 US community sample. Dementia (confirmed by medic), anxiety (caregiver report of 3+ anxious or depressed behaviours).
Snow et al. 2012	Poor	Poor	Excellent		Fair	N=32. M age=78.6. %F=59 US Primary care/community day hospital Mild-moderate dementia (confirmed by medic), anxiety (MINI)
GAI						
Bradford et al. 2013	Fair	Fair	Poor			N=41. M age=79.1. %F=23 US community sample Mild-moderate dementia (notes review), CDR score= 0.5-2. Anxiety disorders in 63.4% of the sample (MINI)
Measure	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Population (demographics, setting, diagnoses)

STAI-S

Ward et al. 1994	Poor		Poor		Fair	N=40. *M* age=68.8. %F: 0. US veterans' medical centre. Dementia diagnosis (DSM -III)
Twelftree & Qazi 2006	Poor		Poor		Fair	See RAID
PAS (Gibbons et al. 2006)	Poor		Poor		Good	See RAID
PSWQ-A Bradford et al. 2013	Fair	Fair	Poor			See GAI
The Worry Scale (La Barge, 1993)	Fair	Fair	Poor	Fair	Good	N=182 *M* age=74.8. %F=64 Very mild dementia (CDR=0.5), mild dementia (CDR=1)
BADS (Mansbach et al. 2015)	Poor		Excellent	Good	Fair	N=224 *M* age=80.5 %F=62 US residential care sample. Mixed sample: Mild, moderate, severe Dementia (54%), no dementia, incl. MCI (46%) (CDR).
HADS (Stott et al. 2016)				Excellent		N=268. *M* age=69.1 %F=52 UK community sample. Mild-moderate dementia (DSM-IV).
E-BEHAVE-AD (Auer et al., 1996)		Poor			Fair	N=44. *M* age= %F=44.9 For reliability analysis N=20 US community sample MCI (12%), AD (61%) VaD (14%), Other (13%). DSM III

N=number of participants, *M* age= mean age of participants, % F= percentage of females, DSM III, IV=Diagnostic and Statistical Manual of Mental Disorders, MMSE=Mini Mental State Exam, MINI= Mini International Neuropsychiatric Interview, CDR= Clinical Dementia Rating CIND: Cognitive Impairment, Not Dementia. MCI=Mild Cognitive Impairment. VaD= Vascular Dementia. AD= Alzheimer's Disease

Table 6: Levels of evidence synthesis: Quality of measurement properties per instrument

Instrument	Internal consistency	Reliability	Content Validity	Structural Validity	Hypothesis testing
RAID	na	-	+++	++	++
PAS	na	na	na	na	--
STAI-S	na	na	na	na	++
GAI	+	?	na	na	+
PSWQ-A	+	-	na	na	na
The Worry Scale	+	+	na	+	--
HADS	na	na	na	+++	na
BADS	na	na	+++	++	+
E-BEHAVE-AD	na	+	na	na	-

See Table 2 for levels of evidence descriptors

DISCUSSION

The first aim of this study was to systematically review the quality of instruments which claim to measure self-reported anxiety in dementia, considering the methodological quality of studies. The second aim was to establish whether informant or self-report measures are more reliable. The COSMIN checklist was used to evaluate the measurement properties and the methodological quality of nine instruments. Internal consistency, reliability, content validity, structural validity, and hypothesis testing were assessed across the included studies.

Summary of findings

The RAID was the most frequently evaluated instrument and had the highest level of evidence in terms of quality of measurement properties (content validity, structural validity, and hypothesis testing) and methodological quality. The BADS showed initial *strong* evidence for content validity as trials involving individuals with dementia formed part of the item selection process. This newly developed instrument also demonstrated *moderate* evidence for structural validity. The HADS showed some initial evidence for structural validity but requires further research to establish an evidence base for a range of measurement properties when using it with a dementia population. The remaining instruments provided *limited* evidence in terms of the quality of measurement properties and the methodological qualities of the studies.

Limitations

The methodological quality of many of studies in this review has been compromised by the problem of small sample sizes. This has implications for the degree of statistical power in these studies and whether this is sufficient to detect true effects, inflating the risk of a Type II error. Small sample size also affects the stability of obtained factor structures, e.g. despite a sample size of less than 100, the original PCA for the RAID (Shankar et al. 1999) was the basis for all further validation studies. Finally, this highlights the problem of representativeness and reduces the precision and level of confidence in extrapolating these results to the dementia population as a whole.

Content validity analysis was lacking across many of the studies reviewed here. The inclusion of individuals with dementia in the process of item testing and selection is an issue which needs to be addressed. This is particularly important given that prior research suggests that individuals with dementia have the capability to participate in research using standardised measures to explore their illness experience (e.g. Dawson et al., 2012, Snow et al. 2005, Clark et al., 2008). In addition, selection of which measure to use must also consider ease of administration, completion time, and the needs of the target population. Some studies have made attempts at adapting their measures for individuals with dementia, e.g. For the GAI, participants were asked to report on the occurrence of experiences during the past week. It is possible that memory difficulties associated with cognitive impairment have an impact on the accuracy of self-reported symptoms of anxiety when the recall period is longer. A number of studies have demonstrated that individuals with memory impairments

have the ability to provide valid data about their current emotions (Feinstein et al., 2010; Ready, O Carvalho, Green, Gavett, & Stern, 2011; Dawson et al., 2012).

An important problem which this review highlights is the lack of consensus about what constitutes the symptoms of anxiety in dementia and the absence of a 'gold-standard' instrument for diagnosing anxiety in dementia. The instruments reviewed here all measure slightly different things. The RAID and BADS, for example, measure anxiety symptoms based on DSM-IV criteria for GAD but the RAID includes additional items associated with somatic and behavioural symptoms (e.g. sleep disturbance, and motor tension). The Worry Scale assesses a range of emotional responses to dementia and coping styles (La Barge, 1993). Informant measures such as the RAID tend to use behavioural items, while self-report measures include more cognitive and affective aspects of anxiety (e.g., BADS; Mansbach et al.). Measures such as the E-BEHAVE-AD look at anxiety symptoms alongside other behavioural and psychological symptoms associated with dementia. Before further validation studies are carried out on the measures reviewed here, research in this field should focus on arriving at a consensus as to which specific symptoms constitute the concept of anxiety in dementia. This could be achieved by conducting studies which use Confirmatory Factor Analysis (CFA) with larger samples of individuals with dementia. Siegnourel et al. (2008) highlighted the problem of how research in this field tends to focus on initial validation of a measure followed by use of the instrument in research and clinical practice without further validation. However, this problem seems to persist, with the emergence of studies validating new measures, e.g. BADS (Mansbach et al., 2015).

Conceptual and methodological issues

The results summarized in the current review suggested considerable overlap between symptoms of anxiety and depression in dementia. These results correspond with rates of co-morbidity in adult clinical samples, e.g. in a large clinical sample (N=1127), 57% of individuals with a diagnosis of a MDE also met criteria for GAD (Brown et al., 2001). The Worry Scale, RAID, GAI and STAI-S all demonstrated moderate to strong correlations with instruments that measure depression. There is also evidence of overlap with other related constructs, e.g. The Worry Scale was associated with both state and trait anger and the RAID showed an association with a measure of agitation.

For studies where both participant and informant ratings of anxiety are considered, the RAID, GAI, and PSWQ-A indicated discrepancies. Typically, informants tended to rate anxiety at higher levels. It is difficult to establish the reasons for these discrepancies as there do not appear to be studies where priori hypotheses are outlined and sufficiently tested. A number of different factors have been proposed to influence careers 'higher ratings, e.g. a focus on cognitive and functional status of the patient (Snow et al., 2005), caregiver mood (Pearson, Teri, Wagner, Truax, & Logsdon, 1993) and caregiver burden (Teri & Traux, 1994). Patient self-reported mood has been linked to deficit awareness which can impact on reliability (Snow et al. 2005). One exception is where care-givers are employed workers and not family members (Gerolimatos et al., 2015). Both caregiver burden and depression have been shown to negatively impact on caregiver's ratings of patient's quality of life (Karlawish, Casarett, Klocinski, & Clark, 2001; Schultz et al., 2004). It may be that professional caregivers experience less burden than family

caregivers. Family caregivers have been shown to have higher levels of distress on a neuropsychiatric inventory, when compared with professional caregivers (Tan, Wong, & Allen, 2005). The authors speculated that family caregivers' perceptions of their own helplessness and concerns over their abilities to control difficult behaviours may contribute to their higher levels of distress.

The RAID, GAI, STAI-S and BADS all demonstrate significant negative correlations with various measures of cognitive impairment which indicates that these anxiety measures may demonstrate divergent validity. An alternative explanation for this result is that having more cognitive impairment is associated with less anxiety. Studies have shown that individuals with dementia who retain insight into their difficulties, experience greater levels of anxiety (e.g., Shankar et al. 1999; Ballard, Boyle & Bowler, 1996).

One of the strengths of the current review is the use of the COSMIN rating tool which introduces rigour into the process of evaluating measurement properties and the methodological quality of studies that report on them. A possible limitation is the appropriateness of the COSMIN rating tool in assessing small scale research studies. The psychometric properties that are impacted by sample sizes of less than 50 participants, e.g. internal consistency, reliability and hypothesis are rated as *fair* or *poor* using the COSMIN criteria. Alternatively, Siegnourel and colleagues (2008) used four qualitative criteria for assessing the studies in their review (1) independent assessment of anxiety, distinguishing it as a construct separate from depression or agitation. (2) Instruments should contain items assessing symptoms that are less likely to be affected by dementia symptoms (3) instruments where possible should be

rated by both patients and informants and (4) instruments should have strong psychometric properties.

There is sufficient evidence to indicate the use of self-report assessments of anxiety in individuals with mild-moderate dementia. It is difficult to conclude which measure is the most suitable given no instrument shows overall *good* or *excellent* psychometric properties. Face validity is a key factor in choosing a measure and this should be taken into consideration. Evidence from instruments that rate informant reports show that there may be important behavioural aspects of anxiety in dementia which can be under-reported by patients, depending on the degree of cognitive impairment. Where possible, informant reports should also be considered. Anxiety has been identified as a predictor of worse outcomes for individuals with dementia, therefore screening and subsequent treatment is necessary.

Conclusion

There is a lack of high quality, high powered factor analytic studies in this area. Stott et al. (2016) is the only example of a comprehensive factor analysis study with a sufficient sample size. Validation of existing instruments with a self-report element for individuals with dementia should be prioritised to improve utility, detection and treatment as well as content validity with larger sample sizes. In conclusion, instruments which screen for anxiety in dementia have the potential to be an important means for detection and subsequent treatment of anxiety in this population. Instruments have relied heavily on content analysis based on studies of older adults and long-term care residents. Qualitative methods such as semi-structured interviews and focus groups are required to involve individuals with dementia and their carers'

REFERENCES

Alexopoulos, G.S., Abrams, R.C., Young, R.C. and Shamoian, C.A. (1988a). Cornell scale for depression in dementia. *Biological Psychiatry, 23*, 271–284.

American Psychiatric Association (1980) *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, D.C.: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Auer, S. R., Monteiro, I. M., & Reisberg, B. (1996). The Empirical Behavioral Pathology in Alzheimer's Disease (E-BEHAVE-AD) Rating Scale. International Psychogeriatrics, 8, 247-266.

Ballard, C., Boyle, A., & Bowler, C. (1996). Anxiety disorders in dementia sufferers. *International Journal of Geriatric Psychiatry, 11*, 987−990.

Ballard, C., Neill, D., O'Brien, J., McKeith, I.G., Ince, P., & Perry, R. (2000). Anxiety, depression and psychosis in vascular dementia: prevalence and associations. *Journal of Affective Disorders*, *59*(2), 97-106.

Barnes, J., Whitwell, J.L., Frost, C., Josephs, K., Rosser, M. & Fox, N.C. (2006). Measurement of the amygdala and hippocampus in pathologically confirmed Alzheimer disease and frontotemporal lobar degeneration. Archives of Neurology, 63, 1434-1439.

Beaudreau, S. A., & O'Hara, R. (2008). Late-life anxiety and cognitive impairment: A review. *The American Journal of Geriatric Psychiatry, 16*(10), 790-803.

Beekman, A.T.F., de Beurs, E., van Balkom, A., Deeg, D., van Dych, R., & van Tilburg, W. (2000) Anxiety and Depression in Later Life: Co-occurrence and Communality of Risk Factors. *Am J Psychiatry, 157*, 89-95.

52

Boddice, G., Pachana, N.A., & Byrne, G.J. (2008). The clinical utility of the geriatric

    anxiety inventory in older adults with cognitive impairment. *Nursing Older*

    People, 20, 36.

Bradford, A., Brenes, G.A., Robinson, R.A., Wilson, N., Snow, A.L., Kunik, M.E.,

    & Amspoker, A. B. (2013). Concordance of self- and proxy-rated worry and

    anxiety symptoms in older adults with dementia. *Journal of Anxiety*

    Disorders, 27, 125-130.

Brown, T.A., Campbell, L.A., Lehman, C.L., Grisham, J.R., & Mancill, R.B. (2001).

    Current and Lifetime Comorbidity of the DSM-IV Anxiety and Mood

    Disorders in a Large Clinical Sample. *Journal of Abnormal Psychology, 110*,

    585-599.

Burke, W.J., Roccaforte, W.H., Wengel, S.P., McArthur-Miller, D., Folks, D.G.&

    Potter, J.F. (1998). Disagreement in the reporting of depressive symptoms

    between patients with dementia of the Alzheimer type and their collateral

    sources. *American Journal of Geriatric Psychiatry, 6*, 308-319.

Byrne, G.J., Pachana, N.A., Arnold, L., Chalk, J.B., & Appadurai, K. (2008).

    Performance characteristics of the geriatric anxiety inventory in memory

    clinic attendees. Alzheimer's & Dementia: *The Journal of Alzheimer's*

    Association, 4, 441-442.

Byrne, G.J., & Pachana, N.A. F. (2011). Development and validation of a short form

    of the Geriatric Anxiety Inventory-the GAI-S. *International Psychogeriatrics,*

    23, 125-131.

Cohen-Mansfield, J., Marx, M.S., Rosenthal, A. S. (1989). A description of agitation

    in a nursing home. *Journal of Gerontology, 44*, 77-84.

Clark, P.A., Tucke, S.S. & Whitlatch, C.J. (2008). Consistency of information from

    persons with dementia: An analysis of differences by question type.

    Dementia: The International Journal of Social Research and Practice, 7,

    341-358.

Cummings, J.L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D.A., &

    Gornbein, J. (1994). The neuropsychiatric inventory: Comprehensive

    assessment of psychopathology in dementia. Neurology, 44, 2308-2314.

Dawson, N.T., Powers, S., Krestar, M., Yarry, J., & Judge, K.S. (2012). Predictors of

    Self-Reported Psychosocial Outcomes in Individuals with Dementia. *The*

    Gerontologist, 53, 748-759.

De Vet, H.C.W., Terwee, C.B., Mokkink, L.B., & Knol, D.L. (2011). *Measurement*

    in Medicine. New York: Cambridge University Press.

Department of Health (2013). *Dementia: A state of the nation report on dementia*

    care and support in

    *England.*https://www.gov.uk/government/uploads/system/uploads/attachment

    _data/file/262139/Dementia.pdf

Diefenbach, G.J., Bragdon, L.B., & Blank, K. (2014). Geriatric anxiety inventory:

    factor structure and associations with cognitive status. *American Journal of*

    Geriatric Psychiatry, 22, 1418-1426.

Dunbar, M., Ford, G., Hunt,K., Der, G. (2000). A confirmatory factor analysis of the
Hospital Anxiety and Depression scale: Comparing empirically and
theoretically derived structures. *British J Clin Psychol, 39,* 79-94.

Feinstein, J. S., Duff, M. C., & Tranel, D. (2010). Sustained experience of emotion
after loss of memory in patients with amnesia. *Proceedings of the National
Academy of Sciences of the United States of America*, *107*, 7674–7679.

Free, S.M.,& Overall, J.E. (1977). The Brief Outpatient Psychopathology Scale
(BOPS). *Journal of the American Geriatrics Society, 33*, 677-688.

Gerolimatos, L.A., Gregg, J.J.,& Edelstein, B.A. (2013). Assessment of anxiety in
long-term care: examination of the Geriatric Anxiety Inventory (GAI) and its
short form. *International Psychogeriatrics, 25*, 1533-1542.

Gerolimatos, L.A., Ciliberti, C.M., Gregg, J.L.,Nazem, S., Bamonti, P.M., Cavanagh,
C.E.,& Edelstein, B.A. (2015). Development and preliminary evaluation of
the anxiety in cognitive impairment and dementia (ACID) scales.
International Psychogeriatrics, 27, 1825-1838.

Gibbons, L.E., Teri, L., Logsdon, R.& McCurry, S.M., Kukull, W., Bowen,
J…Larson, E. (2002). Anxiety symptoms as predictors of nursing home
placement in patients with Alzheimer's disease. Journal of Clinical
Geropsychology, 8, 335-342.

Gibbons, L.E., Teri, L., Logsdon, R.& McCurry, S.M. (2006).Assessment of anxiety
in dementia: An investigation into the associateion of different methods of
measurement. *Journal of Geriatric Psychiatry and Neurology, 19,* 202-208.

Heidenblut, S.& Zank, S. (2014). Screening for Depression with the Depression in

Old Age SDScale (DIA-S) and the Geriatric Depression Scale (GDS-15). *GeroPsych, 27*, 41-49.

Hopko, D.R., Stanley, M.A., Reas, D.L., Wetherell, J.L., Beck, J.G., Novy, D.M., & Averill, P.M. (2003). Assessing worry in older adults: Confirmatory factor analysis of the Penn State Worry Questionnaire and psychometric properties for an abbreviated model. *Psychological Assessment, 15*, 173-183.

Hwang, T.J. Masterman, D.L., Oritz, F., Fairbanks, L.A., & Cummings, J.L. (2004). Mild cognitive impairment is associated with characteristic neuropsychiatric symptoms. *Alzheimer Disease and Associated Disorders, 18*, 17-21.

Karlawish, J.H.T., Casarett, D., Klocinski, J., & Clark, C. (2001). The relationship between caregivers' global ratings of Alzheimer's disease patients' quality of life, disease severity, and caregiving experience. *American Geriatrics Society, 49,* 1066-1070.

Kolanowski, A., Hoffman, L. & Hofer, S.M. (2007*). Journal of Geroontology: Psychological Sciences, 62,* 20-27.

La Barge, E. (1993). A preliminary scale to measure the degree of worry among mildly demented alzheimer disease patients. Physical & Occupational Therapy in Geriatrics, 11, 43-57.

Lai, C. K. Y. (2014). The merits and problems of Neuropsychiatric Inventory as an assessment tool in people with dementia and other neurological disorders. *Clinical Interventions in Aging*, *9*, 1051–61.

Mansbach, W. E., Mae, R. A. & Clark, K. M. (2014). The Brief Anxiety and

Depression Scale (BADS): A new instrument for detecting anxiety and depression in long-term care residents. *International Psychogeriatrics, 27,*673-681.

McClive-Reed, K. P.& Gellis, Z.D. (2011). Anxiety and related symptoms in older persons with dementia: Directions for practice. *J Gerontol Soc Work, 54*, 6-28.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P. W., Knol, D.L…De Vet, H.C.W. (2010). The COSMIN study reached international consensus on taxonomy, terminology and, definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*, 737-745.

Meyer, T.J., Miller, M.L.,Metzger, R.L.,& Borkovec, T.D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*, 487-495.

Neville, C.& Teri, L. (2011). Anxiety, anxiety symptoms, and associations among older people with dementia in asisted-living facilities. *International Journal of Mental Health Nursing, 20*, 195-201

Park, T., Reilly-Spong & Gross, C.R. (2013). Mindfulness: a systematic review of instruments to measure emergent patient-reported outcome (PRO*). Quality of Life Research, 22*, 2639-2659.

Pachana, N.A., Byrne, G.J., Siddle, H., Koloski, N., Harley, E., & Arnold, E. (2007). Development and validation of the Geriatric Anxiety Inventory. *International*

Psychogeriatrics, 19, 103-114.

Pearson, J. L., Teri, L., Wagner, A,, Truax, P., & Logsdon, K. G. (1993). The
relationship of problem behaviors in dementia patients: to the depression and
burden of caregiving spouses. *The American Journal of Alzeimer's Disease
and Related Disorders and Research, 8*, 15-22

Prince, M. Knapp, P. Guerchet, M. McCrone, P.; Prina, M., Comas-Herrera M.
Wittenberg, A., Adelaja, R., Hu, B., King, D. Rehill, D. and Salimkumar, D.
(2014) Dementia UK: update *Alzheimer's Society 2014.*

Razavi, D., Delvaux, N., Farvacques, C., Robaye, E. (1990). Screening for
adjustment disorders and major depressive disorders in cancer in-patients. *Br
J Psychiatry, 156*, 79-83.

Ready, R.E., O Carvalho, J.O., Green, R.C., Gavett, B.E., & Stern, R.A. (2011). The
structure and validity of self-reported affect in mild cognitive impairment and
mild Alzheimer's disease. *International Psychogeriatrics, 23*, 887-898.

Reisberg, B., Auer, S.R., & Monteiro, I.M. (1996). Behavioral pathology in
Alzheimer's disease (BEHAVE-AD) rating scale. *International
Psychogeriatrics, 8,* 301-308.

Reisberg, A.B., Monteiro, I., Torossian, C., Auer, S., Shulman, M.B., Ghimire,
S.,…Xu, J. (2014). The BEHAVE-AD assessment system: A perspective, a
commentary on new findings and a historical review. *Dementia and Geriatric
Cognitive Disorders, 38*, 89-146.

Schultz, S.K., Pharmd, A.H. & Buckwalter (2004). Anxiety and impaired social
function in the elderly. *Annals of Clinical Psychiatry, 16*, 47-51.

Shankar, K.K, Walker, M., Frost, D., & Orrell. (1999). The development of a valid and reliable scale for rating anxiety in dementia (RAID). *Aging & Mental Health, 3*, 38-49.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janays, J., Weiller, E., & Dunbar, G.C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry,* 59, 22–33.

Seignourel, P.J., Kunik, M.E., Snow, L., Wilson, N. & Stanley, M. (2008) Anxiety in dementia: A critical review. *Clinical Psychology Review, 18*, 1071-1082.

Sinoff, G., Ore, L., Zlotogorsky, D., Tamir, A (2002). Does the presence of anxiety affect the validity of a screening test for depression in the elderly? *International Journal of Geriatric Psychiatry. 16*, 1071-1077.

Sinoff, G., & Werner, P. (2003). Anxiety disorder and accompanying subjective memory loss in the elderly as a predictor of future cognitive decline. *International Journal of Geriatric Psychiatry*, *18*, 951–959.

Speilberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y).* Palo Alto, CA: Consulting Psychologists Press.

Spitzer, R.L., Kroenke, K., Williams, J.B.W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Arch Intern Medicine, 166*, 1092-1097.

Snow, L.A., Kunik, M.E., Molinari, V.A., Orengo, C.A., Doody, R., Graham, D.P., & Norris, M. (2005). Accuracy of self-reported depression in persons with of

a structured interview guide for the Rating for Anxiety in Dementia (RAID-SI). *The American Geriatrics Society, 53,* 389-396.

Snow,L.A., Huddleston, C., Robinson, C., Kunik, M.E., Bush, A.L., Wilson, N…Stanley, M.A. (2012). Psychmetric properties *Aging Mental Health, 16*, 592-602.

Starkstein, S.E., Jorge, R., Petracca, G. & Robinson, R.G. (2007). The construct of generalized anxiety disorder in Alzheimer disease. *Am J Geriatr Psychiatry, 15*, 42-49.

Steinberg, M., Shao, H., Zandi, P., Lyketsos, C.G., Welsh-Bonner, K.A., Norton, M.C…Tschanz, J.T. (2008). Point and 5-year period prevalence of neuropsychiatric symptoms in dementia: The Cache County Study. *International Journal of Geriatric Psychiatry, 23*, 170-177.

Stott, J., Spector, A., Orrell, Scior, K., Sweeney, & Charlesworth (2016). Limited validity of the Hospital Anxiety and Depression Scale (HADS) in dementia: evidence from a confirmatory factor analysis. *International Journal of Geriatric Psychiatry, 32*, 805-813.

Tan, L.L., Wong, H.B., & Allen, H. The impact of neuropsychiatric symptoms of dementia on distress in family and professional caregivers in Singapore (2005). *International Psychogeriatrics, 17*, 253-263.

Teri, L., & Traux, P. Assessment of depression in dementia patients: Association of caregiver mood with depression ratings. *Gerontologist, 34*, 231-234.

Terwee, C.B., Bot, S.D.M., de Boer, M.R., van der Windt, D., Knol, D.L., Dekker…de Vet, H.C.W. (2007). Quality criteria were proposed for

measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60,* 34-42.

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & De Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research, 21*, 651–657.

Twelftree, H. & Qazi, A. (2006). Relationship between anxiety and agitation in dementia. *Aging and Mental Health, 10*, 262-367.

Van Tulder, M. Furlan, A., Bombardier, C., Boutter, L., & Editorial Board of the Cochrane Collaboration Back Review Group. (2003). Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine, 28*, 1290-1299.

Ward, L.C., Wadsworth, A.P., & Peterson, L.P. (1994). Concurrent validity of measures of anxiety, depression, and somatization in elderly, demented, male patients. *Clinical Gerontologist, 15*, 3-13.

Westhuis, D., &Thyer, B.A. (1989).Development and validation of the clinical anxiety scale: a rapid assessment instrument for empirical practice. *Educational and Psychological Measurement, 49*, 153-163.

World Alzheimer's Report. (2016). *Improving healthcare for people living with dementia: Coverage, quality and costs now and in the future.* Retrieved from: https://www.alz.co.uk/research/world-report-2016

Zigmond, A.S., & Snaith, R.P. (1983). The hospital anxiety and depression scale. *Acta Psychiatr Scand, 67*, 361-370.

PART 2: EMPIRICAL PAPER

The Clinical Utility of the Four Mountains Test in the Diagnosis of Dementia:

Relationship to Hippocampal and Medial Temporal Lobe Atrophy

ABSTRACT

Aims: The current study aimed to investigate the clinical utility of a spatial memory test (The Four Mountains Test; 4MT) in the determining dementia subtype. In a previous study, the 4MT did not find significant differences between Alzheimer's (AD) and non-Alzheimer's dementia. Structural MRI data was analysed to investigate if neurodegeneration in the hippocampus and medial temporal lobe structures accounted for these non-significant differences**. Method**: Data were extracted for fifteen memory clinic patients with dementia diagnoses (AD, vascular and mixed dementia). This included structural MRI scans, 4MT scores, and other neuropsychological measures. Freesurfer image analysis suite was used for automated analysis of the critical neuronal structures involved in AD. The relationships between volumetry, 4MT performance, and cognitive abilities was compared across diagnostic groups. **Results**: Contrary to prediction, there were no positive associations between these variables. It was not possible to conduct statistical analysis to compare the AD, VaD, and the mixed dementia groups due to the restricted size of the final sample. **Conclusions** The clinical utility of the 4MT in distinguishing AD from other dementia types in heterogeneous groups of memory clinic patients has not been established. It is possible that overlapping patterns of atrophy and other cognitive impairments associated with dementia have a confounding impact on 4MT performance. However, the current study has low power and a lack of inferential statistics and as such the findings are tentative and must be interpreted with caution. However, the introduction of structural MRI data has contributed to further understanding of a tool that may support in the early diagnosis of dementia.

INTRODUCTION

Dementia

It is estimated that approximately 46.8 million people are living with dementia worldwide and this figure is expected to rise to 131.5 million in 2050, if age related prevalence remains as it is (World Alzheimer's Report, 2016). In the UK, it is estimated that there are 850,000 people living with dementia at a total cost of £26 billion per year. This figure incorporates NHS, social care funding, and the cost burden on individuals with dementia and their families and is expected to rise to £55 billion by 2040 (Prince et al., 2014). Dementia is a syndrome characterized by deterioration in cognitive abilities beyond that which is expected in normal ageing. The symptoms include deficits in memory, thinking, orientation, comprehension, calculation, learning capacity, language and judgement (WHO, 2015). A number of dementia subtypes are now categorized in *The Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM–5*; American Psychiatric Association, 2013).

Alzheimer's disease

Alzheimer's disease (AD) is the most common dementia accounting for approximately 55-70% of cases (McKeel, Burns, Meuser & Morris, 2007). The primary symptom is memory loss with mild impairments in language and praxis evident at the early stage of disease progression. (McKeel et al., 2007). This is followed by progressive neocortical damage over the course of the disease (Wood et al., 2016). The AD brain is characterised by the build-up of amyloid plaques made up of beta proteins, and neurofibrillary tangles made up of the tau proteins. The accumulation of these proteins causes the degeneration of neurons (Bloom, 2014).

While all AD diagnoses are considered probable until confirmed by autopsy (Agamanolis, 2014), these AD biomarkers can be detected via amyloid PET scanning or CSF studies.

Vascular Dementia

Vascular dementia (VaD) is the second most common subtype and encompasses all instances of dementia associated with ischemic cardiovascular disease and haemorrhagic and hypoxic-ischemic cerebral lesions. (Roman et al., 1993). It is characterised by "an acute onset, stepwise decline, focal neurological signs, gait impairment and urinary difficulties" (Camicioli, 2006, pp.4). Some patients with AD develop symptoms of VaD over the course of the disease, usually after a stroke, which may result in a sudden worsening of dementia (Roman et al. 1993).

Mixed dementia

Dementia diagnosis is further complicated by the presence of mixed dementia where abnormal protein deposits associated with AD coexist with blood vessel problems linked to VaD (Langa, Foster & Larson, 2004). The symptoms may vary dependent on the brain region affected and may be like those of either AD or VaD. Research suggests that prevalence rates vary markedly among neuro-pathological studies from 0% to 55%, however much of the variance is related to recruitment biases, geographic factors and conceptual differences. True prevalence most likely approximates 20% to 40% of dementia cases (Zekry, Hauw & Gold, 2002).

Mild Cognitive Impairment

Mild cognitive impairment (MCI) is a trajectory marked by a decline in cognitive function beyond that associated with typical aging (Peterson, 2011). The DSM-5

recognizes the importance of diagnosing this level of cognitive decline, where individuals may benefit from adaptations and strategies to minimise the impact on daily living skills. MRI data suggest that hippocampal atrophy in amnestic MCI that falls below the 25th percentile predicts risk of progression to a dementia diagnosis over a two-year period (Clifford et al. 2010).

Other dementias

Fronto-temporal dementia (FTD) involves degeneration of the frontal lobes resulting in progressive changes in behaviour, personality, executive function, or language. There are different forms of FTD, e.g. Pick's disease; characterized by pathology and early onset and behavioral variant FTD (BVFTD); characterized by progressive changes in personality, emotional blunting, and/or loss of empathy. Dementia with Lewy bodies includes a range of cognitive deficits, including visual hallucinations. Huntington's disease is an inherited dementia with features of mood, cognitive and gait/co-ordination deficits.

Alzheimer's disease and topographical disorientation

One of the earliest clinical features of AD is topographical disorientation. Serino, Cipreso, Morganti, & Riva (2014) discussed the various ways this can manifest. There can be difficulties using primary environmental features for orientation (landmark agnosia), integration of object location with respect to self (egocentric disorientation), memory for heading direction with respect to objects in the environment (heading direction), and laying down new representations of the environment in memory (anterograde disorientation). Topographical orientation is dependent on encoding, storage and retrieval of spatial information.

Spatial Memory and the Hippocampus

The hippocampus plays a pivotal role in spatial memory and supports encoding of the context in which events occur (Ritchie et al., 2017). This is achieved via the storage of an internal memory model of our surroundings or a 'cognitive map'. The cognitive map theory of hippocampal functioning developed out of animal experiments whereby place-related neuronal firing was observed in the hippocampal cells of freely moving rats at particular locations (O'Keefe and Dostrovsky, 1971) and was subsequently extended to humans (O'Keefe & Nadel, 1978; Ekstrom et al., 2003). Two different reference frames for organising spatial information have been defined; egocentric and allocentric (Klatzky, 1998). Egocentric representations refer to memory for locations by drawing on information about the location of objects in relation to the self (self-centered). Allocentric representations refer to memory for object-to-object relations or relations between environmental characteristics and are unrelated to the individual's orientation or viewpoint (world-centered). There is evidence to suggest that the integration of ego and allo-centric processing support the long-term encoding of spatial context and navigation of real world environments (Burgess, Trinkler, King, Kennedy & Cipolotti, 2006; Vann, 2009; Land, 2014).

There is a prevalence of allocentric deficits among patients with MCI and AD with evidence suggesting that hippocampal and medial temporal lobe (MTL) degeneration are associated with the ability to maintain long-term allocentric representations of surrounding environments (Serino et al. 2014). King, Burgess, Hartley, Vargha-Khadem & O'Keefe (2002) used a virtual reality paradigm to test memory for object locations where the participants viewed objects from both static and shifted viewpoints. It was demonstrated that hippocampal damage impaired the ability to successfully retrieve memory for object location in the shifted viewpoint

67

condition. Results from studies support the conclusion that the hippocampus is necessary for flexible allocentric memory for object locations (Maguire & Cipolotti, 1998; Kalova et al. 2005; Burgees et al. 2006). In addition to the hippocampal role in allocentric processing, functional neuroimaging studies have shown that multiple brain regions are activated in spatial navigation, including the left hippocampus, posterior cingulate gyrus, precuneus, parahippocampi via retrosplenial cortex, and the parieto-occipital sulcus, (Burgess, Maguire, Spiers, & O'Keefe, 2001; Iaria, Chen, Guariglia, Ptito, & Petrides, 2007). These regions are likely involved in transformations between egocentric and allocentric representations, in both directions.

Allocentric representations and Alzheimer's disease

The MTL network activated during topographical memory tasks overlap with the pattern of neural degeneration in the early stages of AD (Pengas et al., 2012). While there is involvement of both egocentric and allocentric representations in spatial memory function, there is a prevalence of allocentric impairments associated with early AD atrophy in the MTL region (Burgees et al., 2006; Maguire & Cipolotti, 1998). The Four Mountains Test (4MT) is a spatial memory test developed by Hartley et al. (2007), designed to tap the function of the human hippocampus in topographical processing. Participants are presented with computer-generated landscapes displaying four mountains in different configurations and asked to match one of the four images to a sample image. The test involves a spatial task where the target scene is presented followed by presentation of the same scene rotated to give a different view alongside three foil landscapes. This assesses the ability to recognize places from their layout even when the viewpoint changes (see Figure 1; Chan et al., 2016). A non-spatial task is also given, where participants match to sample based on

non-spatial elements e.g. cloud cover, lighting, and colour of vegetation. To assess both spatial perception and memory, the match-to-sample task is administered immediately after presentation of the target image and again after a delay. Hartley et al. (2007) found that participants with focal hippocampal damage had significant difficulty with spatial memory but showed spared abilities in spatial perception, non-spatial perception, and non-spatial memory.

Figure 1: The Four Mountains Test (4MT)
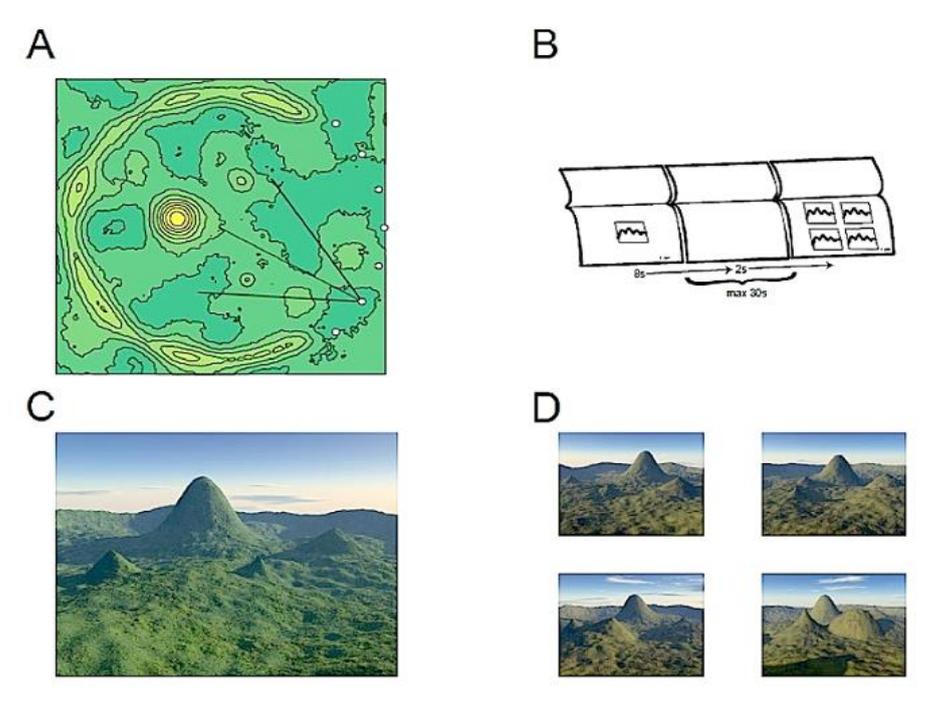
(A) All 4MT stimuli are based on computer-generated heightfields containing 4 mountains as illustrated by a sample contour map **(see A)**. Each landscape is made up of similar topographical features: the ground plane with small scale undulations, a semi-circular mountain range (defining the horizon in each image), and 4 prominent mountains of varying shapes and sizes. An example is shown as a contour map in. Images are rendered using a virtual camera placed at one of the indicated 7 locations. **(B)** Participants see a sample image which they study before seeing four different images (one target showing the same place from a different viewpoint, and 3 foils showing different places). Their task is to identify the target. **(C)** Example of a sample image. **(D)** Corresponding target and foil images (the target is seen top-left). Note that all images are shown at the same scale in the test, and that viewpoint and other non-spatial features are systematically varied between sample and test images (Chan et al., 2016, p 2)

Spatial memory testing and distinguishing dementia type

Spatial memory testing has demonstrated the ability to discriminate between different dementia subtypes, e.g., the 4MT has been demonstrated to discriminate between AD and non-AD dementias (Hartley et al., 2007; Pengas et al., 2012). Topographical memory has been shown to be preserved in FTD (Bird et al., 2010) and in semantic dementia (Pengas et al., 2010). A virtual supermarket task of spatial orientation was able to discriminate between AD and bvFTD, independent of episodic memory performance (Tu et al., 2015). Impaired orientation was associated with the integrity of the parietal rather than temporal lobes indicating that this novel task tapped the egocentric framework. These results suggest that neuropsychological testing of spatial memory may provide diagnostic specificity and indicates a possible utility for routine testing of these abilities in clinical assessments.

There is evidence to suggest that the 4MT has clinical utility in determining whether patients with MCI will develop AD in the future. Patients with amnestic MCI, thought to be in a prodromal AD stage showed similar performance on the 4MT as those with a diagnosis of AD, indicating utility of the 4MT in detecting the presence of prodromal AD (Bird et al., 2010). Performance on the 4MT was incorporated with neuroimaging results to demonstrate that 4MT scores of <8 were associated with 100% sensitivity and 90% specificity for detection of MCI patients with AD biomarker status and those without (Moodley et al., 2015). In addition, 4MT performance correlated with hippocampal volume and cortical thickness of the precuneus and posterior cingulate gyrus, which is consistent with the role of these regions both in spatial memory and the brain pathology of early AD. Most recently, the 4MT predicted conversion from MCI to AD with a 93% accuracy in a group of 15 patients followed up over a 24-month period (Wood et al., 2016), which is

comparable to the predictive power of more invasive methods such as CSF studies of amyloid and tau.

The utility of the 4MT task may be less clear in a real-world diagnostic setting. Gore (2015) examined the clinical utility of the 4MT in differentiating between dementia type, including AD, VaD, mixed dementia, and MCI in participants recruited in a memory service. Unexpectedly, there was no significant difference in 4MT scores between the AD group and other dementia types, a finding that is not in line with previous studies, (Hartley et al., 2007; Pengas et al., 2010; Moodley et al., 2015). In addition, other neuropsychological measures, including memory performance, did not demonstrate significant differences between the AD and the VaD groups as hypothesized.

Current Study

Research has identified that hippocampal atrophy in AD is associated with the ability to represent and remember allocentric spatial information (Maguire et al., 1998; Burgess, et al., 2006; Hartley et al. 2007). Initial evidence for the clinical utility of the 4MT in distinguishing between different types of dementia has been demonstrated (Hartley et al., 2007; Pengas et al., 2012), particularly in predicting disease progression from MCI to early AD (Bird et al. 2010; Moodley et al., 2015; Wood et al. 2016). However, these results have not been replicated in more heterogeneous groups of dementia patients, for example, when individuals with VaD and mixed dementia are included (Gore, 2015). An explanation for this discrepancy is the possible overlap in hippocampal and MTL atrophy in AD, VaD, and mixed dementia. In AD there is a stereotypical pattern of MTL degeneration (entorhinal and hippocampal) and neuroimaging can support in differential diagnosis (Frisoni, Fox,

72

Jack, Scheltens, & Thompson, 2010). Investigating the relationship between 4MT and MTL atrophy typical of AD has the potential to increase understanding of the utility of the 4MT in diagnosis and subsequent treatment of dementia. Furthermore, previous research suggests that other brain regions surrounding the hippocampus are implicated in spatial memory processes e.g. the volume of the entorhinal and parahippocampal regions predicted familiarity memory in older adults, MCI, and early AD patients (Wolk, Dunfee, Dickerson, Aizenstein, & DeKosky, 2011). Entorhinal volume has been shown to be associated with recognition of the spatial environment based on familiarity (Yonelinas et al., 2007).

There are a number of cognitive abilities required to perform the 4MT including, visuo-spatial, language, memory, verbal fluency, executive functioning, and pre-morbid functioning abilities. Deficits in these abilities have the potential to confound 4MT performance. The neuropsychological measures used to assess these abilities are described in the method section.

Aims

Some studies have found an association between 4MT and volumetric changes in specific brain regions (Hartley et al., 2007; Pengas et al., 2012) and found that scores on this test are predictive of conversion from MCI to AD, but when investigated in a mixed dementia sample (Gore, 2015), it was not found to discriminate. The current study aims to explore the volumetric changes to search for an explanation for this inconsistency. A volumetric analysis of critical brain structures associated with 4MT performance was conducted. The regions of interest (ROI's) investigated were the volume of the hippocampus, entorhinal cortex, and parrahippocampal gyrus. The rationale for exploring these regions comes from functional neuroimaging studies

73

which implicate these structures in spatial navigation tasks (Burgess, Maguire, Spiers, & O'Keefe, 2001; Iaria, Chen, Guariglia, Ptito, & Petrides, 2007). The thickness of the precuneus and posterior cingulate gyrus were also investigated as per results from Moodley et al. (2015), where correlations between the volumerty of these regions and 4MT performance were demonstrated. Furthermore, the current study replicated the methodology of Moodley et al. (2015) using a more heterogeneous clinical sample; firstly, by including participants with VaD and mixed dementia and secondly by using a sample with an older average age (78.80 years) compared with Moodley's UK study average age (65.63 years).

Hypotheses

1) The 4M task was designed to specifically depend on the function of the network of ROIs stated in the aims above. Thus, we predict positive associations between the volume and thickness measurements of the ROIs and 4MT scores across the sample.

2) It is predicted based on the Gore (2015) finding of no difference on the 4MT task between the different dementia subtypes, that in the current sample there will be similar levels of volumetric atrophy across these subtypes. Differences in the volumes of the critical structures in the main diagnostic categories (AD, VaD, and mixed dementia) will be investigated.

3) Evidence in this field is limited to "pure" AD and MCI samples rather than mixed samples from diagnostic services (e.g. Hartley et al, 2007; Moodley et al. 2015). An investigation of associations between volume and thickness measurements and the neuropsychological measures available may clarify our understanding of the relationships between dementia subtype, volumetry, and psychometric findings.

METHOD

Design

A cross-sectional observational design was used to explore how brain volumes and thickness of the ROI's related to 4MT scores and to performance on other cognitive measures.

Setting

Scans were available for participants across two NHS Memory Services and their associated Dementia Advisor services provided by the Age Concern Charity. These services were based in West London. Research appointments were carried out in participant's homes or at the memory service. The cognitive measures used in the current study were administered as part of a study by Gore (2015). Permission to use the data was obtained from the primary researcher (see Appendix 1). The participants involved were also part of a wider dementia study about accessibility of Cognitive Behavioural Therapy (CBT) for older people with dementia. The measures described in the current study were used as a means of assessing the cognitive abilities required to benefit from CBT treatment. The first published study from this wider project was published by Stott, Scior, Mandy, & Charlesworth (2017).

Participants

Inclusion/Exclusion Criteria

All participants who had been referred to memory services or were involved with the dementia advisor services were initially considered eligible for the study. Participants invited to participate had met the following criteria:

- Fluent in English language and did not require use of an interpreter.

- Aged 50 years or over.

- Scored above 70 on the Addenbrooke's Cognitive Examination- III at the initial assessment (ACE-III).

- No current significant mood or anxiety disorders, psychotic symptoms, substance misuse problems or a premorbid learning disability.

- No sensory difficulties that would interfere with completion of neuropsychological measures i.e. problems with sight.

- Deemed to have capacity to consent to take part in the study.

The sample included patients with a range of memory difficulties and subsequent dementia diagnoses were heterogeneous. Scores below 70 on the ACE-III tend to be indicative a more moderate to severe dementia. More global deficits in functioning can be expected regardless of dementia type. Capacity to consent in accordance with the Mental Capacity Act (2005) was also re-assessed at the research appointment. Eligibility for the current study was determined by the availability of an MRI brain scan.

Diagnostic assessment

ICD 10 criteria were used to make dementia diagnoses. A clinician from the memory service gathered information by conducting a clinical interview and administering the ACE-III as a test of cognitive abilities. In most cases an MRI scan was used as supplementary information in arriving at a diagnosis. In some cases, participants were referred for further neuropsychological testing if their presentation was more complex. Results from all assessments were discussed in a multidisciplinary team

and a diagnosis was determined. Following this, patients were informed about diagnosis by an assigned clinician.

Ethics

The work was covered by ethical approval for the original study which was granted after review by the City Road and Hampstead National Research Ethics Service Committee. An amendment to the original ethics application was submitted to confirm neuro-radiological data could be analysed under the original consent procedure and this was approved by the committee in July 2015 (see Appendix 2 for letter granting ethical approval).

Sample Size

As this study consists of a secondary data analysis, the sample had been recruited and tested in advance so using a power analysis to specify a sample size was not possible. It is now apparent that the current study is underpowered for parametric analysis due to the unexpected lack of availability of MRI data for some of the participants. While this study is underpowered for the research hypotheses, it provides preliminary data in a field of research with few prior studies. As a result, it was decided to proceed with a more descriptive approach in addition to tentative inferential analyses. The availability of a heterogeneous clinical sample warranted an exploratory analysis to help establish the likely magnitude of effects. The use of small available samples is well-established in neuropsychological research, and this study aimed to make the best possible use of difficult-to-collect data.

Measures

The neuropsychological measures were administered in the original study. This data was used to characterise the sample and examine relationships between brain region

volumes and thickness, 4MT, and other cognitive abilities. The researcher who administered the neuropsychological tests was trained and experienced in this field. Demographic information was also collected both during the testing and retrospectively using the NHS electronic patient database. The orders of the tests in the battery were randomized using Qualtrics, an online survey system designed for administering research protocols. This enabled the researchers to allow for the potential impact of fatigue and carry-on effects on performance. The following measures were used to explore the research hypotheses in the current study:

The Test of Premorbid Functioning **(TOPF)** (Wechsler, 2011) was administered to assess pre-morbid intellectual ability prior to the onset of dementia. The TOPF consists of a list of seventy words irregular words and takes less than 10 minutes to complete. The TOPF demonstrates good internal consistency (Cronbach's alpha = 0.95). The test-retest reliability is also good (corrected correlations between r=.89 and r=.95; Wechsler, 2011). The TOPF correlates with the Wechsler Adult Intelligence Scale- Fourth Edition (WAIS-IV) Full Scale IQ scores. The raw score, sex, years of education, and other demographic variables can be combined to calculate an estimated premorbid IQ score. The measure can be used to predict sub-scale scores on the WAIS-IV and Wechsler Memory Scale (WMS). Research has established the TOPF as a valid method for assessing change between premorbid and current cognitive functioning with a clinical dementia sample (Duff, Chelune, & Dennett, 2011). In a recent small-scale study (N=33), the TOPF underestimated pre-morbid IQ when compared with a demographic equation model (McDonald, 2015).

The Addenbrooke's Cognitive Examination-III (ACE-III; Mioshi, Dawson,
Mitchell, Arnold, & Hodges, 2006) is the updated version of the ACE-R. It assesses
the following cognitive functions: attention/orientation, memory, verbal fluency,
language and visuo-spatial abilities (see Appendix 3). Minor adaptations to ACE-R
items have been made (e.g. repetition, comprehension and visuospatial). The test is
scored out of 100 and higher scores reflect better ability. Internal reliability of the
ACE-III, measured by Cronbach's alpha was 0.88 and there was a strong correlation
with the ACE-R (Hsieh, Schubert, Hoon, Mioshi & Hodges, 2013). The ACE-III also
shows high sensitivity and specificity with cut-off scores at 88 (sensitivity=1.0;
specificity=0.96) and 82 (sensitivity=0.93; specificity1.0) (Velayudhan et al. 2015).

Trail Making Test (TMT) (Reitan, 1992) is a measure of visual search, scanning,
speed of processing, mental flexibility, and executive functions (Tombaugh, 2004). It
is a pen-and-paper task containing two parts; the task of Part A is to connect a set of
twenty-five numbers in consecutive order as fast as possible while maintaining
accuracy. This task measures visual search and motor skill speed. In Part B,
participants alternate between a series of numbers and letters, maintaining
consecutive order of each set. This task measures mental flexibility (Bowie &
Harvey, 2006). Test-retest reliability is reported at r=0.80 (Spreen & Strauss, 1991)
and validity r= 0.59 (Delis, Kaplan, & Kramer, 2001). The TMT is susceptible to the
effects of the normal ageing process as demonstrated by increased time taken to
complete Parts A and B in the absence of sensory or motor deficits (Wahlin,
Bäckman, Wahlin, & Winblad, 1996). When comparing MCI and AD patients to
controls, errors demonstrated a greater sensitivity to diagnostic category than to the

presence of an impaired time to completion score (Ashendorf et al., 2008). This suggests that Parts A and B can be considered separately.

The 4MT is a spatial memory test designed to measure hippocampal-dependent topographical memory processing in humans (see Hartley et al., 2007). The validity of the 4MT has been demonstrated by research which indicates poor performance on the 4MT in individuals with focal hippocampal damage, despite preserved abilities in spatial perception, non-spatial perception and non-spatial memory (Hartley et al. 2007). Poor performance on the 4MT has also been shown in Alzheimer's disease patients (Bird et al. 2010; Moodley et al., 2015), where hippocampal atrophy is a defining feature of this dementia subtype. Formal psychometric properties of the 4MT have not yet been established as it is an experimental instrument that has only been used in a series of laboratory-based research studies.

Participants completed the topographical memory subtest of the 4MT (see Hartley et al. 2007 for further details). The aim on this subtest is to retain the topographical layout information of a computer-generated landscape and to identify this target image among a selection of 4 images (3 foils and 1 target image). A non-spatial task is also given, where participants match-to-sample based on non-spatial elements e.g. cloud cover, lighting, and colour of vegetation. Participants were presented with a target image on a computer screen for 10 seconds and then a blank screen for approximately 2 seconds. On the next computer slide, 4 alternative landscape scenes were presented that were arranged randomly in a 2 by 2 grid. Participants had twenty seconds to select the correct target image. Answers were recorded independently by

the participant using a grid sheet. The test was administered on a laptop but the researcher controlled the laptop and the timings of the images. This was to reduce the need for participants to interact with the computer interface and to minimize potential confounds.

Procedure

The Freesurfer image analysis suite was used to analyse T1 structural images. This software is documented and freely available for download online: (http://surfer.nmr.mgh.harvard.edu/). Automated analysis was used to extract left and right volumes of the hippocampi, parrahippocami and entorhinal cortices and the thickness measurements of the posterior cingulate gyri and precunei.

Cortical Reconstruction

The first stage of this this process performed all of the cortical reconstruction. This involved motion correction and averaging of the T1 structural images (Reuter et al. 2010). There may be variations in anatomical intensity induced by artefacts such as the radio-frequency coil, the acquisition pulse sequence, and by the nature and geometry of the sample itself (Belaroussi, Milles, Carme, Zhu, & Benoit-Cattin, 2006). It also involved removal of non-brain tissue (Segonne et al., 2004), for instance removing the skull from the MRI images. This was followed by automated Talairach transformation and segmentation of the subcortical white matter and deep grey matter volumetric structures (including hippocampus, amygdala, caudate, putamen, ventricles) (Fischl et al., 2002; Fischl et al., 2004a). Talairach transformations refer to a map of human brain structures (known as an 'atlas'). These are used to map the location of brain structures independent from individual

81

differences in the size and overall shape of the brain (Lancaster et al., 2000). The automatic subcortical segmentation involves several stages of processing (see http://surfer.nmr.mgh.harvard.edu/ for further details). (1) CGA linear registration: aligning the initial registration to a template. (2) Canonical Normalization: a further normalization process. (3) Canonical Registration: computes a nonlinear transform to align with the atlas. (4) Neck removal: differentiating and removing the neck from the MRI volume. (5) Registration with skull: computes transform to align MRI volume with the atlas volume possessing the skull. (6) Subcortical labelling: labels the subcortical structures according to the atlas. Freesurfer morphometric procedures have been demonstrated to show good test-retest reliability across scanner manufacturers and across MRI field strengths (Han et al., 2006; Reuter et al., 2012). Out of the eighteen images processed during the cortical reconstruction stage, seventeen images were viable. Two scans were excluded as they did not contain enough DICOM (Digital Imaging and Communications in Medicine) images for processing.

Brain mask

When the cortical reconstruction process finishes, it is possible that errors may have occurred in distinguishing white and pial surfaces. The brain mask function displays the skull-stripped surfaces to allow for manual inspection of coronal, sagittal and horizontal slices (see figure 2 for sample image). The white surface (blue line) is used to calculate total white matter volume and should accurately follow the boundary between white matter and grey matter. The pial surface is used to calculate cortical grey matter volume and should accurately follow the boundary between the grey matter and the CSF. Images were excluded for one participant at this stage as the brain mask process failed.

82

Quality control checks

A quality control check of all the MRI images was undertaken. The Enigma instructions for visual quality control were followed (http://enigma.ini.usc.edu/protocols/imaging-protocols/protocol-for-quality-control-and-summary-statistics/). Enigma is an international consortium which has pooled the human brain images and genome-wide scans of 21,000 participants. The Enigma Consortium aims to gain a greater understanding of brain structure and function, based on MRI, DTI, fMRI, and genetic data. Examples of successful and poor labelling were obtained from the Enigma website and the study images were compared against these as a first step in the visual quality control process (see figure 3 and 4). The primary researcher (M.S.) identified six images which deviated from the successful labelled examples due to the presence of deep sulci and severe atrophy. These images were reviewed by two researchers (S.C. and J.K.) and a consensus was reached to include these six images. Through making comparisons with the Enigma examples of successful and poor labelled images, it was evident that the software had located the anatomical boundaries in brains with significant atrophy and no anatomical mislabelling was identified. More importantly, no obvious mislabelling was evident in the regions of interest, i.e. the medial temporal lobes.

*Subcortical segmentation and parcellation*

Statistical output files were generated for each participant during the initial stage of processing. These were generated for the subcortical segmentation (aseg) and the cortical parcellation (aparc). The statistical output from the subcortical segmentation contained the volumes of specific structures, while the statistical output from the cortical parcellation contained the thickness of specific structures. Tables were

83

generated to include measures of area, volume, and thickness of the labelled regions

for individual participants.

Figure 2: A sample output brain mask image generated by Freesurfer.
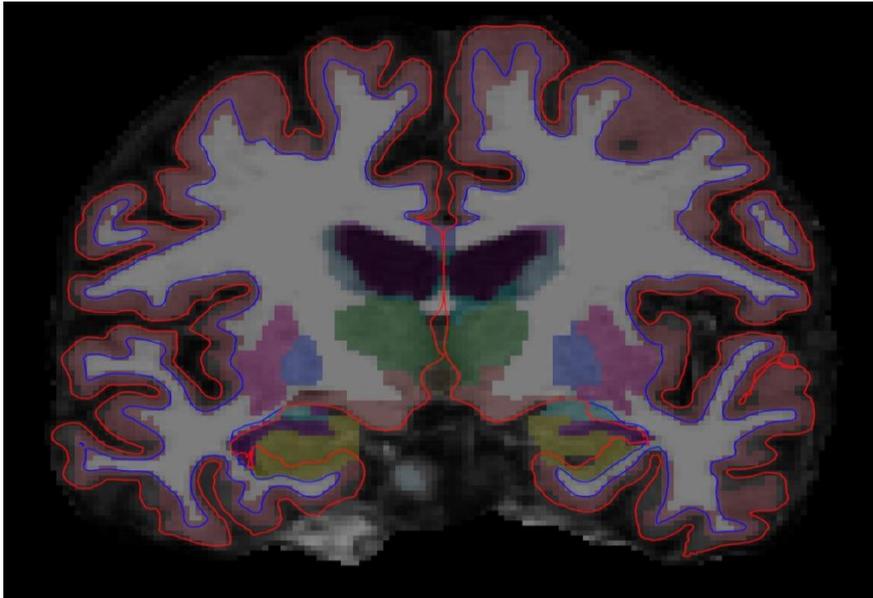
Figure 3: Visual quality control of FreeSurfer results: Examples of successful

reconstructions.
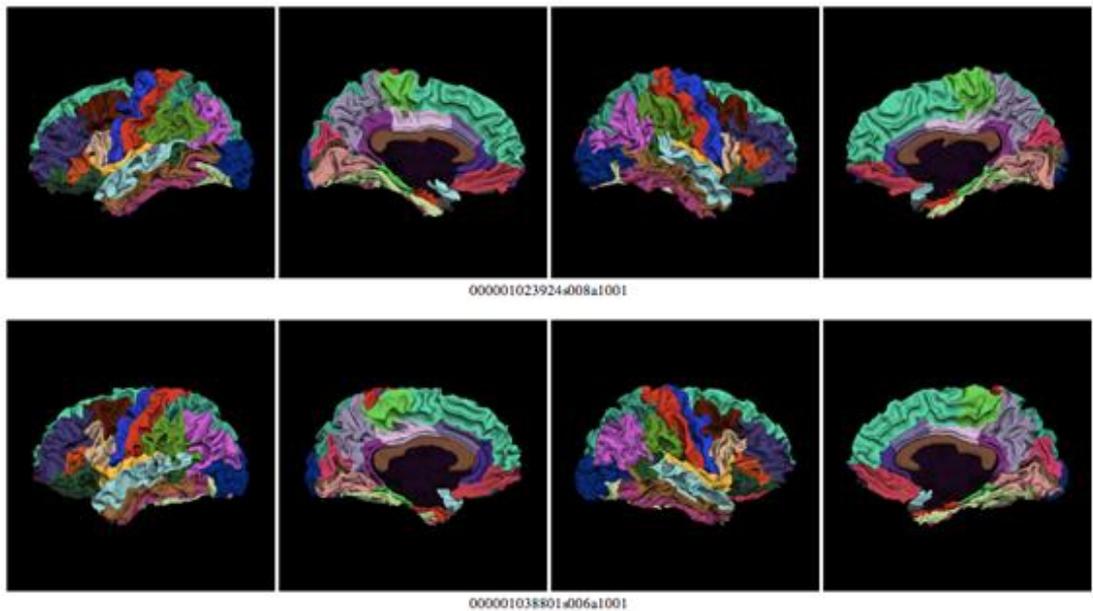
Figure 4: Visual quality control of FreeSurfer results: examples of unsuccessful reconstructions.

000010400245s012a1001
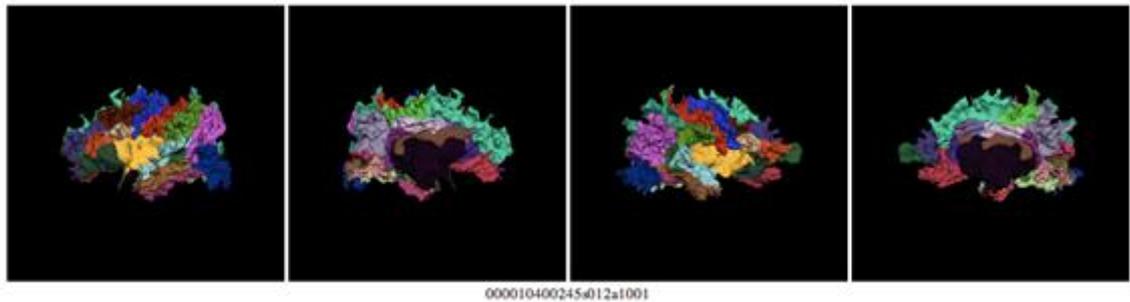
Data Analysis

Data were entered and graphed with scatterplots using Microsoft Office Excel 2016 to examine the distribution of volumes and thickness measurements across the dementia groups. Statistical Package for the Social Sciences (SPSS) version 24.0 was used to run a series of Pearson's Correlation Coefficients to explore relationships between scores on the 4MT, scores on neuropsychological tests, and different dementia diagnoses.

RESULTS

Thirty-two participants for whom there was 4MT data available met the inclusion criteria for the current study. Eighteen structural MRI brain scans were available for analysing. Two participants were excluded at this stage as their scan files did not contain enough images for processing. One scan failed at the brain mask stage and was therefore excluded from further analysis. Eleven participants had both viable scans, neuropsychological data, and 4MT scores available for analysis. There were a further 4 participants who had not been administered the 4MT as part of their research assessment and had available scans with corresponding neuropsychological data. It was decided to include these participants in the study in order to investigate correlations between brain atrophy and neuropsychological measures.

Participant Characteristics

The demographic information and baseline neuropsychological data for the sample (N=15) are presented in Table 1. Descriptive statistics (means and standard deviations) are reported in Table 2 a. The scores for Trails B were not analysed as participants with time completion scores exceeding 300 seconds were not recorded. Pearson correlation coefficients were used to investigate relationships between 4MT, neuropsychological measures, and volumetry data throughout.

Table 1: *Demographic and clinical characteristics of the participants*

	N	%	Mean	SD
Age (years)	15	-	78.80	6.26
Years of Education	15	-	13.53	3.38
ACE score	14	-	75.21	10.66
Dementia Diagnosis				
AD	8	53.3	-	-
Vascular Dementia	3	20.0	-	-
Mixed Vascular and AD	4	26.7	-	-
Gender				
Male	8	53.3	-	-
Female	7	46.7	-	-
Ethnicity				
White British	9	60.0	-	-
Irish	1	6.7	-	-
White Other	4	26.7	-	-
Black Caribbean	1	6.7	-	-
Baseline scores				
Estimated Premorbid Functioning	15	-	48.47	14.45
Memory	14	-	15.36	5.40
4MT	11	-	9.82	3.31
Visuo-spatial	14	-	13.07	1.64

Mean scores with standard deviations and the range. Neuropsychological data based on raw scores for ACE and TOPF.

Table 2: *Means and standard deviations of the neuropsychological measures for each dementia subtype*

	Dementia Diagnosis		
	AD **N=8**	**VaD** **N=3**	**Mixed AD & VaD** **(N=4)**
Age (Years)	78.63 (7.89)	81.00 (4.36)	77.50 (4.20)
Education (Years)	13.63 (4.31)	14.00 (2.65)	13.00 (2.16)
4MT	9.67(N=6) (3.89)	9.67 (3.79)	10.50 (N=2) (2.12)
ACE-III Total	76.14(N=7) (10.29)	69.67 (9.07)	77.75 (13.57)
ACE-III Memory	15.00 (N=7) (5.80)	14.00 (2.65)	17.00 (6.98)
ACE-III Attention	15.57 (N=7) (2.73)	15.33 (1.53)	14.25 (4.27)
ACE-III Language	22.57 (N=7) (2.23)	21.00 (3.46)	21.50 (1.00)
ACE-III Visuo-spatial	13.57 (N=7) (1.40)	11.00 (1.00)	13.75 (1.26)
ACE-III Fluency	9.43(N=7) (2.23)	8.33 (2.08)	11.25 (1.50)
Trails Making A (seconds)	41.14 (N=7) (24.50)	49.47 (10.69)	45.75 (18.30)

Raw scores reported for ACE, Trails Making Test and TOPF.

Hypothesis 1

The 4M task was designed to specifically depend on the function of the network of

ROIs investigated. Thus, we predict positive associations between the volume and

thickness measurements of the ROIs and 4MT scores across the sample.

Table 3: Pearson's Correlations between ROIs and 4MT scores across all dementia groups (N=11)

	r	*p* value
Volumes		
Left hippocampus	-.03	.94
Right hippocampus	-.07	.83
Left entorhinal	.15	.66
Right entorhinal	.27	.42
Left parahippocampus	.09	.80
Right parahippocampus	.06	.87
Thickness		
Left posterior cingulate gyrus	.10	.78
Right posterior cingulate gyrus	-.20	.95
Left precuneus	.32	.34
Right precuneus	.38	.12

**Correlation is significant at the 0.05 level (2-tailed)*

P=participant

Scatterplot observations

It was not possible to statistically analyse differences between AD, VaD, and mixed dementia due to the small sample size. To explore these data, a series of scatter plots were graphed (see Figure 5). This method was adopted by Chan et al. (2016) and Moodley et al. (2015) to describe observations in the dataset. A cluster of AD participants (N=5) showed relatively higher 4MT scores alongside right posterior cingulate thickness measurements on the lower end of the scale. A group of four AD participants cluster together with 4MT scores ranging from 9 to 11 and left precuneus thickness measurements ranging from 1.94mm- 2.32mm. Three AD participants show 4MT scores ranging from 9-11 and left entorhinal volume between 1127mm and 1408mm.

Figure 5: Scatterplots demonstrating 4MT scores and volume and thickness measurements of the ROIs (N=11)

Left precuneus thickness x 4MT score

Left enthorinal vol x 4MT score
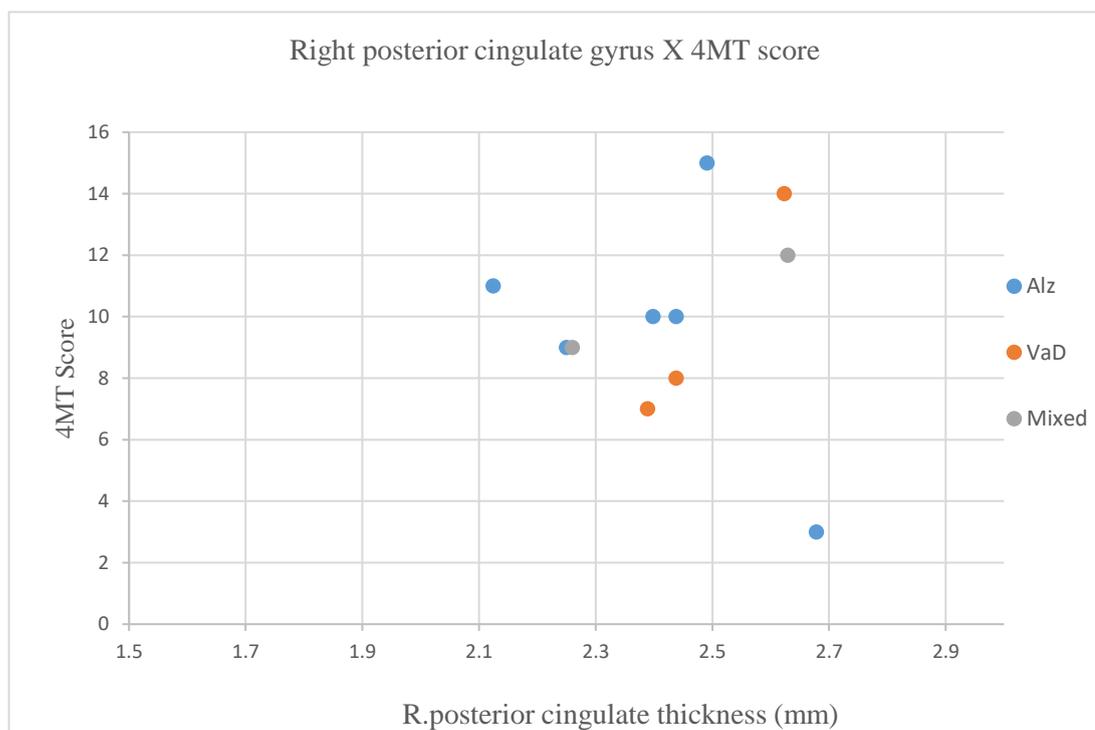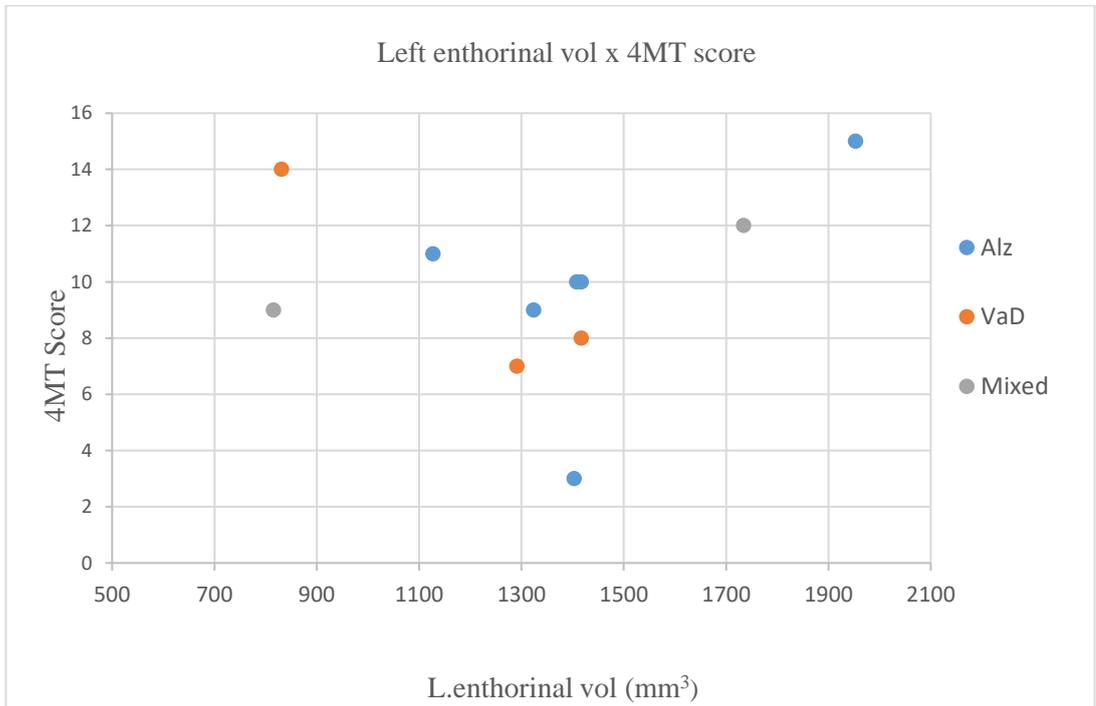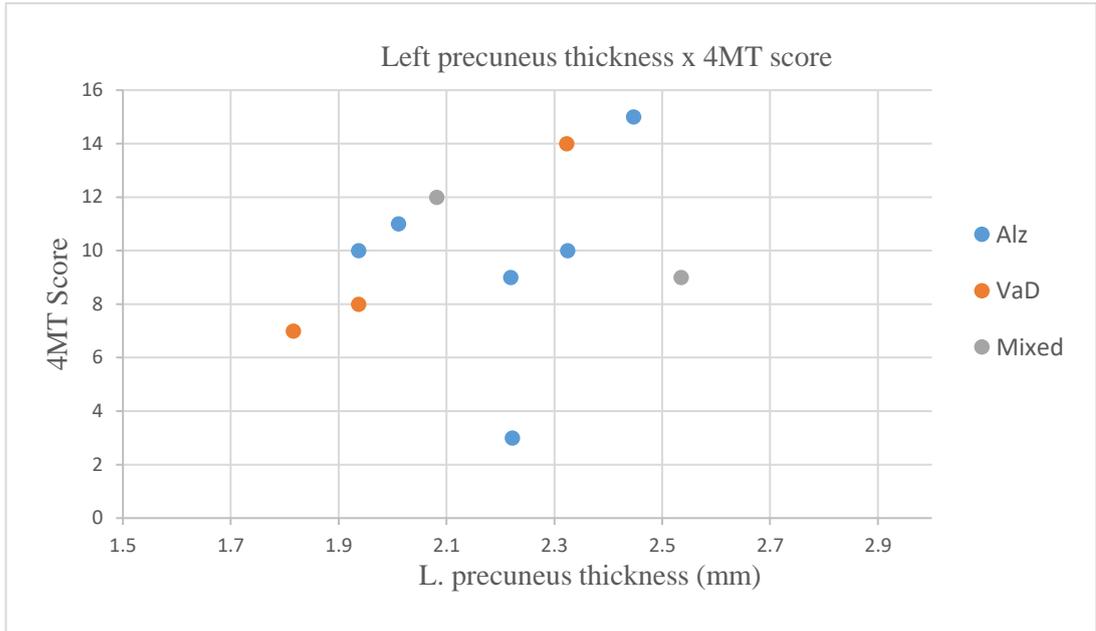
Hypothesis 2

It is predicted based on the Gore (2015) finding of no difference on the 4MT task between the different dementia subtypes, that in the current sample there will be similar levels of volumetric atrophy across these subtypes. Differences in the volumes of the critical structures in the main diagnostic categories (AD, VaD, and mixed dementia) will be investigated.

The mean volume of the left hippocampus was lower in the AD group vs. both the VaD, and mixed groups (see table 4 for means and standard deviations). For the right hippocampus, the mixed dementia group showed the lowest mean volume. The VaD group had the lowest mean volumes for left entorhinal, right entorhinal, left parahippocampus, right parahippocampus, left posterior cingulate gyrus and left precuneus thickness. The AD group had the lowest mean thickness measurement for the right posterior cingulate gyrus. For right precuneus thickness, the VaD and the mixed dementia groups had lower mean volumes.

|		Dementia Diagnosis		
		AD **N=8**	**VaD** **N=3**	**Mixed AD &** **VaD** **(N=4)**
Volume (mm3)				
Left Hippocampus	Mean SD	2723.75 (535.69)	2898.07 (225.58)	2926.43 (343.98)
Right Hippocampus	Mean SD	3011.39 (544.32)	3045.20 (293.96)	2849.52 (318.00)
Left entorhinal	Mean SD	1427.50 (291.07)	1179.67 (308.46)	1496.75 (454.91)
Right entorhinal	Mean SD	1493.75 (294.07)	1194.67 (270.06)	1491.75 (225.97)
Left parahippocampus	Mean SD	1647.38 (322.77)	1362.00 (300.49)	1474.75 (360.93)
Right parahippocampus	Mean SD	1627.00 (250.84)	1302.33 (51.79)	1726.75 (275.64)
Thickness (mm)				
Left posterior cingulate gyrus	Mean SD	2.48 (.24)	2.40 (.23)	2.57 (.22)
Right posterior cingulate gyrus	Mean SD	2.33 (.212)	2.48 (.123	2.40 (.160)
Left precuneus	Mean SD	2.15 (.19)	2.03 (.27)	2.23 (.21)
Right precuneus	Mean SD	2.15 (.15)	2.13 (.24)	2.13 (.09)

Hypothesis 3

Evidence in this field is limited to "pure" AD and MCI samples rather than

mixed samples from diagnostic services (e.g. Hartley et al, 2007; Moodley et al.

**2015). An investigation of associations between volume and thickness**

measurements and the neuropsychological measures available may clarify our understanding of the relationships between dementia subtype, volumetry, and psychometric findings.

There were no significant correlations between ACE memory scores and volume and thickness measurements in the sample. In terms of diagnostic category, there was a small cluster of AD participants (N=3) with relatively reduced right posterior cingulate gyrus thickness, ranging from 2.05mm to 2.17mm coupled with a scores of 13 for ACE memory (see figure 6).

There was no correlation between ACE attention scores and volume of the left entorhinal cortex. There was a positive correlation between ACE fluency scores and the volume of the left (r=0.58, p=.04) and right entorhinal cortices (r=.66, p=.11). There was also a positive correlation between ACE fluency scores and the volume of the right parahippocampus (r=.59, p=.030). There were no significant correlations between TOPF scores and volume and thickness measurements. Looking to diagnostic category, there was a cluster of AD participants with reduced volumes of the left hippocampus and the posterior cingulate gyri who had higher estimated IQ scores when education, age and gender were accounted for (see figure 6).

Figure 6: Scatterplot demonstrating volumes and thickness measurements of the ROIs and ACE memory, ACE attention and TOPF scores.

Right posterior cingulate thickness x ACE memory

Left enthorinal volume x ACE attention

Left posterior cingulate gyrus X TOPF (adjusted for education, age and gender)

- Alz
- VaD
- Mixed

L posterior cingulate thickness (mm)

Estimated premorbid IQ

Right posterior cingulate thickness x TOPF (adjusted for education, age and gender)

- Alz
- VaD
- Mixed

R. posterior cingulate thickness (mm)

Estimated premorbid IQ

Comparisons with normative data

There are no available normative data for the 4MT measure, therefore comparisons

are made with the group means from studies that included a comparison group of

healthy participants. The overall sample mean (*M*=9.82), for 4MT was lower than that of a group of healthy controls in a UK sample (*M*=11.10, *N*=20), but marginally higher than an Italian sample (M=9.00, *N*=10) in the Moodley et al. (2015) study. A further study by Bird et al. (2010) demonstrated a mean 4MT score of 10.70 for a group of healthy controls (*N*=25). In the current study, the AD (M=9.67), VaD (*M*=9.67), and mixed dementia groups (M=10.50) average 4MT scores were lower than the means for healthy control samples (Bird et al. 2010; Moodley et al., 2015). The average ACE scores for the full sample and for each dementia type are compared with normative data in Table 5.

Table 5: ACE scores for the full sample and as per dementia type compared with the normative sample.

	ACE mean scores (*N*=14)	AD (N=7)	VaD M (N=3)	Mixed AD & VaD (N=4)	Normative data 95% CI
ACE-III Total	75.21	71.00	71.00	82.50	96.0 (95.2-96.6)
ACE-III Memory	15.36	13.00	13.00	19.00	24.60 (24.30-25.00)
ACE-III Attention	15.14	15.00	15.00	16.00	17.60 (17.5-17.8)
ACE-III Language	21.93	23.00	23.00	21.00	25.30 (25.20-25.60)
ACE-III Visuo-spatial	13.07	13.00	11.00 10-12	14.00 12-15	15.5 (15.3-15.7)
ACE-III Fluency	9.71	10.00	9.00 6-10	11.00 10-13	12.6 (12.3-12.9)

M=mean scores, 95% CI= 95% confidence intervals

DISCUSSION

Summary of results

This study aimed to explore associations between 4MT performance and volumetric changes in a heterogeneous sample of memory clinic patients. It was predicted that there would be positive associations between the volume and thickness measurements of ROIs and 4MT scores across the sample. Contrary to prediction, there were no positive associations between these variables. It was not possible to conduct statistical analysis to compare the AD, VaD, and the mixed dementia groups due to the restricted size of the final sample. The data showed that higher 4MT scores coincided with relatively reduced right posterior cingulate gyrus thickness and 4MT scores between 9-11 coincided with left precuneus thickness and left entorhinal volume.

It was predicted based on the Gore (2015) finding of no difference on the 4MT task between AD, VaD and mixed dementia, that in the current sample there would be similar levels of volumetric atrophy across these subtypes. It was not possible to conduct statistical analyses to investigate differences in volumes and thickness measurements between the diagnostic groups due to small sample size. Through examination of means and standard deviations, the AD group did not show higher degrees of atrophy across the ROIs. The AD group had the lowest mean volume for the left hippocampus and right posterior cingulate gyrus thickness. The VaD group showed the lowest mean volume for left entorhinal, right entorhinal, left parahippocampus, right parahippocampus, left posterior cingulate gyrus, and left precuneus thickness. The mixed dementia group had the lowest mean volume for the right hippocampus.

Evidence from studies exploring the utility of spatial memory testing in predicting dementia subtype is limited to "pure" AD and MCI samples rather than mixed samples from diagnostic services (e.g. Hartley et al, 2007; Moodley et al. 2015). An investigation of associations between volume and thickness measurements and the neuropsychological measures available attempted to clarify our understanding of the relationships between dementia subtype, volumetry, and psychometric findings. There were no significant correlations between ACE memory scores and volume and thickness measurements in the sample. The mean ACE-memory scores were below the normative cut-offs and therefore it was expected that memory performance may be related to neurodegeneration in the hippocampus and MTL, but this was not the case in this sample. There was a positive correlation between ACE fluency scores and the volume of the left and right entorhinal cortices and the volume of the right parahippocampus. Looking to diagnostic category, there was a cluster of AD participants with reduced volumes of the left hippocampus and the posterior cingulate gyri who had higher estimated IQ scores when education, age and gender were accounted for.

Comparison with previous research

The current findings do not support the evidence from previous studies which have demonstrated that impairments in hippocampal-dependent allocentric memory can distinguish AD from non-AD dementia (Hartley et al., 2007; Pengas et al., 2012; Tu et al., 2015).

Interpretation of findings

The current study is underpowered and therefore these results must be cautiously interpreted. There is a risk of falsely rejecting the null hypothesis in underpowered

studies (Cohen, 1988, 1992), therefore it cannot be concluded that there is no relationship between 4MT performance and the structural measurements of the MTL regions. When p values >0.05, it can neither be concluded that the null hypothesis is false, nor can it be assumed that the null hypothesis is true (Vadillo, Konstantinidis & Shanks, 2016). Underpowered studies have lower probability that any observed effect will pass the $p < 0.05$ threshold (Button et al. 2013). The current sample size was not large enough to detect associations between the variables. A power analysis was conducted prior to commencement of this research project and this was based on a sample size of 32 participants from the Gore (2015) study. For exploring the possible association between 4MT and volumetry of the ROIs, the study was powered for a medium to large effect. However, only eleven viable MRI scans were available for comparison with 4MTscores. It was not possible to collect further data as the available sample were memory clinic patients for whom structural brain regions and performance on the 4MT and other neuropsychological measures were obtained at a pre-diagnostic phase. This underpowering is a result of unexpected problems with the quality of the scans obtained which was only detectable after a commitment had been made to the project. Due to the time restriction of doctorate research, it was not possible to recruit and test a new sample. There is a wider problem in terms of the prevalence of underpowered studies in psychological research and more specifically within neuroscience research. For example, by extracting data from meta-analyses of structural brain abnormality studies between 2006 and 2009, Ioannidis (2011) showed that the median statistical power of 461 studies was 8 per cent. Furthermore, the dichotomous nature of Null Hypothesis Significance Testing (NHST) sets an arbitrary threshold which can lead to

100

researchers concluding the null hypothesis is true if a study fails to reject the null (Hoekskra, Finch, Kiers & Johnson, 2000; Vadillo et al., 2016).

Exploring relationships between brain pathology and cognitive functions using structural MRI assumes a modular understanding of human cognitive functioning. Exploring relationships between 4MT scores and degrees of atrophy of specific ROI's excludes the possible effect of disrupted functional connectivity between different regions, which is an established phenomenon in the literature (Sporns, Chial, Kaiser & Hilgetag, 2004). It is likely that a complex range of inter-connected systems are involved in topographical disorientation (see review by Serino, Cipresso, Morganti, & Riva, 2014). A number of different techniques have been adopted to investigate functional networks of various brain regions, e.g. EEG, PET and resting state fMRI (Tomasi & Volkow, 2011; Delbeuck, Van Der Linden & Collette, 2003). These studies support the hypothesis of Alzheimer's disease as a neocortical 'disconnection syndrome' that compromises both structural and functional connectivity of cortical white matter tracts (Leuchter et al., 1992, Rose et al. 2000). In AD patients, resting state fMRI studies have demonstrated disrupted connectivity between the hippocampal formation and the medial prefrontal cortex, ventral anterior cingulate cortex, right infrotemporal cortex, right cuneus and precuneus, left cuneus, right superior and middle temporal gyri and the posterior cingulate cortex (Wang et al., 2006). Evidence for a disconnection syndrome in AD extends more broadly across whole brain networks, e.g. an anterior-posterior disconnection between pre-frontal and parietal lobe regions using resting brain state analysis (Wang et al., 2007) and when engaged in cognitive tasks, such as immediate and delayed memory for faces (Grady et al. 2001).

Low statistical power means that it is not possible to draw meaningful conclusions from these results, however previous studies with adequate power to detect effects have shown no significant association between the thickness of the posterior cingulate gyrus and 4MT performance, e.g. Moodley et al. (2015). In the current study, contrary to predictions, higher 4MT scores coincided with relatively reduced right posterior cingulate gyrus thickness. The posterior cingulate gyrus is typically involved in early AD (Minoshima et al., 1997; Scheff et al., 2015) and has been linked to a broader region known as the brain's 'default mode network' which facilitates free-thinking and the generation of self- relevant mental explorations, e.g. anticipating and mental rehearsal of possible future events (Lehmann et al., 2010). The role of the 'default mode network' in AD is evidenced by fMRI and PET studies (Buckner, Andrews-Hanna & Schacter, 2008). This network includes "the posterior cingulate cortex (PCC), precuneus, dorsal and ventral medial prefrontal, lateral (mainly inferior) parietal cortices, and medial temporal lobes" (Mevel. Chetelet, Eustace & Desanges, 2011, p.2). Episodic memory impairment, the hallmark of MCI and AD has been shown to be induced by disruption to functional connectivity between the PCC and the cingulum bundle (Chtelat et al., 2003; Villain et al., 2008). The links between declining cognitive functions typical in AD and functional connectivity between different brain regions provides a further rationale for investigating connections between MTL regions rather than MTL atrophy in isolation.

Another possible explanation for the absence of expected effects is the validity of dementia diagnoses. Diagnoses of dementia cannot be confirmed until post-mortem analysis is undertaken as the current methods available for making a clinical diagnosis of dementia are only an estimation of suspected underlying

pathology. Rates of diagnostic accuracy vary from 52 to 100% for AD (Molsa et al., 1985, Victoroff et al., 1995) and between 21 and 95% for VaD (Knopman et al., 2001). In a more recent study by Gay et al. (2008), increasing age was associated with neuropathological diagnoses via post-mortem. Out of a sample of 221 older adults 67.8 % of the clinically diagnosed patients received a definitive diagnosis of AD, VaD or mixed dementia. The sensitivity for AD was 75.9 % and specificity was 60.6 %.

Correlations between memory deficits and atrophy were not borne out in the results. This finding is not in keeping with the established function of the hippocampal and MTL structures in episodic memory functions. One possible reason for this is the validity of ACE-III as a tool for detecting neural change over time in dementia (Larner & Mitchell, 2014). ACE fluency scores were correlated with the entorhinal cortices and right parrahippocampal volume. This result is in keeping with the established role of the temporal lobes in language abilities. A major contributing factor to the non-significant correlations between 4MT, volumetry, and other cognitive functions is likely due to a smaller sample size than expected. This means the study is underpowered. The implications of this are discussed in detail below (see limitations section).

Limitations

The main limitation of this study is the smaller than anticipated sample size and consequent lack of statistical power to detect effects. The implications of these limitations have been addressed in the interpretation of findings section.

The exclusion of a healthy control group for comparison with the dementia groups is a limiting factor in the research. Longitudinal designs are recommended for future research in this area. Furthermore, a number of cognitive factors may have had an influence on 4MT performance, e.g. the correlation between verbal fluency and entorhinal and right parahippocampal volumes may indicate that language difficulties associated with cognitive decline may have had a confounding influence on 4MT performance. The possible influence of premorbid abilities was highlighted by observations in the data. These results demonstrate a cluster of AD participants with greater pre-morbid intellectual abilities. This trend in the data highlights the possible need for premorbid IQ matching of participants as higher premorbid IQ may be a protective factor in dementia. The cognitive reserve hypothesis proposes that higher IQ, education, occupational attainment, or participation in leisure activities act as protective factors that may modulate the clinical expression of AD pathology (Stern, 2006).

The use of structural MRI has associated methodological weaknesses e.g. the Freesurfer tool automatically segments between white matter and grey matter which is used to map the location of brain structures independent from individual differences in the size and overall shape of the brain (Lancaster et al., 2000). It is possible that errors can occur during this process. There is a possibility that the one scan which failed at the skull strip stage of processing could have been investigated further and if appropriate, manual adjustment of the parameters could have been applied. This was not possible due to the sole researcher's limited skills and training in this area. The complexity of volumetric analysis may also lead to variations in findings; for example, decisions about adjustments for intracranial volume may be taken earlier or later in the process leading to systematic differences. Moodley et al.

104

(2015) corrected for total intracranial volume after averaging between left and right hemispheres before running their correlations between 4MT and the ROI's.

Clinical Implications

The findings of the current study are not sufficient to determine the reliability of the 4MT for use in dementia diagnosis with heterogeneous groups that attend memory clinics. The finding that hippocampal and MTL atrophy is common to AD, VaD, and mixed dementia in this sample may account for the non-significant differences in 4MT performance, however lack of statistical power does not allow for statistical conclusions to be drawn from the results. Previous research has demonstrated that the 4MT is sensitive and specific in predicting conversion from MCI to AD, identifying individuals with biomarkers and in distinguishing AD from rarer forms of dementia e.g., FTD and semantic dementia. However, previous studies have suggested that overlap in structural degeneration between the most common types of dementia may mean that the 4MT is not specific enough in differential diagnosis and clinicians may need to rely on other clinical features of each of the dementia subtypes (see Gore for references) Subtyping in dementia is important in terms of treatment decisions. It is estimated that approximately 60% of patients with AD demonstrate subjective improvements with memory drug treatments, whereas these drugs are not effective in VaD and not indicated due to possible side effects (Department of Health, 2014).

Difficulties with spatial memory identified by individuals with dementia and their carers include; difficulty or inability to remember familiar and unfamiliar environments, learn new routes, use maps and recognize places. The 4MT is a laboratory developed test based on research which pinpoints a specific deficit in allocentric memory in AD patients. These homogenous samples have been tested in

controlled environments using rigorous diagnostic assessment procedures. It is likely that in a heterogeneous clinical sample other cognitive deficits apart from allocentric representations may be at play when it comes to spatial memory deficits.

Future research

It would be beneficial to replicate this study with a larger sample size to address the issues with power previously discussed. This would also allow group differences to be compared statistically and provide more robust evidence about possible differences or similarities between the dementia subtypes. A larger sample would also allow for the use of logistic regression using a stepwise approach to select the best cognitive and neuronal predictors of dementia type (AD versus non-AD). The inclusion of patients with MCI would allow investigation of the utility of the 4MT at an earlier stage of the disease process, where previous research has demonstrated the 4MT's sensitivity and specificity in predicting conversion to AD.

Future research in this field may benefit from the use of neuroimaging techniques that identify metabolic activity in the brain, e.g. fMRI, PET, and SPECT. In AD, resting-state metabolic activity in the brain's 'default mode network' has been shown to correlate with the distribution of amyloid plaques (Buckner et al., 2008; Mintun et al., 2006a). These neuro-imaging techniques may contribute to a greater understanding of how spatial representations are mapped in different dementias and therefore support the refinement of behavioural measures for use in differential diagnosis. It is important to note the challenges and ethical considerations when balancing the potential research benefits of using invasive imaging procedures with the clinical needs of participants with dementia. In addition, further studies are warranted using measures that can separate out and control for the influence of

egocentric processing, e.g. the virtual reality Starmaze navigation task (Bellassen et al. 2012). This would also allow for an investigation of the possible deficits in the translation between these two reference frames.

Conclusions

It was not possible to determine the clinical utility of a test of hippocampal-dependent spatial memory, the 4MT in distinguishing AD from other dementia types in heterogeneous groups of memory clinic patients. This was due to lack of statistical power to detect possible effects that have been demonstrated in previous studies. It is possible that overlapping patterns of atrophy across different dementia types in clinical samples confound the use of this test, which claims to specifically measure hippocampal-dependent allocentric representations. Other cognitive impairments associated with dementia may also have a confounding impact on the performance of spatial memory tests. This is an exploratory study with low power and a lack of inferential statistics and as such the findings are tentative and must be interpreted with caution. However, the introduction of structural MRI data has contributed to further understanding of a tool that may support in the early diagnosis of dementia.

REFERENCES

Agamanolis, D. P. (2014, June). *Alzheimer's disease*. Retrieved from

http://neuropathology-web.org/agamanolis.html.

Albert M.S., DeKosky S.T., Dickson D., Dubois, B., Feldman, H.H., Fox, N.C., et al.

(2011) The diagnosis of mild cognitive impairment due to Alzheimer's

disease: recommendations from the National Institute on Aging-Alzheimer's

Association workgroups on diagnostic guidelines for Alzheimer's disease.

Alzheimers Dement (2011) 7:270–9.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of*

mental disorders (5th ed.). Washington, DC: Author.

Andersen, R.A., Snyder, L. H., Bradley, D.C., & Xing, J. (1997). Multimodal

representation of space in the posterior parietal cortex and its use in planning

movements. *Annual. Review of. Neuroscience, 20*, 303–330.

Ashendorf, L., Jefferson, A. L., O'Connor, M.K., Chaisson, C., Green, R.C.& Stern,

R.A. (2008). Trail Making Test errors in normal ageing, mild cognitive

impairment and dementia. *Archives of Clinical Neuropsychology, 23*, 129-

137.

Belaroussi, B., Milles, J., Carne, S., Zhu, Y.M. & Benoit-Cattin, H. (2006). Intensity

non-uniformity correction in MRI: Existing methods and their validation.

Medical Image Analysis, 10, 234-246.

Bellassen, V., Igloi, K., de Souza, L. C., Dubois, B., & Rondi-Reig, L. (2012).

Temporal order memory assessed during spatiotemporal navigation as a

behavioral cognitive marker for differential Alzheimer's disease diagnosis.

The Journal of Neuroscience, 32, 1942-1952.

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

Bloom, G.S. (2014). The trigger and bullet in Alzheimer disease pathogenesis. *Journal of the American Medical Association, 71*, 505-508.

Bowie, C.R., & Harvey, P.D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols, 1,* 2277-2281.

Buckner, R.L., Andrews-Hanna, J.R., & Schacter, D.L. (2008). The Brain's Default Network Anatomy, Function, and Relevance to Disease. *Annals of the New York Academy of Science, 1124*: 1–38.

Burgess, N., Maguire, E.A., Spiers, H.J., O'Keefe, J. (2001) Temporoparietal and Prefrontal Network for Retrieving the Spatial Context of Lifelike Events. *Neuroimage, 14*, 439-53.

Burgess, N., Trinkler, I., King, J., Kennedy, A., & Cipolotti, L. (2006). Impaired allocentric spatial memory underlying topographical disorientation. *Reviews in the Neurosciences*, *17*(1-2), 239.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., & Munafo (2013). Power Failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews, 14,*365-376.

Cavanna, A. & Trimble (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain, 129*, 564-583.

Chan, D., Gallaher, L.M., Moodley, K., Minati, L., Burgess, N. & Hartley, T. (2016). The 4 Mountains Test: A short Test of Spatial Memory with High Sensitivity for the Diagnosis of Pre-dementia Alzheimer's Disease. *Journal of Visualised Experiments, 116,* 1-11.

Chtelat, G., Desgranges, B., De La Sayette, V., Viader, F. Berkouk, K. Landeau, B.,… Eustache, F. (2006). Dissociating atrophy and hypometabolism impact on episodic memory in mild cognitive impairment. *Brain, 126*, 1955-1967.

Camicioli, R. (2006). Distinguishing Different Dementias. *Canadian Review of Alzheimer's Disease and Other Dementias 9*(3), 4–11.

Clifford, J., Wiste, H., Vemuri, P., Weigand, S., Senjem, M., Zeng, G… Knopman, D. (2010). Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. *Brain, 133*(11), 3336-3348*.*

Cohen, J. (1988). Statistical Power Analysis for the behavioural sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Delbeuck, X., Van der Linden, M. & Collette, F. (2003). Alzheimer's Disease as a Disconnection Syndrome. *Neuropsychology Review, 13*, 79-92

Delis, D. C., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan Executive Function System*. San Antonio TX: Psychological Corporation.

Department of Health. (2014). *Dementia Revealed: What Primary Care Needs to Know.* Retrieved from https://www.england.nhs.uk/wp-content/uploads/2014/09/dementia-revealed-toolkit.pdf

Duff, K., Chelune, G. J., & Dennett, K. (2011). Predicting estimates of premorbid memory functioning: Validation in a dementia sample. *Archives of Clinical Neuropsychology, 26,* 701-705.

Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T., Isham, E.A., Newham, E.L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature, 425*, 184-188.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C. ... & Dale, A.M. (2002) Whole brain segmentation: automated labelling of neuroanatomical structures in the human brain. *Neuron 33*, 341-355.

Fischl, B., Salat, D.H., Van der Kouwe, A.J., Makris, N., Segonne, F., Quinn, B.T. & Dale, A.M. (2004a). Sequence-independent segmentation of magnetic resonance images. *Neuroimage 23*, 69-84.

Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., & Thompson, P.M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nature Reviews Neurology, 6,* 67-77.

Gore, L.M. (2015) The Clinical Utility of the Four Mountains Test in the Early Diagnosis of Alzheimer's Disease: A Measure of Allocentric Memory Ability. Doctoral thesis, University College London.

Grady, C.L., Furey, M.L., Pietrini, P., Horwitz, B., Rapoport, S.I. Altered brain

    functional connectivity and impaired short-term memory in Alzheimer's

    disease. *Brain*, *124*, 739–756.

Halligan, P.W., Fink, G.R., Marshall, J.C., & Vallar, G. (2003). Spatial Cognition:

    evidence from visual neglect. *Trends in Cognitive Science, 7*, 125-133.

Han, X., Jovicich, J., Salat, D., Van der Kouwe, A., Quinn, B., Czanner, S.,…Fischl,

    B. (2006). Reliability of MRI-derived measurements of human cerebral

    cortical thickness: the effects of field strength, scanner upgrade and

    manufacturer. *Neuroimage 32*, 180-194.

Hanakawa, T., Immisch, I., Toma, K., Dimyan, M.A., Van Gelderen, P., & Hallett,

    M. (2003). Functional properties of brain areas associated with motor

    execution and imagery. *Journal of Neurophysiology 89*, 989–1002.

Hartley, T., Bird, C. M., Chan, D., Cipolotti, L., Husain. M., Vargha-Khadem, F., &

    Burgees, N. (2007). The hippocampus is required for short-term topographical

    memory in humans. *Hippocampus*, *17*, 34–48.

Hassabis, D., Kumaran, D. & Maguire, E.A. (2007). Using Imagination to

    Understand the Neural Basis of Episodic Memory. *Journal of Neuroscience,*

    27 (52), 14365-14374.

Hoekstra, R., Finch, S., Kiers, H.A.L.& Johnson A. (2006). Probability as certainty:

    Dichotomous thinking and the misuse of *p*-values. *Psychonomic Bulletin and*

    *Review, 21,*1157-1164.

Hsieh S., Schubert S., Hoon C., Mioshi E. & Hodges J. R. (2013). Validation of the

Addenbrooke's Cognitive Examination III in Frontotemporal Dementia and Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, 36, 242–250.

Iaria, G.,. Chen, J.K., Guariglia, C., Ptito, A., & Petrides, M. (2007). Retrosplenial and hippocampal brain regions in human navigation: complementary functional contributions to the formation and use of cognitive maps. *Eur J Neurosci, 25*, 890-899.

Ioannidis, J.P.A., Munafo, M. R., Fusar-Poli, P., Nosek, B.A., &David S.P. (2014). Publication and other reporting biases in cognitive sciences: Detection, prevalence, and prevention. *Trends in Cognitive Neurosciences, 18*, 235-241.

King J, A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). The human hippocampus and viewpoint dependence in spatial memory. *Hippocampus*, *12*, 811–820.

Kalová, E., Vlcek, K., Jarolímová, E., & Bures, J. (2005). Allothetic orientation and sequential ordering of places is impaired in early stages of Alzheimer's disease: Corresponding results in real space tests and computer tests. *Behavioural Brain Research*, *159*, 175–186.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections In C. Freksa & C. Habel (Eds.), *Spatial Cognition*: *An Interdisciplinary Approach to Representing and Processing Spatial Knowledge* (pp. 1-17). Heidelberg: Springer.

Knopman, D.S., DeKosky, S.T., Cummings, J.L., Chui H., Corey-Bloom, J., Relkin N, Small. G.W., Miller B., Stevens, J.C. (2001). Practice parameter: diagnosis

of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology. 56*, 1143-53.

Langa, K.M., Foster, N.L., & Larson, E. B. (2004). Mixed dementia: Emerging concepts and therapeutic implications. *Journal of the American Medical Association, 292*, 2901-2908.

Lancaster, J. L., Woldoroff, M.G., Parsons, L., Liotti, M. Freitas, C.S., Rainey, L.,…Fox, P.T. (2000). Automated Talairach Atlas Labels for Functional Brain Mapping. *Hum Brain Mapp*, *10*, 120-131.

Land, M. F. (2014). Do we have an internal model of the outside world? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 1636.

Larner, A.J. and Mitchell, A.J. (2014). A meta-analysis of the accuracy of the Addenbrooke's Cognitive Examination (ACE) and the Addenbrooke's Cognitive Examination-Revised (ACE-R) in the detection of dementia. *International Psychogeriatrics, 26*, 555-563.

Lehmann, M., Rohrer, J.D., Clarkson, M.J., Ridgway, G.R., Scahill, R. I., Modat, M.,… Fox, N.C. (2010). Reduced cortical thickness in the posterior cingulate gyrus is characteristic of both typical and atypical Alzheimer's disease. *Journal of Alzheimer's Disease*, *20* (2) 587 - 598.

Leuchter, A.F., Newton, T.F., Cook, I.A., Walter, D.O., Rosenburg-Thompson & Lachenbruch, P.A. (1992). *Brain, 115*, 1543-61.

Maguire, E.A. & Cipolotti, L. (1998) Selective sparing of topographical memory. *J Neurosurg Psychiatry, 65*, 903-909.

Malouin, F., Richards, C.L. Jackson, P.L., Dumas, F., & Doyon, J. (2003). Brain activations during motor imagery of locomotor-related tasks: a PET study. *Human Brain Mapping*, *19*, 47–62.

McDonald, A. (2015). Assessing the robustness of the Test of Premorbid Functioning (TOPF) as a measure of premorbid intelligence in Alzheimer's and vascular dementia and clinical research portfolio. D Clin Psy thesis: http://theses.gla.ac.uk/6793/

McKeel, D. W., Burns, J. M., Meuser, T. M., & Morris, J. C. (2007). *Chapter 1: Cognitive aging and dementia: An overview.* Oxford: Clinical Publishing, An Imprint of Atlas Medical Publishing Ltd. Retrieved from http://search.proquest.com/docview/189247532?accountid=14511

Minoshima, S., Giordani, B., Bernet, S., Frey, K.A., Foster, N.L., & Kuhl, D. E. (1997). Metabolic Reduction in the Posterior Cingulate Cortex in Very Early Alzheimer's Disease. *Annals of Neurology, 42*, 85-94.

Mintun, M.A., Sacco, D., Snyder, A. Z., Couture, L., Powers, W. J., Hornbeck, R.,…Morris, J. C.(2006a). Distribution of glycolysis in the resting healthy human brain correlates with distribution of beta-amyloid plaques in Alzheimer's disease. *Soc. Neurosci. Abstr.*, 707.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges J, R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *Internal Journal of Geriatric Psychiatry, 21*, 1078–1085.

Misaki, M., Matsumoto, E., & Miyauchi, S. (2002). Dorsal visual cortex activity elicited by posture change in a visuo-tactile matching task. *NeuroReport, 13*, 1797–1800.

Molsa P.K., Paljarvi, L. Rinne, J.O., Rinne U.K., Sako, E. (1985). Validity of clinical diagnosis in dementia: a prospective clinicopathological study. *J Neurol Neurosurg Psychiatry ,48*, p.1085-90.

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

Morganti, F., Stefanini, S., & Riva, G. (2013). From allo- to egocentric spatial ability in early Alzheimer's disease: a study with virtual reality spatial tasks. *Cognitive Neuroscience*, *4*, 171-180.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research, 34*, 171–175.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford University Press, New York.

Pengas, G., Patterson, K., Arnold, R.J., Bird, C.M., Burgess, N., and Nestor, P.J. (2010). Lost and Found: Bespoke Memory Testing for Alzheimer's Disease and Semantic Dementia *Journal of Alzheimer's Disease, 21*, 1347-1365.

Pengas, G., Williams, G. B., Acosta-Cabronero, J., Hong, Y. T., Izquierdo-Garcia, D., Fryer, T. D., &Nestor, P.J. (2012). The relationship of topographical memory

performance to regional neurodegeneration in Alzheimer's disease. *Frontiers in Aging Neuroscience*, *4*, 1–10.

Peterson, R. (2011) Mild Cognitive Impairment. *The New England Journal of Medicine, 364* (23), 2227-2234.

Prince, M. Knapp, P. Guerchet, M. McCrone, P.; Prina, M., Comas-Herrera M. Wittenberg, A., Adelaja, R., Hu, B., King, D. Rehill, D. and Salimkumar, D. (2014) Dementia UK: update *Alzheimer's Society 2014*.

Reitan. R. M. (1992). *Trail making test: Manual for administration and scoring*. Neuropsychology Laboratory, Tucson, AZ.

Ritchie, K., Ropackic, M., Albalad, B., Harrisone, J., Kayeg, J., Kramerh, J., …& Ritchie, C.W. (2017). Recommended cognitive outcomes in preclinical Alzheimer's disease: Consensus statement from the European Prevention of Alzheimer's Dementia project. *Alzheimer's & Dementia*, *13*,186-195.

Roman, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., …& Scheinberg, P. (1993). Vascular dementia: Diagnostic criteria for research studies, Report of the NINDS-AIREN International Workshop. *Neurology*, *43*, 250–260.

Rose, S. E., Chen, F. Chalk, J. B., Zelaya, F. O. Strugnell, W E., Benson, M. Semple, J. Doddrell, D.M. (2000). *Journal of Neurology, Neurosurgery & Psychiatry 69* (4), p.528.

Reuter, M., Rosas, H.D., Fischl, B., (2010). Highly Accurate Inverse Consistent Registration: A Robust Approach. *Neuroimage 53*, 1181–1196.

Scheff, W., Price, D.A., Roberts. K. N., Schmitt, F. A., Ikonomovic, M. D. & Mufson, E. J. (2015). Synaptic Change in the Posterior Cingulate Gyrus in the Progression of Alzheimer's Disease. *Journal of Alzheimer's Disease 43*, 1073–1090.

Segonne, F., Dale, A.M., Busa, E., Glessner, M., Salat, D., Hahn, H.K. & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *Neuroimage 22*, 1060-1075.

Serino, S., Cipresso, P., Morganti, F., & Riva, G. (2014). The role of egocentric and allocentric abilities in Alzheimer's disease: A systematic review. *Ageing Research Reviews, 16*, 32-44.

Spreen, O., & Strauss, E. (1991). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* New York, NY: Oxford University Press.

Sporns, O., Chival, D.R., Kaiser, M. & Hilgetag, C.C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences, 8*, 418-425.

Stern, Y. (2006). Cognitive Reserve and Alzheimer Disease. *Alzheimer Disease and associated disorders*, *20*, 69-74.

Stott, J., Scior, K., Mandy, W., & Charlesworth, G. (2017). Dementia Screening Accuracy is Robust to Premorbid IQ Variation: Evidence from the Addenbrooke's Cognitive Examination-III and the Test of Premorbid Function. *Journal of Alzheimer's Disease,* (In Press), 1-10.

Tomasi, D. & Volkow, N. D. (2011). Association between Functional Connectivity

    Hubs and Brain Networks. *Cerebral Cortex*, *21* p 2003-13.

Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by

    age and education. *Archives of Clinical Neuropsychology, 19*, 203-214.

Tu, S., Wong, S., Hodges, J.R., Irish, M., Piguet, O., & Hornberger, M. (2015) Lost

    in spatial translation: A novel tool to objectively assess spatial disorientation

    in Alzheimer's disease and frontotemporal dementia. *Cortex, 67*, 83-94

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial

    cortex do? *Nature Reviews Neuroscience, 10,* 792-802.

Velayudhan, L., Ryu, S., Raczek, M., Philpot, M. Lindesay, J., Critchfield, M. &

    Livingston, G. (2014). Review of brief cognitive tests for patients with

    suspected Dementia*. International Psychogeriatrics, 26*, 1247-1262.

Victoroff, J., Mack, W.J., Lyness, S.A., Chui, H.C. Multi-center clinicopathological

    correlation in dementia (1995*). American Journal of Psychiatry,152*, 1476-

    84.

Villain, N., Desgranges, B., Viader, F., De La Sayette, V. Mézenge, F., Landeau,

    …Chételat, G. (2008). Relationships between hippocampal atrophy, white

    matter disruption, and gray matter hypometabolism in Alzheimer's disease.

    The Journal of Neuroscience, 28, 6174-8.

Vogley, K., May, M., Ritzl. A., Falkai, P., Zilles. K., & Fink, G.R. (2004). Neural

    Correlates of First Person Perspective as One Constituent of Human Self-

    Consciousness. *Journal of Cognitive Neuroscience, 16*, 817-826.

Wang, K.J., Jiang, T., Liang, M., Wang, L., Tian, L., Zhang, X.,…Li, K. (2006).
Changes in hippocampal connectivity in early stages of Alzheimer's disease:
Evidence from resting-state fMRI. *Neuroimage, 31*, 496-504.

Wang, K. Liang, M., Wang, L., Tian, L., Zhang, X. Li, K., Jiang, T. (2007). Altered
functional connectivity in early Alzheimer's disease; A resting-state fMRI
study. *Human Brain Mapping, 28*, 967-978.

Wahlin, T.B.R. Backman, L., Wahlin, A., & Winbald, B. (1996). Trail Making Test
performance in a community-based sample of healthy very old adults: Effects
of age on completion time, but not on accuracy. *Archives of Gerontology and
Geriatrics, 22*, 87-102.

Wechsler, D. (2011). *Test of Premorbid Functioning. UK Version (TOPF UK)*.
Bloomington, MN: Pearson.

Wood, R.A., Moodley, K.K., Lever, C., Minati, L., & Chan, D. (2016). Allocentric
spatial memory testing predicts conversion from mild cognitive impairment to
dementia: An initial proof of concept study. *Frontiers in Neurology 7*.

Wolk, D. A., Dunfee, K. L., Dickerson, B. C., Aizenstein, H. J., & DeKosky, S. T.
(2011). A medial temporal lobe division of labor: Insights from memory in
aging and early Alzheimer disease. *Hippocampus, 21*, 461–466.

World Alzheimer's Report. (2016). *Improving healthcare for people living with
dementia: Coverage, quality and costs now and in the future*. Retrieved from:
https://www.alz.co.uk/research/world-report-2016

World Health Organisation (2015). *Dementia*. Retrieved from
http://www.who.int/mediacentre/factsheets/fs362/en/.

Yonelines, AP., Widaman, K., Mungas, D., Reed B., Weiner, MW., Chui, H.C.
Memory in the ageing brain: Doubly dissociating the contribution of the
hippocampus and entorhinal cortex. *Hippocampus, 17*, 1134-1140.

Zekry, D., Hauw, J., & Gold, G. (2002). Mixed Dementia: Epidemiology, Diagnosis,
and Treatment. *Journal of the American Geriatrics Society, 50*, 1431-1438.

PART 3: CRITICAL APPRAISAL

Introduction

This critical appraisal will be a personal reflection on my experiences of undertaking a secondary analysis of data and a discussion of the conceptual, ethical, and personal issues associated with research into early diagnosis of dementia. Firstly, I will begin by discussing my motivations for the choice of both the systematic literature review and the empirical project. This will be followed by a discussion about the challenges of conducting a secondary analysis of data and overcoming technical challenges. I will reflect on the limitations of the analysis and findings and discuss implications for clinical practice. I will conclude this critical appraisal with discussing the broader conceptual issues of early diagnosis, stigma and lived experiences of dementia.

Motivations for choice the of research topics

The World Health Organisation (2012) highlighted dementia as a global challenge. A broad public health approach is recommended to improve care and enhance quality of life for people living with dementia and their family caregivers. The G8 summit outlined ambitious global plans for a cure or a disease modifying therapy for dementia by 2025. There are a raft of similar global initiatives and national campaigns with a focus on future goals. Meanwhile, there are a large number of individuals currently living with dementia who may benefit from early diagnosis and intervention to improve health outcomes and maintain or improve quality of life. I have observed the pivotal role that clinical psychologists play in the assessment and treatment of people living with dementia, for example, pre-diagnostic counselling, obtaining informed consent for assessment, cognitive and neuropsychological assessment, communicating about dementia diagnosis, and delivering post-diagnostic

psychosocial interventions (BPS, 2014). In addition, I have experience working with individuals with dementia and have carried out a broad range of neuropsychological assessments in my clinical work. As such, I was motivated to undertake my own research and contribute to the growing body of evidence aiming to improve care and outcomes for individuals with dementia and their caregivers.

The choice of my systematic literature topic differed from that of the empirical paper as I wanted to capture a different aspect of the lived experience of individuals with dementia. Evidence suggests that anxiety is prevalent (Seignourel, Kunik, Snow, Wilson, & Stanley (2008) and CBT treatment with adaptations has been shown to be effective in the mild-moderate stages of dementia (NICE, 2006). It was important for me to choose a topic to highlight mental health difficulties in this population. In my clinical work as a trainee psychologist, I have witnessed the impact of behavioural and psychological symptoms associated with dementia, how these are perceived and treated and the impact on the individual and their family caregivers. A greater recognition that individuals with dementia suffer co-occurring mental health difficulties which lead to worse health outcomes was important for me to highlight and investigate in the systematic literature review.

When I first started to explore options for the major research project the idea of taking on a secondary analysis of data was unfamiliar but appealing. From previous experiences of recruiting clinical and non-clinical samples for undergraduate and MSc research, I was struck by the time, effort, and willingness of the participants to contribute. This was often part of fulfilling an aim to further scientific knowledge or improve health outcomes for themselves and others. These aims are often not achieved because of a long history in psychological research of

prioritising publication of significant findings to the detriment of important

replications and null findings (Laws, 2013). For this reason, a secondary analysis

struck me as a more responsible and ethical approach; using data that had already

been collected to answer different but related questions and further knowledge in the

field of early diagnosis of dementia.

Secondary analysis and technical challenges

The decision to analyse neuroimaging data as the main task of this project presented

a number of challenges. Getting to grips with the technicalities of Freesurfer

processes presented a steep learning curve for me. I had initially anticipated that I

could learn these skills with support from my research supervisor and collaborator

and apply them independently to analyse the data. In reality, this process was more

complex and required a collaborative approach with my research supervisor and the

external collaborator working together to figure out the process and troubleshoot

problems. Due to my lack of experience and previous knowledge, there were times

where I felt out of my depth with this process. It helped to discuss these challenges

with my research supervisor who helped me to appreciate which elements of the

process were more important to understand, for example, understanding the

technicalities of Freesurfer code was unnecessary and my efforts were better placed

in understanding the segmentation and mapping processes so that I could make

informed decisions about the analysis.

I also questioned the reliability and accuracy of the Freesurfer automated

analysis. While automated methods provide more time-efficient analysis of data

compared with manual investigations, they still require operator input and vigilance

with regards to quality control checks (Bigler, 2015). It was beneficial to spend time

with my supervisor and external collaborator discussing observations of the processed scans in order to carry out quality control checks. Furthermore, caution must be exercised when interpreting automated segmentation in older people with dementia as both age-related atrophy and vascular disease have an impact on the delineation of white, grey and CSF boundaries, which can sometimes be less well defined and irregular in the older adult population (Wenger et al.;2014; Clerx et al. 2015).

In the initial stages of the research project, the technicality of the language used to describe the Freesurfer processes and the medical terminology in the research literature were difficult to understand and I noticed that I was at risk of not being able to critique the methods until I had a greater understanding of how they operated. Learning about these concepts was essential to interpreting the results and generalising the findings. The generalizability of findings in this research field is compromised by inconsistencies between research studies, e.g. automated analysis may differ as a result of using different operating systems (Gronenschild et al.,2012) and the degree of smoothing and image modulation used (Scarpazza et al., 2015). Conducting this research has highlighted the importance of understanding the methods used in research studies in order to critically appraise the findings. In the past I have found myself skimming over complex method sections, particularly in neuropsychology research papers that use imaging techniques. This research has highlighted the need to examine research methods more carefully and critically.

Learning these new skills has given me insight into working with the methods and techniques of neuroimaging and the need to collaborate with disciplines outside of clinical psychology to pursue neuroimaging research. I have acquired skills and

knowledge in neuro-imaging techniques and functional neuroanatomy. As a trainee

clinical psychologist, I feel better prepared to liaise and consult with multi-

disciplinary colleagues in working with dementia and other neurological conditions.

Limitations

Deriving a sample size was the main difficulty encountered during this process and

subsequent issues with power was the main limitation of the study. Undertaking a

secondary analysis of previously collected data meant that further recruitment was

not possible. The first stage of cortical reconstruction was not possible for three of

the participants due to the lack of images available for processing. At the brain mask

stage, it was not possible to manually inspect the skull stripped images for one

participant as Freesurfer identified an error. A solution to Freesurfer segmentation

problems is to manually define the parameters between subcortical white matter and

deep grey matter structures (Perlman, 2007), however this task was beyond my skills

in this area and so it was decided to exclude this participant from further analysis.

The lack of available scans to match the neuropsychological dataset meant that there

were no scans available for MCI participants and this dementia subgroup were not

represented in this study. Finally, four participants for whom MRI scans had been

processed and analysed did not have available 4MT scores, as the measure was not

completed as part of the diagnostic assessment. These participants were included for

comparison with other neuropsychological measures. These developments with

availability of data were disappointing as they impacted on the sample size and

power to detect possible significant relationships between the variables. Historically,

a criticism of neuropsychological research is the use of small sample sizes, hence

studies that are statistically underpowered for the hypothesis being investigated

(Millis, 2003). These setbacks highlighted the challenges of conducting research in

clinical settings and the problems with retrospectively obtaining data for secondary analysis.

Exploring relationships between brain pathology and cognitive functions using structural MRI assumes a degree of functional specialization and a modular understanding of human cognitive functioning. The suggestion that a test like the 4MT correlates with the structural make-up of a specific brain region runs the risk of being reductionist as anatomical and functional connectivity are well established phenomena in the literature (Sporns, Chial, Kaiser & Hilgetag, 2004). It is likely that a complex range of inter-connected systems are involved in topographical disorientation (see review by Serino, Cipresso, Morganti, & Riva, 2014). Recruitment of a larger sample size would have allowed for the selection of data-driven regions of interest across diverse neuronal networks followed by statistical modelling where different structural measurements are incorporated into a single statistical model that relates to performance on the 4MT. This approach was adopted by Cook et al. (2014) for investigating the diverse neural correlates of verbal fluency in FTD.

Implications for clinical practice

Carrying out this project has highlighted the gap between experimental research and clinical practice in dementia care. There is a strong focus in the literature of identifying dementia at the preclinical and early stages of the disease and this in stark contrast with the diagnostic rates in the UK where only 46% of individuals with dementia will receive a formal diagnosis at any stage (Department of Health, 2013). The utility of spatial memory testing in differentiating dementia subtypes e.g. FTD and AD, or predicting conversion from MCI to AD has been shown to be possible in

homogenous research samples, however these studies do not account for the high prevalence of vascular disease in the dementia population. Clinical studies in western memory clinics show varying prevalence rates from 4.5 to 39 % and these figures are likely to be an underestimation as diagnostic criteria only show moderate sensitivity (approximately 50%) and variable specificity (64-98%) (Mc Aleese et al., 2016). For cognitive tests to have utility in memory clinics, tools that aid in diagnosis need to be tested on more heterogeneous groups of individuals with dementia and inclusion of VaD.

The time gap between presentation at primary care services and disclosure of diagnoses in a memory clinic setting represents a possible 'window of opportunity' for investigating progression from MCI to AD by conducting neuropsychological testing and MRI imaging at this stage. However, this methodological design raises ethical considerations. It is possible that there may be a risk of undue psychological distress associated with participation in a study investigating diagnosis of dementia without the individual and family receiving a confirmed diagnosis and access to support around this.

Early diagnosis of dementia

The process of investigating the utility of the 4MT in early diagnosis of dementia has encouraged me to reflect on the pros and cons of early diagnosis. There is a general consensus among health care professionals regarding the benefits of early diagnosis, e.g., reducing uncertainty, planning support and avoiding crises, and organising future support plans and legal arrangements (Iliffe, Manthorpe, & Eden, 2003). Early detection improves access to treatments which are indicated in the earlier stages of the disease process. NICE (2006) recommend that people with mild-to-moderate

dementia of all types should be given the opportunity to participate in a structured group cognitive stimulation programme (CST). CST has demonstrated evidence of significant improvements in measures of cognitive function and improved quality of life (Spector et al., 2003). In a qualitative study exploring the psychological impact of early diagnosis, over half of participants highlighted benefits including; appreciation and acceptance of life; less concerns about failure; self-reflection, tolerance of others, and courage to face problems in life; strengthened relationships and new opportunities to meet people (Moore et al., 2016).

There are ethical considerations with early diagnosis of dementia and concerns over accuracy of diagnostic tools, particularly in the MCI stage. Even when tests have diagnostic accuracy of approximately 90% (e.g. CSF studies), this still results in a large number of misdiagnosed persons, considering that the prevalence rate of AD in MCI cohorts is 50% (Mattson, Brax & Zetterberg, 2010). There is an increased risk of suicide in dementia, which may be linked to co-morbid mood disorders or as a result of associated stigma (Draper, Peisah, Snowdon, & Brodaty, 2010). The stigmatizing reactions of others, e.g., the individual with dementia being accused, restricted, ignored or patronized by others (Steeman, De Casterle, Godderis, & Grypdonck, 2006) is another important factor. There is a need to balance the potential benefits of research into early diagnostic tests with the risks and ethical considerations for individuals with dementia and their families.

Stigma

This project has encouraged me to reflect on the issue of stigma in dementia. This stigma is rooted in misconceptions about incapacitation and dependency (Batsch & Mittelman, 2012). Stigma has been highlighted as a key contributing factor to delays

in the diagnosis of dementia (Vernooij-Dassen et al., 2005; Burgener & Berger, 2008) and a reluctance to participate in research studies (Garand, Lingler, Connor, & Dew, 2009), with less than 4% of individuals with dementia participating in clinical research studies in the UK (Department of Health, 2012a). Furthermore, GP's perceptions of dementia have been shown to map on to conceptualisations of stigma and hinder timely diagnosis (Gove, Downs, Vernooij-Dassen & Small, 2016). National and international policy publications on dementia tend to focus on the consequences of growing prevalence and report statistics in terms of care and cost 'burden' The descriptions in these documents portray a sense of urgency and fear around dementia. The use of language in these policy documents reflects some of the dominant, stigmatizing narratives about dementia. Through my experience of working with older people, I have had conversations with clients about their fears about dementia and becoming a 'burden' on others. These reflections have highlighted for me the need to have externalising conversations to support older people and their families and carers in separating what is defined as a presenting problem from the person's identity (Morgan, 2000). Carrying out this research has caused me to reflect on my own perceptions of dementia and to challenge my own susceptibility to these widely-held assumptions.

Lived experience of dementia

As I did not meet the participants whose data is reported on in this study, as a researcher I was far removed from their lived experiences. There is a dominance of studies investigating aetiology and pathology in dementia research and considerations of the lived experience of individuals with dementia are often neglected. An exception to this is the body of research investigating quality of life in dementia, however these studies are often informed by proxy caregiver reports and

the voices and opinions of individuals with dementia are less often heard. Predictors of quality of life in dementia include the quality of relationship with the carer as perceived by the person with dementia (Clare et al. 2014). O'Rourke, Duggleby, Fraser, & Jerke (2015) asked people living with dementia about their perspectives on what affects quality of life which included connection in relationships, agency in life, wellness and a sense of belonging. While early diagnosis of dementia is important for access to evidence-based treatments, listening to the perspectives of people experiencing dementia may be more beneficial in identifying where to intervene. It may be necessary to broaden the conceptualisation of dementia as a socially-embedded phenomenon. Kitwood (1990) suggested that dementia is composed of interactions between the neuro- logical impairment and life history, health status, personality and malignant social psychology. This conceptualisation may serve to focus efforts on improving care and quality life for persons with dementia and their carers.

Conclusion

I was motivated to assess the utility of a spatial memory tool for the purposes of improving the early diagnosis of dementia. I was also keen to investigate tools which assist clinicians in detecting and treating anxiety in the dementia population. Analysing MRI data was a challenging undertaking as the main task of a DClinPsy research project. It required close working with an experienced collaborator and the learning of novel skills. The limitations of working with a pre-determined sample were further hindered by the technical challenges of the analysis. For me this project has highlighted the gap between research and clinical practice in the early diagnosis of dementia and has helped me to more closely consider the lived experience of

individuals with dementia and how this knowledge is essential in future

developments of diagnostic assessments and treatment.

REFERENCES

Bigler, E.D. (2015). Structural image analysis of the brain in neuropsychology using magnetic resonance imaging (MRI) techniques. *Neuropsychology Review, 25*, 224-249.

Batsch, L.N., & Mittelman, M.S. (2012). International World Alzheimer Report 2012. Overcoming the Stigma of Dementia. Retrieved from http://www.alz.org/documents_custom/world_report_2012_final.pdf

Burgener, S.C., & Berger, B. (2008). Measuring perceived stigma in persons with progressive neurological disease. *Dementia, 7*, 31-53.

BPS (2014). A guide to psychosocial interventions in early stages of dementia. Retrieved from https://www.bps.org.uk/system/files/user-files/DCP%20Faculty%20for%20the%20Psychology%20of%20Older%20People%20(FPoP)/public/a_guide_to_psychosocial_interventions_in_dementia.pdf.

Cook, P.A., McMillan, C.T., Avants, B.B., Peele, J.E., Gee, J.C. & Grossman, M. (2014). Relating brian anatomy and cognitive ability using a multivariate multimodal framework. *Neuroimage, 99,* 477-486.

Clare, L. Woods, R.T., Nelis, S.M., Martyr, A., Markova, I.S., Roth, I., Whitaker, C.J., & Morris, R.G. (2014). Trajectories of quality of life in early-stage dementia: Individual variations and predictors of change. *International Journal of Geriatric Psychiatry, 29*, 616-623.

Clerx, L., Gronenschild, E.H., Echavarri, C., Verhey, F., Aalten, P., & Jacobs, H.I. (2015). Can FreeSurfer compete with manual volumetric measurements in Alzheimer's disease? *Current Alzheimer Research, 12*, 358-367.

Department of Health (2012a). Prime Minister's challenge on dementia: delivering major improvements in dementia care and research by 2012: Annual report on progress. Retrieved from:https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/200030/9535-TSO-2900951-PM_Challenge_Dementia_ACCESSIBLE.PDF

Department of Health (2013). Alzheimer's Society publishes data on dementia diagnosis rates. Retrieved from: https://www.gov.uk/government/news/alzheimer-s-society-publishes-data-on-dementia-diagnosis-rates.

Draper, B., Peisah, C., Snowdon, J., & Brodaty, H. (2010). Early dementia diagnosis and the risk of suicide and euthanasia. *Alzheimer's and Dementia, 6*, 75–82.

Garand, L., Lingler, J.H., Conner, K.O., & Dew, M.A. (2009). Diagnostic labels, stigma, and participation in research related to dementia and mild cognitive impairment. *Research in Gerontological Nursing, 2*, 112-121.

Gove, D., Downs, M., Vernooij-Dassen, M., & Small, N. (2016). Stigma and GP's perceptions of dementia. *Aging & mental health, 20*, 391-400.

Gronenschild, E.H., Habets, P. Jacobs, H.I., Mengelers, R., Rozendaal, N., van Os, J., & Marcelis, M. (2012). The effects of FreeSurfer version, workstation

type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. *PLoS ONE, 7* (6).

Iliffe, S., Manthorpe, J., & Eden, A. (2003). Sooner or later? Issues in the early diagnosis of dementia in general practice: a qualitative study. *Family Practice, 20*, 376-381

Kitwood, T. (1990). The dialectics of dementia: with particular reference to Alzheimer's disease. *Ageing and Society, 10*, 177-196.

Laws, K.R. (2013). Negativland- a home for all findings in Psychology. *BMC Psychology, 1*: 2.

Mattsson, N., Brax, D., & Zetterberg (2010). To know or not to know: Ethical issues related to early diagnosis of Alzheimer's disease. *International Journal of Alzheimer's disease, 841941.*

McAleese, K.E., Alafuzoff, I., Charidimou, A., De Reuck, J., Grinberg, L.T.,…Atterns, J. (2016). Post-mortem assessment in vascular dementia: advances and aspirations. *BMC Medicine, 14* (1).

Millis, S. (2003). Statistical practices: the seven deadly sins. *Child Neuropsychology 9*, 221−233.

Moore, C., Holmes, S.D., Bardach, S., Caban-Holt, A.M., Murphy, R., Gregory, A. J. (2016). Even dementia may have a 'silver-lining': Half of MCI and early dementia patients report positive life experiences as a result of their diagnosis on the silver lining questionnaire. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 12*, 785-786.

Morgan, A. (2000). *What is narrative therapy*?:An easy-to-read introduction. Adelaide S. Australia: Dulwich Centre.

NICE (2006). Dementia: Supporting people with dementia and their carers in health and social care. Retrieved from https://www.nice.org.uk/guidance/CG42/chapter/1-Guidance#interventions-for-cognitive-symptoms-and-maintenance-of-function-for-people-with-dementia

O' Rourke, H.M., Duggleby, W., Fraser, K., & Jerke, L. (2015). Factors what affect quality of life from the perspective of people with dementia: A meta-synthesis. *Journal of the American Geriatrics Society, 63*, 24-38

Perlman, D. (2007). Cortical Thickness: Practicalities and Comparisons. Retrieved from http://brainimaging.waisman.wisc.edu/~perlman/Perlman692report.pdf

Scarpazza, C., Togin, S., Frisciata, S., Sartori, G., & Mechelli, A. (2015). False positive rates in voxel-based morphometry studies of the human brain: should we be worried? *Neuroscience and Biobehavioral Reviews, 52*, 49-55.

Seignourel, P.J., Kunik, M.E., Snow, L., Wilson, N. & Stanley, M. (2008) Anxiety in dementia: A critical review. *Clinical Psychology Review, 18*, 1071-1082.

Serino, S., Cipresso, P., Morganti, F., & Riva, G. (2014). The role of egocentric and allocentric abilities in Alzheimer's disease: A systematic review. *Ageing Research Reviews, 16*, 32-44.

Spector, A., Thorgrimsen, L., Woods, B., Royan, L., Davies, S., Butterworth M., & Orrell, M. (2003). Efficacy of an evidence-based cognitive stimulation

therapy programme for people with dementia: randomised controlled trial. *British Journal of Psychiatry*, *183*, 248-254.

Sporns, O., Chival, D.R., Kaiser, M. & Hilgetag, C.C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences, 8*, 418-425.

Steeman, E., De Casterle, B.D., Godderis, J. & Grypdonck. M. (2006).Living with early-stage dementia: A review of qualitative studies. *Journal of Advanced Nursing*, 54, 722-738.

Vernooij-Dassen, M.J., Moniz-Cook, E.D., Woods, R.T., De Lepeleire, J., Leuschner, A., Zanetti, O., & Iliffe, S. (2005).Factors affecting timely recognition and diagnosis of dementia across Europe: From awareness to stigma. *International Journal of Geriatric Psychiatry, 20*(4), 377-386.

Wegner, E., Martensson, J., Noack, H., Bodammer, N.C., Kuhn, S., & Lovden, M. (2014). Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. Human Brain Mapping, 35, 4236-4248.

World Health Organisation (2012). *Dementia: A public health priority.* Retrieved from

http://www.who.int/mental_health/publications/dementia_report_2012/en/

APPENDICES

Appendix 1: Data permission letter

Maura Scanlon
Trainee Clinical Psychologist
UCL Division of Psychology and Language Sciences
1 - 19 Torrington Place
London. WC1E 7HB

19/04/2017

Dr Joshua Stott
Research department of clinical, educational and health Psychology
University College London,
London, UK,
WC1E 6BT

Dear Dr Stott,

Running Title of Project: The clinical utility of the Four Mountains Test in the diagnosis of dementia: relationship to hippocampal atrophy.

I am writing to request permission for the use of data from the cohort of participants recruited for your study:

Stott, J., Scior, K., Mandy, W., & Charlesworth, G. (2017). Dementia Screening Accuracy is Robust to Premorbid IQ Variation: Evidence from the Addenbrooke's Cognitive Examination-III and the Test of Premorbid Function. *Journal of Alzheimer's Disease*, (Preprint), 1-10.

This data will be analysed and used as part of a thesis submission for a Doctorate in Clinical Psychology (DClinPsy) at University College London.

Yours Sincerely,

Maura Scanlon

Trainee Clinical Psychologist

Appendix 2: Letter granting ethical approval

NHS
Health Research Authority

Dear Dr Stott

Study title   **Exploring cognitive mediation ability in people with dementia: the factors that influence it and effects of difference in ability**

REC reference: **14/LO/0554**
Amendment number: **Substantial Amendment 1**
Amendment date: **17 April 2015**
IRAS project ID: **147241**

The above amendment was reviewed on 19 June 2015 by the Sub-Committee in correspondence.

Ethical opinion

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

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Interview schedules or topic guides for participants [MBAT script]		
Notice of Substantial Amendment (non-CTIMP) [including more routine clinical data, addition of mindfulness]	Substantial Amendment 1	17 April 2015
Other [Sections of original submission affected by proposed amendments]		
Participant information sheet (PIS) [Clinical PIS Stage 1 CB]	5 (clean)	10 April 2015
Participant information sheet (PIS) [Clinical PIS Stage 1 CB]	4-5 (tracked)	17 June 2015
Validated questionnaire [CAMS R questionnaire with instructions]		

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

14/LO/0554:	Please quote this number on all correspondence

Yours sincerely

pp *EAHearn*

Dr Koula Asimakopoulou Acting Alternate Vice Chair

E-mail: nrescommittee.london-cityroadandhampstead@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

Copy to: *Dr   Tumi Kaminskas, North Central London Research Consortium Portfolio Suzanne Emerton*

NRES Committee London - City Road & Hampstead Attendance at Sub-Committee of the REC meeting on 19 June 2015

Committee Members:

Name	Profession	Present	Notes
Dr Koula Asimakopoulou	Senior Lecturer in Health Psychology	Yes	Acting Alternate Vice Chair [Meeting Chair]
Mr Thomas Dowe	Tissue Collection Officer	Yes	

Also in attendance:

Name	Position (or reason for attending)
Miss Maeve Groot Bluemink	REC Manager
Dr Imran Jawaid	General Practitioner
Mr Alex Shortt	MRC Clinical Research Fellow

Written comments received from:

Name	Position
Miss Maeve Groot Bluemink	REC Manager

Appendix 3: ACE III

ADDENBROOKE'S·COGNITIVE·EXAMINATION·—·ACE-III
English·Version·A·(2012)

Name:	Date·of·testing: ..../....../.......
Date·of·Birth:	Tester's·name:
Hospital·No.·or·Address:	Age·at·leaving·full-time·education:
	Occupation:
	Handedness:

A·T·T·E·N·T·I·O·N

| Y → Ask: What·is·the | Day | Date | Month | Year | Season | **Attention**
[Score·0-5] |
| Y → Ask: Which | No./Floor | Street/Hospital | Town | County | Country | **Attention**
[Score·0-5] |

A·T·T·E·N·T·I·O·N

Y·Tell: "I'm·going·to·give·you·three·words·and·I'd·like·you·to·repeat·them·after·me:·lemon,·key·and·ball."·After·subject·
repeats, say "Try·to·remember·them·because·I'm·going·to·ask·you·later".
Y → Score·only·the·first·trial·(repeat·3·times·if·necessary).
Y → Register·number·of·trials: _____

Attention
[Score·0-3]

A·T·T·E·N·T·I·O·N

Y → Ask·the·subject: "Could·you·take·7·away·from·100?·I'd·like·you·to·keep·taking·7·away·from·each·new·number·until·
I·tell·you·to·stop."
Y → If·subject·makes·a·mistake,·do·not·stop·them.·Let·the·subject·carry·on·and·check·subsequent·answers·(e.g.,·93,·84,·
77,·70,·63 — score·4).
Y → Stop·after·five·subtractions·(93,·86,·79,·72,·65): _____

Attention
[Score·0-5]

M·E·M·O·R·Y

Y → Ask: "Which·3·words·did·I·ask·you·to·repeat·and·remember?" _____

Memory
[Score·0-3]

F·L·U·E·N·C·Y

Y → **Letters**
Say: "I'm·going·to·give·you·a·letter·of·the·alphabet·and·I'd·like·you·to·generate·as·many·words·as·you·can·beginning·with·
that·letter,·but·not·names·of·people·or·places.·For·example,·if·I·give·you·the·letter·"C",·you·could·give·me·words·like·"cat,·
cry,·clock"·and·so·on.·But,·you·can't·give·me·words·like·Catherine·or·Canada.·Do·you·understand?·Are·you·ready?·You·
have·one·minute.·The·letter·I·want·you·to·use·is·the·letter·"P".

Fluency
[Score·0 — 7]

≥18	7
14-17	6
11-13	5
8-10	4
6-7	3
4-5	2
2-3	1
0-1	0
total	correct

Y → **Animals**
Say: "Now·can·you·name·as·many·animals·as·possible.·It·can·begin·with·any·letter."

Fluency
[Score·0 — 7]

≥22	7
17-21	6
14-16	5
11-13	4
9-10	3
7-8	2
5-6	1
<5	0
total	correct

M·E·M·O·R·Y

Y → Tell: "I'm·going·to·give·you·a·name·and·address·and·I'd·like·you·to·repeat·the·name·and·address·after·me.
So·you·have·a·chance·to·learn,·we'll·be·doing·that·3·times.·I'll·ask·you·the·name·and·address·later."

Score·only·the·third·trial.

Memory
[Score·0·—·7]

	1ˢᵗ·Trial	2ⁿᵈ·Trial	3ʳᵈ·Trial
Harry·Barnes 73·Orchard·Close· Kingsbridge·Devon			

M·E·M·O·R·Y

Y···Name·of·the·current·Prime·Minister...Y···Name·of·the·woman·who·was·Prime·Minister...Y···Name·of·the·USA·president...Y···Name·of·the·USA·president·who·was·assassinated·in·the·1960s.................................

Memory
[Score·0·—·4]

L·A·N·G·U·A·G·E

Y → Place·a·pencil·and·a·piece·of·paper·in·front·of·the·subject.·As·a·practice·trial,·ask·the·subject·to·"Pick·up·the·pencil·and·then·the·paper."·If·incorrect,·score·0·and·do·not·continue·further.

Y → If·the·subject·is·correct·on·the·practice·trial,·continue·with·the·following·three·commands·below.
    • → Ask·the·subject·to·"Place·the·paper·on·top·of·the·pencil"
    • → Ask·the·subject·to·"Pick·up·the·pencil·but·not·the·paper"
    • → Ask·the·subject·to·"Pass·me·the·pencil·after·touching·the·paper"·Note:·
Place·the·pencil·and·paper·in·front·of·the·subject·before·each·command.

Language
[Score·0-3]

L·A·N·G·U·A·G·E

Y → Ask·the·subject·to·write·two·(or·more)·complete·sentences·about·his/her·last·holiday/weekend/Christmas.·Write·in·complete·sentences·and·do·not·use·abbreviations.
Give·1·point·if·there·are·two·(or·more)·complete·sentences·about·the·one·topic;·and·give·another·1·point·if·grammar·and·spelling·are·correct.

Language
[Score·0-2]

L·A·N·G·U·A·G·E

Y → Ask·the·subject·to·repeat:·'caterpillar';·'eccentricity';·'unintelligible';·'statistician'
Score·2·if·all·are·correct;·score·1·if·3·are·correct;·and·score·0·if·2·or·less·are·correct.

Language
[Score·0-2]

Y→ Ask the subject to repeat: 'All that glitters is not gold'

Language
[Score 0-1]

Y→ Ask the subject to repeat: 'A stitch in time saves nine'

Language
[Score 0-1]

Y→ Ask the subject to name the following pictures:

Language
[Score 0-12]
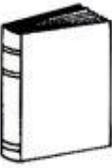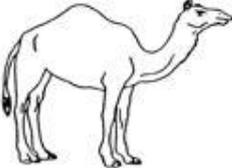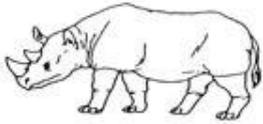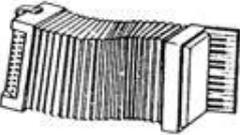
Y→ Using the pictures above, ask the subject to:

Language
[Score 0-4]

- Point to the one which is associated with the monarchy → ....
- Point to the one which is a marsupial → ....
- Point to the one which is found in the Antarctic → ....
- Point to the one which has a nautical connection → ....

L·A·N·G·U·A·G·E

Y→ Ask the subject to read the following words: (Score 1 only if all correct)

sew·
pint·
soot·
dough·
height¤

Language
[Score 0-1]

V·I·S·U·O·S·P·A·T·I·A·L··A·B·I·L·I·T·I·E·S

Y→ Infinity Diagram:···Ask the subject to copy this diagram
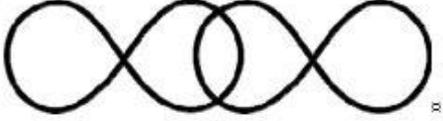
Visuospatial
[Score 0-1]

Y→ Wire cube: Ask the subject to copy this drawing (for scoring, see instructions guide).
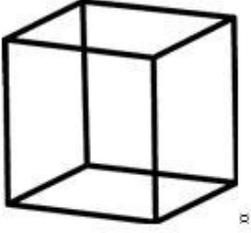
Visuospatial
[Score 0-2]

Y→ Clock: Ask the subject to draw a clock face with numbers and the hands at ten past five. (For scoring see instruction guide: circle = 1, numbers = 2, hands = 2 if all correct).

Visuospatial
[Score 0-5]

Section Break (Next Page)

Y→ Ask·the·subject·to·count·the·dots·without·pointing·to·them¤

V·I·S·U·O·S·P·A·T·I·A·L ··A·B·I·L·I·T·I·E·S

Y→ Ask the subject to identify the letters

Visuospatial
[Score 0-4]
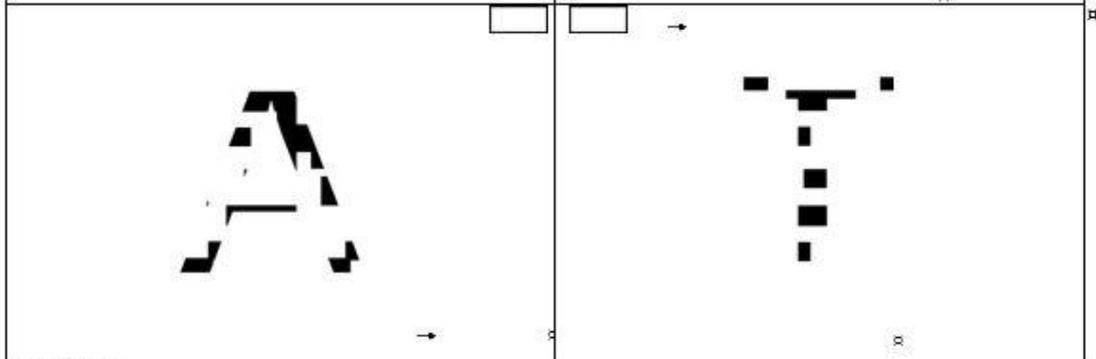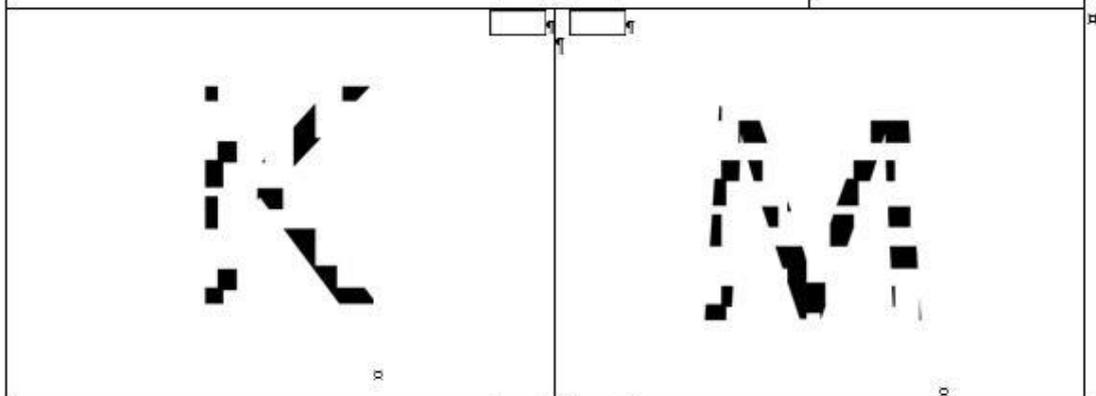
M·E·M·O·R·Y

Y→ Ask "Now tell me what you remember about that name and address we were repeating at the beginning"

Harry Barnes
73 Orchard Close
Kingsbridge
Devon

Memory
[Score 0-7]

M·E·M·O·R·Y

Y→ This test should be done if the subject failed to recall one or more items above. If all items were recalled, skip the test and score 5. If only part was recalled start by ticking items recalled in the shadowed column on the right hand side; and then test not recalled items by telling the subject "ok, I'll give you some hints: was the name X, Y or Z?" and so on. Each recognised item scores one point, which is added to the point gained by recalling.

Memory
[Score 0-5]

Jerry Barnes		Harry Barnes		Harry Bradford		recalled	
37		73		76		recalled	
Orchard Place		Oak Close		Orchard Close		recalled	
Oakhampton		Kingsbridge		Dartington		recalled	
Devon		Dorset		Somerset		recalled	

S·C·O·R·E·S

TOTAL ACE-III SCORE	/100
Attention	/18
Memory	/26
Fluency	/14
Language	/26
Visuospatial	/16