

An exploration of the experience of caregiving in relation to dementia subtype diagnosis, and of the potential utility of the Short Parallel Assessments of Neuropsychological Status (SPANS) in the assessment of older adult cognition.

Thesis submitted in partial fulfilment of the degree of
Doctorate in Clinical Psychology
(DClinPsy)
University of Leicester

By
Louise Crossley-Zels
Supervised by: Dr Gerald Burgess

April 2019
Resubmitted post-viva in September 2019

Declaration

I confirm that the research reported in this thesis is my own original work, and has not been submitted for any other academic award. Exceptions to this relate to the inclusion of additional data collected through previous research using the Short Parallel Assessments of Neuropsychological Status (SPANS), conducted by Dr Gerald Burgess and published in 2014. Where this data is included in analysis it is clearly acknowledged within the main text.

I additionally confirm that this thesis has been checked for completion prior to submission.

An exploration of the experience of caregiving in relation to dementia subtype diagnosis, and of the potential utility of the Short Parallel Assessments of Neuropsychological Status (SPANS) in the assessment of older adult cognition.

Louise Crossley-Zels

Thesis Abstract

Within an ageing population, there is continued need for timely and accurate dementia diagnosis, and provision of support for family members who increasingly provide informal care. Understanding of care recipients' cognitive profile, and greater knowledge of the experience of caregiver burden, may facilitate the provision of such support.

Literature Review

Characteristics of carer burden reported by caregivers of persons diagnosed with Alzheimer's Disease, Vascular dementia, Dementia with Lewy Bodies, and Frontotemporal dementias were reviewed. An integrative approach was taken to combine both quantitative and qualitative findings from 15 published papers. Inconclusive findings are presented. Further research incorporating a mixed-methods design and matched participant groups would enable greater understanding.

Research Report

An initial exploration of the reliability and validity of the Short Parallel Assessments of Neuropsychological Status (SPANS) was undertaken within a sample of healthy and functionally able adults over the age of 65 years. The psychometric properties of the SPANS with working-age adults has been previously documented (Burgess, 2014), and acted as a source of comparison for current findings. Internal reliability was explored, in addition to confirmation of good alternate version reliability. Convergent validity was explored in relation to the Addenbrooke's Cognitive Assessment (ACE-III), and indicated that the SPANS likely performs at least to an equivalent level, with potential to offer additional clinical benefits beyond those conveyed by the ACE-III. Findings were limited by insufficient statistical power and lack of variance within the sample, leading to the recommendation of further research to confirm and expand upon current findings.

Acknowledgements

Firstly, I would like to thank the participants who gave their time to take part in this research – thank you, for your motivation and your kindness. I would also like to thank the organisations that allowed us to advertise this study to their members and service users, and whose support was integral to the completion of this work.

To the volunteer research assistants who committed their time to supporting this research - I am hugely grateful for your efforts and dedication throughout the duration of this study.

Thanks to Dr Gerald Burgess, for your constant support and round-the-clock guidance, the thoughtfulness of your responses, and your general dedication to this piece of work and to your role as my supervisor. I hope that our paths will cross again in the future. Thank you also to Dr Inga Stewart, for your thoughts, your enthusiasm and your support of this project.

Finally, thank you to my parents for teaching me the value of hard work, and to never give up on a dream. Thank you to my husband for enduring the stressful chaos and emotional rollercoaster that have characterised the past few months, and for knowing exactly when to supply cheesecake and hugs. And thank you to all three of you for always believing in me.

Word count

Thesis Abstract	250
Literature Review	8,090
Abstract	208
Main Text	7,882
Research Report	10,250
Abstract	261
Main Text	9,989
Non-mandatory Appendices:	3,407
Total Word Count:	21,997

Note: All values exclude contents lists, tabulated numerical data, figures / diagrams and reference lists.

Contents

Part One: Literature Review	10
Abstract.....	12
1. Introduction	13
2. Method	21
3. Results.....	24
4. Discussion	36
References	43
Part Two: Research Report	51
Abstract.....	52
1.Introduction.....	53
2. Method	60
3. Results.....	70
4. Discussion	83
References	93
Appendices	100

List of Tables

Part One: Literature Review

Table 1: Characteristic profiles of dementia subtype presentations (Alzheimer's Association, 2016; Alzheimer's Society (n.d); Niu et al., 2017; Podcasy & Epperson, 2016; Schoenberg & Scott, 2011)	15
Table 2. Summary of databases searched and search terms employed.....	23
Table 3. Self-reported intensity of aspects of caregiver burden, and total burden experienced, by carers of persons with AD and VaD, as measured by the CBI (<i>M</i>, <i>SD</i>)	32
Table 4. Self-reported intensity of aspects of caregiver burden, and total burden experienced, among carers of persons with AD as measured by the CB Scale (Pudelewicz et al. (2018).....	33
Table 5: A comparison of self-reported intensity of aspects of caregiver burden, and total burden experienced, by carers of persons with AD and DLB as measured by the RSS (Svendsboe et al., 2016)	33
Table 6. A comparison of self-reported intensity of aspects of caregiver burden, and total burden experienced, by carers of persons with AD, FTD and DLB as measured by the ZBI	34

Part Two: Research Report

Table 1: Table 1: Self-reported participant demographic, health and ability characteristics	61
Table 2: Demographic characteristics of participants whose data was collected as part of previous SPANS research (Burgess, 2014) and is included where specified in the current analyses	64

Table 3: Internal consistency values (Cronbach's α) for all SPANS index scores and total score, across versions A and B for the current older adult sample, the younger sample described within the manual, and all SPANS data collected to date.....	72
Table 4: Descriptive statistics for SPANS index and total scores derived from the current older adult sample ($N = 108$)	73
Table 5: Correlation coefficients (Kendall's Tau and Pearson's r) demonstrating the relationship between performance on the two alternate versions of the SPANS (versions A and B) for each SPANS index...	74
Table 6: A comparison of SPANS index median scores across the alternate versions (A and B) using a related samples Wilcoxon's signed-rank test (M and SD reported to aid comparison to SPANS manual)	75
Table 7: Correlation (τ) between SPANS indices and total score, and ACE-III domains and total score, derived from all participants who completed both measures ($N = 54$)	78
Table 8: Correlations (τ) between SPANS indices and extracted phonemic and semantic fluency scores derived from all current participants ($N = 72$) who completed both measures.....	81
Table 9: Correlations (τ) between ACE-III domains and extracted phonemic and semantic fluency scores derived from all current participants ($N = 50$) who completed both measures.....	81
Table 10: Median SPANS and ACE-III total scores for all current participants for each level of educational achievement as reported in completion of the DHLQ.....	83

List of Figures

Part One: Literature Review

Figure 1. Suggested process of integrative review (Whittemore & Knafl, 2005).....	22
---	----

Part Two: Research Report

Figure 1: Histogram with bell curve for SPANS (a), ACE-III (b) total scores.....	76
Figure 2: Frequency distribution and bell curve of Mendez et al (1992) clock drawing score (a), and phonemic (b) and semantic (c) fluency raw scores.....	79

List of Appendices (mandatory appendices denoted by *)

Appendix A*: Identification of intended journal for publication.....	101
Appendix B: Inclusion criteria and search filters.....	102
Appendix C: Visual representation of systematic search process.....	103
Appendix D: Data extraction pro forma.....	104
Appendix E: Quality Appraisal Pro Forma.....	106
Appendix F: Quality appraisal scores summary.....	108
Appendix G: Summary of papers reviewed.....	109
Appendix H: Supplementary information regarding quantitative measures of burden employed within the reviewed literature.....	118
Appendix I: Summary of SPANS indices.....	122
Appendix J: A comparison of the SPANS with other brief screening assessments.....	124
Appendix K: Exploration of predictive validity.....	125
Appendix L*: Visual representation of the chronology of the research process.....	127
Appendix M*: Ethical approval and sponsorship confirmation.....	128
Appendix N: Ethical issues considered.....	130
Appendix O: Gatekeeper contact standard email template.....	132
Appendix P*: Demographic, Health and Lifestyle Questionnaire (DHLQ).....	133
Appendix Q: Visual representation of the research procedure.....	137
Appendix R: Participant contact sheet.....	138
Appendix S*: Participant information sheet.....	139
Appendix T*: Participant consent form.....	143
Appendix U*: Checklist to ensure anonymity of clients/services.....	144

Part One: Literature Review

Does the experience of caregiver burden differ according to dementia subtype diagnosis? An integrative literature review

(Intended journal for publication is detailed in Appendix A¹)

¹ Note that the word count and reference restrictions required for submission to the target journal were felt to be incompatible with the University's requirements of this thesis. As such, this thesis is formatted and referenced in accordance with University guidance and will be amended post-viva for publication.

Abstract

Introduction

In the context of an ageing population, there are estimated to be 850,000 people currently diagnosed with dementia in the UK. Many are supported at home by spouses or other relatives, for whom caregiving can be associated with significant levels of burden and distress. Given the differences in symptoms and demographic characteristics reflected by dementia subtype diagnoses, it may be that characteristics of carer burden also differ depending on the specific dementia diagnosis of the care recipient.

Method

A systematic search of the PsycINFO, Medline and Scopus databases was undertaken to identify literature relevant to the current review. A total of 15 papers were identified for inclusion, based upon specific criteria. These were subject to quality assessment and critical review.

Results

Findings were subject to a number of limitations and remain inconclusive. Some preliminary support for the hypothesis that characteristics of caregiver burden may differ by subtype diagnosis was identified however alternative explanations may be available. Additional shared aspects of the burdensome experiences of caregiving were also indicated. Further research is recommended to address this question in greater detail.

Discussion

The present review was subject to several limitations. Overall, findings suggest that further research in this area is warranted, and recommendations for such work are presented.

1. Introduction

1.1 Dementia

Dementia is an umbrella term for a collection of usually progressive conditions, generally occurring later in life, and characterised by a decline in cognitive function beyond that which would be expected as part of the normal ageing process (World Health Organisation, 2017). The most common dementia subtypes are: Alzheimer's Disease (AD), Vascular Dementia (VaD), Lewy Body Dementia (DLB), and Frontotemporal Dementia (FTD), which together account for around 90% of all cases of dementia in the UK (Alzheimer's Society, 2014; National Institute for Health and Care Excellence (NICE), 2017). These subtypes each have a characteristic profile, including symptom progression, and duration (summarised in Table 1). Greatest distinction between subtype presentations is evident in the earlier stages, while towards the mid- and late-stages of progression there is considerably greater overlap (Alzheimer's Society, n.d).

In 2017 it was estimated that 50 million people worldwide were living with dementia (World Health Organisation, 2017), including an estimated 850,000 people in the UK, with a cost of care of £23 billion per year (House of Commons Library, 2018).

1.2 Informal dementia caregiving

In the UK there are estimated to be 700,000 informal (generally family) carers of people with dementia (Alzheimer's Research UK, 2016a), with 1 in 3 people expected to care for a person with dementia at some point in their lives (Department of Health, 2013). While the number of male caregivers is increasing (Robinson et al., 2014), currently around 66% of primary informal dementia caregivers are female. Women in the UK are more than twice as likely as men to care for someone with dementia over an extended period, or to provide 24-hour care (Alzheimer's Research UK, 2015; Global Alzheimer's and Dementia Action Alliance, 2017).

1.3 Carer burden

For many, this caring role is associated with significant physical, financial and psychosocial stresses, the experience of which has been termed 'carer burden' (Adelman et al., 2014; D'Onofrio et al., 2015; Etters et al., 2008). To date, research into

carer burden has been multifaceted, with much focus on the degree of burden experienced, patient and caregiver-related factors which are predictive of the presence, or greater levels, of burden (e.g. Etters et al., 2008; Lou et al., 2015), and factors associated with

Table 1: Characteristic profiles of dementia subtype presentations (Alzheimer's Association, 2016; Alzheimer's Society (n.d); Niu et al., 2017; Podcasy & Epperson, 2016; Schoenberg & Scott, 2011)

Diagnostic subtype	Characteristic profile	Typical age of onset (years)	Gender differences	Average post-diagnostic life expectancy (years)
AD	Gradual onset. Early stages are characterised by memory difficulties and impairment in at least one other cognitive domain. Progression may include emotional lability, increasing confusion, agitation and restlessness. Later stages often see difficulties in communication, orientation, judgement, and speech, in addition to delusions and hallucinations Slow and progressive nature.	65 +	Up to twice as common among women as men	10
VaD	Sudden or gradual onset. Early stages often characterised by impaired decision-making and organisation, in addition to some recall difficulties or other cognitive impairments. Slowed thought processes, apathy and attentional difficulties may also be noted. Falls linked to changes in gait and balance are common features. Stepwise progression, with confusion, disorientation and delusions emerging over time	60-75	Slightly more common in men	Variable
DLB	Characterised by early sleep disturbances, hallucinations, and parkinsonian movement changes (often leading to falls), in addition to visuo-spatial and attentional difficulties often as the main initial signs of cognitive impairment. Fluctuations in cognitive ability and level of alertness are present. Agitation, restlessness and worsening cognitive impairment (which may broaden to include memory and language difficulties) are seen as the disease progresses.	50-70	Up to 3 times more common among men	5-7
FTD	Umbrella term for a collection of diagnoses, of which the behavioural variant (bvFTD) is included here. This is characterised by early changes in personality or behaviour impacting social functioning. This can include apathy, withdrawal, and reduced empathy. Memory impairment typically emerges later in the progression, however executive functioning impairment may be seen relatively early in the disease course, along with repetitive or compulsive behaviours	50-60	3-4 times more common among men	5-7

more successful coping (e.g. Etters et al., 2008), or effective intervention (e.g. Jackson & Browne, 2017).

This is unsurprising given that research has indicated significant negative correlates of carer burden and distress, ranging from links to likelihood of long-term care placement (Cepoiu-Martin et al., 2016), through to links with depression and suicidal ideation among caregivers (Etters et al., 2008; O'Dwyer et al., 2016).

1.3.1 Caregiver factors associated with burden

The literature indicates that carers who live with the care recipient describe significantly higher levels of burden compared to those who live externally (Raccichini et al., 2015). This remains the case when mediating factors such as time spent caring, and the level of cognitive or functional impairment of the person with dementia, are held constant (Raccichini et al., 2015).

It has been demonstrated that female caregivers report greater levels of burden than their male counterparts when groups are matched for factors such as relationship to the care recipient, financial status, and duration of caring role (Akpınar et al. 2011; Thompson et al., 2004). Furthermore, personality traits of caregivers have been found to mediate the impact of the caring role, with neuroticism and extraversion being linked to self-reported health-related quality of life, burden and depression (Kim et al., 2017; Melo et al., 2011).

It seems that greater burden is experienced by caregiving offspring than spouses, with greater guilt reported by care-recipients' children in response to caregiving (Melo et al., 2011). This trend may appear independently of patient factors such as level of functional or cognitive impairment (Chappell et al., 2014; Conde-Sala et al., 2010). One possible understanding of this could reflect the different external stressors associated with caregiver age and relationship to the person with dementia (PwD). For example, spousal caregivers of those with FTD diagnoses (which typically sees a younger age of onset), or adult-child carers of any dementia may be working and raising a family in addition to their caregiving role.

It may be hypothesised that the demographic composition of caregivers may vary between dementia subtypes, and that dimensions of carer burden associated with caregiver demographics could therefore also vary in a broad sense between subtype diagnoses. For example, dementias such as DLB, which disproportionately impact men, may see a greater number of female caregivers, and therefore the qualitative experiences of burden described in the literature may in part be reflective of gendered differences in carer burden experience.

1.3.2 Patient factors associated with burden

Varied conclusions are presented in the literature with regard to the impact of cognitive impairment upon levels of carer burden, with authors such as Ferreira de Oliveira et al. (2015) report no relationship between cognitive functioning and burden. Conversely, researchers such as Dauphinot et al. (2015) have indicated that cognitive impairment was more closely linked to carer burden than dementia subtype. Such mixed findings may indicate that other factors perhaps moderate the relationship between degree of cognitive impairment and severity of caregiver burden experienced.

Evidence does however suggest a strong impact of impaired ability to complete activities of daily living (ADLs; Schoenmakers et al., 2010) upon the level of burden experienced, which itself may be related to the degree of cognitive impairment. ADL independence also interacts with presence and severity (Cheng, 2017) of behavioural and psychological symptoms of dementia (BPSD, see below). As such, it may be that with increasing duration and disease severity will come a greater level of support required, and therefore greater levels of burden experienced by informal caregivers. While disease profiles appear more distinct in the earlier stages, progression brings greater overlap (Alzheimer's Society, n.d), which, it could be hypothesised, may create more similar experiences of burden in later disease stages. However, it must also be considered that the experience of burden in the later stages may also reflect what has come before – for example, burden experienced by those caring for PwD's with late-stage AD may have been in the caring role for longer than those caring for PwD's with late-stage FTD.

The literature has been clearer regarding the impact of non-cognitive symptoms of dementia upon carer burden. Often referred to as 'behavioural and psychological symptoms of dementia' (BPSD), this encompasses symptoms such as apathy, agitation, delusions, hallucinations, depression and disinhibition (e.g. Cerejeira et al., 2012). Such elements are present in up to 90% of dementia cases, and are related to negative outcomes such as carer distress, reduced quality of life among both patients and carers, and need for long-term care placement (Cerejeira et al., 2012; Finkel, 2000; Burns & Rabins, 2000). Links between BPSD severity and level of carer burden in those dementia subtypes where BPSD may be considered an inherent feature of the disease profile (DLB, and FTD) have been indicated (Lee et al., 2013; Liu et al., 2017).

In keeping with the literature (Liu et al., 2017; Raccichini et al., 2015) it is clear that patient-related factors associated with the presence and degree of carer burden vary in their prominence and presentation, both between dementia subtype diagnoses, and within each subtype diagnosis over the course of the disease progression (see Table 1).

1.3.3 Previous literature reviews

Reflecting the available literature, previous reviews in this area have cited heterogeneity in the approaches to research, measures used, and reporting of findings, as resulting in difficulty in achieving comparison or integration of results (e.g. Feast et al., 2016; Terum et al., 2017). The majority of reviews have focussed upon patient or carer factors predictive of the level of carer burden reported (e.g. Cheng et al., 2017; Etters et al., 2008; Feast et al., 2016; Nunnemann et al., 2012; Orgeta & Leung, 2015; Schoenmakers et al., 2010; Terum et al., 2017), rather than qualities of the burden experience.

Evidence for the impact of the degree of cognitive impairment upon carer burden appears mixed, with Etters et al. (2008) concluding little impact on the basis of the studies reviewed, and suggesting instead that diagnostic subtype and relationship to the care recipient appeared to be a more important factors among the studies reviewed – where female carers and closeness of kinship appeared to correspond with greater levels of reported burden. The authors highlight the gendered differences in

coping strategies evident within the studies reviewed, therefore indicating that gender differences in the level of carer burden expressed may be based in personality or coping mechanisms. Further support for the role of caregiver personality and coping style can be seen in the review work by Orgeta and Leung (2015), and Merrilees (2016), who suggests that carers who demonstrate higher levels of extroversion and agreeableness report lower level of burden than those who demonstrate high levels of neuroticism. It may be hypothesised that these factors mediate the impact of patient-related factors such as degree of cognitive impairment, and would cut across diagnostic subtypes, therefore possibly also mediating any diagnostic trends in the burden experience.

Furthermore, Merrilees (2016) indicates that guilt and the impact of caring upon the carer's life were more prominent aspects of burden among younger adult caregivers, compared to their older adult counterparts. This may correspond to the importance of kinship ties noted by Etters and colleagues (2008), and could give rise to hypotheses regarding the predominant demographic characteristics of carers, which may vary by subtype diagnosis, and if so could contribute to differences in qualitative experiences of burden for carers of different dementia subtype diagnoses.

Some references to qualities of carer burden have been indicated. Etters et al (2008) highlight differences in the subjective appraisals of burden made by those caring for persons with AD and FTD diagnoses, respectively. They cite work by de Vugt et al. (2006) which reported greater impact upon personal life for AD carers, and lower levels of satisfaction with the self as a caregiver among FTD carers. This led these authors to suggest that diagnostic subtype may be a useful focus for future work in this area.

Lastly, with regard to BPSD, it has been difficult to establish which BPSD are most associated with caregiver distress and burden, with met-analytic studies reaching mixed conclusions (Feast et al., 2016). Terum et al (2017) reported that individual BPSD symptoms make unique and unequal contributions to level of carer burden, perhaps due to mediating effects of other factors (such as carer personality and coping style) as discussed above. In their review, Cheng (2017) cite research by Bass et al (2012), in which cognitive, functional (ADL) and behavioural (BPSD) symptoms of

dementia were explored in terms of their contribution to dimensions of caregiver burden as measured by the Zarit Burden Inventory (ZBI; Zarit et al., 1980). These dimensions of burden corresponded to relationship strain, social isolation strain, emotional strain, and physical health strain. It was found that relationship strain was primarily impacted by behavioural symptoms, while social isolation strain was most impacted by functional, followed by cognitive, and then behavioural difficulties (Bass et al. 2012). Emotional strain was most associated with behavioural, followed by cognitive and then functional aspects, while physical strain was equally associated with behavioural and functional difficulties. It could therefore be hypothesised that the presentation of the person with dementia may impact upon the characterisation of the experience of burden.

1.4 Aims and hypotheses of the current review

Given the distinct profiles of each dementia subtype, it is hypothesised that unique challenges may be faced by caregivers of different dementias, reflective of the initially unique symptom combinations faced, and of any demographic trends which may exist in relation to caregiver profiles for each subtype. It is therefore hypothesised that dementia subtype diagnoses, viewed as reflecting the combination and interaction of these characteristics, may be associated with qualitatively different experiences of burden. Furthermore, diagnostic subtype may further reflect social elements of burden relating to the level of knowledge, familiarity and understanding of the public which varies by diagnosis (Alzheimer's Research UK, 2016b), and could result in more or less accommodative responses to difficulty, and impact upon feelings of being understood, or upon availability of peer support and a feeling of connectedness with others facing similar difficulties. The current review therefore aimed to explore the possibility of differences in the characterisation and experiences of caregiver burden among those caring for the most common dementia subtypes (AD, VaD, FTD, DLB).

Identification of such differences would be important in guiding the provision of appropriate interventions for caregivers, in order to address the negative impact of carer burden upon caregivers and patients alike. A summary of findings will be presented, along with recommendations for future research in this field.

2. Method

2.1 Review process

The current review was undertaken in line with guidance set out by Whitemore and Knafl (2005) for integrative review work, the stages of which are summarised in Figure 1.

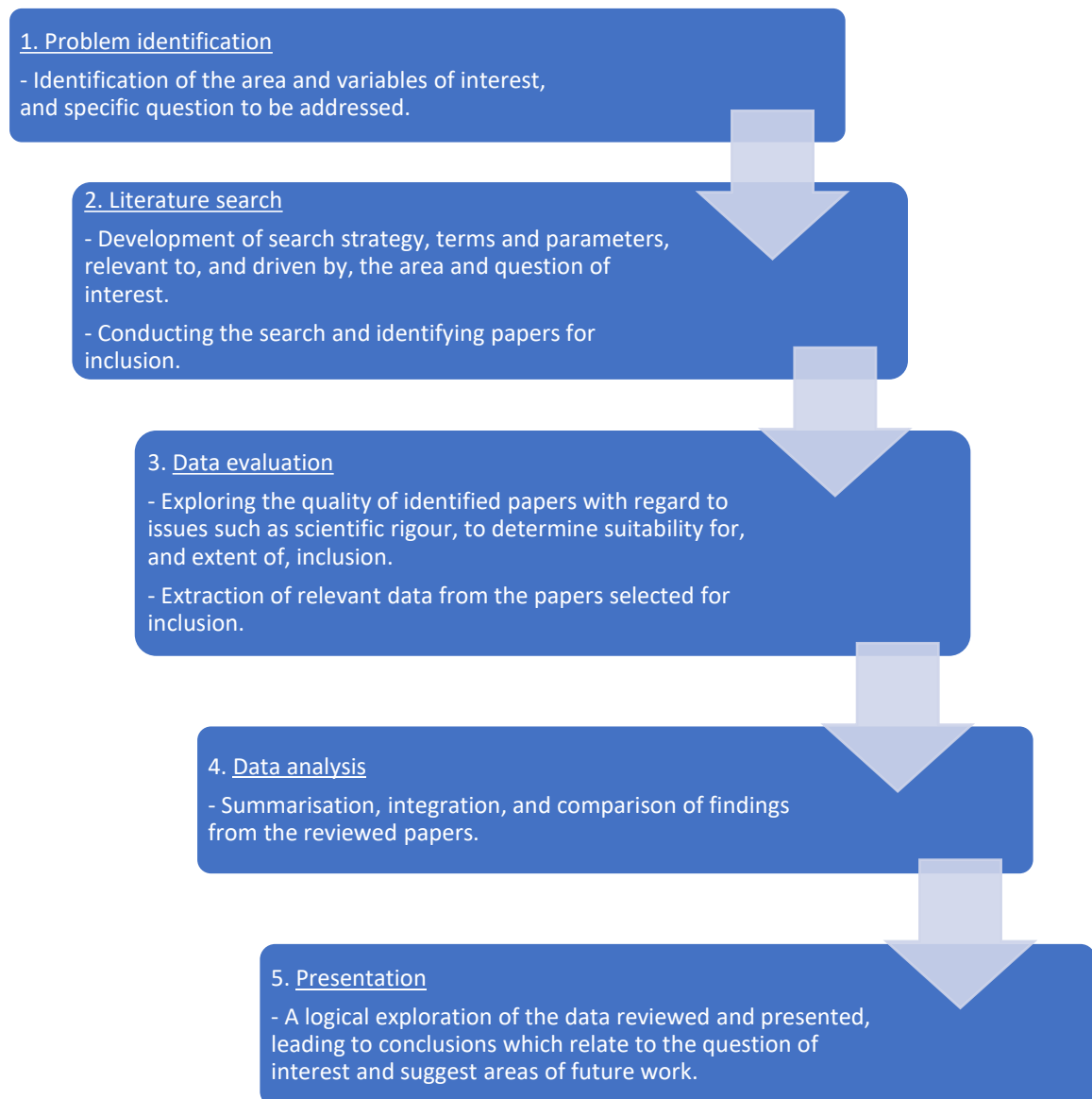
2.2 Systematic search

A systematic approach was taken to identify literature relevant to the focus of the current review. Three databases were selected as appropriate search locations: PsycINFO, which is specific to the field of psychology; Scopus, which incorporates both psychological and wider multidisciplinary resources; and Medline which includes articles from the medical field.

An initial scoping search, using the above databases and Google Scholar was conducted in July 2018, to refine the area of interest, and develop accurate search terms (Table 2) from the existing literature. A second scoping search was conducted in August 2018 using amended search terms and parameters. This was extended to include the Cochrane Database of Systematic Reviews to maximise the likelihood of identifying existing reviews within this subject area, and none were found.

Finally, a systematic search was completed between 22nd November 2018 and 6th December 2018, for relevant articles published since 2008 (see Appendix B for inclusion criteria). This yielded 9,319 results, which were first subject to the application of relevant filters (Appendix B) before removing duplicates and screening titles. Abstracts, and subsequently full articles, were screened in accordance with pre-determined inclusion criteria (Appendix B). Any articles not meeting the criteria were removed at each successive stage of the process. This resulted in the identification of 15 articles for inclusion. The references of each identified article were explored, from which no further suitable studies were found.

Figure 1. Suggested process of integrative review (Whittemore & Knafl, 2005)



A search of Google Scholar was also conducted in order to identify any potential grey literature, however this did not yield any unique results of relevance to the current review. For a visual representation of the search process, see Appendix C.

Table 2. Summary of databases searched and search terms employed

Databases	Search Term
PsycINFO	(dementia* OR Alzheimer*)
Scopus	AND
Medline	car*
	AND
	(burden OR *stress)
	AND
	(experience* OR character* OR 'subjective apprais*' OR meaning OR narrative)
Google Scholar	'Experience of carer burden dementia'
	'Appraisal of carer burden dementia'
	'Subjective experience carer burden dementia'
	'Carer burden characterisation dementia'

2.3 Quality assessment and data extraction

Data extraction was completed using a pro forma designed to specifically elicit details of relevance to the current review (Appendix D).

Of the 15 articles identified for inclusion, 14 were assessed using a specifically designed quality appraisal tool (Appendix E). Items were informed by the Critical Appraisal Skills Programme (2018a, 2018b) checklists for qualitative and quantitative studies, as well as criteria outlined by Hawker et al (2002). One article (Wegierek, 2012) was unsuitable for this process, taking the form of a narrative account rather than a research study. A percentage score was calculated for each paper, which demonstrated that all except one fell within a range of scores from 85.7% - 100% (Appendix F). One study (Nicolaou et al., 2010) scored below this range (78.6%), and demonstrated a lack of information felt to be of specific importance to the current review, therefore reducing the potential for comparison with other papers. This study was retained for inclusion, however was limited to only one area of the final review.

Overall it was felt that the range of quality scores among the remaining papers represented only marginal differences, and as such all were weighted equally for inclusion.

For the purpose of the current review, it was felt that two particular aspects of each article – reporting of participant characteristics, and use of appropriate data collection methods – were of specific relevance to the review question. Predetermined participant characteristics were searched for within the identified papers. All papers provided fully or mostly sufficient information, with the exception of Nicolaou et al. (2010), described above, and Oyeboode et al. (2013). With regard to the use of appropriate and clearly described data collection methods and tools, all papers were found to be fully or mostly sufficient, with the exception of Nicolaou et al. (2010). One paper (Zucchella et al., 2012) was found to be fully concordant across both of these criteria, scoring 100% overall. With this in mind, this study perhaps delivered the most robust contributions to the current review.

2.4 Approach to review

An exploratory approach was taken, with emphasis placed upon developing a rudimentary understanding of possible differences in the characterisation of carer burden experienced by those caring for persons with the most common dementia subtype diagnoses. It was aimed that this work could inform further research in this field. In the case of many of the studies reviewed here, additional areas such as correlates or predictors of burden were also examined in terms of the level of burden with which they were associated. It was outside of the aims and scope of the present exploratory review to incorporate these elements, however their contribution to carer burden has been reviewed and explored within the literature previously and a summary of previous findings has been provided in the introduction.

3. Results

3.1 Overview

A summary of sample characteristics and key findings of the reviewed papers can be found in in Appendix G.

A total of 3,477 caregivers took part in the studies reviewed, however large variation in the individual sample sizes of each study was noted. In accordance with the present inclusion criteria, the majority of all caregiver participants were relatives of the person for whom they were caring, largely spouses ($N = 1894$) or children ($N = 1314$). A small number of non-family caregivers ($N = 64$) were also represented across two studies (D'Onofrio et al., 2015; Galvin et al., 2010). A significant proportion of carers lived with the person with dementia ($N = 1537$), however slightly more than half of the total caregivers were living separately, or living arrangements were not disclosed (D'Onofrio et al., 2015; Iavarone et al., 2014; Nicolaou et al., 2010; Svendsboe et al., 2016).

The number of hours spent caregiving varied largely, from a mean of 2.5 hours per week (Nicolaou et al., 2010) up to a mean of 17.9 hours per day (Bursch & Butcher, 2012). Finally, there was observed to be significant variation in disease stage and severity of the persons with dementia (mild through to severe), accompanied by variable reporting of this information, and variation in measures used to determine severity or stage.

All quantitative papers employed measures of caregiver burden for which psychometric properties have been previously explored and documented. Measures used were: the Caregiver Burden Interview (CBI; Novak & Guest, 1989); the Zarit Burden Inventory (ZBI; Zarit et al., 1980); the Relative Stress Scale (RSS; Greene et al., 1982); and the Caregiver Burden Scale (CB Scale; Elmståhl et al., 1996). A description of these can be found in Appendix H.

3. 2 Alzheimer's Disease

A summary of quantitative findings is provided in Tables 3, 4, 5 and 6 (below), while qualitative themes arising from the literature are summarised in Appendix G.

Using the CBI, two studies (Iavarone et al., 2014; Raggi et al., 2015) reported identical patterns of burden. This was characterised by greatest burden arising from 'time strains', followed by 'developmental' and then 'physical' strains, with least burden arising from 'emotional' and 'social' strains. Authors found similar levels of total burden, placing their caregivers in the 'in need of respite' range, in the context of similar disease duration and severity (moderate to severe). A third study (Zucchella et

al., 2012) reported an almost identical pattern, but with social strain proving more burdensome than emotional strain. Participants in this cohort reported low levels of total burden, and were caring for family members in the early stages of AD, with disease severity in the mild range.

Pudelewicz and colleagues (2018) used the CB Scale to explore caregiver burden among a cohort of AD carers with similar dementia severity (moderate) to those in the studies by Iavarone et al. (2014) and Raggi et al. (2015), described above. These authors identified a 'medium' level of overall burden, with scores indicating greatest feelings of burden in the 'general strain', 'social isolation' and 'disappointment' domains. It appeared that 'environmental strain' was least burdensome for their participants. Examination of the content of each subscale indicated some commonalities between the CB Scale's indices of 'disappointment' and 'isolation', and the CBI's 'developmental' domain. Further similarities were seen between the CB Scale's 'general strain' index, and the CBI's 'time strain' and 'physical strain' domains – those which caregivers in the studies by Iavarone et al. (2014) and Raggi et al. (2015), also identified as most burdensome.

Contrasting findings were reported by D'Onofrio et al. (2015), whose cohort described the greatest overall burden of those reviewed here, using the CBI. Carers reported that 'physical strain' was most burdensome, with 'time strain' appearing least burdensome for this cohort, and all indices resulting in notably increased indications of strain in comparison to the other studies described using this measure. PwDs in this cohort were moderately impaired, making them similar to those of Raggi et al. (2015) and Pudelewicz et al. (2018), with relatively short disease duration but high daily hours of care received.

Lastly, research by Svendsboe et al. (2016), whose caregivers reported low levels of burden using the RSS, indicated that 'emotional' and 'social' aspects of distress were more burdensome for them than 'negative feelings'. These concepts had appeared less burdensome to Zucchella et al.'s (2012) cohort (above), despite a similar degree of overall carer strain, as well as similarities in dementia severity, and disease duration. However, it may be possible that while the constructs of the RSS share similar labels to those of the CBI, the underlying concepts may differ.

Understanding of these experiences of caregiving was further elaborated by qualitative work. Here, themes which had some links to the developmental element of the CBI, and the isolation and disappointment elements of the CBS, were evident and extended beyond that which could be measured by these scales. More than a feeling of missing out on life as it should be, caregivers described a loss of personal identity in the context of an all-consuming experience of caregiving demands:

"I felt devastated more and more as my life became hers... over time my body became hers. She needed my hands, my legs, eventually my mind, with nothing left for me" (Wegierek, 2012 p.4564.)

A similar theme, 'experiencing loss of authenticity after assuming the role of caregiver' was identified by Bursch and Butcher (2012), where caregivers reported: "I am grieving my loss of self" (p.212.)

Välimäki et al. (2012) described this within themes of 'living in a tunnel' which could give rise to the negative feelings and emotional elements of burden that had appeared less important when exploring the quantitative literature alone. Feelings of guilt, anger and resentment emerged:

"Is a feeling of guilt one of the most prominent feelings of a caregiver?" (Välimäki et al., 2012, p.483.)

"...It makes me angry because I don't think it's fair for me to do this..." (Shim et al., 2012, p.225.)

Caregivers reported feeling conflicted as a result of their changing role, and with it their changing relationship to the care recipient:

"I hate being his mother and he dislikes it even more" (Bursch & Butcher, 2012, p.212.)

"I can also no longer be his daughter, now that I've become his caregiver against my own will" (Välimäki et al., 2012, p.484.)

The loss of the relationship as it had been was additionally described:

"J. truly does not like having me talk to him and says so.... The isolation is killing me. I feel a terrible void." (Bursch & Butcher, 2012, p.211.)

Finally, uncertainty appeared as a significant aspect of the experience. For many caregivers, this was fear of what the future may hold (Wegierek, 2012; Bursch & Butcher, 2012; Välimäki et al., 2012), evidenced by thoughts such as: "...what is left? Pain for sure, sorrow, loss – the worst is yet to come..." (Bursch & Butcher, 2012, p.214.)

3.3 Vascular Dementia

D'Onofrio and colleagues (2015) also collected CBI data from carers of persons diagnosed with VaD. These caregivers reported overall levels of burden which placed them in the 'at risk of burnout' category. This was underpinned by highest scores in the 'physical burden' domain, with moderate burden also reported in the 'developmental', 'social', and 'time-related' domains. The 'emotional' domain appeared least burdensome for this group. Care recipients in this study fell within the 'mild' dementia severity range, with a relatively short associated duration of caregiving, but with a higher number of daily hours of caregiving in comparison to other studies reviewed.

No qualitative works were available for the diagnosis of VaD, however a summary of quantitative findings can be found in Table 3.

3.4 Dementia with Lewy Bodies

Leggett et al. (2010) employed the 12-item ZBI in their exploration of burden among DLB caregivers who reported a severe level of burden. Three factors were described, with the third factor, 'worry about performance', carrying the greatest burden. This was followed by 'role strain', and finally 'personal strain'. This cohort was similar in duration since diagnosis, severity of dementia, and percentage of caregivers living with the PwD to work by Galvin et al. (2010). Galvin and colleagues used a shorter, 11-item version of the ZBI in their exploration of the experience of burden among caregivers of persons with moderate-to-severe DLB. Severe levels of total burden were reported, however factors of burden within the scale were not explored. For the purposes of comparison, responses from this cohort were mapped onto the factor structure defined by Leggett et al (2010). This indicated that within Galvin et al.'s (2010) study,

performance strain items were again most burdensome, followed by role strain, and finally personal strain which was notably less burdensome overall.

The RSS has also been used (Svendsboe et al., 2016) to explore burden among carers of individuals with 'mild' DLB, and a shorter time since diagnosis than the studies described above. This group reported an overall level of burden falling within the 'low risk' range. Greater burden was perceived in the 'emotional distress' and 'social distress' domains, than in the 'negative feelings' domain.

No qualitative literature was identified in relation to DLB caregiving. Quantitative data are summarised in Tables 5 and 6.

3.5 Frontotemporal dementia

A recent study by Besser and Galvin (2018) involved caregivers for those with FTD diagnoses ranging from mild to terminal in severity (with the majority of patients in the moderate to severe range). Within this group, total levels of strain fell within the 'severe' range (Table 6). A three-factor structure was described using the 12-item ZBI, in which 'role strain' appeared the most burdensome factor, followed by 'performance strain', and finally 'personal strain'.

A qualitative exploration of caregiving experiences in frontal variant FTD (fvFTD) was conducted by Oyeboode et al. (2013), looking broadly at caregiving experience, and as such incorporating negative aspects including carer strain. Similarities to the above concept of 'role strain' emerged, with the caring role described as being one of wearing many hats, in the context of conflicting roles and emotions:

"I'm treating him like a child three-quarters of the time, and then I've got to reverse my roles to be a wife again, and a lover, and one thing and another" (Oyeboode et al., 2013, p. 162).

Within the theme of 'defending, asserting and explaining', the need to advocate for the recipient was addressed. Quotes provided here also seemed to indicate a belief that more could, or should, be done: "I should be more forceful... I should keep pestering..." (p. 160). This again appeared to partially reflect the concept of 'role strain', the importance of which was highlighted by Besser and Galvin (2018).

Experiences of sorrow and grief in relation to the perceived loss of the person they once knew, and the loss of imagined futures, were also captured. Quotes here gave rise to a sense of isolation *within* the relationship, and detachment from a shared past which is no longer shared:

“[She is] nothing like the person I married...” (pp.161.)

A summary of the themes identified by Oyebode et al (2013) is presented in Appendix G.

3.6 Comparing subtype diagnoses

In their comparison of AD and VaD caregivers (Table 3), D’Onofrio et al. (2015) found significantly higher levels of burden among the AD group across all factors of burden except ‘developmental’, for which no significant difference was found. While both groups identified ‘physical’ aspects as most burdensome in their experience, ‘emotional’ aspects were least distressing for VaD caregivers, while ‘time-related’ burden was least distressing for the AD caregivers (D’Onofrio et al., 2015). Upon first glance, it appears that VaD caregivers may experience a unique pattern of burden as measured by the CBI in comparison to AD. However, AD caregivers in this study also demonstrated a unique pattern of burden in comparison to other studies reviewed, indicating that this should be viewed with caution and may be an artefact of cohort characteristics. Groups were not matched in relation to the severity of cognitive impairment or stage of dementia. AD participants demonstrated significantly greater degree of impairment as measured by the Mini-Mental State Examination (Folstein et al., 1975), and severity stage as measured by the Clinical Dementia Rating (CDR; Morris, 1993), than their VaD counterparts (for both, $p < .001$).

Caregiver burden among those caring for AD patients has also been compared to FTD caregivers (Table 6) in Nicolaou et al.’s 2010 paper using the ZBI. Those caring for individuals with a diagnosis of FTD reported proportionately greater levels of carer burden than those caring for someone with AD. However, both groups reported relatively higher levels of ‘personal strain’ than ‘role strain’ as underpinning their experience of burden (Nicolaou et al., 2010). Statistical exploration of group means was not reported. Information regarding severity of cognitive impairment was not

provided, therefore time since diagnosis has been included in Table 6 as an alternative indication of group differences in disease severity.

Lastly, a comparison of caregiver strain among those caring for DLB and AD groups (Table 5) was provided by Svendsboe et al. (2016), who employed the RSS to measure and characterise caregiver burden among their participants. Groups were matched for age, dementia severity, and severity of cognitive impairment. Carers of the DLB group reported significantly higher levels of total burden and greater levels of burden across all subscales of the RSS (at the $p = .05$ level and above). Both groups reported greater burden in relation to 'emotional' and 'social' distress than in relation to 'negative feelings'.

A comparison of FTD and DLB was achieved through comparison of the findings of Besser and Galvin (2018) and Leggett et al. (2010), summarised in Table 6. This indicated that carer burden may have a similar, although not identical, composition for these groups when measured using the ZBI. These studies reported an almost identical factor structure (see Appendix G) and little difference in responses to the most highly scoring dimensions of 'performance' and 'role' strain. Slightly higher 'performance strain' was seen in Leggett et al.'s (2010) cohort of DLB caregivers, while slightly higher 'role strain' in Besser and Galvin's (2018) FTD cohort, with 'personal strain' appearing least burdensome to both groups.

Table 3. Self-reported intensity of aspects of caregiver burden, and total burden experienced, by carers of persons with AD and VaD, as measured by the CBI (M, SD)

Dementia subtype	N	MMSE	CDR	CBI Time	CBI Developmental	CBI Physical	CBI Social	CBI Emotional	CBI Total
AD (Iavarone et al., 2014)	86	-- ^a	-- ^a	12.25	8.25	6	5.5	4.75	37.05 (20.37)
AD (Raggi et al., 2015)**	73	16 (11 – 19)	2 (1-2)	11 (6-15.5)	7 (4 – 13.5)	4 (2-11)	3 (0.5-5.5)	3 (1-5)	33 (16.5-47.5)
AD (Zucchella et al., 2012)	126	22 (2.86)	-- ^a	33% ^b	25% ^b	17% ^b	12% ^b	13% ^b	19.95 (15.96)
AD (D'Onofrio et al., 2015)	253	18.48 (3.83)	1.32 (0.51)	12.04* (3.10)	12.63 (2.60)	14.64* (2.16)	12.90* (2.64)	12.50* (2.04)	64.71* (11.45)
VaD (D'Onofrio et al., 2015)	253	20.46 (3.20)	1.15 (0.50)	11.48 (3.14)	12.46 (2.60)	13.44 (2.78)	12.21 (2.55)	11.24 (2.55)	60.83 (12.86)

^a Statistics not reported; ^b Percentage of total reported burden; *Indicates CBI scales for which differences between AD and VaD caregiver mean scores (D'Onofrio et al., 2015) were statistically significant at the $p < .05$ level or above; **Reported data as median and interquartile range

Table 4. Self-reported intensity of aspects of caregiver burden, and total burden experienced, among carers of persons with AD as measured by the CB Scale (Pudelewicz et al. (2018))

Study	<i>N</i>	Disease severity	General strain <i>M</i> (SD)	Isolation <i>M</i> (SD)	Disappointment <i>M</i> (SD)	Emotional Involvement <i>M</i> (SD)	Environmental <i>M</i> (SD)	Total <i>M</i> (SD)
Pudelewicz et al. (2018)	55	Mild – severe*	3.0 (0.59)	2.7 (0.76)	2.7 (0.66)	2.4 (0.58)	1.8 (0.5)	2.5 (0.73)

*Majority = moderate

Table 5: A comparison of self-reported intensity of aspects of caregiver burden, and total burden experienced, by carers of persons with AD and DLB as measured by the RSS (Svendsboe et al., 2016)

Dementia subtype	<i>N</i>	CDR <i>M</i> (SD)	MMSE <i>M</i> (SD)	RSS emotional <i>M</i> (SD)	RSS social <i>M</i> (SD)	RSS negative feelings <i>M</i> (SD)	RSS total <i>M</i> (SD)
AD	100	4.9 (2.0)	23.6 (2.4)	5.9 (4.2)*	5.7 (4.6)*	3.1 (2.0)	15.0 (9.7)*
DLB	86	5.4 (2.9)	23.4 (3.2)	9.03 (5.3)	8.1 (5.6)	3.5 (2.2)	19.9 (11.2)

*Indicates RSS subscales for which differences between AD and DLB caregiver mean scores were statistically significant at the $p = .05$ level or above.

Table 6. A comparison of self-reported intensity of aspects of caregiver burden, and total burden experienced, by carers of persons with AD, FTD and DLB as measured by the ZBI

Study	<i>N</i>	Months since diagnosis <i>M</i> (<i>SD</i>)	Personal strain <i>M</i> (<i>SD</i>)	Role strain <i>M</i> (<i>SD</i>)	Performance strain <i>M</i> (<i>SD</i>)	Total <i>M</i> (<i>SD</i>)
AD (Nicolaou et al., 2010)	30	45.4 (24.5)	23.2 (6.6)	14.1 (5.2)	-- ^a	46.0 (13.4)
FTD (Nicolaou et al., 2010)	30	38.8 (23.1)	19.9 (9.8)	10.9 (6.3)	-- ^a	39.3 (17.8)
FTD (Besser & Galvin, 2018)	674	56.4 (42)	<i>M</i> = 1.88 ^b	<i>M</i> = 2.67 ^b	<i>M</i> = 2.39 ^b	27.8 (8.0)
DLB* (Galvin et al., 2010)	962	79.2 (48)	Least endorsed items	Most endorsed items	Second most endorsed items	37.6 (8.8)
DLB (Leggett et al., 2010)	611	72 ^b	<i>M</i> = 2.47 ^b	<i>M</i> = 3.33 ^b	<i>M</i> = 3.4 ^b	25.6 (8.8)

^a Not applicable; ^b SD not reported; *Individual percentages reported per response category, per item. Data does not lend itself to tabulation.

With regard to the qualitative literature reviewed, it can be observed from the limited data presented here that themes of loss in terms of carers' relationship with the care recipient, and the strain of conflicting roles, were evident across both AD and FTD. Other themes, such as loss of the self and the experience of negative emotions were predominantly observed in the narratives of AD caregivers in this review.

3.7 Methodological limitations of the reviewed studies

Several methodological limitations were common to the reviewed studies. Firstly, the cross-sectional nature meant that a 'snapshot' of carer burden was taken at one timepoint in the disease process. There was a failure by some studies (D'Onofrio et al., 2015) to match participant groups with respect to level of cognitive impairment or severity, and by others to assess or report dementia stage or severity (Leggett et al., 2010; Nicolaou et al., 2010; Oyebode et al., 2013; Shim et al., 2012). This has meant that the differences in respective burden scores cannot be reliably understood as reflecting disease specific, rather than stage or severity specific, experiences of burden.

Additionally, opportunity sampling methods left the majority of studies open to non-response bias, whereby those who chose to participate may have been qualitatively different from those who declined. Within some studies, insufficient information was provided regarding the recruitment process, meaning that such sources of potential bias could not be identified (Oyebode et al., 2013; Pudelewicz et al., 2018; Välimäki et al., 2012). Finally, the use of single-centre recruitment strategies (Donofrio et al., 2015; Iavarone et al., 2014; Nicolaou et al., 2010; Oyebode et al., 2013; Raggi et al., 2015; Zucchella et al., 2012), may carry the impact of unknown locality-specific factors, such as availability of carer support, which could have uniquely impacted upon findings.

Two quantitative papers reviewed (Nicolaou et al., 2010; Pudelewicz et al., 2018) may be limited by their small sample sizes, however this was difficult to assess as no details of power calculation processes were provided. Furthermore, none of the studies reviewed here reported effect sizes. Only one study (Iavarone et al., 2014) assessed the statistical significance of differences between subscale scores within their chosen measure of caregiver burden. As such, it was not possible to assess the relative

differences between subscale scores beyond a visual examination of the means and standard deviations, which remains problematic due to differences in the chosen reporting of statistics for each study.

One further paper (Wegierek, 2012) presented the retelling of a single caregiver's experience, who ultimately required support for decline in her own mental health as a result of the burden of care. While providing a valuable insight into what may be the upper extremes of caregiver burden and distress, the extent to which this may reflect common caregiving experiences is uncertain.

The use of self-report measures alone by all of the quantitative studies reviewed makes findings vulnerable to response biases such as social desirability bias, or potential over-reporting to illicit help. Additionally, no discussion is provided of the ethical procedure or protocol for caregivers who expressed clinically concerning levels of burden.

Overall, these methodological issues highlight difficulties in the generalisability of the majority of studies reviewed here, which in turn have implication for the conclusions of this review.

4. Discussion

4.1 Summary and interpretation of findings

The findings of the present review indicate that carers of persons with dementia experience often high levels of caregiver stress and burden, with group differences evident both within and between the dementia subtypes explored. These results are in keeping with previous findings, which have suggested that the level of distress experienced may vary by subtype diagnosis (Liu et al., 2017), and across the course of disease progression (e.g. Kamiya et al., 2014; Raccichini et al., 2015). Varied indication of differences in the patterns of burden characteristics experienced by carers of different dementia aetiologies was identified. Three studies of AD caregivers (Iavarone et al., 2014; Raggi et al., 2015; Zucchella et al., 2012) yielded an almost identical pattern of burden using the CBI. This may be seen to indicate that early-to-mid stages of AD progression (Iavarone et al., 2014; Raggi et al., 2015) are characterised by burden which is most heavily comprised of time pressures, the developmental impact

of caregiving upon congruence of life stage and expectations, and physical aspects of burden such as the impact upon the health and sleep of the caregivers. Taking into account the findings of Zucchella et al. (2012), it may be that social burden is least problematic in the early stages of the disease, when overall level of burden is lower, but later increases to overtake emotional burden as dementia severity and duration of caregiving increase. As such, a shifting pattern may be indicated from that reported by Zucchella et al (2012), to that seen by Iavarone et al. (2014) and Raggi et al. (2015). In comparison, VaD caregivers (D'onofrio et al., 2015) reported a distinct pattern of burden using the CBI in the only study of VaD caregivers reviewed here. Given the lack of additional studies, and the finding by the same paper of an again distinct burden profile among AD caregivers using this measure (in the context of similar disease severity to other studies reviewed), it may also be possible that these patterns remain an artefact of some characteristic of the study itself rather than reflecting a truly distinct burden experience.

Among carers of those with DLB diagnoses, findings (Galvin et al., 2010; Leggett et al., 2010) gave a tentative indication of a discernible pattern, however differences in the statistics reported prevent this from being clear. It appeared that 'performance' and 'role strain' may make similar contributions to the experience of burden among these caregivers, with 'personal strain' appearing notably less burdensome. Disease severity information for Leggett et al's (2010) cohort was unavailable, however the two cohorts were equivalent in their time since diagnosis, and contained a similar proportion of live-in caregivers. This pattern of burden was similar but not identical to that reported by Besser and Galvin's (2018) cohort of FTD caregivers, for whom 'role strain' appeared most burdensome, followed closely by 'performance strain', and finally 'personal strain' again appearing much less prominent. Similarities in the patterns of burden experienced by caregivers of these dementia subtypes may be conceptually understood as reflecting similarities in disease presentation (Schoenberg & Scott, 2011).

The identical pattern of burden described by Svendsboe et al. (2016) for their AD and DLB cohorts is perhaps less in keeping with the hypotheses of the current review, and provides a note of caution to the findings thus far. It should be noted however that

while the same pattern of burden was identified, DLB caregivers reported greater levels of distress arising from this (Svendsboe et al., 2016), despite participants in the AD and DLB cohorts being matched for disease severity and time since diagnosis. In contrast to other DLB cohorts reported here (Galvin et al., 2010; Leggett et al., 2010), PwDs in Svendsboe et al's (2016) cohort were notably earlier in their dementia process, and as such, one possibility may be that burden profile as assessed by the RSS is similar to AD in the earlier stages of the disease. Alternatively, in the absence of any other articles reporting the RSS measure, the possibility that this finding is reflective of characteristics of the internal structure of the measure itself, or of unknown sociodemographic characteristics, cannot be ruled out. It is not currently possible to conclude further, and so rather the results of this study must be taken as an indication of a possible lack of difference in caregiving experiences across dementia subtypes. Similarly, the findings of Nicolaou et al. (2010), in which the same pattern of burden was found among AD and FTD caregivers, whose total measured burden also fell within the same clinical range, adds to this possibility. Nicolaou et al. (2010) did however highlight that this was inconsistent with the literature. Groups were relatively well matched in terms of caregiving input, duration of caregiving and time since diagnosis, however information regarding dementia severity was not provided, and so again this unexpected finding cannot be further explored at this time.

On the basis of such mixed findings, it may be that where differences were indicated these are perhaps better explained by alternative hypotheses than by diagnostic subtype. As discussed above, caregiver characteristics including gender, personality and coping style (Etters et al., 2008; Merrilees, 2016; Orgeta & Leung, 2015) arguably transcend diagnostic subtype and likely mediate the impact of patient-related factors which would be reflected by subtype diagnosis. Alternatively, it may be that differences in burden characterisation observed are reflective of variations in disease stage, for which data was not consistently available or directly comparable. As such, a number of alternative interpretations may be likely, and cannot be further clarified on the basis of the literature reviewed.

From the small number of qualitative papers reviewed here, it appears that aspects burden related to role conflict and changing relationships may be common across

dementia subtypes, featuring in the narratives of both AD and FTD caregivers. Qualitative data were unavailable for DLB and VaD caregivers, and as such, wider conclusions cannot be drawn. Overall, the qualitative data reviewed here was important to expand upon the findings of quantitative research, allowing for exploration beyond the constructs of psychometric measures, however was not sufficient in volume or breadth to lend itself to unique conclusions.

4.2 Limitations of the present review

The present review is subject to several limitations. Firstly, the majority of studies reviewed were conducted outside of the UK, although were restricted to 'western' cultures. It is still possible however that cultural differences may exist in the experience of caregiver burden (e.g. Janevic & Connell, 2001; Mekala et al., 2013), which could be uniquely impacted by, for example, varying levels of support, or differences in attitudes and approaches to dementia or to caregiving more broadly.

Differences in cohort attributes, both between and within the studies reviewed here resulted in difficulty in achieving a meaningful comparison of findings. Variables such as length of time in the caring role, hours per week spent in caregiving activities, living with the care recipient, level of cognitive impairment, time since diagnosis, and dementia stage severity, are all likely to impact upon the level and characteristics of burden experienced. It cannot be assumed that these sociodemographic factors will covary, as evident in Nicolaou et al.'s 2010 study which reported the greatest length of time since diagnosis, and longest period spent in the caring role, but interestingly also reported the lowest mean number of hours spent on caregiving activities per week. In relation to this, it is additionally acknowledged that hypotheses regarding the interaction of caregiver burden and disease progression could not be fully explored. This is a disappointing limitation of the current review, and is reflective of the limited literature available.

Additionally, in the context of such variation the author must reflect on the outcome of the quality appraisal process, in which it was felt appropriate to allocate equal weighting to all but one paper, on the basis of the percentage score achieved. Contained within many of the quality appraisal criteria, and particularly those deemed

of specific importance to the current review, were a number of factors for consideration, compliance with which was then reduced to an overall score using a 4-point scale. As such, papers may have received the same score for any given criterion while fulfilling different factors within this. While the same could be said of other quality appraisal tools, such as the CASP and Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018), for which the outcome of each criteria is reduced further to response categories of 'yes/no/can't tell', the creation of a bespoke tool for use here presented the opportunity to move away from such a limitation. It is therefore acknowledged that the quality appraisal tool employed also represented a limiting factor of the current review by constraining the extent to which compliance with quality criteria, and differences within this, could be reflected in the resulting scores.

Furthermore, exploration of other variables of interest, such as the gender and age of the caregiver, relationship to the care recipient, or other demographic factors which could vary by diagnostic subtype, would likely have provide significant contributions to the understanding of differences in caregiver burden observed here. Similarly, the findings of the present review may have been strengthened if patterns of burden could also have been explored in relation to variables such as cognitive function, ADL independence, and BPSD. This would have allowed further comparison to the existing body of literature, as well as further exploration of the extent to which diagnostic subtype may be seen to reflect overarching patterns of burden, or useful ways of grouping similar burden experiences. With hindsight, it is recognised that failure to incorporate these elements within the current view created an inability to determine whether diagnostic subtype diagnosis does indeed capture and reflect the unique combinations of factors associated with burden as is hypothesised. As a result, this review had to assume that this would be the case, and findings are then partially based upon this assumption. As experienced by previous reviews in this area, differences in measures of carer burden, identified factor structures, and reported statistics, similarly restricted the comparison process. It is acknowledged that this represents a significant limitation of the current review.

Lastly, it is clear that AD was over-represented here, while VaD was only minimally represented. Additionally, both DLB and VaD diagnoses lacked representation among

the qualitative research presented. This has meant that for these dementias, the investigation of carer burden characteristics was confined to the factor structure of the measures employed, therefore presenting a significant limitation to the depth of learning about characteristics of carer burden which could be achieved.

4.3 Clinical implications

Given the exploratory nature and acknowledged limitations of this review, no formal clinical implications can be outlined. However, an awareness of the potential for different burden experiences to be associated with dementia subtype diagnoses may be beneficial when planning interventions for carers of persons with dementia, especially in the case of mixed-diagnosis group interventions.

4.4 Recommendations

Further research is recommended to add depth and rigour to the understanding of the qualitative experience of carer burden, and to determine whether diagnostic subtype diagnoses, and the distinct combination of factors which they may reflect, do indeed capture differences within this experience. Ideally this would encompass a mixed-methods approach, with carefully matched participants in order to avoid the confounding effects discussed above. Additional emphasis should also be placed upon understanding the relative impact of specific symptoms, and caregiver demographic variables such as age, gender and relationship to the PwD. Incorporating one or more objective measures of caregiver burden, such as the ZBI or CBI seen here, supplemented by a qualitative investigation of narrative themes, would allow the investigation of burden to expand beyond the dimensionality of questionnaire measures.

Longitudinal research would also be beneficial, allowing for changes in burden characterisation to be tracked over the course of disease progression. This would provide a rich and informative investigation of the differences in caregiver burden experienced by carers of different dementia aetiologies. Such understanding could prove invaluable to the effective planning and delivery of interventions aimed at preventing or reducing caregiver burden among this subset of carers.

4.5 Conclusions

In the context of the above limitations, the findings of the present review must be interpreted with caution. While the current results appear to indicate the possibility of group differences with regard to characteristics of burden experienced by carers of different dementias, these findings are inconclusive and may be explained by alternative hypotheses. As such, the work presented here can be viewed only as a platform to encourage and inform future research in this field.

References

(*indicates studies included in the current review)

Adelman, R. D., Tmanova, L. L., Delgado, D. et al. (2014). Caregiver burden: A clinical review. *JAMA - Journal of the American Medical Association*, 311(10), 1052-1059.

Akpınar, B., Küçükgülü, Ö, & Yener, G. (2011). Effects of gender on burden among caregivers of Alzheimer's patients. *Journal of Nursing Scholarship*, 43(3), 248-254.

Alzheimer's Association. (2016). 2016 Alzheimer's disease facts and figures. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 12(4)459-509. Alzheimer's Research UK. (2015). *Women and dementia: A marginalised majority*. Retrieved January, 2019, from <https://www.alzheimersresearchuk.org/about-us/our-influence/reports/women-dementia/>

Alzheimer's Research UK. (2016a). *10 THINGS YOU NEED TO KNOW about the impact of dementia on people, carers and the economy*. Retrieved March, 2017, from <http://www.alzheimersresearchuk.org/wp-content/uploads/2015/05/3.-10-things-impact-2016.pdf>

Alzheimer's Research UK. (2016b). *Boost in funding and awareness for rare dementias*. Retrieved August, 2019, from <https://www.alzheimersresearchuk.org/boost-in-funding-and-awareness-for-rare-dementias/>

Alzheimer's Society. (2014). *Dementia UK: second edition – overview*. Retrieved September 2018, from http://eprints.lse.ac.uk/59437/1/Dementia_UK_Second_edition_-_Overview.pdf

Alzheimer's Society. (n.d). *How dementia progresses*. Retrieved August 2019, from <https://www.alzheimers.org.uk/about-dementia/symptoms-and-diagnosis/how-dementia-progresses>

Bass, D.M., Judge, K.S., Snow, A. L., et al. (2012). Negative caregiving effects among caregivers of veterans with dementia. *American Journal of Geriatric Psychiatry*, 20, 239-247.

- *Besser, L.M., & Galvin, J.E. (2018). Perceived burden among caregivers of patients with frontotemporal degeneration in the United States. *International Psychogeriatrics*. [Advance online publication] doi:10.1017/S104161021800159X
- Burns, A., & Rabins, P. (2000). Carer burden in dementia. *International Journal of Geriatric Psychiatry*, 15, S9-S13.
- *Bursch, H.C., & Butcher, H.K. (2012). Caregivers' deepest feelings in living with Alzheimer's disease: A Ricoeurian interpretation of family caregivers' journals. *Research in Gerontological Nursing*, 5(3), 207.
- Cepoiu-Martin, M., Tam-Tham, H., Patten, S., et al. (2016). Predictors of long-term care placement in persons with dementia: A systematic review and meta-analysis. *International Journal of Geriatric Psychiatry*, 31(11), 1151-1171.
- Cerejeira, J., Lagarto, L., & Mukaetova-Ladinska, E. (2012). Behavioral and psychological symptoms of dementia. *Frontiers in Neurology*, 3, 73.
- Chappell, N. L., Dujela, C., & Smith, A. (2014). Spouse and adult child differences in caregiving burden. *Canadian Journal on Aging: La Revue Canadienne Du Vieillissement*, 33(4), 462-472.
- Cheng, S. (2017). Dementia caregiver burden: A research update and critical analysis. *Current Psychiatry Reports*, 19(9):64.
- Conde-Sala, J., Garre-Olmo, J., Turró-Garriga, O., et al. (2010). Differential features of burden between spouse and adult-child caregivers of patients with Alzheimer's disease: An exploratory comparative design. *International Journal of Nursing Studies*, 47(10), 1262-1273.
- Critical Appraisal Skills Programme (2018a). *Qualitative Checklist*. Retrieved on 18th January 2017 from <https://casp-uk.net/casp-tools-checklists/>
- Critical Appraisal Skills Programme (2018b). *Cohort Study Checklist*. Retrieved on 18th January 2017 from <https://casp-uk.net/casp-tools-checklists/>

- Dauphinot, V., Delphin-Combe, F., Mouchoux, C. et al. (2015). Risk factors of caregiver burden among patients with Alzheimer's disease or related disorders: a cross-sectional study. *Journal of Alzheimer's Disease*, 44(3), 907-916.
- de Vugt, M. E., Riedijk, S. R., Aalten, P. et al. (2006). Impact of behavioural problems on spousal caregivers: A comparison between Alzheimer's disease and frontotemporal dementia. *Dementia and Geriatric Cognitive Disorders*, 22(1), 35-41.
- Department of Health. (2013). *Dementia: A state of the nation report on dementia care and support in England*. Retrieved August 2018 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/262139/Dementia.pdf
- *D'Onofrio, G., Sancarolo, D., Addante, F., et al. (2015). Caregiver burden characterization in patients with Alzheimer's disease or vascular dementia. *International Journal of Geriatric Psychiatry*, 30(9), 891-899.
- Elmståhl, S., Malmberg, B., & Annerstedt, L. (1996). Caregiver's burden of patients after stroke assessed by a novel caregiver burden scale. *Archives of Physical and Medical Rehabilitation*, 77, 177-182.
- Etters, L., Goodall, D., & Harrison, B. E. (2008). Caregiver burden among dementia patient caregivers: A review of the literature. *Journal of the American Academy of Nurse Practitioners*, 20(8), 423-428.
- Feast, A., Moniz-Cook, E., Stoner, C. et al (2016). A systematic review of the relationship between behavioral and psychological symptoms (BPSD) and caregiver well-being. *International Psychogeriatrics*, 28(11), 1761-1774.
- Ferreira de Oliveira, F., Wajman, J.E., Bertolucci, P.H.F., et al. (2015). Correlations among cognitive and behavioural assessments in patients with dementia due to Alzheimer's disease. *Clinical Neurology and Neurosurgery*, 135, 27-33.
- Finkel, S. (2000). Introduction to behavioural and psychological symptoms of dementia (BPSD). *International Journal of Geriatric Psychiatry*, 15(S1), S2-S4.

- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189-198.
- *Galvin, J.E., Dud, J.E., Kaufer, D.I. et al. (2010). Lewy body dementia: Caregiver burden and unmet needs. *Alzheimer Disease and Associated Disorders*, 24(2), 177.
- Global Alzheimer's and Dementia Action Alliance. (2017). *Women and dementia: A global challenge*. Retrieved January, 2019, from https://www.gadaalliance.org/wp-content/uploads/2018/09/Women__Dementia_AGlobalChallenge_GADAA_Report_2018_Online.pdf
- Greene, J. G., Smith, R., Gardiner, M., & Timbury, G. C. (1982). Measuring behavioural disturbance of elderly demented patients in the community and its effects on relatives: A factor analytic study. *Age and Ageing*, 11(2), 121-126.
- Hong, Q.N., Pluye, P., Fàbregues, S., et al. (2018). Mixed Methods Appraisal Tool (MMAT). *Education for Information*, 34(4), 285-291.
- House of Commons Library. (2018). *Dementia: Policy, services and statistics*. Retrieved August 2018 from <https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN07007>
- Hawker, S., Payne, S., Kerr, C. et al. (2002). Appraising the evidence: Reviewing disparate data systematically. *Qualitative Health research*, 12(9), 1284-1299.
- *Iavarone, A., Ziello, A.R., Pastore, F, et al. (2014). Caregiver burden and coping strategies in caregivers of patients with Alzheimer's disease. *Neuropsychiatric Disease and Treatment*, 10, 1407-1413.
- Jackson, G.A., & Browne, D. (2017). Supporting carers of people with dementia: What is effective? *BJPsych Advances*, 23, 179.
- Janevic, M.R., & Connell, C.M. (2001). Racial, ethnic, and cultural differences in the dementia caregiving experience: Recent findings. *The Gerontologist*, 41(3), 334-347.

- Kamiya, M., Sakurai, T., Ogama, N. et al. (2014). Factors associated with increased caregivers' burden in several cognitive stages of Alzheimer's disease. *Geriatrics & Gerontology International*, 14(S2), 45-55.
- Kim, S. K., Park, M., Lee, Y. et al. (2017). Influence of personality on depression, burden, and health-related quality of life in family caregivers of persons with dementia. *International Psychogeriatrics*, 29(2), 227-237.
- Lee, D. R., McKeith, I., Mosimann, U. et al. (2013). Examining carer stress in dementia: The role of subtype diagnosis and neuropsychiatric symptoms. *International Journal of Geriatric Psychiatry*, 28(2), 135-141.
- *Leggett, A.N., Zarit, S., Taylor, A., & Galvin, J.E. (2010). Stress and burden among caregivers of patients with Lewy Body Dementia. *The Gerontologist*, 51(1), 76-85.
- Liu, S., Jin, Y., Shi, Z. et al. (2017). The effects of behavioral and psychological symptoms on caregiver burden in frontotemporal dementia, Lewy body dementia, and Alzheimer's disease: Clinical experience in china. *Aging and Mental Health*, 21(6), 651-657.
- Lou, Q., Liu, S., Huo, Y.R. et al. (2015). Comprehensive analysis of patient and caregiver predictors for caregiver burden, anxiety and depression in Alzheimer's disease. *Journal of Clinical Nursing*, 24, 2668.
- Mekala, S., Alladi, S., Chandrasekar, K. et al. (2013). Cultural differences are reflected in variables associated with carer burden in FTD: A comparison study between India and Australia. *Dementia & Neuropsychologia*, 7(1), 104-109.
doi:10.1590/S1980-57642013DN70100016
- Melo, G., Maroco, J., & de Mendona, A. (2011). Influence of personality on caregiver's burden, depression and distress related to the BPSD. *International Journal of Geriatric Psychiatry*, 26(12), 1275-1282.
- Mirrilees, J. (2016). The impact of dementia on family caregivers: What is research teaching us? *Current Neurology and Neuroscience Reports*, 16:88.

- Morris, J. C. (1993). The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, 43(11), 2412-2414.
- National Institute of Health and Care Excellence. (2017). *Clinical knowledge summaries: Dementia*. Retrieved September 2018 from <https://cks.nice.org.uk/dementia#!backgroundsub:1>
- *Nicolaou, P. L., Egan, S. J., Gasson, N., & Kane, R. T. (2010). Identifying needs, burden, and distress of carers of people with frontotemporal dementia compared to Alzheimer's disease. *Dementia: The International Journal of Social Research and Practice*, 9(2), 215-235.
- Niu, H., Álvarez- Álvarez, I., Guillén-Grima, F., & Aguinaga-Ontoso, I. (2017). Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. *Neurología*, 32(8), 523-532.
- Novak, M., & Guest, C. (1989). Application of a multidimensional caregiver burden inventory. *The Gerontologist*, 29(6), 798-803.
- Nunnemann, S., Kurz, A., Leucht, S., & Diehl-Schmid, J. (2012). Caregivers of patients with frontotemporal lobar degeneration: A review of burden, problems, needs, and interventions. *International Psychogeriatrics*, 24(9), 1368-1386.
- O'Dwyer, S. T., Moyle, W., Zimmer-Gembeck, M., & De Leo, D. (2016). Suicidal ideation in family carers of people with dementia. *Aging & Mental Health*, 20(2), 222-230.
- Orgeta, V., & Leung, P. (2015). Personality and dementia caregiving: A review and commentary. *Current Opinion in Psychiatry*, 28, 57-65.
- *Oyebode, J.R., Bradley, P., & Allen, J.L. (2013). Relatives' experiences of frontal-variant Frontotemporal dementia. *Qualitative Health Research*, 23(2), 156-166.
- Podcasy, J.L. & Epperson, C.N. (2016). Considering sex and gender in Alzheimer's disease and other dementias. *Dialogues in Clinical Neuroscience*, 18(4), 437-446.
- *Pudelewicz, A., Talarska, D., & Bączyk, G. (2018). Burden of caregivers of patients with Alzheimer's disease. *Scandinavian Journal of Caring Sciences* [Advance online publication]. doi: 10.1111/scs.12626

- Raccichini, A., Spazzafumo, L., Castellani, S., et al. (2015). Living with mild to moderate Alzheimer patients increases the caregiver's burden at 6 months. *American Journal of Alzheimer's Disease and Other Dementias*, 30(5), 463-467.
- *Raggi, A., Tasca, D., Panerai, S. et al. (2015). The burden of distress and related coping processes in family caregivers of patients with Alzheimer's disease living in the community. *Journal of the Neurological Sciences*, 358, 777-81.
- Robinson, C.A., Bottorf, J.L., Pesut, B., Oliffe, J.L. et al. (2014). The male face of caregiving: A scoping review of men caring for a person with dementia. *American Journal of Men's Health*, 8(5), 409-426.
- Schoenberg, M. R., & Scott, J. G. (2011). *The little black book of neuropsychology: A syndrome-based approach*. Boston, MA: Springer.
- Schoenmakers, B., Buntinx, F., & Delepeleire, J. (2010). Factors determining the impact of care-giving on caregivers of elderly patients with dementia. A systematic literature review. *Maturitas*, 66(2), 191-200.
- *Shim, B., Barosso, J., & Davis, L.L. (2012). A comparative qualitative analysis of stories of spousal caregivers of people with dementia: Negative, ambivalent and positive experiences. *International Journal of Nursing Studies*, 49, 220-229.
- *Svendsboe, E., Terum, T., Testad, I. et al. (2016). Caregiver burden in family carers of people with dementia with Lewy bodies and Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 31(9), 1075-83.
- Terum, T.M., Andersen, J.R., Rongve, A., et al. (2017). The relationship of specific items on the Neuropsychiatric Inventory to caregiver burden in dementia: A systematic review. *International Journal of Geriatric Psychiatry*, 32, 703-717.
- Thompson, R. L., Lewis, S. L., Murphy, M. R., et al. (2004). Are there sex differences in emotional and biological responses in spousal caregivers of patients with Alzheimer's disease? *Biological Research for Nursing*, 5(4), 319-330.

- * Välimäki, T., Vehviläinen-Julkunen, K., Pietilä, A., & Koivisto, A. (2012). Life orientation in Finnish family caregivers' of persons with Alzheimer's disease: A diary study. *Nursing and Health Sciences*, 14, 480-487.
- *Wegierek, A.M.J. (2012). Taking care of a loved one who has Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, 27(7), 463-467.
- Whittemore, R. & Knafl, K. (2005). The integrative review: updated methodology. *Methodological Issues in Nursing Research*, 52(5). 546-553.
- World Health Organisation. (2017). *Dementia fact sheet*. Retrieved September, 2018, from <http://www.who.int/news-room/fact-sheets/detail/dementia>
- Zarit, S. H., Reever, K. E., & Bach-Peterson, J. (1980). Relatives of the impaired elderly: Correlates of feelings of burden. *The Gerontologist*, 20(6), 649-655.
- *Zucchella, C., Bartolo, M., Pasotti, C. et al. (2012). Caregiver burden and coping in early-stage Alzheimer disease. *Alzheimer Disease and Associated Disorders*, 26(1), 55-60.

Part Two: Research Report

An initial exploration of the reliability and validity of the Short Parallel Assessments of Neuropsychological Status (SPANS) for use with older adults

(Intended journal for publication is detailed in Appendix A²)

² Note that the word count and reference restrictions required for submission to the target journal were felt to be incompatible with the University's requirements of this thesis. As such, this thesis is formatted and referenced in accordance with University guidance and will be amended post-viva for publication.

Abstract

Introduction

The Short Parallel Assessments of Neuropsychological Status (SPANS) is a brief, comprehensive assessment battery, designed to screen for a range of neuropsychological conditions within the working-age adult population. A number of features, including the ability to distinguish between cortical and subcortical profiles, indicate that the SPANS may be a clinically useful tool in the assessment of suspected of dementia. This study aimed to provide an initial exploration of the convergent validity of the SPANS in relation to an existing assessment of older adult cognition (the Addenbrooke's Cognitive Assessment), as well as the internal and alternate-form reliability of the SPANS when administered to a sample of adults over the age of 65 years.

Method

A total of 103 participants were recruited across diverse geographical areas in the U.K, through a combination of personal connection, and organisations working with older people. A within-subjects, cross-sectional design was employed, within which the internal consistency, alternate version reliability, and convergent validity of the SPANS were explored. Convergent validity was assessed in relation to the Addenbrooke's Cognitive Assessment (ACE-III).

Results

Results indicated that the SPANS may continue to offer a useful measure of cognitive ability within a sample of older adults. A high level of alternate version re-test reliability was upheld, and encouraging convergent validity was demonstrated. Findings indicated that the SPANS may also offer additional benefits beyond those currently associated with the ACE-III.

Discussion

Findings were subject to a number of limitations, however despite this have demonstrated that further exploration within the older adult population is warranted. Recommendations for future research were provided.

1.Introduction

1.1 Clinical neuropsychology and neuropsychological assessment

Neuropsychology is the scientific field of study exploring the relationship between neuroanatomy and neural functioning, and cognition, behaviour, and emotion (Colman, 2009; Frerichs, 2004). Clinical neuropsychology refers to the application of this field of knowledge, and is concerned with the implications of neuroanatomical change (through injury or disease processes) upon psychological domains such as attention, memory, language, and executive function, as well as emotion and personality (Frerichs, 2004; Stirling & Elliot, 2008).

Neuropsychological assessment refers to the use of standardised psychometric tests to identify, and quantify, changes which may have occurred in relation to a person's cognitive functioning across domains such as those described above. This is in aid of diagnosis and understanding of a patient's presentation, and to inform rehabilitative intervention through the identification of cognitive difficulties and areas of preserved ability or relative strength (Frerichs, 2004; Stirling & Elliot, 2008). Supplementary information is gathered from a range of sources, including (but not limited to) medical history, observation, self-reported and informant-reported concerns, indicators of premorbid levels of functioning, and broader biopsychosocial history of the individual (Frerichs, 2004; Goldstein & McNeil, 2004).

1.2 Clinical neuropsychology in the context of older adult populations

Within older adult populations, one of the most common cause of referral for neuropsychological assessment relates to cognitive decline (Frerichs, 2004), often in the context of suspected dementia.

1.2.1 Dementia

Dementia is an umbrella term for a collection of usually progressive conditions, generally occurring later in life, and characterised by a decline in cognitive function beyond the normal ageing process (World Health Organisation, 2017). In 2017, the World Health Organisation (WHO) estimated that 50 million people were living with

dementia worldwide, with an estimated 850,000 of those living in the UK (House of Commons Library, 2018).

A number of dementia subtypes exist, the most common of which (Alzheimer's Research UK, 2018) are Alzheimer's Disease (AD), Vascular dementia (VaD), Dementia with Lewy Bodies (DLB), and Frontotemporal dementias (FTDs), each of which has a documented clinical profile and trajectory (e.g. Schoenberg & Scott, 2011). Early and accurate diagnosis enables individuals to plan for their future while they retain the capacity to do so, facilitates access to services such as Dementia Care Advisors (Alzheimer's Society, n.d.), and ensures that the most appropriate treatment interventions are provided (e.g. Robinson, 2015). Accurate differential diagnosis of dementia subtype is of further importance in selecting appropriate pharmacological interventions, for which contraindications exist in relation to dementia aetiology – for example, acetylcholinesterase (AChE) inhibitors, prescribed to stall disease progression in AD and DLB, are contraindicated for those with FTD diagnoses (NICE, 2018).

In the case of suspected dementia, the neuropsychological assessment process can contribute to diagnosis, and can additionally inform differential diagnosis of dementia subtype (Jacova et al., 2007).

1.3 Neuropsychological assessment tools for use with older adult populations

Three cognitive screening assessments were found to be recommended, and known to the author to be routinely used, in the assessment and diagnosis of suspected dementia (Alzheimer's Society, 2015; East Midlands Clinical Networks, 2017; Quinn, Elliott & Langhorne, 2018). These were the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), the Addenbrooke's Cognitive Assessment in its former ('revised'; ACE-R; Mioshi et al., 2006) and current (3rd; ACE-III; Hsieh et al., 2013) editions, and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998).

1.3.1 The Montreal Cognitive Assessment (MoCA)

The MoCA (Nasreddine et al., 2005) is a brief cognitive screen, designed for use in the assessment and diagnosis of mild cognitive impairment (MCI). The MoCA incorporates items designed to reflect memory, language, attention, visuospatial, and executive skills (Nasreddine et al., 2005), and has recently been recommended for use in memory clinic settings by the International Consortium for Health Outcome Measures (ICHOM, 2017). However, the literature indicates that caution in interpretation of the MoCA is required, with disagreement evident regarding the optimal cut-off score, and the recommended cut-off yielding high sensitivity, but low specificity (e.g. Coen et al., 2016), therefore offering detection of impairment in a broad sense but without sufficient ability to differentiate between underlying causes. Additionally, some authors (e.g. Coen et al., 2011) recommend caution in interpretation of failure within the domains, as insufficient knowledge is gained from this assessment to determine the mechanism underlying this failure.

1.3.2 The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)

The RBANS (Randolph, 1998) was designed for use in dementia assessment. Cognition is assessed across the domains of immediate memory, visuospatial and constructional ability, language, attention, and delayed memory, with an overall total score also calculated. However, the RBANS is subject to several limitations with regard to the interpretability and test-retest reliability of its subtests (Strauss et al., 2006), and indication of unacceptably low sensitivity to mild cognitive impairment, in the context of good specificity (Duff et al., 2010).

1.3.3 The Addenbrooke's Cognitive Examination 3 (ACE-III)

Currently, the ACE-III (Hsieh et al., 2013) appears to be the most widely selected assessment in British memory clinics (Sajjadi & Brown, 2015) and may therefore be considered the current 'gold standard' dementia screening assessment. This is likely due to the comprehensive nature of the ACE-III in comparison to other brief bedside assessments (Cheung et al., 2015). Total score and domain sub-scores are calculated,

relating to attention & orientation, memory, language, verbal fluency, and visuospatial skills. As such, it facilitates both the detection of cognitive impairment, and the creation of a basic cognitive profile (Matias-Guiu, Cortés-Martínez et al., 2017). The ACE-III also contains two items indicated to be of specific clinical utility in the diagnosis of dementia – clock drawing, and verbal fluency (Strauss et al., 2006).

The ACE-III has been shown to possess impressive psychometric properties (discussed below), and superior diagnostic accuracy in comparison to other brief screening measures (Matias-Guiu, Valles-Salgado et al., 2017). It is therefore recommended for use in dementia assessment by the Alzheimer's Society (2015).

The ACE-III is not without limitation, and the publishers acknowledge that their normative sample was small (NeuRA, 2016), relatively young (mean age of 68.7 ± 7.0), and well educated ($14.1y \pm 2.8$ mean years in education; NeuRA, 2016). As such, the reliability and validity of the proposed clinical cut-off score (88/100) cannot be guaranteed for use with individuals of greater age or lower educational level. Indeed, several studies have reported that their analyses demonstrate greater appropriateness of lower cut-off scores for use with their older participant groups (Cheung et al., 2015; Jubb & Evans, 2015). Furthermore, the presence of only one cut-off score means that interpretation may appear very rigid. For better clinical utility, some authors have suggested that stratified norms, indicating degrees of impairment, may prove useful (Velayudhan et al., 2014). Finally, it must be noted that the ACE-III is weighted towards memory and language domains, and appears insensitive to behavioural and executive impairments (Sajjati & Brown, 2015).

1.4 The SPANS and rationale for the current research

Initially developed for use in inpatient brain injury rehabilitation (Burgess, 2014), the SPANS is a brief but comprehensive neuropsychological assessment battery, developed to screen for a range of neuropsychological conditions. The SPANS incorporates seven indices, assessing cognition across the domains of orientation (ORI), attention and concentration (ACI), memory and learning (MLI), visuo-motor

performance (VPI), efficiency (ECI), and conceptual flexibility (CFI). A more detailed description of each domain can be found in Appendix I.

The potential utility of the SPANS in dementia assessment and diagnosis is noted to have become apparent during its initial development (Burgess, 2014), since the range of subtests and cognitive domains assessed mirror those indicated by the Diagnostic and Statistical Manual (DSM—IV; APA, 2013) criteria for dementia diagnosis (Burgess, 2014). The SPANS may also contribute to the identification of dementia subtype diagnosis through screening of both cortical and subcortical cognitive functions (Burgess, 2014) which may aid in the differentiation of cortical (e.g. AD) and subcortical (e.g. VaD) dementias.

Demonstrating psychometric properties which fall within the ‘satisfactory’ to ‘excellent’ range (Kline, 2000; discussed below), the SPANS has been shown to be a valid and reliable assessment of adult cognition, sensitive to, and able to differentiate between, degrees of cognitive impairment (Burgess, 2014). The SPANS was developed in part to address some of the difficulties of existing neuropsychological assessment tools in relation to length, suitability for bedside use, and ability to screen across the range of cognitive functions (Burgess, 2014). As such it may be reasonable to expect that the SPANS could address similar difficulties in the assessment of older adults. Furthermore, the greater number and range of subtests and indices in the SPANS (in comparison to the MoCA, RBANS, and ACE-III) may offer a more detailed and descriptive assessment of cognitive functioning than these existing screening assessments, while requiring minimal additional time to administer and score. For a comparison of the SPANS with other existing short cognitive assessments, see Appendix J.

The SPANS also offers an alternative version, possessing strong alternate version reliability in working-age adult populations (Burgess 2014). This allows for re-assessment while minimising the influence of practice effects – the impact of having previously completed an assessment upon future test performance (Strauss et al., 2006). Since dementia is a progressive condition, repeated assessment is common in

order to determine whether observed cognitive impairment is static or declining over time. Therefore, alternate versions conveying high test-retest reliability are particularly beneficial.

To date, norming data for the SPANS have been published for groups between the ages of 18 to 74 years. Additionally, the internal reliability, alternate version reliability, and inter-rater reliability of the SPANS have been published for the same age group. Convergent and divergent validity data have also been published for this age range in relation to the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997a), Wechsler Memory Scale (WMS-III; Wechsler, 1997b), Rey Complex Figure (RFC; Meyers & Meyers, 1995), and the Trail Making Test (TMT; Reitan & Wolfson, 1995). Further investigation is required to explore the psychometric properties of this assessment within an older population, and to provide an initial indication as to whether the SPANS may subsequently be of clinical use in the diagnosis of dementia. The current research represented the initial stages of this investigation, and chose to employ the ACE-III as a tool for comparison.

1.5 Rationale for selecting the ACE-III over other available and widely accepted assessments

As discussed, the ACE-III may be considered a ‘gold standard’ for use in the assessment of dementia in British memory clinics. The ACE-III is more directly comparable to the SPANS than the MoCA, both in terms of defined domains and overall duration, and affords greater and more balanced sensitivity and specificity (Velayudhan et al., 2014). The MoCA, with a total score of 30 points and estimated completion time of just 10 minutes, is arguably more comparable with other brief screening tools than it is with an assessment battery such as the SPANS.

Alternatively, the RBANS (Randolph, 1998) is both similar in length and domain structure to the SPANS. However, area under ROC curve (AUC) analysis appears to indicate similarity between the RBANS and ACE-III (Duff et al., 2008; Matias-Guiu, Cortés-Martínez et al., 2017). As such, the RBANS may not offer any more to the investigation of SPANS validity than the ACE-III. Furthermore, use of the ACE-III, as a

preferred memory clinic assessment tool, reflects 'real-world' assessment of older adult cognition. Furthermore, the shorter administration time required for the ACE-III in comparison to the RBANS reduces the cognitive demand placed upon participants, and therefore reduces the negative impact of participation such as fatigue.

Finally, the ACE-III is freely available in the public domain with open permission granted for use in research.

1.6 Research aims

The primary aim of the current research was to complete a preliminary investigation of whether the SPANS may be a reliable and valid tool for assessing cognition in adults over the age of 65 years. In pursuit of this, internal reliability (the extent to which items within each index measure aspects of the same, single construct) and alternate form reliability (the extent to which the two forms of the SPANS (versions A and B) are equivalent), were explored. Additionally, convergent validity - the extent to which the SPANS total and index scores co-vary with the ACE-III, and pertinent subtests (verbal fluency and clock drawing) was also evaluated. Exploration of predictive validity was planned, however could not be realised due to restrictions within the data (Appendix K).

One secondary aim was also defined; to explore the relationship between age and educational history, and SPANS performance.

1.7 Research questions

- a) Do the SPANS indices continue to demonstrate good internal reliability when used with older adults?
- b) Do the alternate versions of the SPANS continue to offer good alternate version reliability when used with older adults?
- c) Does the SPANS co-vary with another neuropsychological test used with older adults (the ACE-III), and / or constituent parts known to add value in the assessment of dementia (clock drawing and verbal fluency)?
- d) How do age and educational history interact with SPANS total score?

2. Method

2.1 Design

The research design was developed by the principal investigator in collaboration with the chief investigator. A within-subjects, cross-sectional design was employed in which quasi groupings were created based upon level of education, in order to form an independent variable. Dependent variables were comprised of SPANS domain and total scores across versions A and B, ACE-III domain and total scores, clock drawing and verbal fluency scores.

2.2 Ethical approval

A process of review was undertaken through the University of Leicester, including peer and service-user reviewers (see Appendix L for an outline of the chronology of the research process), for which the principal investigator prepared proposal documents tailored to each audience. Following this, sponsorship and ethical approval were sought and granted through a process of co-application by the University of Leicester Research Ethics Panel (Appendix M). Specific ethical issues considered by the principal investigator in the planning and implementation of the current research are discussed in Appendix N.

2.3 Participants

Convenience sampling was used to recruit 68 predominantly healthy, community-dwelling older adults, across a range of geographical locations throughout the U.K. Recruitment took place over a 12-month period beginning in February 2018. Participant ages ranged from 65 to 90 years ($M = 73.4$, $SD = 5.9$), with the majority reporting British ethnicity (94.2%), and completion of higher education (66%). Participant characteristics are summarised in Table 1. Sixteen of these participants were recruited and assessed by the principal investigator.

Table 1: Self-reported participant demographic, health and ability characteristics

Variable	Number of participants (%)
Gender	
Male	34 (49.3)
Female	35 (50.7)
Age	
64-74	62 (60.2)
75+	41 (39.8)
Education	
Less than secondary	5 (7.2)
Secondary school completed	19 (27.5)
College or vocational training	17 (24.6)
University degree	28 (40.6)
Employment	
Unskilled labour	11 (15.9)
Skilled labour	14 (20.3)
Management / business, no degree	15 (21.7)
Professional with degree	27 (39.1)
Stay home / family	2 (2.9)
Health	
ABI	0 (0)
MCI	1 (1.4)
Dementia	0 (0)
Stroke	3 4.3
Epilepsy	1 (1.4)
Movement Disorder	2 (2.9)
Diabetes	8 (11.6)
Heart problem	19 (27.5)
Respiratory problem	10 (14.5)

Current or previous use of alcohol exceeding 10 units per day	7 (10.1)
Some level of assistance required with:	
Meal preparation	5 (7.2)
Personal care	2 (2.9)
Housework	7 (10.1)
Transport	5 (7.2)
Managing finances	2 (2.9)
Shopping	5 (7.2)
'Poor' ability level reported in relation to:	
Memory	3 4.3
Fine motor skills	3 (4.3)
Walking	6 (8.1)
Getting up from a chair	3 (4.3)
Vision (distance)	6 (8.7)
Vision (close up)	9 (13.0)
Hearing	7 (10.1)

Inclusion criteria were developed by the principal investigator aimed to be broad in order to maximise variance in the data and therefore increase the extent to which the sample could be considered representative of the wider older adult population. As such, while participants were predominantly healthy individuals, those with a history of mild cognitive impairment, early dementia, stroke or similar neurological diagnosis were deemed eligible for participation. Inclusion criteria required participants to be aged 65 years and over, community-dwelling, and to possess the mental capacity to provide informed consent for participation. Additionally, participants were required to possess a sufficient level of English to adequately comprehend verbal instructions and make their responses understood. Those who were blind or significantly visually impaired, deaf or extremely hard of hearing, paralysed or had an orthopaedic injury

that would have prevented them from using their dominant hand, were excluded from participation.

Participants were relatives or friends of the research team in some instances, or were recruited through facilities and organisations local to the team member (such as local libraries, Age UK, or University for the Third Age). Advertising was via the use of posters, flyers, newsletters, and talks at group meetings by members of the research team. Permission was sought from relevant 'gatekeeper' staff at any organisation that was approached prior to the commencement of advertising. Gatekeepers were generally identified through the organisation's website, and were approached using a standardised email developed by the principal investigator (see Appendix O).

Data from a further 73 participants (Table 2) who met the current inclusion criteria were available from a previous study conducted by the Chief Investigator. Their data were incorporated partially or wholly in analyses for research questions a, b, and d. This resulted in a final total of 141 participants, and 180 SPANS administrations.

Table 2: Demographic characteristics of participants whose data was collected as part of previous SPANS research (Burgess, 2014) and is included where specified in the current analyses

Variable	Number of participants (%)
Gender	
Male	34 (44.2)
Female	43 (55.8)
Age	
64-74	48 (62.3)
75+	29 (37.7)
Education	
Less than secondary	0 (0)
Secondary school completed	41 (53.2)
College or vocational training	18 (23.4)
University degree	18 (23.4)

2.4 Prospective power analysis

In order to determine required sample size, the principal investigator conducted a priori power analysis using 'GPower' software, for two anticipated analyses. In line with Cohen's (1992) guidance, power analyses were based on an alpha value of 0.05, and power of 0.80. It was determined that to detect medium effect sizes, a total sample of $n = 180$ would be required for analysis of variance (ANOVA) with 4 groups, and $n = 64$ for correlational analyses.

2.5 Measures

2.5.1 The SPANS

The SPANS (Burgess, 2014) was designed to be comprehensive, brief, and diverse in the range of tests and subtests presented. As such, it offers quick administration (30-

45 minutes) and scoring (10 minutes). The SPANS comprises 30 subtests across 7 indices - attention & concentration (ACI), orientation (ORI), memory & learning (MLI), language (LAI), visuo-motor performance (VPI), efficiency (ECI), and conceptual flexibility (CFI).

Previous validation studies have included healthy adults (Burgess, 2014), brain-injured adults (Burgess, 2014), adults with diagnosed intellectual disability (Harker-Brown, 2018), and adults for whom English is a second language (Haddlesey, 2016), or who represent culturally diverse backgrounds (Tan & Burgess, 2018). The SPANS demonstrates good construct validity and 'satisfactory' to 'excellent' levels of internal consistency within each index ($\alpha = .73$ and above). The assessment offers two parallel versions, demonstrating 'satisfactory' to 'excellent' alternate-version test-retest reliability ($r = .89$), and excellent inter-rater reliability ($r = .95$). Additionally, the SPANS offers high levels of sensitivity and specificity, therefore successfully detecting and discriminating between those with an acquired brain injury, a neurological condition, and healthy control participants (Burgess, 2014).

2.5.2 ACE-III

The ACE was developed as a brief cognitive assessment tool, taking roughly 20 minutes to administer, in order to detect dementia and differentiate between Alzheimer's Disease and the Frontotemporal Dementias (Mathuranath et al., 2000). Cognition is assessed across the domains of attention, memory, verbal fluency, language, and visuospatial ability. The current version (ACE-III) conveys moderate to strong convergent validity (Hsieh et al., 2013) in the absence of internal reliability information for each domain. Convergent validity was explored in comparison to the Frontotemporal Dementia Functional Rating Scale (FTDFRS; Mioshi et al., 2010), Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1997) digit span subtest, Rey Osterreith Complex Figure (ROCF; Meyers & Meyers, 1995), Rey Auditory Verbal Learning Task (RAVLT; Schmidt, 1996), and the Sydney language Battery (SYDBAT; Savage et al., 2013). The most recent revision, the ACE-III, offers three parallel versions, however alternate version test-retest reliability statistics could not be found in the literature. ACE-III total score demonstrates high levels of internal reliability ($\alpha =$

.88), sensitivity (1.00) and specificity (.96) at the published cut-off score of 88/100 (Hsieh et al., 2013).

In order to gain full utility of the clock drawing and verbal fluency tasks contained within the ACE-III, both tasks were extracted for further exploration as discussed below.

2.5.3 Verbal fluency

Verbal fluency requires the generation of as many words as possible within one minute, in accordance with a specific instruction. Tasks may take the form of phonetic fluency (where the requirement is to list words beginning with a certain letter), or semantic fluency (in which words must be generated in accordance with a category). It has been suggested that the inclusion of verbal fluency tasks may be beneficial in increasing sensitivity when assessing cognition in the context of suspected dementia (Strauss et al., 2006). Additionally, some evidence suggests that discrepancies in phonemic and semantic fluency performance potentially aid in the diagnosis of dementia subtype (Strauss et al., 2006). As the ACE-III yields a converted score for verbal fluency performance, total raw scores (total number of permissible responses according to the ACE-III scoring criteria) were extracted and explored separately in the current analyses.

2.5.4 Clock drawing

The clock drawing task requires a drawing to be constructed of a round, analogue clock face, containing the numbers one through to twelve, and two hands of distinctly different length to be placed in accordance with the instruction delivered. It has been suggested that clock drawing provides an indicator of general cognitive functioning and likelihood of decline, as well as contributing to differential diagnosis, thanks to the task's reliance upon multiple cognitive domains (Hubbard et al., 2008; Strauss et al., 2006). A number of variations in instruction and scoring criteria exist (Strauss et al., 2006), with the ACE-III utilising a 5-point scoring system for this item, therefore scoring performance in a broad sense. In order to explore convergent validity of the SPANS in greater detail within the current analyses, this task was additionally scored in

accordance with the 20-point criteria set out by Mendez et al. (1992). These criteria were selected by the principal investigator and were amended only to reflect the different time requested in the ACE-III version of this task.

Mendez et al's (1992) criteria have been shown to demonstrate high levels of sensitivity in the context of moderate specificity, and have been found to predict dementia more accurately than other scoring systems (Strauss et al., 2006). Additionally, these scoring criteria have been shown to demonstrate high levels of inter-rater reliability ($r = .94$), test-retest reliability ($r = .76$ after 24 weeks) and convergent validity when compared to a number of neuropsychological tests including the Rey Osterreith Complex Figure (ROCF; Meyers & Meyers, 1995), and Mini Mental State Examination (MMSE; Folstein et al., 1975).

2.5.5 Demographic, general health and daily living questionnaire (DHLQ)

This questionnaire (Appendix P) was developed by the principal investigator to gather study-relevant participant information in relation to demographic characteristics of participants which may be expected to co-vary with cognitive ability. This included age, education and employment history, general health, and level of independence in performing activities of daily living (ADLs).

The DHLQ was used predominantly used here to acquire subjective evidence of existing cognitive impairment, and to expand knowledge of the sample composition.

2.6 Procedure

A visual representation of the study procedure is presented in Appendix Q.

2.6.1 Recruitment of volunteer research assistants

An advert was sent to pre-qualified clinical psychologist groups across the U.K. Shortlisted applicants completed a telephone interview, conducted by the Chief Investigator. Offers were subject to background checks via the Disclosure and Barring Service (DBS), and completion of 'Good clinical practice' training. Thirteen VRAs were initially recruited, with a further two recruited at a later stage to replace those who

were no longer able to fulfil the role. The Chief Investigator was ultimately responsible for the VRAs and, together with the Principal Investigator, provided supervision and guidance throughout the duration of the study.

VRAs were trained by the Chief Investigator in a one-day seminar covering study-specific processes, and all necessary aspects of the SPANS, ACE-III, and DHLQ administration and scoring procedures. Training additionally incorporated general assessment principals, which included, but were not limited to: informed consent processes; the importance of building a rapport; and best-practice standards for the administration of neuropsychological assessments (Mitrushina et al., 2005).

Each VRA completed two practice administrations of all measures included in this study, which were reviewed by the Chief Investigator. When deemed fully competent in the administration and scoring of the materials, they were signed off to begin participant sessions. Practice data were added to the general SPANS norm database, but were not used within the present analysis due to the younger age range of practice participants.

2.6.2 Data collection

VRAs, in addition to the Principal Investigator, aimed to recruit nine participants each and to complete the full range of assessments (SPANS A, SPANS B, and ACE-III) with as many of these participants as possible.

At the time of expressing interest in the study, prospective participants completed a 'Participant Contact Sheet' (Appendix R), to facilitate further communication. Having expressed interest, participants were visually and verbally presented with the participant information sheet (PIS; Appendix S), and encouraged to ask questions. Those who wished to participate were consented into the study (Appendix T). They were then assigned a unique Participant Identification (PID) code, which was used to label their data in order to maintain anonymity.

Testing took place largely within participants' own homes (for which a lone working policy and procedure was provided to VRAs), or public spaces deemed suitable - such as rooms within local libraries and universities.

On the day of testing, the DHLQ was administered prior to commencing the cognitive assessments. The SPANS A was then administered followed by clock drawing and verbal fluency tasks. Following this, VRAs used their judgement (based on perceived levels of fatigue and motivation) to decide whether to seek consent for further participation through completion of the remainder of the ACE-III. Where consent was given, participants were offered a 10-minute break before further testing. The potential to counterbalance the order of administration was considered in order to reduce concerns regarding the impact of fatigue, and practice effects. However, doing so would have created the possibility for a participant to withdraw after completing the ACE-III, but before completing the SPANS A. This was felt to represent an unethical situation, due to the resultant data being unusable. Administering the SPANS first in all occurrences maximised the likelihood of gathering useable data relevant to the research question and, on balance, was deemed to be the preferable approach.

Upon completion of the initial testing session, participants were offered the opportunity to complete the alternate version of the SPANS (SPANS B) in a second testing session. This was required to take place within the subsequent 30-day period. A total of 37 participants went on to complete the SPANS B.

Assessments were administered and scored in accordance with their standardised instructions, with the additional scoring of the clock drawing task occurring as detailed above. Scoring took place after testing sessions, and never in the presence of a participant. As the ACE-III is a well validated and widely used measure, with a known clinical 'cut-off score' (88/100; Hsieh et al., 2013) recommended by the test's developers, it was recognised that some participants may score below this clinically relevant level. At the point of providing consent, participants were given the option to 'opt-in' to receive a standard, written notification from the Chief Investigator in the event that they completed the ACE-III and their score fell below the cut-off. Where participants opted-in and subsequently triggered such feedback, their contact

information was shared with the Chief Investigator to facilitate this. No specific detail of their score, or comment regarding their performance, was shared.

VRAs entered anonymised scores into an Excel spreadsheet which was shared electronically with the Chief and Principal investigator upon completion of data collection. All data were stored securely, in line with General Data Protection Regulation (GDPR) requirements.

A website was set up to offer participants the opportunity to view updates about the progress of the research, and a summary of findings upon completion.

3. Results

Exploration of data and subsequent statistical analyses were conducted by the Principal Investigator using IBM SPSS Statistics (version 24) for Windows.

3.1 Exploration of assumptions

The data were initially explored in relation to assumptions of parametric analysis, using a combination of graphical representation and statistical measures. Statistical assessment of normality was conducted using the Kolmogorov-Smirnov test with Lilliefors significance correction, in conjunction with the Shapiro-Wilks test since this is considered to be more powerful but requires caution in interpretation with larger sample sizes (Field, 2005). A number of outliers were identified by box plots, while histograms indicated negatively skewed data which deviated from the normal distribution. This was confirmed by values of skewness and kurtosis, Q-Q plots and statistical assessment of normality ($p < .05$). These findings were in line with theoretically driven expectations, since the SPANS was designed to fit within the remit of a screening assessment, and therefore intended to be only moderately challenging when presented to healthy, non-clinical populations (Burgess, 2014). Furthermore, this distribution was in keeping with that of earlier studies examining the use of the SPANS with healthy populations (Burgess, 2014).

In light of the above, non-parametric statistics were preferred for analysis. Of note, Kendall's Tau was selected for correlational analyses, chosen for its ability to manage

the high proportion of tied ranks within the data. However, this statistic is conservative in nature (Field, 2005), and as such may underestimate the magnitude of the observed relationship (Göktaş & İşçi, 2011). Therefore, while effect sizes were considered in relation to widely accepted thresholds (Cohen, 1992), true effect sizes may be larger than conventional guidance indicated, and should be interpreted with this in mind.

Where multiple comparisons were conducted, the Bonferroni correction was applied (Armstrong, 2014) in order to reduce type 1 error (false positive) rate. This is a conservative correction (Field, 2005), the use of which can increase the risk of type 2 error (false negative), however it was felt on balance that this risk was preferable to the risks of omitting the correction.

3.2 Statistical analyses

Results of statistical analyses are presented in relation to the specific research question to which they refer.

3.2.1 Research question A: Do the SPANS indices continue to demonstrate good internal reliability when used with older adults?

The extent to which the SPANS indices continue to provide a meaningful measure of cognitive function when used with older adults was assessed through evaluation of internal consistency. Assessed using Cronbach's alpha (Cronbach & Meehl, 1955), this statistic yields an indication of the extent to which items comprising a given index combine to measure a single construct, where values of .70 and above are considered sufficient (Kline, 2000). Table 3 shows the Cronbach's alpha value of each index score for combined version A and B data within the older adult sample ($N=180$), the younger adult sample reported by the SPANS manual, and all SPANS data held by Dr Burgess. The encompasses data gathered across a number of studies and clinical uses of the SPANS since its publication in 2014 and includes data from participants who took part in the current study.

Table 3: Internal consistency values (Cronbach's α) for all SPANS index scores and total score, across versions A and B for the current older adult sample, the younger sample described within the manual, and all SPANS data collected to date.

	Older adult sample (age 65 to 90, $N = 180$)	Manual sample (age 18 to 74, $N = 318$)	All data (age 18 to 90, $N = 609$)
Orientation Index (ORI)	0.33*	0.79	0.77
Attention / Concentration Index (ACI)	0.66	0.83	0.80
Language Index (LAI)	0.57	0.86	0.83
Memory and Learning Index (MLI)	0.78	0.90	0.88
Visuo-motor Performance Index (VPI)	0.71	0.85	0.83
Efficiency Index (ECI)	0.65	0.85	0.83
Conceptual Flexibility Index (VPI)	0.62	0.73	0.71
SPANS total	0.85	Unavailable	0.91

*Two items ('person' and 'condition') were excluded due to zero variance

In comparison to the internal consistency values reported by the SPANS manual for working-age adults, values obtained here were substantially lower in at least five of the seven indices, and failed to reach the .70 threshold in five of the seven indices. These differences were understood to reflect the highly able, homogeneous nature of the current sample, which consequently lacked the variance required for achieving satisfactory levels of internal consistency. This is in contrast to the greater variation within the sample reported by the SPANS manual, for which a much broader age range, and combination of clinically diverse and healthy participants was reported. The addition of the older adult data to the existing database of clinical and healthy 18 to 74-year-olds resulted in only a very small adjustment to values of internal consistency within the SPANS. This may provide tentative indication that the constructs assessed are maintained at a reliable level into older age.

Internal consistency, and descriptive statics (Table 4) indicated that the acknowledged lack of variance was particularly evident within ORI scores, with two items from this index being removed from alpha calculations due to zero variance. Indeed, internal consistency values of .70 and above were achieved in the present study only by the MLI and VPI indices, which descriptive statistics indicated contained greatest variance, as demonstrated by the values of the standard deviation.

Table 4: Descriptive statistics for SPANS index and total scores derived from the current older adult sample (*N* = 108)

	SPANS							
	ORI	ACI	LAI	MLI	VPI	ECI	CFI	Total
Maximum achievable score	22	46	53	67	70	48	28	334
<i>M</i> (SD)	21.55 (.80)	41.38 (4.41)	50.14 (2.98)	54.64 (8.88)	61.37 (7.10)	42.18 (5.32)	26.04 (3.18)	297.29 (26.76)
Median	22.00	43.00	51.00	57.00	63.00	43.00	27.00	304.50
Minimum	17.00	25.00	33.00	24.00	36.00	26.00	12.00	182.00
Maximum	22.00	46.00	53.00	67.00	70.00	48.00	28.00	329.00

3.2.2 Research question B: Do the alternate versions of the SPANS continue to offer good alternate version reliability when used with older adults?

Descriptive statistics were employed in conjunction with correlational and comparative analyses to explore the reliability of the alternate versions of the SPANS. These within-subjects analyses necessarily included only those participants who completed both SPANS A and SPANS B, providing an overall sample of *N* = 38.

Correlational analyses (one-tailed) were conducted using Kendall's Tau. It was theoretically expected that significant positive correlations would be found between versions A and B across all indices. The Bonferroni correction was applied and determined a critical value for significance of $p < .007$. Analysis (Table 5) supported this expectation, demonstrating large effect sizes for all indices except ORI ($\tau = -.005$, $p > .01$). While theoretically unexpected, this may be a reflection of the lack of variance within this index. Since the SPANS manual reports Pearson's r for correlational analyses of test-retest reliability, Pearson's correlations are additionally reported here for comparison. Test-retest coefficients of $r = .7$ and above are suggested to convey an acceptable level of retest reliability (Kline, 2000), and are achieved here by all indices other than ORI. While τ coefficients fall below $.7$, the conservative nature of this statistic (Field, 2005) may justify a reduction in acceptability thresholds.

Differences in median score did not exceed 1.00, and to aid comparison the manual, mean score differences between versions did not exceed 0.89 (for which the younger cohort described in the manual saw a maximum difference between scores of $.28$). A series of related samples Wilcoxon's signed-rank tests with Bonferroni correction applied (adjusted critical value of $p < .007$) further confirmed that scores were not significantly different (Table 6) between the two versions when the adjusted p value was implemented.

Table 5: Correlation coefficients (Kendall's Tau and Pearson's r) demonstrating the relationship between performance on the two alternate versions of the SPANS (versions A and B) for each SPANS index

SPANS Index	τ	r
ORI	.005	-.004
ACI	.652*	.737*
LAI	.653*	.818*
MLI	.519*	.706*
VPI	.640*	.790*
ECI	.641*	.813*
CFI	.566*	.615*

* $p < .007$

Table 6: A comparison of SPANS index median scores across the alternate versions (A and B) using a related samples Wilcoxon's signed-rank test (*M* and *SD* reported to aid comparison to SPANS manual)

Index	<i>M</i> (<i>SD</i>)		<i>Median</i>		<i>Z</i>
	SPANS A	SPANS B	SPANS A	SPANS B	
ORI	21.32 (1.08)	21.76 (.47)	22.00	22.00	2.17*
ACI	41.97 (3.63)	42.14 (3.21)	43.00	42.00	.245
LAI	50.65 (2.75)	50.78 (2.32)	52.00	52.00	.365
MLI	56.51 (5.45)	56.32 (6.75)	57.00	58.00	-.134
VPI	62.92 (6.05)	63.81 (5.08)	65.00	64.00	1.552
ECI	42.95 (4.89)	43.81 (4.11)	44.00	44.00	1.635
CFI	26.35 (3.30)	26.89 (1.73)	28.00	28.00	.729

* $p < .05$; ** $p < .007$

All indices except MLI indicated an increase in mean score between versions A and B, while MLI demonstrated a (non-significant) decrease. Overall, the lack of significant difference between versions provides support for the alternate version equivalency and reliability of the SPANS.

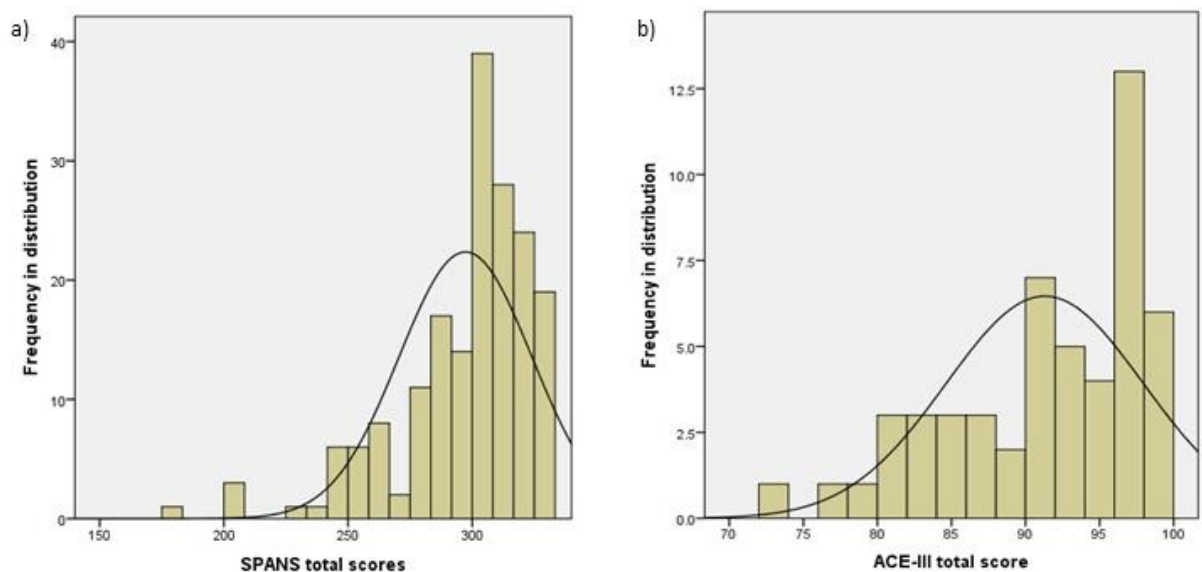
3.2.3 Research question C: Does the SPANS co-vary with another neuropsychological test used with older adults (the ACE-III), and / or constituent parts known to add value in the assessment of dementia (clock drawing and verbal fluency)?

The following within-subjects analyses were completed using data from all participants who completed both the SPANS and the ACE-III ($N = 54$), or SPANS, clock drawing, and fluency ($N = 72$), or ACE-III, clock drawing and fluency ($N = 50$), as required.

3.2.3.1 SPANS and ACE-III

As anticipated, comparison of distributions demonstrated similarly negatively skewed patterns between SPANS and ACE-III data (Figure 1), supporting the expectation that both are based upon a screening paradigm. Values of skewness and kurtosis were noted to indicate a stronger and steeper negative skew in SPANS data (skewness = -1.53, kurtosis = 2.85) compared to ACE-III data (skewness = -.79, kurtosis = -.18). In the context of the healthy and able sample reported here, this may indicate a tendency for the ACE-III to incorrectly classify more people as having dementia than would occur using the SPANS. As such, this may provide initial indication of the potential for the SPANS to offer greater specificity in detection of impairment, however base rates are not known for either assessment, and would need to be referenced in order to properly evaluate this hypothesis

Figure 1: Histogram with bell curve for SPANS (a), ACE-III (b) total scores



Correlational analyses (one-tailed) were conducted using Kendall's Tau (Table 7) to explore the extent to which ACE-III and SPANS scores covaried. It was anticipated that total scores, and theoretically similar domain and index scores, would demonstrate the greatest degree of correlation. The Bonferroni correction was applied, which indicated a critical value of $p < .0001$ for statistical significance. A significant, positive correlation of large effect size was found between SPANS and ACE-III total score ($\tau =$

.54, $p < .01$). Furthermore, all SPANS indices were significantly positively correlated with ACE-III total score, and conversely all ACE-III domains were significantly positively correlated with SPANS total score. This further indicates that the SPANS performed well in comparison to this widely used measure of cognition.

Additionally, theoretically similar constructs (highlighted in Table 7 to aid the reader) were also found to be positively correlated at medium-to-large effect sizes, with the exception of ORI and ACE-III attention for which no significant correlation was found. That MLI was most highly correlated with ACE-III memory, ACI with ACE-III attention, and LAI with ACE-III fluency and language, was of particular note. A significant, large correlation between VPI and the ACE-III visuospatial domain was also encouraging support for the convergent validity of the SPANS indices. Additionally, with only minimal correlations between CFI and ACE-III domains and total, this may indicate that CFI taps into an aspect of cognitive ability which is less well measured by the ACE-III.

While the large effect sizes seen within correlations of the corresponding memory and visuospatial domains convey clear support for convergent validity, the moderate effect sizes seen in relation to attention and language domains may be considered encouraging, but perhaps insufficient, to conclude a satisfactory level of convergent validity.

Some support for divergent validity was also identified, with the lowest correlations existing between ORI and ACE-III fluency and visuospatial domains. However, within other seemingly distinct domains, expected support for divergent validity was not seen. For example, the language and visuospatial domains of each assessment demonstrated unexpected significant correlations of medium effect size. Further anomalous results were also identified. Firstly, it was surprising that ORI was most associated with ACE-III memory, and not more significantly correlated with ACE-III attention, given that both share orientation items. This again likely reflects the lack of variance within ORI. Within this index, the item demonstrating most variance ($SD = .50$) was the 'Prime Minister' item, which appears under the memory domain of the ACE-III. This provides a context within which to understand the covariance between

these domains. It was also surprising that ECI was not correlated to a higher degree with ACE-III attention, given the nature of items contained within this domain.

Finally, large correlations found between the memory and visuospatial elements of each assessment were also unexpected, as were the associations between these elements and total scores. These findings suggest commonalities between the memory and visuospatial elements of each assessment, as well as indicating that both the SPANS and ACE-III total scores may place greater weight upon these domains.

Table 7: Correlation (τ) between SPANS indices and total score, and ACE-III domains and total score, derived from all participants who completed both measures ($N = 54$)

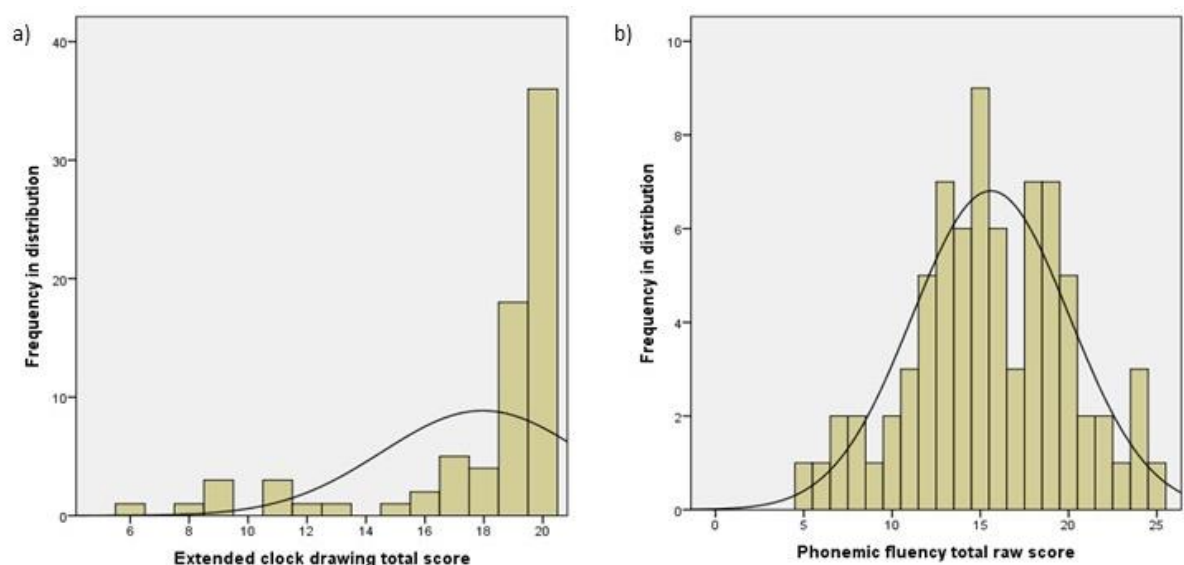
	ACE-III					
	Attention	Memory	Fluency	Language	Visuospatial	Total
ORI	0.19	0.23*	0.09	0.06	0.03	0.17
ACI	0.41***	0.36***	0.15	0.20*	0.19*	0.31***
LAI	0.21*	0.30***	0.33***	0.31***	0.28**	0.40***
MLI	0.33***	0.49***	0.36***	0.26**	0.43***	0.49***
VPI	0.30***	0.52***	0.29***	0.29***	0.49***	0.51***
ECI	0.29***	0.47***	0.22***	0.35***	0.35***	0.43***
CFI	0.16	0.27**	0.22*	0.20*	0.21*	0.29***
Total	0.40***	0.56***	0.33***	0.34***	0.46***	0.54***

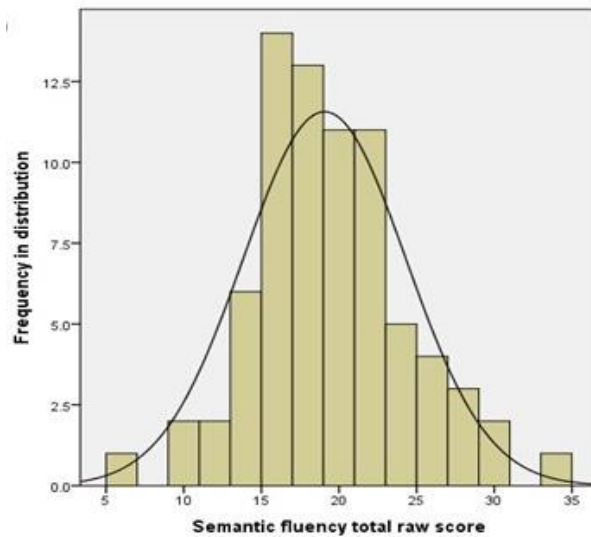
* $p < .05$; ** $p < .01$; *** $p < .0001$; correlation coefficients of theoretically similar constructs are highlighted.

3.2.3.2 SPANS and ACE-III in comparison to Mendez et al.'s (1992) clock drawing score, and verbal fluency raw scores

A comparison of score distributions was conducted, which revealed that clock drawing scores shared the same negative skew as the ACE-III and SPANS, while both phonemic and semantic fluency were normally distributed (Figure 2). It was therefore decided to focus further comparisons upon the verbal fluency tasks, in order to explore and compare the covariance of the SPANS and ACE-III with normally distributed data, which may be sensitive to the range of cognitive abilities existing within the general population.

Figure 2: Frequency distribution and bell curve of Mendez et al (1992) clock drawing score (a), and phonemic (b) and semantic (c) fluency raw scores.





Kendall's Tau correlational analyses (two-tailed) were conducted in order to explore the relationship of both the SPANS (Table 8) and ACE-III (Table 9) with the raw phonemic and semantic fluency scores. The Bonferroni correction was applied, following which statistical significance was determined at the level of $p < .006$ and $p < .008$, for SPANS and ACE-III analyses, respectively. Since the raw semantic and phonemic fluency data explored here are extracted from the ACE-III, correlations between the ACE-III total and extracted fluency raw scores are inherently inflated.

Table 8: Correlations (τ) between SPANS indices and extracted phonemic and semantic fluency scores derived from all current participants ($N = 72$) who completed both measures

	SPANS indices							
	ORI	ACI	LAI	MLI	VPI	ECI	CFI	Total
Phonemic fluency	-.046	-.002	.289***	.232***	.231***	.149	.172	.209*
Semantic fluency	.199*	.295***	.264***	.292***	.377***	.368***	.178	.397***

* $p < .05$; ** $p < .01$; *** $p < .006$

Table 9: Correlations (τ) between ACE-III domains and extracted phonemic and semantic fluency scores derived from all current participants ($N = 50$) who completed both measures

	ACE-III domains					
	Attention	Memory	Fluency	Language	Visuospatial	Total
Phonemic fluency	-.015	.191	.604***	.369***	.282*	.400***
Semantic fluency	.190	.474***	.619***	.294***	.392***	.582***

* $p < .05$; ** $p < .01$; $p < .008$

Of particular interest were statistically significant correlation coefficients indicating a large effect size ($\tau = .5$) or greater. As expected with regard to the ACE-III data, such correlations were identified between both sets of raw fluency scores and the ACE-III fluency domain, and between semantic fluency and ACE-III total score. Moderate and significant correlations were also seen between semantic fluency and both ACE-III memory and ACE-III visuospatial scores, as well as between phonemic fluency and ACE-III language.

Turning to the SPANS, further moderate and significant correlations were noted between semantic fluency and both VPI and ECI, as well as SPANS total score. Meanwhile, phonemic fluency produced correlations of only small to medium effect size with SPANS indices and total score, correlating most highly with LAI. As such, a similar pattern of relationships was observed between semantic and phonemic fluency for both the SPANS and ACE-III, indicating equivalence in convergent validity in relation to the fluency items.

3.2.4 Research Question D: How do age and education interact with the SPANS and ACE-III scores?

3.2.4.1 Age

A Kendall's Tau correlation of SPANS ($N = 180$) and ACE-III ($N = 54$) total scores with participant age was conducted to assess the nature of the relationship between age and SPANS performance. It was theoretically expected that scores would decrease with increasing age, in line with the natural ageing process. In the case of both SPANS and ACE-III, a small negative effect was indicated, which reached statistical significance in relation to SPANS ($r = -.21, p < .01$) but not ACE-III ($r = -.18, p < .05$) scores, supporting expectation.

3.2.4.2 Education

An initial look at median SPANS and ACE-III total scores across the four levels of education recorded (Table 10) indicated that greater scores were associated with higher levels of education.

Table 10: Median SPANS and ACE-III total scores for all current participants for each level of educational achievement as reported in completion of the DHLQ

	Less than secondary		Secondary education		College or vocational		University degree	
	<i>N</i>	<i>Mdn</i>	<i>N</i>	<i>Mdn</i>	<i>N</i>	<i>Mdn</i>	<i>N</i>	<i>Mdn</i>
SPANS total	7	298	28	299	28	305	44	318.5
ACE-III total	5	83	15	87	11	90	23	96

The interaction of education with SPANS and ACE-III total scores was explored using an independent samples Kruskal-Wallis tests to compare scores across the levels of educational achievement. With medium and large effect sizes respectively, this analysis indicated that both SPANS total score ($H(3) = 29.04, p < .01, r = .39$) and ACE-III total score ($H(3) = 16.18, p < .01, r = .48$) were significantly affected by level of education. Dunn's pairwise comparisons were conducted for the 6 pairs of groups (with Bonferroni correction automatically applied). This indicated a significant difference in SPANS scores between those with a university degree and all other levels of education ($p < .05$). Within the ACE-III total scores, a significant difference was reported only between those with secondary education and those with a university degree ($p < .05$).

4. Discussion

4.1 Summary of findings

This study was an initial exploration of the potential validity and reliability of the SPANS when administered to adults over the age of 65 years. Internal consistency, alternate version test-retest reliability, and convergent validity were explored. Key findings, their interpretation, and implications, are discussed below.

4.1.1 Internal consistency

The internal consistency of the SPANS within the older adult sample was assessed using the Cronbach's α statistic. Resulting alpha values were notably lower than had been previously documented with younger age groups. One possible hypothesis may be that the SPANS indices cease to represent useful or meaningful constructs when assessing older adult cognition. It is recommended that this could be further explored by future research, using factor analytic methods. However, in the context of the robust nature of SPANS indices documented previously – noted to endure across the adult age range, and a number of neurological diagnoses of varying severity (Burgess, 2014) - a further hypothesis may be that this finding reflects the highly able, independent, and homogenous nature of the study's participants, whose scores subsequently lacked sufficient variance to support internal consistency. Of particular note was the ORI, which contained the least variance of all, and which consequently demonstrated poorest internal consistency ($\alpha=.33$). This index demonstrated a 'ceiling effect' within the healthy older adult sample, whereby the vast majority of participants score at the maximum possible level. This can be viewed as conveying a great deal of clinical utility since, within a clinical setting, a moderate-to-low score within this index may be seen as clear indicator of impairment.

When the current data were combined with existing SPANS data from previous studies, internal consistency was only minimally impacted, and remained in the 'good' range (Kline, 2000). One interpretation could be that this provided tentative support (in conjunction with findings of convergent validity analyses) to the hypothesis that the constructs assessed by the SPANS at the point of creation endure into older age. It is however acknowledged that, in the context of the large total dataset ($N = 609$) created, it may be unsurprising that the addition of the relatively small older adult data produced little difference. As such, this should be interpreted cautiously, with further exploration of internal consistency of the SPANS in older adult populations warranted.

4.1.2 Alternate version reliability

Correlational analyses provided strong support for alternate version re-test reliability, with large significant effects identified for all indices except ORI, again due to the lack

of variance within this index. No significant differences in scores between versions A and B were identified for any index when ranked values were compared. Therefore, the equivalence of SPANS alternate versions within the older adult population was further supported.

A slight difference in scores (not exceeding .89) was seen in all domains, and constituted an improvement for all indices except MLI. The small magnitude of this difference appeared to indicate limited practice effects among this sample of older adults, which may represent a deviation from the literature. Statistically significant improvements in cognitive assessment scores, with large effect sizes, have been documented among healthy older adult participants, particularly in relation to memory assessment tasks (e.g. Gavett et al., 2016; Duff, 2014) when exploring test-retest and alternate-form reliability across both short (one week) and longer (annual) terms. This finding therefore warrants further investigation, with a possible hypothesis being the existence of a ceiling effect which, although not supported by a visual examination of scores, requires further exploration.

It must also be acknowledged that the minimal differences between scores across the alternate versions of the SPANS were in keeping with the SPANS manual's guidance that up to a one-point difference may be seen at retest (Burgess, 2014). This was based upon the minimal difference also observed among the working age adult sample reported there. Overall differences in score from version A to B were slightly greater among the current older adult sample than those previously reported in a younger cohort (Burgess, 2014), which appeared to indicate greater practice effects within the older adult sample. This would be unusual, and so it is wondered whether, within the current very homogeneously capable, healthy, and independent sample, were individuals who were most able to benefit from repeated administrations. This would be in contrast to the younger cohort described within the SPANS manual, which incorporated participants with a range of diagnoses associated with cognitive impairment (e.g. ABI), and who would not be expected to improve at re-test. It may also be the case that since mean scores on completion of SPANS A were lower within the older adult sample, this left greater 'room for improvement' within this group when considering the potential range of achievable scores. Nonetheless, this finding

represents a deviation from the literature, in which it is documented that practice effects may be reduced among older adults in comparison to younger adults (Salthouse, 2010; Weschler, 1997b).

The observed decrease in MLI scores was additionally of interest, as this was also incongruent with the SPANS manual. However, it is known that alternate content within two items contributing to MLI was designed to be slightly more difficult in version B in order to compensate for practice effects (Burgess, 2014). Within the younger cohort described by the manual, these items demonstrated a small reduction in mean score from version A to B, providing some context within which to understand this finding.

4.1.3 Convergent validity

Convergent validity of the SPANS was assessed through comparison to the ACE-III, which may be considered a current 'gold standard' assessment for use with older adult populations. Further exploration of convergent validity was conducted using raw verbal fluency scores in order to compare the convergent validity of the SPANS and the ACE-III. This allowed for exploration of convergent validity in relation to normally distributed assessment scores, which indicate that this task may assess the full range of potential ability.

Within the current sample, both the SPANS and ACE-III total scores demonstrated 'excellent' internal consistency, in keeping within earlier studies (Burgess, 2014; Hsieh et al., 2013). Comparison of distributional properties identified similarly negatively skewed distributions for both SPANS and ACE-III scores, demonstrating that both are based around a 'screening' paradigm.

Correlational analyses revealed both theoretically anticipated, and unexpected results. Convergent validity was supported by findings that SPANS indices correlated most highly with theoretically matched ACE-III domains, however a mixture of moderate and large effect sizes meant that support for the convergent validity of indices such as LAI, ECI and ACI was of insufficient magnitude to conclude clear achievement of sufficient convergent validity.

Of further interest were the comparatively lower correlations between CFI and all ACE-III scores, for which a number of possible explanations could be hypothesised. It may be that the CFI ceases to represent a meaningful construct within the assessment of older adult cognition – a possibility which again could be further explored through application of factor analysis. Alternatively, this finding may indicate that this SPANS index assesses an aspect of cognition which is less well measured within the ACE-III. This may be supported when considering that items within the CFI may be thought of as reflecting ‘executive functioning’ skills, which are not directly assessed within the ACE-III (Sajjati & Brown, 2015).

Perhaps the most theoretically incongruous finding was the significant and substantial relationships between the memory and visuospatial elements of both assessments. This may reflect the broader nature of memory and visuospatial items within the SPANS in comparison to the ACE-III. For example, while ACE-III memory is entirely verbal, the SPANS MLI incorporates assessment of both verbal and visual learning and recognition. Similarly, the VPI contains recall of visually presented items.

Comparing the convergent validity of the SPANS and ACE-III in relation to verbal fluency, a similar pattern was noted across both assessments indicating further equivalency of these assessments. Semantic fluency demonstrated significant correlations of medium effect size with the visuospatial elements of both assessments, as well as with ACE-III memory and SPANS ECI. This is in keeping with the literature (e.g. Strauss et al., 2006) which indicates that fluency tasks may reflect a range of abilities. Correlations with visuospatial items may be understood as reflecting access to visual representations of responses. The ACE-III memory domain contains items requiring retrieval of well consolidated semantic knowledge, which may underpin this relationship. Finally, the need to pace oneself and manage time effectively during completion of verbal fluency tasks may underpin the relationship to ECI. From these findings, it seems that the visuospatial and memory or efficiency domains of the SPANS and ACE-III may be most sensitive to the full range of existing ability.

4.1.4 Impact of age and education

A significant negative correlation of small effect size was seen between age and SPANS scores, indicating that scores decreased with increasing age, as would be expected. This supports the utility of creating age stratified normative data and cut-off scores for older age ranges, as are currently in place for younger adults. A similar finding was not seen for the ACE-III, indicating that this assessment is less sensitive to the effect of ageing, and therefore supporting the use of a single cut-off score within the older adult population, as is currently recommended (Hsieh et al., 2013).

With regard to education, a comparison of ranked ACE-III and SPANS total scores indicated a significant impact of education upon assessment score. Descriptive statistics and pairwise comparisons indicated that scores increased with higher levels of education, for which achievement of a university degree produced the greatest impact. This effect was greater in relation to ACE-III scores, indicating that this test was more sensitive to the impact of higher levels of education. A lesser effect of education in relation to the SPANS may indicate that education has less of a mediating effect upon SPANS scores in comparison to the ACE-III. This would be clinically beneficial, as lower scores could be seen to more reliably reflect cognitive decline, rather than also reflecting lower levels of education.

4.2 Limitations

This research was the first exploration into the potential reliability and validity of the SPANS with older adults, and has demonstrated promising results. Nonetheless, this work was subject to limitations which have been held in mind when evaluating the findings presented above, and which limit the ability to generalise conclusions from this work to the wider older adult population.

A significant limitation stemmed from difficulties in the recruitment of participants – both in terms of numbers, and diversity of ability. In a number of cases, gatekeeper approval was granted to advertise to members of organisations, however no successful recruitment occurred. In some cases, it seemed that recruitment from a distance – that is through posters and flyers alone – may have acted as a barrier. Face-to-face

recruitment is naturally more engaging, with the opportunity to ask questions and become familiar with the researcher. This may put potential participants at ease, and foster a feeling of 'wanting to help', as well as removing the need for potential participants to 'take the first step' of contacting the researcher. Those participants who did respond to a flyer or poster likely represented those who were most able and independent, and perhaps most confident in their ability to take on the task of participation. As a result, the present sample consisted largely of very well able, functionally independent individuals. This culminated in much lower variance within the current data than might reasonably be anticipated within the wider older adult population. As previously discussed, the lack of variance had a repeated impact throughout data analysis.

Overall, difficulties in recruitment meant that the target sample size was not achieved, resulting in insufficient statistical power for a number of the reported analyses. This was further problematic in the context of the inherently negatively skewed data, as the required non-parametric analyses were also less powerful than their parametric alternatives (Field, 2005). Due to the high number of tied ranks within the data, a particularly conservative method of correlation was warranted, which may have tended towards underestimation of the magnitude of the observed effect. A further conservative correction was also applied (Bonferroni's correction) to manage the potential for inflation of false-positive error rates in the context of multiple comparisons. As such, findings here may be an underestimate of true effects within the data.

Within the design of the study, the potential impact of order effects has been discussed. The approach taken - maintaining the same order of SPANS and ACE-III presentation across all participants - was felt to be the most ethical. The consequences of this in terms of impact upon ACE-III performance are acknowledged. Furthermore, although the decision to administer the ACE-III only once was taken to reduce the cognitive load placed upon participants, it is further acknowledged that administering this assessment alongside both SPANS A and SPANS B may have facilitated the counterbalancing of administration sequence, which could have been implemented at the second testing session

A further limitation with regard to study design is reflected in the limited opportunities to explore divergent validity within the current study. Inclusion of a further assessment measure, for the specific purpose of divergent validity analysis would have conveyed a useful contribution to the understanding of SPANS performance within an older adult sample.

It is further acknowledged that adopting a mixed methods design would have enabled a formal gathering of feedback from participants regarding their experience of completing the SPANS. This would have allowed for discussion of the acceptability and face validity of the SPANS among older adults, both of which would provide useful insight into the ability of the SPANS to engage this client group.

Additionally, it was not possible for record sheets to be 'cross checked' by the Chief or Principal Investigator. Therefore, while all VRAs demonstrated competence in administration and scoring when practice administrations were completed, no ongoing check of interrater reliability was performed. As such, although the SPANS has previously demonstrated good inter-rater reliability between trained raters (Burgess, 2014), the possibility of researcher-introduced error within the data cannot be fully ruled out.

Finally, alternative explanations for the reduced internal consistency of the SPANS seen within this older adult sample could also have been explored, in addition to the idea of the lack of variation as being a causal factor. It may be possible that the SPANS may measure different cognitive constructs within the older adult population in comparison to the working age population – a possibility supported by evidence of structural change and functional reorganisation within the brains of older adults (e.g. Meunier, Stamatakis & Tyler, 2013). This could have been addressed through factor analysis of the older adult SPANS data, to identify statistically driven structures, rather than accepting the theory-driven index structure currently employed by the SPANS (Burgess, 2014). Such analysis was not undertaken here for several reasons – firstly that the SPANS was not originally developed in this way, and so conducting a factor analysis in relation to only the older adult sample felt inappropriate, in the absence of a similar exploration of the working-age adult data. Secondly, as convergent validity

was supported in relation to an existing measure of older adult cognition, this was felt to support the existing structure of the SPANS as applied within the older adult population. Thirdly, a number of other cognitive assessments employ theory-driven, as opposed to statistically-driven, internal structures (e.g. the RBANS; Schmitt et al., 2010) for which exploration of convergent validity was deemed sufficient evidence for the measurement of the constructs described.

4.3 Implications of the current findings

The current findings provide initial indication that the SPANS and its constituent indices may continue to represent a useful measure of cognitive ability within a sample of older adults. Alternate version reliability was upheld, indicating that repeated assessment could be undertaken as clinically required. Support for continued convergent validity was encouraging in relation to the ACE-III, while the SPANS and ACE-III demonstrated similar patterns of performance and interaction with verbal fluency tasks, which may represent a broader assessment of cognitive ability. Furthermore, preliminary findings suggest that the CFI within the SPANS may reflect an aspect of cognitive ability (such as executive functioning skills) which is less well assessed within the ACE-III.

Support was also identified for the creation of age-stratified norms within this population, in line with SPANS' current offering for younger adult groups.

Overall, this constitutes a promising indication that the SPANS may continue to be a useful and clinically beneficial tool in the assessment of cognition in older adult populations. The SPANS may perform at similar levels to an existing 'gold standard' measure, and could offer additional benefits beyond those currently associated with the ACE-III.

4.4 Recommendations for future research

In light of the above implications and limitations, further research is recommended to expand the initial exploration of reliability and validity conducted here, and to determine whether these promising results remain with a larger and more diverse

sample. As such, future research should include a mixed clinical and healthy normative sample, from diverse clinical and socioeconomic groups, in order to increase variance. This would additionally facilitate comparison to the younger age sample described by the SPANS manual (Burgess, 2014), and investigation of sensitivity and specificity for detection and differentiation of conditions impacting cognition in older populations, such as dementia, in comparison to existing measures such as the ACE-III.

Such work would also allow for exploration of discriminate validity, and understanding of the SPANS' potential contribution to differential diagnosis in older adult populations. This could be followed by further exploration of discriminate validity, specifically in relation to dementia subtype diagnosis.

Lastly, a factor analytic approach may be beneficial to confirm the reliability of the existing index structure within an older adult population, which could be further supported through the inclusion of a measure of divergent validity.

4.5 Conclusions

This study joins a growing number of explorations of the SPANS across a range of settings and populations. Provided here is an initial exploration of the reliability and validity of the SPANS within a healthy and functionally able older adult sample. Findings were subject to a number of limitations, however despite this have demonstrated that further exploration within the older adult population is warranted. On the basis of learning from the current work, suggestions to guide future research have been provided. Continued study will be vital in order to fully understand the potential contribution of the SPANS to the field of older adult clinical neuropsychology.

References

- Armstrong, R.A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, 34, 502-508.
- Alzheimer's Research UK. (2018). Different types of dementia. Retrieved January 2019 from <https://www.dementiastatistics.org/statistics/different-types-of-dementia/>
- Alzheimer's Society. (n.d.). Dementia advice – Help for people with dementia and their families. Retrieved November 2018 from <https://www.alzheimers.org.uk/get-support/your-support-services/dementia-advice>
- Alzheimer's Society. (2015). Helping you to assess cognition: A practical toolkit for clinicians. Retrieved April 2017, from www.alzheimers.org.uk/download/downloads/id/3475/alzheimers_society_cognitive_assessment_toolkit.pdf
- American Psychiatric Association, DSM-5 Task Force. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Arlington, VA, US: American Psychiatric Publishing, Inc.
- Burgess, G. (2014). *Short parallel assessments of the neuropsychological status: Manual*. Oxford: Hogrefe.
- Cheung, G., Clugston, A., Croucher, M. et al. (2015). Performance of three cognitive screening tools in a sample of older New Zealanders. *International Psychogeriatrics*, 27(6) 981-989.
- Coen, R.F., Cahill, R., & Lawlor, B.A. (2011). Things to watch out for when using the Montreal cognitive assessment (MoCA) [letter to the editor]. *International Journal of Geriatric Psychiatry*, 26, 106-108.
- Coen, R.F., Robertson, D.A., Kenny, R.A., & King-Kallimanis, B.L. (2016). Strengths and limitations of the MoCA for assessing cognitive functioning: Findings from a large representative sample of Irish older adults. *Journal of Geriatric Psychiatry and Neurology*, 29(1), 18-24.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.

- Colman, A. (2009). *Oxford dictionary of psychology* (3rd edn). Oxford: Oxford university press.
- Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
- Duff, K. (2014). One-week practice effects in older adults: Tools for assessing cognitive change. *The Clinical Neuropsychologist*, 28(5), 714-725.
- Duff, K., Humphreys Clark, J.D., O'Bryant, S.E. et al. (2008). Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's Disease: Sensitivity, specificity, and positive and negative predictive powers. *Archives of Clinical Neuropsychology*, 23, 603 – 612.
- Duff, K., Hobson, V.L., Beglinger, L.J., & O'Bryant, S.E. (2010). Diagnostic accuracy of the RBANS in mild cognitive impairment: Limitations on assessing milder impairments. *Archives of Clinical Neuropsychology*, 25, 429-441.
- East Midlands Clinical Networks. (2017). Guide to the provision of psychological support following stroke. Retrieved February 2019 from <https://www.england.nhs.uk/mids-east/wp-content/uploads/sites/7/2018/04/10-stroke-psychological-support.pdf>
- Field, A. (2005). *Discovering statistics using SPSS* (2nd edn). London: Sage Publications Ltd.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189-198.
- Frerichs, R. (2004). When should an older adult be referred to neuropsychology? *The Canadian Alzheimer's Disease Review*, 7(1), 4-9.
- Gavett, B.E., Gurnani, A.S., Saurman, J.L., et al. (2016). Practice effects on story memory and list learning tests in the neuropsychological assessment of older adults. *PLoS ONE*, 11(10), e0164492. doi:10.1371/journal.pone.0164492

- A. & İşçi, O. (2011). A Comparison of the Most Commonly Used Measures of Association for Doubly Ordered Square Contingency Tables via Simulation. *Metodološki Zvezki*, 8(1), 17-37.
- Goldstein, L.H., & McNeil, J.E. (Eds.) (2004). *Clinical neuropsychology: A practical guide to assessment and management for clinicians*. Sussex: John Wiley & Sons Ltd.
- Haddlesey, C. (2016). *An exploration of the impact of language and culture on a neuropsychological test*. Unpublished doctoral thesis, Canterbury Christ Church University.
- Harker-Brown, D. (2018). *An exploration of the quality of life of people labelled with intellectual disabilities, and their capabilities as assessed by the Short Parallel Assessments of Neuropsychological Status (SPANS)*. Unpublished doctoral thesis, University of Leicester.
- House of Commons Library. (2018). *Dementia: Policy, services and statistics*. Retrieved August 2018 from <https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN07007>
- Hsieh, S., Schubert, S., Hoon, C. et al. (2013). Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 36, 242-250.
- Hubbard, E.J., Santini, V., Blankevoort, C.G. et al. (2008). Clock drawing performance in cognitively normal elderly. *Archives of Clinical Neuropsychology*, 23, 295-327.
- International Consortium for Health Outcome Measures. (2017). *Dementia data collection reference guide*. Retrieved September 2017 from <http://www.ichom.org/download/dementia-reference-guide/>
- Jacova, C., Kertesz, A, Blair, M. et al. (2007). Neuropsychological testing and assessment for dementia. *Alzheimer's & Dementia*, 3(4), 299-317.

- Jubb, M.T. & Evans, J.J. (2015). An investigation of the utility of the Addenbrooke's Cognitive Examination III in the early detection of dementia in memory clinic patients aged over 75. *Dementia and Geriatric Cognitive Disorders*, 40, 222-232.
- Kline, P. (2000). *A Psychometrics Primer*. London: Routledge.
- Mathuranath, P.S., Nestor, P.J., Berrios, G.E. et al. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology*, 55, 1613-1620.
- Matias-Guiu, J.A., Cortés-Martínez, A., Valles-Salgado, M. et al. (2017) Addenbrooke's cognitive examination III: Diagnostic utility for mild cognitive impairment and dementia and correlation with standardized neuropsychological tests. *International Psychogeriatrics*, 29(1), 105-113.
- Matias-Guiu, J.A., Valles-Salgado, M., Rognoni, T. et al. (2017). Comparative diagnostic accuracy of the ACE-III, MIS, MMSE, MoCA and RUDAS for screening of Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, 43, 237-246.
- Mendez, M.F., Ala, T., & Underwood, K.L. (1992). Development of scoring for the clock drawing task in Alzheimer's Disease. *Journal of the American Geriatric Society*, 40, 1095-1099.
- Meunier, D., Stamatakis, E.A., & Tyler, L.K. (2013). Age-related functional reorganization, structural changes, and preserved cognition. *Neurobiology of Ageing*, 35, 42-54.
- Meyers, J.E., & Meyers, K.R. (1995). *Rey Complex Figure Test and recognition trial professional manual*. Odessa: Psychological Assessment Resources.
- Mioshi, E., Dawson, K., Mitchell, J. et al. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21(11), 1078-1085.
- Mioshi, E., Hsieh, S., Savage, S. et al. (2010). Clinical staging and disease progression in frontotemporal dementia. *Neurology*, 74(20), 1591-1597.

- Mitrushina, M., Boone, K.B., Razani, J. et al. (2005). *Handbook of normative data for neuropsychological assessment* (2nd edn). New York: Oxford University Press.
- Nasreddine, Z.S., Phillips, N.A., Bédirian, V. et al. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699.
- National Institute for Health and Care Excellence. (2018). Dementia: assessment, management and support for people living with dementia and their carers. Retrieved February 2019 from <https://www.nice.org.uk/guidance/ng97/resources/dementia-assessment-management-and-support-for-people-living-with-dementia-and-their-carers-pdf-1837760199109>
- NeuRA (2016). The Addenbrooke's Cognitive Examination III frequently asked questions. Retrieved April 2017 from www.neura.edu.au/wp-content/uploads/2016/06/ACE-III-FAQ-June-2016.pdf
- Quinn, T.J., Elliott, E., & Langhorne, P. (2018). Cognitive and mood assessment tool for use in stroke. *Stroke*, 49, 483-490.
- Randolph, C., Tierney, M.C., Mohr, E. et al. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Preliminary clinical validity. *Journal of Clinical and Experimental Neuropsychology*, 20(3), 310-319.
- Reitan, R.m. & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery*. San Antonio TX: Psychological Corporation.
- Robinson, L. (2015). Dementia: timely diagnosis and early intervention. *British Medical Journal*, 350.
- Sajjadi, S.A. & Brown, J. (2015). Clinical assessment of patients with dementia. *Advances in Clinical Neuroscience and Rehabilitation*, 15(4), 10-13.
- Salthouse, T.A. (2010). Influence of age on practice effects in longitudinal neurocognitive change. *Neuropsychology*, 24(5), 563-572.

- Savage, S., Hsieh, S., Leslie, F. et al. (2013). Distinguishing subtypes in primary progressive aphasia: Application of the Sydney Language Battery. *Dementia and Geriatric Cognitive Disorders*, 35, 208-218.
- Schmidt, M. (1996). *Rey Auditory Verbal Learning test: A handbook*. Los Angeles: Western Psychological Services.
- Schmitt, A.L., Livingston, R.B., Smernoff, E.N. et al. (2010). Factor analysis of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in a large sample of patients suspected of dementia. *Applied Neuropsychology*, 17(1), 8-17.
- Schoenberg, M. R., & Scott, J. G. (2011). *The little black book of neuropsychology: A syndrome-based approach*. Boston, MA: Springer.
- Stirling, J., & Elliot, R. (2008). *Introducing Neuropsychology* (2nd edn). Hove: Psychology Press.
- Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological tests* (3rd edn). New York: Oxford University Press.
- Tan, Y.W. & Burgess, G.H. (2018). Multidimensional effects of acculturation at the construct or index level of seven broad neuropsychological skills. *Culture and Brain*, doi.org/10.1007/s40167-018-0075-1
- Velayudhan, L., Ryu, S., Raczek, M. et al. (2014). Review of brief cognitive tests for patients with suspected dementia. *International Psychogeriatrics*, 26(8), 1247-1262.
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale* (3rd Edn). San Antonio: Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale* (3rd Edn). San Antonion: Psychological Corporation.

World Health Organisation. (2017). *Dementia fact sheet*. Retrieved September, 2018,
from <http://www.who.int/news-room/fact-sheets/detail/dementia>

Appendices

(Mandatory Appendices denoted by *)

Appendix A*:

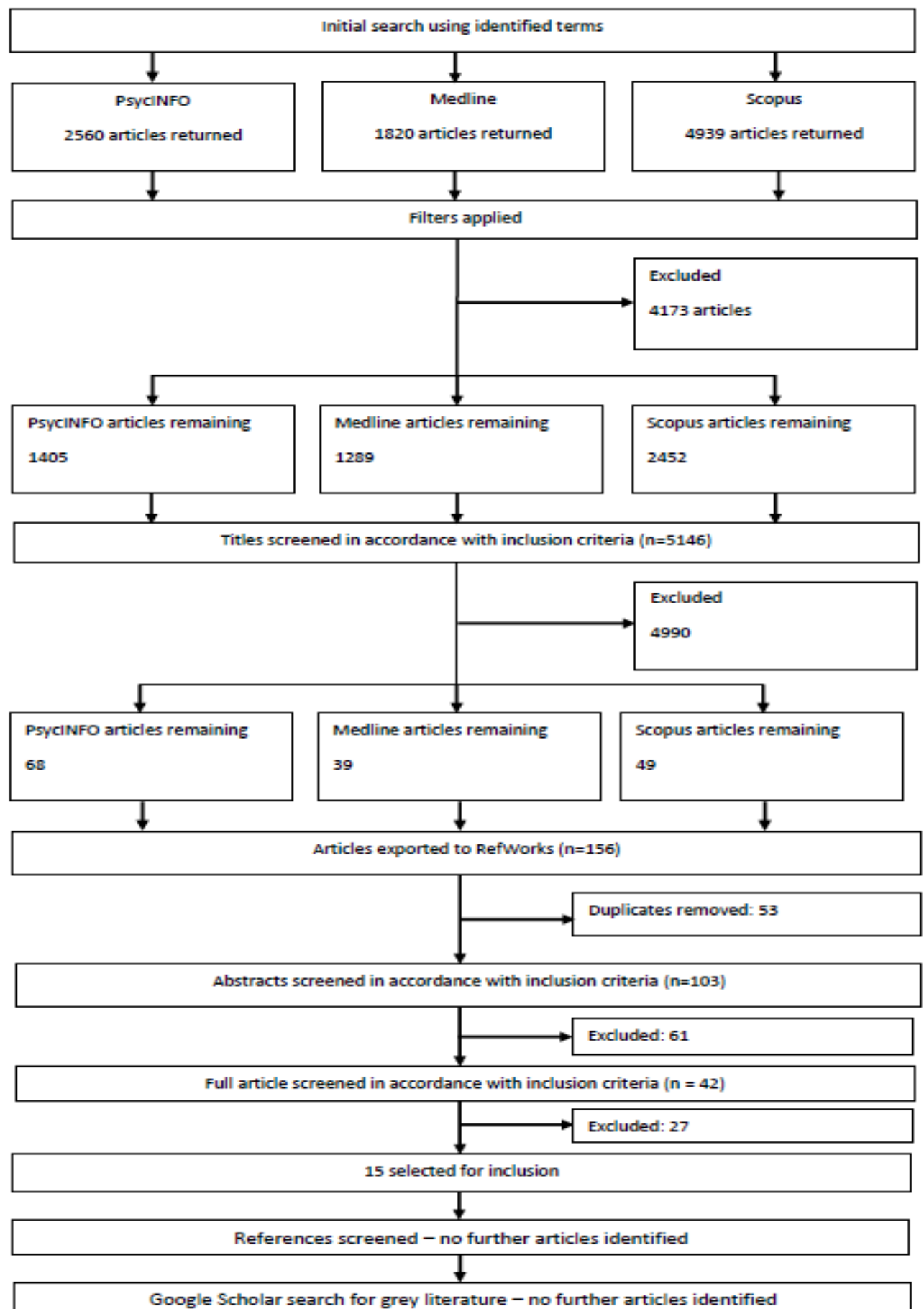
It is anticipated that this thesis will be submitted to the Journal of Neuropsychology for publication. The thesis will be amended post-viva to comply with the 'Author Guidance' provided by this journal, which can be found at: https://pericles.pericles-prod.literatumonline.com/page/journal/17486653/homepage/forauthors.html#_1._SUBMISSION

Appendix B: Inclusion criteria and search filters

	Search filters applied
1.	Published from 2008 onwards
2.	English language
3.	Journal article or thesis

	Inclusion criteria	Exclusion criteria
1.	Conducted within western cultural setting	Focus on coping
2.	Focus on specified dementia subtype(s), limited to AD, VAD, FTD and DLB	Young onset dementia diagnosis
3.	Focus on informal caregivers	Focus on intervention efficacy
4.	Focus on community-dwelling persons with dementia	Comparison to caregiver burden among the caring for non-dementia conditions and MCI
5.	Incorporates elements of, or focus on, understanding or exploring the experience of carer burden	Focus on predicting carer burden

Appendix C: Visual representation of systematic search process



Appendix D: Data extraction pro forma

Title:						
1st Author:						
Year of publication:						
Location:						
Background						
Aims / research question:						
Methodology						
Quantitative				Qualitative		
Sampling:						
Dementia subtype(s):						
Carer relationship(s):						
	<i>N</i>	Gender	Age (range, mean, SD)	Living arrangement	Time spent caring	Duration of caring
Carer						
	<i>N</i>	Gender	Age (range, mean, SD)	Time since diagnosis	Disease stage / severity	
PwD						
Design						
Design:						
Data collection and outcome measurement:						
Burden measure:						

Results and analysis	
<p>Quantitative</p> <p>Level of burden (<i>M</i>, <i>SD</i>):</p> <p>Total and index / subscale scores:</p> <p>Power:</p> <p>Effect size:</p> <p>Statistical comparisons:</p> <p>Additional detail:</p>	<p>Qualitative</p> <p>Key themes:</p> <p>Further detail:</p>
Discussion	
<p>Conclusions:</p>	
<p>Limitations:</p>	

Appendix E: Quality Appraisal Proforma

Criteria and questions to consider		Fully sufficient (score 3)	Mostly sufficient (score 2)	Lacking detail (score 1)	Insufficient detail / Cannot tell (score 0)
Clear description in abstract Was a clear description of the study provided within the abstract?					
Clear aims Were the research aims clearly stated? Was sufficient rationale / background provided? Was a clearly focussed issue / research question stated?					
Appropriate and clear methodology Was the chosen methodology clearly stated? Was the chosen methodology appropriate to the aims?					
Appropriate and clear design Was the research design clearly stated? Was the design appropriate to the research aims?					
Appropriate and clear sampling Was the recruitment / sampling strategy clearly stated? Was the strategy appropriate to the research aims? Was the resulting sample representative of the target population?					
Issues of specific relevance to the current study	Sufficient detail of PwD characteristics Sufficient reporting of e.g. age, gender, diagnosis, disease duration, disease severity etc.				
	Sufficient detail of caregiver characteristics Sufficient reporting of e.g. age, gender, relation to care-recipient, duration of caregiving, time spent caregiving, additional support, living arrangement etc.				
	Clear and appropriate data collection and measures Were data collection methods explicitly stated?				

	If quantitative, was a reliable, validated measure used? Were data collection methods appropriate to the research question? Was rationale for data collection methods justified?				
Identifies sources of bias or confounds Were confounding factors identified and discussed? If qualitative, was the relationship between researcher and participants considered / discussed?					
Ethical issues considered Were ethical considerations discussed? Was ethical approval sought from a relevant committee?					
Rigorous analysis Are details of analysis provided? Are contradictory findings addressed? If quantitative, are elements such as statistical power, or effect size reported?					
Clear findings Are results / findings clearly stated? Was there adequate discussion of findings in the context of the wider literature?					
Potential for generalisability Could the findings generalise to the wider population? Is this discussed?					
Valuable / useful research Is the contribution / impact of the research discussed? How useful is the research to the context within which it sits? Are recommendations for further research provided?					

Note: Criteria derived and adapted from the Critical Appraisal Skills Programme (CASP) criteria for qualitative and cohort studies, and from criteria outlined by Hawker et al (2002).

Appendix F: Quality appraisal scores summary

Study	Sum of key criteria (x/9)	% score
Nicolaou et al. (2010)	5	78.6
Bursch & Butcher (2012)	8	85.7
Oyebode et al. (2013)	4	85.7
Besser and Galvin (2018)	7	88.1
Galvin et al. (2010)	7	88.1
Iavarone et al. (2014)	7	88.1
Leggett et al. (2010)	6	88.1
Pulelwich et al. (2018)	7	88.1
Raggi et al. (2015)	8	88.1
Svendsboe et al. (2016)	8	90.5
Välimäki et al. (2012)	7	90.5
D'Onofrio et al. (2015)	7	95.2
Shim et al. (2012)	7	95.2
Zucchella et al. (2012)	9	100.0
Weigierrek (2012)	N/A	N/A

Appendix G: Summary of papers reviewed

Part 1: Qualitative papers

Author (Year)	Aims	Data source and methodology	Sample					Location / Setting	Resulting themes
			<i>N</i>	Dementia subtype	Carer relation to PwD	Carer input (<i>M</i>)	Dementia severity & cognitive impairment (CDR and MMSE, <i>M</i> (SD))		
Bursch & Butcher (2012)	To develop understanding of caregivers' deepest feelings	Diary study Ricoeurian interpretation	27	AD	Spouse 60%, 33% child, 7% other	81.5% living with PwD 125.5 (58.2) hours per week	Unknown	USA Sample taken from larger study – details unavailable.	Theme: Feeling connected or disconnected in friendship Subthemes: The experience of losing connection to the care recipient; strategies to stay connected to the care recipient; experiencing the presence or absence of friendship with family and friends; questioning own self-worth Theme: Trying to find authenticity for the self and care recipient Subthemes: witnessing the care recipient's loss of authenticity; experiencing loss of authenticity after assuming the role of caregiver Theme: Struggling for self esteem Subtheme: strategies to find self- esteem; Experiencing threats to self-esteem

									Theme: suffering impaired capacity to act Subthemes: the experience of not being able to act autonomously; the experience of autonomy in changing care settings; worrying about the future; trying to reclaim capacity to act.
Oyebode et al. (2013)	To obtain a detailed description of caregiver experience	Face-to-face interview Interpretive Phenomenological analysis	6	FTD	Spouse 66%, 16.5% child, 16.5% brother	83.3% living with PwD	Unknown	UK Setting unknown	Changes in appetite and drives; lack of forward planning; loss of inhibition leading to socially embarrassing behaviours; risky behaviour because of lack of common sense and judgement; communication problems; taking on tasks and roles; defending, asserting and explaining; promoting quality of life; working around lack of awareness; loss of the person and relationship, and heartbreak; sources of support, solace and hope.
Shim et al. (2012)	To explore caregiving experiences	Interview data Manifest and latent content analysis	21	AD	Spouse	100% living with PwD	Unknown	USA Secondary analysis of data from Project ASSIST, which recruited	Caregiver's narratives were used to group them into negative, ambivalent and positive subgroups, reflecting their experience of caregiving.

								through memory clinic, medical clinics and care agencies.	
Välimäki et al. (2012)	To describe life orientation of caregivers in the first year following diagnosis	Diary study Inductive content analysis	83	AD	Spouse 77%, child 22%, other 1%	83.1% living with PwD	<u>CDR-SB</u> 4.4 (1.8) <u>MMSE</u> 21.5 (3.3)	Finland Setting unknown	<p>Core theme: meaning of the onset of AD for the lives of family caregivers</p> <p>Themes: time of doubt; inner conflict; a conception is set</p> <p>Subthemes: becoming aware; contradictory feelings; coming up against a brick wall; a need for information.</p> <p>Core theme: changes in caregivers' personal milieu</p> <p>Themes: transition to being a family caregiver; living in a tunnel; finding harmonizing elements</p> <p>Subthemes: daily care; responsibilities outside the home; the foundations of life are wobbling; life narrows to the home; a desperate atmosphere</p> <p>Core theme: familial cohesion</p> <p>Themes: ensuring familial cohesion</p> <p>Subthemes: embracing what is intact in the family; unexpected changes in the family</p>

									Core theme: Creating a new future Themes: hopeless future; confident about the future Subthemes: My own life is threatened by the disintegration of the family; getting to grips with the disease; hope for recovery; we will manage in the future
Wegierek (2012)	To highlight the impact of caregiving upon the carer	Retelling of narrative. No analysis.	1	AD	Daughter	100% living with PwD 24 hour support	Unknown	USA Narrative told over the course of therapy sessions	Narrative was simply retold, and not analysed for themes. The present author noted ideas around loss of the self to the other; fear and uncertainty; and negative emotions.

Note: Inaccuracies in percentage reporting may be present due to hand calculation

*Reported in the absence of CDR and MMSE data

PwD = Person with dementia; AD = Alzheimer's Disease; CDR-SB = Clinical Dementia Rating Scale sum of boxes; MMSE = Mini Mental State Examination; FTD = Frontotemporal Dementia; FTD = Frontotemporal Dementia

Part 2: Quantitative papers

Author (Year)	Aims	Design	Sample					Location / Setting	Relevant Measures	Key Findings
			<i>N</i>	Dementia subtype	Dementia severity & cognitive impairment (CDR and MMSE, <i>M</i> (SD))	Carer relation to PwD	Carer input (<i>M</i>)			
Besser & Galvin (2018)	To examine the domains of FTD caregiver burden, and the associated caregiver and patient characteristics.	Cross- sectional	674	FTD	<u>QDRS</u> 18.4 (7.6)	82% spouse; 11.6% child; 6.4% other. 68.4% living with PwD	Unknown	USA Recruitment and data collection through Association for FTD website	ZBI 12-item version	Caregivers reported burden in the severe range, which did not differ significantly across mild to terminal disease stage. 'Role strain' was indicated to be most burdensome, followed by 'performance strain', and finally 'personal strain' was least burdensome.
D'Onofrio et al. (2015)	To characterise differences in carer burden experiences of those	Cross- sectional	506	AD (<i>N</i> =253) VaD (<i>N</i> =253)	<u>CDR</u> AD 1.32 (SD = 0.51) VaD 1.15 (SD = 0.5)	Spouse or other relative: 97.6% AD and 91.7% VaD.	AD 8.15 (SD = 0.72) hours per day, for 16.28 (SD	Foggia, Italy Recruitment through an Alzheimer's Evaluation Unit	CBI	Caregivers of those with AD reported significantly higher levels carer burden than those caring for VaD patients.

	caring for PwDs of AD and VaD subtypes.				<u>MMSE</u> AD 18.48 (3.83) VaD 20.46 (3.2)	Remainder private carers. Living arrangements not reported.	= 8.18) months VaD 5.72 (SD = 2.20) hours per day, for 14.35 (SD = 8.89) months			Significant differences were found in aspects of carer burden experienced, with AD carers identifying greater levels of objective, physical, social and emotional burden.
Galvin et al. (2010)	To determine the unmet needs of DLB caregivers through exploration of the challenges, burdens and frustrations which they face.	Cross-sectional	971	DLB	Reported as 'moderate to severe' based on questionnaire responses. Mean time since diagnosis*: 6.6 years (SD = 4)	Spouse 40.6%; child 51.7%; other relative 3.8%; friend 3.8% 46.3% living with PwD	Daily contact	USA Recruitment and data collection through Lewy Body Dementia Association website	ZBI 11- item version	Overall burden reported fell within the severe range. Findings reported as % of responses in each 'frequency category' (e.g. never, rarely etc.) per measure item. Most endorsed items related to fear for loved ones' future, stress of balancing responsibilities, impact on social life, and uncertainty regarding what to do.
Iavarone et al. (2014)	To assess burden and anxiety as experienced by caregivers of	Cross-sectional	86	AD	Assessed by CDR however values not reported. Reported as spanning 'mild' to 'severe', with	Spouse 43%; child 45%, other 12%. Living arrangements	Daily contact	Naples, Italy Memory clinic of the neurological unit of AORN	CBI	Caregivers reported burden in the upper range, indicating need for respite. Time-related strain appeared most

	people with AD.				majority in 'moderate' to 'severe' range.	ts not reported.		Cardarelli Hospital		burdensome, followed by developmental strain. Social and emotional elements appeared least burdensome to these caregivers.
Leggett et al. (2010)	To explore levels of burden and emergent factor structure in relation to predictive factors of burden in DLB.	Cross-sectional	611	DLB	Mean time since diagnosis*: 6 years (range 1 – 22 years)	Spouse 42%; child / child-in-law 48%. 49.8% living with PwD.	Daily contact	USA Recruitment and data collection through Lewy Body Dementia Association website	ZBI 12-item version	Caregivers reported overall burden in the 'severe' range. Performance-related strain appeared most burdensome, closely followed by 'role strain', and finally 'personal strain'.
Nicolaou et al. (2010)	To investigate the needs of, and experience of burden among, carers of FTD and AD patients.	Cross-sectional	60	FTD (N =30) AD (N =30)	Mean time since diagnosis*: AD = 38.8 months (SD = 23.1) FTD = 45.4 months (SD = 24.5)	<u>AD</u> Spouse 70%, child 30% 38.3% living with PwD <u>FTD</u> Spouse 93%, child 7%. 46.7% living with PwD.	FTD 2.7 (SD= 0.6) hours per week, for 5.6 (SD = 2.4) years AD 2.5 (SD = 0.7) hours per week, for 5.1 (SD = 3.1) years.	Australia Referred from Alzheimer's Australia Ltd and healthcare professionals	ZBI 22-item version	Carers of FTD patients reported higher levels of burden overall, compared to carers of AD patients. Carers of FTD patients reported higher levels burden related to both personal strain and role strain than those caring for AD patients.

Pudelewicz et al. (2018)	Explore burden among AD caregivers, and relationship with other factors	Cross-sectional	55	AD	Range from mild to severe, majority 'moderate'. Measure not reported.	Spouse 36%, child 49%, other 8%. 82% living with PwD.	11.12 (6.11) hours per day	Poland Greater Poland Association for people with Alzheimer's Disease, and the Greater Poland Alzheimer's Association, Poznan.	CB scale (polish version)	Total burden reported in the 'medium' range. Burden predominantly underpinned by 'general strain', 'disappointment' and 'social strain', with environmental aspects appearing least burdensome.
Raggii et al. (2015)	To explore the factors contributing to burden and the related coping strategies	Cross-sectional	73	AD	<u>CDR</u> Median = 2, interquartile range = 1 – 2 <u>MMSE</u> Median = 16, interquartile range = 11 - 19	Spouse 57.5%, child 38.4%, sibling or nephew 4.1% 87.7% living with PwD	unknown	Italy Morgagni-Pierantoni Hospital of Forlì, and the Institute for Research on mental Retardation and Brain Ageing of Troina.	CBI	Total burden reported fell within the 'need for respite' range. Time-related and developmental aspects of strain were most burdensome, social and emotional aspects appeared less burdensome.
Svendsboe et al. (2016)	To characterise differences in caregiving related distress of those caring for PwDs of	Cross-sectional	186	AD (N =100) DLB (N =86)	<u>CDR sum of boxes</u> AD 4.9 (2.0) DLB 5.4 (2.9) <u>MMSE</u> AD 23.6 (SD = 2.4)	<u>AD</u> Spouse 50%, child 50% <u>DLB</u> Spouse 62%, child 38%	Unknown	Norway Recruited from a longitudinal dementia study cohort and The Norwegian Dementia Register	RSS	Caregivers of people with DLB reported significantly higher levels of caregiver burden than carers of people with AD. DLB caregivers indicated greater levels

	DLB and AD type.				DLB 23.4 (SD = 3.2)					<p>of emotional and social distress than carers of people with AD.</p> <p>Groups did not differ in respect to distress linked to negative feelings.</p>
Zucchella et al. (2012)	To explore caregiver burden and coping in the early stages of AD	Cross-sectional	126	AD	<p>CDR Not utilised</p> <p>MMSE 22 (SD = 2.86)</p>	partner 32%, child = 64% other relative 4%	<p><12 hours per week (82%) >12 hours per week (18%)</p> <p>< 1 year (16%) 1-2 years (61%) 2-3 years (23%)</p>	<p>Pavia, Italy</p> <p>Alzheimer's Disease Assessment Unit</p>	CBI	<p>Carers reported burden in the low range.</p> <p>Time-related burden contributed most to total burden score, followed by developmental, physical, social and then emotional burden.</p>

Note: Minor inaccuracies in percentage reporting may be present for those which have been calculated by hand

*Reported in the absence of CDR and MMSE data

PwD = Person with dementia; AD = Alzheimer's Disease; VaD = Vascular Dementia; CDR = Clinical Dementia Rating Scale; MMSE = Mini Mental State Examination; CBI = Caregiver Burden Inventory; FTD = Frontotemporal Dementia; ZBI = Zarit Burden Interview; PDD = Parkinson's Disease Dementia; BI = Burden Interview; ; DLB =Dementia with Lewy Bodies; RSS = Relative Stress Scale; bvFTD = behavioural variant Frontotemporal Dementia; PCQ = Perceived Control Questionnaire; QDRS = Quick Dementia Rating Scale

Appendix H: Supplementary information regarding quantitative measures of burden employed within the reviewed literature

The Caregiver Burden Inventory (Novak & Guest, 1989)

The CBI consists of 24 items, containing five subscales: 'time-dependent burden', reflecting the impact of the amount of time spent caring; 'developmental burden', reflecting the extent to which caregivers feel they are missing out on experiences common to their peer group; 'physical burden', reflecting the impact of caregiving on physical health; 'social burden', reflecting the impact of caregiving on relationships; and 'emotional burden', reflecting negative feelings experienced by the caregiver towards the care recipient (Novak & Guest, 1989). Scores of 24 or higher are indicative of a need for respite care, while scores of 36 and above represent a risk of burnout (e.g. D'Onofrio et al., 2015). High levels of internal consistency have been demonstrated (Marvardi et al., 2005),

The Zarit Burden Inventory (Zarit et al., 1980)

A review of caregiver burden measures (Van Durme et al., 2012) suggested that the ZBI was the most useful of the measures assessed, displaying good content validity, high levels of internal consistency, and having been validated in longitudinal research studies. A two-factor structure is largely accepted, reflecting 'personal strain' and 'role strain', however individual items explore the emotional, physical and social impacts of caregiving (Van Durme et al., 2012). The original version contains 22 items, with scores of 21-40 indicating little burden, while scores in the 41-60 range point to moderate burden, and scores above 60 represent severe carer burden (Zarit et al., 1980). Several of the articles reviewed here employed the 12-item version, which also demonstrates good levels of reliability and validity (Bédard et al. 2001), for which scores of 17 or higher would be seen to indicate high to severe levels of burden (Branger et al., 2016). Of the articles reviewed here, two (Besser & Galvin, 2018; Leggett et al., 2010) report a three factor structure reflecting 'personal strain', role strain' and 'performance strain' or 'worry about performance', for which the item structure is identical with the exception of one item. A further article (Galvin et al., 2010), used an 11-item short form of the ZBI, which excludes the floating item noted in

the aforementioned papers. Reliability and validity data, and factor structure information could not be found for this version, however it was possible from the data provided to map responses to this 11-item version against the factor structure of Besser and Galvin (2018), and Leggett et al (2010), for the purpose of aiding comparison here. Finally, one study (Nicolaou et al., 2010) employed the 22-item version, and reported a two factor structure but did not provide this structure to the reader, and as such only tenuous comparisons may be drawn as it is not known to what extent the 'personal strain' and 'role strain' reported are similar to that reported by other authors describing identically named constructs.

The Relative Stress Scale (Greene et al., 1982)

The RSS is a 15-item measure, underpinned by a three-factor structure comprising: 'emotional distress'; 'social distress'; and 'negative feelings', similar to the corresponding factors of the CBI discussed earlier (Ulstein et al., 2007). Scores of 23 and above indicate an 'intermediate risk of clinical relevant distress', while scores of 30 and above reflect a high level of such risk (Greene et al., 1982).

Caregiver Burden Scale (Elmståhl et al., 1996)

The CB Scale is a 22-item measure, for which reliability and validity have been well investigated (Elmståhl et al., 1996). Subscales reflect: 'general strain' (referring to impact upon health, time for self, and perception of difficulties); 'isolation' (reduction in social life and activity levels), 'disappointment' (physical, economic and developmental strain), 'emotional involvement' (negative internal states) and 'environment' (obstacles to providing good care, and related concerns). A total burden score is also available by calculating the mean of all items, resulting in a score between 1 and 4, with scores between 1.00-1.99 reflecting low burden; 2.00-2.99 reflecting medium burden; and finally scores of 3.00 and above indicating high levels of burden (Pudelewicz et al., 2018). It should be noted the CB Scale employed within the current review was a Polish version, for which specific psychometric properties could not be found.

References

- Bédard, M., Molloy, W., Squire, L. et al. (2001). The Zarit Burden Interview: A new short version and screening version. *The Gerontologist*, 41(5), 652-657.
- Besser, L.M., & Galvin, J.E. (2018). Perceived burden among caregivers of patients with frontotemporal degeneration in the United States. *International Psychogeriatrics*. [Advance online publication] doi:10.1017/S104161021800159X
- Branger, C., O'Connell, M.E., & Morgan, D.G. (2016). Factor analysis of the 12-item Zarit Burden Interview in caregivers of persons diagnosed with dementia. *Journal of Applied Gerontology*, 35(5), 489-507.
- D'Onofrio, G., Sancarolo, D., Addante, F., et al. (2015). Caregiver burden characterization in patients with Alzheimer's disease or vascular dementia. *International Journal of Geriatric Psychiatry*, 30(9), 891-899.
- Elmståhl, S., Malmberg, B., & Annerstedt, L. (1996). Caregiver's burden of patients after stroke assessed by a novel caregiver burden scale. *Archives of Physical and Medical Rehabilitation*, 77, 177-182.
- Galvin, J.E., Dud, J.E., Kaufer, D.I. et al. (2010). Lewy body dementia: Caregiver burden and unmet needs. *Alzheimer Disease and Associated Disorders*, 24(2), 177. Greene, J. G., Smith, R., Gardiner, M., & Timbury, G. C. (1982). Measuring behavioural disturbance of elderly demented patients in the community and its effects on relatives: A factor analytic study. *Age and Ageing*, 11(2), 121-126.
- Leggett, A.N., Zarit, S., Taylor, A., & Galvin, J.E. (2010). Stress and burden among caregivers of patients with Lewy Body Dementia. *The Gerontologist*, 51(1), 76-85.
- Marvardi, M., Mattioli, P., Spazzafumo, L. et al. (2005). The caregiver burden inventory in evaluating the burden of caregivers of elderly demented patients: Results from a multicenter study. *Aging Clinical and Experimental Research*, 17(1), 46-53.

- Nicolaou, P. L., Egan, S. J., Gasson, N., & Kane, R. T. (2010). Identifying needs, burden, and distress of carers of people with frontotemporal dementia compared to Alzheimer's disease. *Dementia: The International Journal of Social Research and Practice*, 9(2), 215-235.
- Novak, M., & Guest, C. (1989). Application of a multidimensional caregiver burden inventory. *The Gerontologist*, 29(6), 798-803.
- Pudelewicz, A., Talarska, D., & Bączyk, G. (2018). Burden of caregivers of patients with Alzheimer's disease. *Scandinavian Journal of Caring Sciences* [Advance online publication]. doi: 10.1111/scs.12626
- Ulstein, I., Bruun Wyller, T., & Engedal, K. (2007). The relative stress scale, a useful instrument to identify various aspects of carer burden in dementia? *International Journal of Geriatric Psychiatry*, 22(1), 61-67.
- Van Durme, T., Macq, J., Jeanmart, C., & Gobert, M. (2012). Tools for measuring the impact of informal caregiving of the elderly: A literature review. *International Journal of Nursing Studies*, 49(4), 490-504.
- Zarit, S. H., Reever, K. E., & Bach-Peterson, J. (1980). Relatives of the impaired elderly: Correlates of feelings of burden. *The Gerontologist*, 20(6), 649-655.

Appendix I: Summary of SPANS indices

Index	Items	Description
Orientation Index (ORI)	<ul style="list-style-type: none"> • Orientation to person, time, place, condition, and political leadership (current and past) • Time estimation 	Scoring employs a graded system by which greater points are received for responses closer to the correct answer, therefore reflecting a gradient of impairment.
Attention / Concentration Index (ACI)	<ul style="list-style-type: none"> • Digit span forwards • Digit span backwards • Sustained and divided listening – round 1 • Sustained and divided listening – round 2 • Counting backwards • Monetary calculations 	Attention span, sustained and divided attentional ability, and attentional control are assessed. Two items include an additional measure of response time, assessing information processing speed.
Language Index (LAI)	<ul style="list-style-type: none"> • Repetition • Naming • Yes/No questions • Following directions • Reading • Writing Sentences • Similarities 	Designed to detect a range of language disturbances, the LAI assesses comprehension, repetition, confrontation naming, reading and writing.

Index	Items	Description
Visuo-Motor Performance Index (VPI)	<ul style="list-style-type: none"> • Object recognition • Spatial decision • Unusual views • Figures copy • Letter-number coding • Figures recognition • Facial expressions • 3-and-1 concept test 	Includes screening for spatial impairment, object perceptual agnosia, and visual attention difficulties. Visuo-perceptual and visuo-motor difficulties are assessed in addition to visual recognition memory, emotion recognition and visual concept formation.
Efficiency Index (ECI)	<ul style="list-style-type: none"> • Sustained and divided listening - round 2 • Spatial decision • Letter-number coding • Counting backwards • Monetary calculations 	Comprised of the timed items from other indices, this index provides a measure processing efficiency incorporating reaction, thinking, scanning and visuo-motor movement time.
Conceptual Flexibility Index (CFI)	<ul style="list-style-type: none"> • Similarities • 3-and-1 concept test 	Assesses concept formation and cognitive flexibility using visual and verbal stimuli.

Appendix J: A comparison of the SPANS with other brief screening assessments

	SPANS	ACE-III	MoCA	RBANS
Time to administer (minutes)	35	15	10	20-30
Number of domains / indices	7	5	8	5
Total number of items	30	24	13	12
Orientation				
Person	+	-	-	-
Time	+	+	+	-
Place	+	+	+	-
Additional orientation	+	+	-	-
Attention				
Span	+	+	+	+
Sustained	+	-	+	-
Divided	+	-	+	-
Working memory	+	+	+	-
Calculations	+	-	-	-
Language				
Naming	+	+	+	+
Repetition	+	+	+	-
Comprehension	+	+	-	-
Reading	+	+	-	-
Writing sentences	+	+	-	-
Visuospatial				
Visual screening	+	-	-	-
Spatial	+	+	-	+
Object	+	+	-	-
Construction	+	+	+	+
Neglect	+	+	+	-
Processing speed				
Visuomotor	+	-	+	+
Reaction time	+	-	-	-
Memory				
Verbal recall	+	+	+	+
Visual recall	+	-	-	+
Verbal recognition	+	+	+	+
Visual recognition	+	-	-	-
Learning trials	+	+	+	+
Executive function				
Concept formation	+	-	-	-
Reasoning / abstraction	+	-	+	-
Planning / sequencing	-	+	+	-
Judgement	-	-	-	-
Cognitive flexibility	+	-	+	-
Verbal fluency	-	+	+	+

Appendix K: Exploration of predictive validity

In designing the current study, it was hoped that items 8 to 20 DHLQ may provide an indication of functional ability in relation to activities of daily living (ADLs) which could be used to create quasi-experimental groups for the exploration of predictive validity of the SPANS. As such, the following research question was defined: Do SPANS scores covary with functional independence or demographic characteristics which would be expected to covary with cognitive ability?

Insufficient internal consistency of these DHLQ items ($\alpha = .693$) prevented the scale from being taken forward into analysis. On the basis of face validity, the internal consistency of just those items pertaining to ADLs was also assessed, however was again insufficient ($\alpha = .432$). This was understood to reflect the highly able and independent nature of participants, and subsequent lack of sufficient variance to support internal consistency.

In order to determine whether any internally reliable scale could be produced from the DHLQ, principal components analysis was conducted, incorporating items 8 to 20 using orthogonal rotation.

Analysis

The Kaiser-Mayer-Olkin measure of sampling indicated mediocre adequacy overall, but raised concerns in relation to four items individually - 'far away vision', 'close up vision', 'shopping' and 'transport'. Additionally, the 'personal care' item was automatically removed due to zero variance.

Bartlett's test of sphericity indicated that the data were suitable for this type of analysis ($p < .05$). However, a visual scan of the correlation matrix indicated a significant proportion of item pairings for which no significant relationship was found, indicating potentially insufficient relationships between items, and therefore issues of singularity within the data.

Four uncorrelated factors with eigenvalues greater than 1.00 were extracted, which explained 23.23%, 16.30%, 13.20%, and 8.29% of total variance, respectively. Within reproduced correlations matrix, 60% of residuals had values greater than .05, indicating further cause for concern.

Factor structure was as follows:

FACTOR	ITEMS
1	<ul style="list-style-type: none">• Walking• Getting up from a chair• Housework• Meal preparation• Fine motor skills
2	<ul style="list-style-type: none">• Far away vision• Close up vision• Hearing• Memory• Managing finances
3	<ul style="list-style-type: none">• Transport• Driving
4	<ul style="list-style-type: none">• Shopping

Factor 1 demonstrated sufficient internal consistency ($\alpha = .77$), and appeared to reflect gross and fine motor abilities, with perhaps some elements of efficiency. However, it was judged that the issues of concern within this analysis (which were not addressed though removal of the four items identified above) were sufficient to prevent continuation. As such, the DHLQ's contribution to the present research was limited to providing description of sample characteristics and demographic context.

Appendix L*: Visual representation of the chronology of the research process

January 2017	<ul style="list-style-type: none"> •Initial meeting with research supervisor
February 2017	<ul style="list-style-type: none"> •Meeting with research supervisor and Hogrefe (SPANS publisher) to discuss research
March 2017 - May 2017	<ul style="list-style-type: none"> •Development of DHLQ, PIS, consent form, and contact form •Initial draft of procedure
May 2017	<ul style="list-style-type: none"> •Submission of research proposal to the University of Leicester for panel review
June 2017	<ul style="list-style-type: none"> •Panel review meeting with course staff
September 2017	<ul style="list-style-type: none"> •Submission of sponsorship and ethics application to the University Ethics Sub-committee for Psychology
November 2017	<ul style="list-style-type: none"> •Resubmission of research proposal for peer review •Submission of summary proposal for review by Service User Reference Group •Ethical approval and sponsorship granted
December 2017	<ul style="list-style-type: none"> •VRA advert circulated
January 2018	<ul style="list-style-type: none"> •VRA interviews and recruitment •Preparation of VRA guidance and training materials
February 2018	<ul style="list-style-type: none"> •VRA background checks and training day
February 2018 - February 2019	<ul style="list-style-type: none"> •Data collection
July 2018 - January 2019	<ul style="list-style-type: none"> •Literature review
February 2019 - March 2019	<ul style="list-style-type: none"> •Collation and checking of data •Exploration of data and statistical analysis
February 2019 - April 2019	<ul style="list-style-type: none"> •Write-up of research report, including a process of revision following supervision and feedback from drafts
April 2019	<ul style="list-style-type: none"> •Thesis submission

Appendix M*: Ethical approval and sponsorship confirmation



UNIVERSITY OF
LEICESTER

University Ethics Sub-Committee for Psychology

23/11/2017

Ethics Reference: [REDACTED]

TO:

Name of Researcher Applicant: [REDACTED]

Department: Psychology

Research Project Title: An initial exploration of the validity of the Short Parallel Assessments of Neuropsychological Status (SPANS) for use with older adults

Dear [REDACTED],

RE: Ethics review of Research Study application

The University Ethics Sub-Committee for Psychology has reviewed and discussed the above application.

1. Ethical opinion

The Sub-Committee grants ethical approval to the above research project on the basis described in the application form and supporting documentation, subject to the conditions specified below.

2. Summary of ethics review discussion

The Committee noted the following issues:

All suggested revisions have been addressed.

3. General conditions of the ethical approval

The ethics approval is subject to the following general conditions being met prior to the start of the project:

As the Principal Investigator, you are expected to deliver the research project in accordance with the University's policies and procedures, which includes the University's Research Code of Conduct and the University's Research Ethics Policy.

If relevant, management permission or approval (gate keeper role) must be obtained from host organisation prior to the start of the study at the site concerned.

4. Reporting requirements after ethical approval

You are expected to notify the Sub-Committee about:

- Significant amendments to the project
- Serious breaches of the protocol
- Annual progress reports
- Notifying the end of the study

5. Use of application information

Details from your ethics application will be stored on the University Ethics Online System. With your permission, the Sub-Committee may wish to use parts of the application in an anonymised format for training or sharing best practice. Please let me know if you do not want the application details to be used in this manner.

Best wishes for the success of this research project.

Yours sincerely,



Chair

Appendix N: Ethical issues considered

1. Participant considerations

- Exclusion criteria were in place to exclude those who were deemed to lack mental capacity.
- The potential for participation to cause mild to moderate anxiety or negative psychological consequences was considered. It was anticipated that completion of the DHLQ, in which participants were prompted to think about their current level of daily independence, could be experienced as upsetting if a loss of independence was perceived. Furthermore, neuropsychological assessment may be experienced as anxiety provoking or cause participants to feel less cognitively able in comparison to their younger selves, or their expectations of themselves. Finally, it was possible for testing to indicate to a participant the presence of a previously unrecognised cognitive impairment through the experience of unexpected difficulty. All such possibilities were highlighted within the participant information sheet (PIS), and discussed in relation to the right to withdraw at any time.
- With regard to feedback from assessment, participants were clearly advised of the nature and boundaries of the current study as being purely for research purposes in the investigation of the currently unvalidated SPANS assessment, and that feedback regarding overall performance or SPANS score could therefore not be provided. Participants were given the option to be notified by the Chief Investigator in the event that they completed the (validated) ACE-III assessment and their score fell below the known cut-off level. Notification was in writing, using a standard template, to convey this feedback in a sensitive and

unalarming manner and advise that if concerned about memory they may elect to report this to their G.P.

Assessor considerations

- Volunteer research assistants (VRAs) were vetted via the application and interview process, and checks via the Disclosure and Barring Service (DBS).
- VRAs were trained in the administration of the assessments employed so as to feel competent in administering and scoring these.
- VRAs remained in contact with the chief and principal investigator throughout the data collection period, who were to be contacted for questions, problems, or trouble-shooting.
- VRA's went out into the community to seek participants and assessment locations varied depending on availability. To promote personal safety, the chief and principal investigators took the following steps: a) appropriate recruitment and training, and b) implementation of a lone worker policy.

Appendix O: Gatekeeper contact standard email template

Template for initial contact with gatekeepers of organisations whose members may be participants

Dear [*Gatekeeper*],

I am a degree-seeking student who is interested in psychology, and am currently working as a volunteer research assistant (VRA) for a research study exploring how well a new measure of cognitive abilities (e.g. concentration), currently used with working-age adults, performs when used with people over the age of 65. The study is organised and run by [redacted] of [redacted], and has been reviewed and approved by the University's Research Ethics Committee.

I am based in your local area and am writing to ask whether it might be possible for me to advertise the opportunity to participate in this study to the [*members / residents / other*] of your [*organisation / facility / group / other*]? Advertising could involve the use of posters, flyers, or my presence on a given day, depending on your preference, and I would be guided by your advice and requirements around this. I would then meet with anyone interested in taking part to provide further information, answer questions, and obtain informed consent, before arranging participation sessions.

I have attached a copy of our 'Participant Information Sheet' for your reference, however in brief, participants would be asked to complete a demographic questionnaire and at least one cognitive assessment, with the option to complete further tasks should they wish to do so. By participating, they would contribute to our understanding of the cognitive assessment of persons over 65 years, which may then have benefits more broadly, such as making it possible for the NHS to use a more thorough and informative assessment, perhaps contributing to better care.

I would be grateful to hear your thoughts regarding the possibility of advertising participation in this study within your [*organisation / facility / group / other*] and of course would be happy to provide you with any further information which you may like or require.

I look forward to hearing from you.

Kind regards,

[*VRA Name*]

Volunteer Research Assistant

Supervised by [redacted]
[redacted]

Appendix P*: Demographic, Health and Lifestyle Questionnaire (DHLQ)

Demographic, General Health & Daily Life Questionnaire (DHLQ)

Your responses to this questionnaire will inform any conclusions which we may draw from the current research by adding greater context to other findings.

Please answer as many as you are able or willing to let us know about...

1. What is your date of birth?

2. What is your highest level of education completed? (tick one)

- ☐ Left school before getting a qualification
- ☐ Completed secondary school
- ☐ Apprenticeship
- ☐ Some university but did not earn degree
- ☐ Earned bachelor's degree
- ☐ Earned post-graduate degree

3. Please tick the category you would consider is/was your primary employment. (tick just one)

- ☐ Was not able to work during my prime working age years
- ☐ Unskilled labour
- ☐ Skilled labour – 6+ month training/apprenticeship required
- ☐ Management/business – no university degree required
- ☐ Professional – university degree required
- ☐ Stay at-home spouse/parent

4. What is your country of ethnicity?

5. Please tick as many of the following boxes that apply to you:

- ☐ English is my first language.
- ☐ The majority of my primary education (age 5-12) was in a *Western / English-speaking* school.
- ☐ More than 50% of my life has been spent living in *Western/English-speaking* country.

KEY: *The "Western/English-speaking" countries are: U.S.A., Western Canada, Australia, New Zealand, U.K., Republic of Ireland, White South African*

_____ = Total ticks

6. Please tick any of the following conditions you have been diagnosed with in your lifetime. (tick as many as apply to you)

- ☐ Acquired brain injury (e.g. traumatic, bleed, infection)
- ☐ Dementia (e.g. vascular, Alzheimer's)
- ☐ Epilepsy
- ☐ Mild cognitive impairment
- ☐ Movement disorder (e.g. Parkinson's, multiple sclerosis)
- ☐ Stroke
- ☐ Other neurological condition:
- ☐ Diabetes
- ☐ Heart/vascular disease (e.g. cholesterol, hypertension, history of heart attack/MI, arteriosclerosis)
- ☐ Respiratory disease (i.e. diseases that affect your breathing)

7. Have you ever, for 10 years straight or more, drank 5+ units of alcohol most days, or been considered alcoholic at any time in your life?

Yes No

8. How would you consider your vision, or ability to see to be?

a) For far away distances?

Good Fair Poor Nearly blind

b) For things close up, including reading?

Good Fair Poor Nearly blind

9. How would you consider your hearing?

Good Fair Poor Nearly deaf

10. How would you consider your fine motor ability / ability to use your hands?

Good Fair Poor Barely able to

11. How would you consider your ability to walk?

Good Fair Poor Barely able to

12. How would you consider your ability to get out of bed or up from a chair?

Good Fair Poor Barely able to

13. How would you consider your ability to remember things?

Good Fair Poor I don't remember anything

14. Do you still drive and have a full license?

- a) Yes, fully b) Yes, but I limit my driving somewhat
c) No because it may be a risky to
d) There is no way I will drive again N/A) I have never driven

If you no longer drive, or limit your driving, what is the main reason?

.....
.....

Below are listed six activities of daily living.

For each activity please circle the level of assistance which you require to complete the task.

15. Meal preparation	a) Independent, no assistance needed b) Minimal assistance needed c) Mostly dependent on others d) Full assistance required
16. Personal care & dressing	a) Independent, no assistance needed b) Minimal assistance needed c) Mostly dependent on others d) Full assistance required
17. Housework	a) Independent, no assistance needed b) Minimal assistance needed c) Mostly dependent on others d) Full assistance required
18. Managing finances	a) Independent, no assistance needed b) Minimal assistance needed c) Mostly dependent on others d) Full assistance required
19. Shopping	a) Independent, no assistance needed b) Minimal assistance needed c) Mostly dependent on others d) Full assistance required
20. Transport	a) Independent, no assistance needed b) Minimal assistance needed c) Mostly dependent on others d) Full assistance required

If you have indicated 'minimal assistance', 'mostly dependent' or 'full assistance' in response to any of the above activities, please briefly outline below which aspect(s) of the task(s) you require assistance with (for example, physical/mobility aspects, planning / organisation etc.).

.....

.....

.....

.....

For completion by researcher:

Date this form was completed: _____

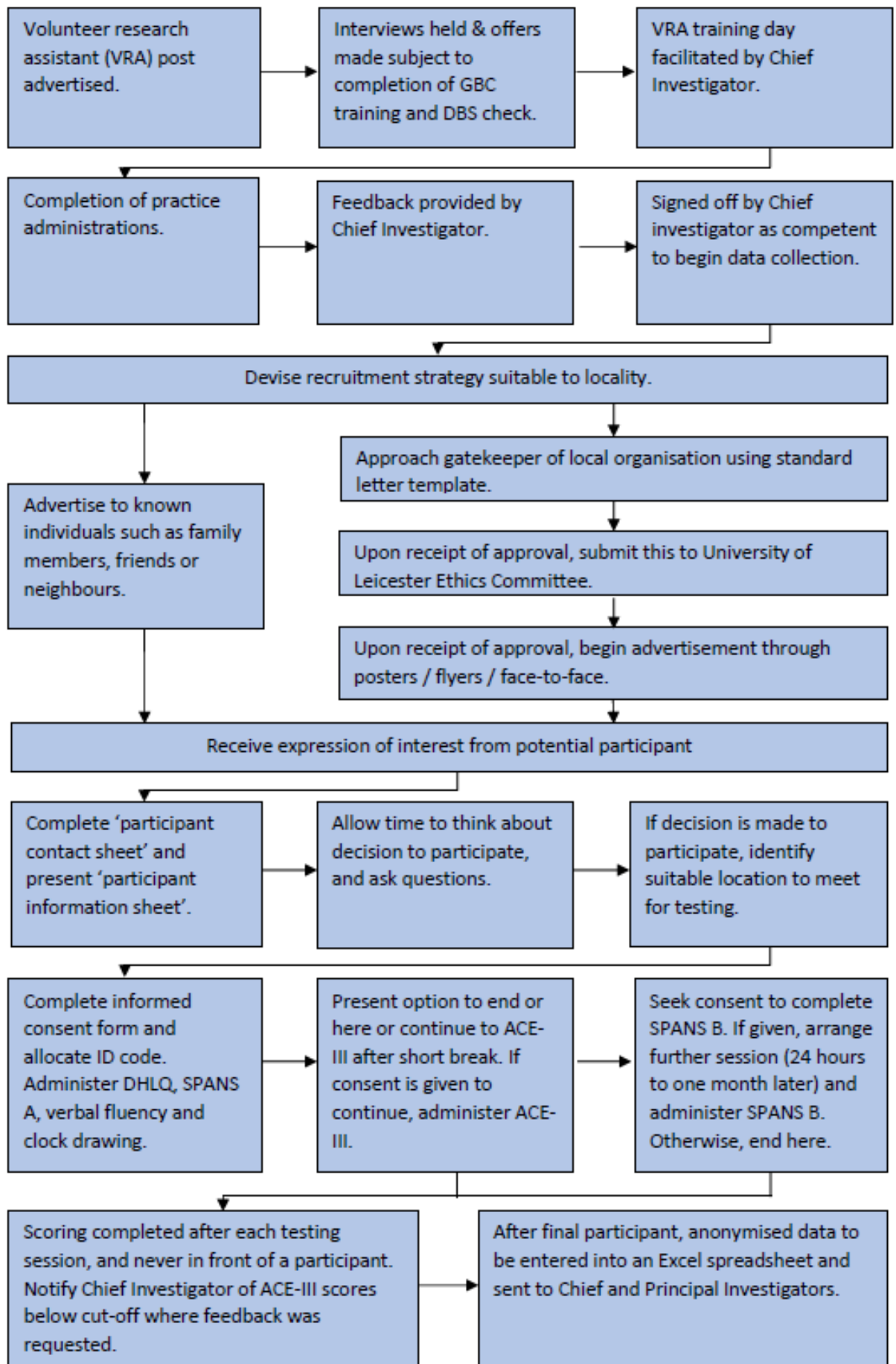
Age of participant: _____

PSR ID: _____

Score items 8 through 20 as 3 to 0 points, with '3' representing 'most capable' and '0' representing 'least capable'. Total possible score range is 0 to 42.

Score achieved: ____ / 42

Appendix Q: Visual representation of the research procedure



PARTICIPANT CONTACT SHEET

For study entitled,

“An initial exploration of the validity of the Short Parallel Assessments of Neuropsychological Status (SPANS) with older adults”

For all who enquire and/or agree to take part in the study, this sheet provides the researcher with your first name or nickname (only), your preferred means of being contacted should this be necessary, and is used to assign those who decide to participate with a PSR (i.e. personal study reference) ID Number, which is used to keep your data anonymised. This sheet will be kept separate from all your other data, and will only be used for study-related administrative purposes, kept safe and private and destroyed as soon as it is no longer needed.

“I am interested in learning more or participating in the above study.”

(please tick) ☐

My first name or nickname (ONLY) is: _____

My preferred means of being contacted is:

By landline: ☐ _____

By mobile: ☐ _____

By email: ☐ _____

By post: ☐ _____

For completion by researcher:

Assigned PSR ID: _____

Gender/sex: M F

Date questionnaire completed: / /

Participant age: _____

Appendix S*: Participant information sheet

Participant Information Sheet

Study title

An initial exploration of the validity of the Short Parallel Assessments of Neuropsychological Status (SPANS) with older adults.

Purpose of this information sheet

In this information sheet we would like to tell you about, and invite you to participate in, a study run by [REDACTED] and sponsored by the [REDACTED]. Please feel free to ask questions at any point and if you would like more information about the study, please contact [REDACTED] or [REDACTED] using the contact information provided below.

Brief overview of the purpose of the study

This study explores how well a new, not yet tried-and-tested measure of cognitive abilities (e.g. concentration), currently used with working-age adults, performs when used with people over the age of 65, and compare it to an already tried-and-tested measure.

Who am I (i.e. the person who invited you to this study)?

I am a psychology graduate, and am working as a volunteer research assistant (VRA) for this study. I am doing so in order to gain further experience in the clinical field of psychology, in pursuit of gaining a place on a clinical psychology doctoral training programme. I am supervised by [REDACTED] from the [REDACTED] and was selected through an application and competitive process, and I have been vetted.

Why have you been asked to be involved?

There is a wide range of inclusion criteria into this study, but it is most likely because you are over the age of 65, speak English well, have the capacity to understand the purpose of this study and consent to take part, are generally in good health, and don't have any major disabilities that would prevent you from completing the majority of the study.

If you decide to take part in the study, what will happen?

You will be asked to do your best while completing a number of tasks which look at a range of thinking skills and require your concentration. These are called 'cognitive tests' and include pictures, words, and numbers. For each task, the VRA will explain the instructions to you, and you will be able to ask questions if you are unsure about anything. It will not be possible to tell you how you're doing along the way or provide you with feedback afterwards, because we are exploring an as yet un-validated measure.

There are four parts to the study, outlined below, and you may do as many as you feel interested in and/or capable of doing. The four parts are:

1. Complete a demographic, health, and activities of daily living questionnaire (*15 minutes*)
2. Be administered the Short Parallel Assessments of Neuropsychological Status (SPANS) Version A (*35 minutes*)
3. Be administered the Addenbrooke's Cognitive Examination (ACE-III) (or just two subtests from it) (*25 minutes OR 10 minutes*)
4. In 24 hours to a month from now, meet again with VRA and be administered the SPANS Version B (*35 minutes*)

Do I have to take part?

It is entirely your choice whether you take part, and how much you do. You may start then stop at a point of your choosing, in which case then we will assume that we may use all data you provided up to the point of stopping with the study. You can also take a break if you need to at any time by letting the VRA know so that she/he can select the nearest convenient point to pause the tests. You may take as much time as you feel you need to consider the information in this sheet, including asking questions, and decide whether you'd like to participate.

What are possible benefits of taking part?

If you decide to take part, you may find it enjoyable and more interesting than not having taken part. You would also contribute to the understanding of the cognitive assessment of persons over 65 years, which may then have benefits more broadly, such as making it possible to provide older adults a more thorough and informative assessment in medical cases, perhaps contributing to better care for them.

What are possible downsides of taking part?

This study requires your concentration and some may find that this causes them fatigue. In some cases if you have lost some levels of independence, dexterity, or visual acuity in recent years and find this upsetting then you may find some of the cognitive tasks unpleasant. It may be the case that by completing the questions and activities that compose this study, relative strengths and weaknesses that you possess may become apparent to you. In this instance, should you feel that you would like to explore this further, we would direct you to discuss this with your GP. These are all natural concerns, but please be reassured that no one is expected to get everything right – some of the assessment tasks are intentionally difficult in order to differentiate between levels of ability. We just ask that you try your best.

Will my responses and/or test performances be used to diagnose me with any condition?

No - all information which you provide, including test performance and scores, will be used solely for the purpose of exploring the SPANS assessment with persons over the age of 65. As this is an as yet un-normed assessment, in final stages of development, no conclusions can be drawn and no feedback given. You may also complete, however, the ACE-III, an already-validated cognitive test, which will act as a comparison for the SPANS in our study. For this test a particular 'cut-off' score is suggested, with scores below this proposed to indicate a lower level of performance than might be expected. There are a number of reasons why a person's performance may fall below this cut-off point. As such, we as a research team are not in a position to draw any conclusions from your score, however you may choose to be notified by us in the event that your ACE-III score is below this cutoff point, and this option will be made available to you on the Participant Consent Form. Notification will be in writing via email (so you would need to provide this information on the Participant Contact Information form) and will not include any specific feedback, or information regarding your SPANS score.

Will I be able to be identified?

The brief answer is 'no', because if you agree to participate you will be assigned a made-up, non-identifiable ID number. All the data you supply to this study will be connected only to this ID number. There will be one sheet (see 'Participant Contact Sheet') that would contain your first name or a nickname only, this ID number, and your preferred means of being contacted (for example to make an appointment to meet the VRA). As soon as your participation in the study is complete, this sheet with your first name or nickname and ID number will be shredded and securely disposed of.

Will my personal information be kept secure?

Yes. We (myself, [REDACTED] and [REDACTED]) will be using information from you in order to undertake this study and will act as the data controllers for this study. This means that we are responsible for looking after your information and using it properly.

We will keep your name, contact details, and date of birth confidential and will not pass this information to [REDACTED]. I, along with [REDACTED] and [REDACTED], will use this information as needed, to contact you about the research study, and to oversee the quality of the study. Certain individuals from [REDACTED] and regulatory organisations may look at our research records to check the accuracy of the research study. They will only receive information without any identifying information. The people who analyse the information will not be able to identify you and will not be able to find out your name or contact details.

Is there any compensation for expenses, or payments?

Taking part in this study would unfortunately be entirely voluntary on your part, calling on your goodwill, as we do not have a way or means to pay you for your time or expenses (e.g. in getting here today).

May I see the results of this study if I wish?

Updates and summaries of findings will be put onto the publisher's (Hogrefe Ltd.) website at hogrefe.co.uk/spans.html and follow the SPANS-X link, you may check there at any time you may wish to for the latest update.

What if something goes wrong or I'm unhappy with the study itself or its conduct?

In the first instance it is hoped that if any problems arise while you are taking part in this study, or that something about it causes you distress, that you speak to the research assistant about what you are experiencing or feeling, and that between you any issues are satisfactorily resolved. If any issues are not able to be resolved, and you need to take the issue further, you may contact the Chief Investigator of this study [REDACTED], using the contact information below.

Researcher contact information

[REDACTED] (Chief Investigator): [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED] (Principal Investigator): [REDACTED]

Appendix T*: Participant consent form



PARTICIPANT CONSENT FORM

Title of Project: An initial exploration of the validity of the Short Parallel Assessments of Neuropsychological Status (SPANS) with older adults

Chief Investigator: [REDACTED]

Researcher: _____ (Please print)

Please tick each box that applies:

- ☐ I confirm that I have read and understand the information sheet dated 30/12/2018 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
- ☐ I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.
- ☐ I understand that I have been assigned a personal ID number which will be kept separate from this consent form, and will be used to anonymise the answers I provide, which will subsequently be stored on a computer database using only this ID number. I understand I will remain anonymous.
- ☐ I understand that I will not be provided with any feedback on my performance in the tasks that I undertake in this research.
- ☐ I agree to take part in the above study.

Name of Participant Date Signature

Name of Researcher Date Signature

Optional feedback:

I request to be notified by [REDACTED] in the event that I complete the ACE-III assessment and my score is below the 'cut-off' point determined by the test's developers. **YES / NO (please circle your choice)**

Please note: notification will be via email and as such requesting this feedback will require you to provide an email address on the Participant Contact Information form.

Appendix U*: Checklist to ensure anonymity of clients/services

	Checked in Executive Summary/ Abstract/ Overview (if included in assignment)	Checked in main text	Checked in appendices
Pseudonym or false initials used	Y	Y	Y
Reference to pseudonym/false initials as a footnote	Y	Y	Y
Removed any reference to names of Trusts/hospitals/clinics/services (including letterhead if including letters in appendices)	Y	Y	Y
Removed any reference to names/specific dates of birth/specific date of clinical appointments/addresses/ location of client(s), participant(s), relatives, caregivers, and supervisor(s). [For research thesis – supervisors can be named in the research thesis “acknowledgements” section]	Y	Y	Y
Removed/alterd references to client(s) jobs/professions/nationality where this may potentially identify them. [For research thesis – removed potential for an individual research participant to be identifiable (e.g., by a colleague of the participant who might read the thesis on the internet and be able to identify a participant using a combination of the participants specific job title, role, age, and gender)]	Y	Y	Y
Removed any information that may identify the trainee (consult with course staff if this will detract from the points the trainee is making)	Y	Y	Y
No Tippex or other method has been used to obliterate the original text – unless the paper is subsequently photocopied and the trainee has ensured that the obliterated text cannot be read	Y	Y	Y
The "find and replace" function in word processing has been used to check the assignment for use of client(s) names/other confidential information	Y	Y	Y