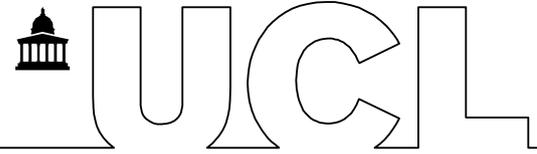


---

RESEARCH DEPARTMENT OF CLINICAL, EDUCATIONAL AND HEALTH  
PSYCHOLOGY



---

Pre-therapy skills required to be ready for  
Cognitive Behavioural Therapy in People Living  
with Dementia

Joshua Charles Hugh Stott

A thesis submitted for the degree of Doctor of Philosophy to the Division of  
Psychology & Language Sciences, University College London

## Declaration

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

Signature:

A solid black rectangular box redacting the signature.

Date: 28/08/2018

## **Abstract**

*Background:* People living with dementia (PLWD) commonly experience depression and anxiety. For the general adult population, cognitive behavioural therapy (CBT) is a recommended treatment. Adapted forms of CBT have been used with PLWD. However, method of adaptation and outcomes are inconsistent across studies. Examining pre-therapy skills required to take part in a core aspect of CBT (cognitive restructuring) in PLWD could inform future adaptation. Given the limited previous work in PLWD, the intellectual disability literature was systematically reviewed and integrated with the dementia literature to inform aims.

*Main aims:* 1) to develop measures of pre-therapy skills (behaviour-thought-feeling discrimination and cognitive mediation) validated for use with PLWD; 2) to compare performance of PLWD and older (OA) and younger (YA) adults without a recognised neurocognitive impairment on these validated pre-therapy skill measures; 3) to examine whether neurocognition mediates observed differences between PLWD and OA in pre-therapy skill performance; 4) to examine neurocognitive correlates of pre-therapy skill measures in PLWD with a focus on memory, language and executive function.

*Main methods:* 102 PLWD, 77 OA and 56 YA were recruited. Measures of pre-therapy skills used in an intellectual disability context were adapted for PLWD using a published framework and subjected to factor analysis and validity checks. Performance on pre-therapy skills measures was compared across groups, mediation of between group differences was assessed (using structural equation modelling) and correlations between pre-therapy skills and neurocognitive functions were examined.

*Main findings:* Tools were developed. PLWD scored lower than OA who scored lower than YA on pre-therapy skills measures. Differences between OA and PLWD but not between OA and YA were mediated by neurocognition. Pre-therapy skill performance was associated with scores on measures of language and, to a lesser extent, executive function. Use of tools within, and implications of findings for, CBT practice and research are discussed.

## Impact Statement

This thesis has several potential benefits within academia, many of which are alluded to within the body of the thesis. Three examples of such benefits are:

- i) The thesis involved validation of previously un-validated measures of CBT pre-therapy skills will be of use to researchers looking to investigate how to optimise CBT for PLWD and individuals with neurocognitive limitations in general.
- ii) The thesis examined performance of PLWD on these measures, which will be helpful in making evidence based decisions as to how to develop future CBT interventions for use in clinical research with this population.
- iii) In examining the neurocognitive correlates of pre-therapy skill performance in PLWD, the results of this thesis will be helpful in starting to understand the relationship between engagement in CBT and core neuropsychological deficits in dementia.

These academic benefits and others alluded to throughout the thesis are enhanced through the publication of a number of thesis related papers in academic journals as detailed in the 'dissemination' section below.

There are also numerous potential clinical benefits of the work and these are discussed in some detail in the text. Four examples of such benefits are (i) the pre-therapy skills measures developed in chapters 5 and 6 could be used by clinicians in CBT practice to idiosyncratically adapt CBT for PLWD and OA on a case by case basis, (ii) the findings as to performance on these measures (reported in chapter 7) will be helpful in adapting future CBT interventions through understanding of which aspects of CBT PLWD might require support to engage with, (iii) the validation of the Hospital Anxiety and Depression Scale (HADS) in

chapter 3 will aid professionals in knowing how to interpret this widely used measure of anxiety and depression (and what the limits of interpretation are), iv) the work on validity of the Test of Premorbid Function (TOPF) in chapter 4 will be useful in informing clinicians as to the limitations of the TOPF in a dementia context.

The clinical benefits of this thesis will be enhanced by incorporation into teaching by the candidate in his role training clinical psychologists on the largest clinical psychology training course in the country. Additionally, many of the academic publications alluded to in dissemination below are widely read by clinicians, aiding uptake of findings. Finally, work incorporated here has been publicised through presentation at a number of clinically (as well as research focussed) workshops and conferences.

## Acknowledgments

I would like to acknowledge the encouragement, support and invaluable advice of my supervisors; Georgina Charlesworth and Katrina Scior. The UCL Clinical Psychology Department and in particular, Professors Peter Fonagy, Steve Pilling and Pasco Fearon have also been incredibly supportive. I would like to acknowledge the work of those involved in data collection through associated projects and research assistant roles: Catherine Bousfield; Lucy Gore; Noor Habib; Eleanor Chadwick; Maya Patel and Janina Brede. I have relied on CBT experts Henry Clements, Neil Ralph, Vyv Huddy, Peter Scragg and Sue Watson in developing the CBT skills measures I discuss below, and four peer support workers with dementia employed by Kensington Chelsea and Westminster memory service provided invaluable advice on making measures suitable for people with dementia. My wife Andry Moustras and parents Robin and Janet Stott have been a constant source of support, and my three children, Martha, Eleni and Zacharias have been remarkably tolerant of this endeavour despite their limited years.

There are also some acknowledgements related to funding: This PhD was conducted as part of a fellowship awarded to Joshua Stott by the Alzheimer's society. Grant number 236 (AS-CTF-14-005).

Chapter 3 was a secondary data analysis of data obtained from the Can't Forget to Worry trial and the Support at Home – Interventions to Enhance Life in Dementia (SHIELD) trial.

Funding for the Can't forget to worry trial is detailed in the following statement: This thesis presents independent research funded by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (grant reference number PB-PG-0609-18230). The views

expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Funding for the Shield trial is detailed in the following statement: This thesis presents independent research funded by the National Institute for Health Research [Programme Grant Reference RP-PG-0606-1083. Support at Home – Interventions to Enhance Life in Dementia (SHIELD)]. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

## **Contribution**

I designed, conceptualised and wrote up this project based on an initial outline used to apply successfully for an Alzheimer's Society research fellowship (see acknowledgments). I have overseen all ethical approval and data collection for the project. In order to collect data for the project I have supervised three students in (successfully completed) clinical psychology doctoral projects (DClinPsys). The aims of my project were unique and completely separate from the DClinPsys, and there was also not complete overlap of data or participants between my project and these projects as none of them separately or combined contain all the participants or measures that are used in my PhD. This was in line with the departmental guidelines on conducting PhDs in conjunction with professional doctoral theses. I have enclosed this guidance (appendix A), a declaration of project overlap for my project, and the declarations on project overlap submitted as part of each of the three DClinPsys (appendix B) as the guidance recommends. For completeness, I have detailed the specific overlap of data and participants for each empirical chapter of my thesis in Table 1 in the main body of the thesis where the datasets used in each empirical chapter are briefly described. I have also supervised one MSc project and two research assistants in collecting data that was used for the project. In agreement with my supervisors, I collected 10% of the data myself.

## Dissemination

Below is a list of papers published based on research conducted as part of this thesis (some of which are reported or referred to in the main body of this thesis, others in the appendix):

**Stott, J.**, Orrell, M., & Charlesworth, G. (2017). The validity and usefulness of the Hospital Anxiety and Depression Scale in carers of people with dementia: Evidence from confirmatory factor analysis, concurrent validity and measurement invariance in a large sample. *American Journal of Geriatric Psychiatry*, 25(11), 1270-1280, (referred to in chapter 3).

**Stott, J.**, Scior, K., Mandy, W., & Charlesworth, G. (2017). Dementia screening accuracy is robust to premorbid IQ variation: evidence from the Addenbrooke's Cognitive Examination-III and the Test of Premorbid Function. *Journal of Alzheimer's Disease*, 57(4), 1293-1302, (published version of chapter 4).

**Stott, J.**, Charlesworth, G., & Scior, K. (2017). Measures of readiness for cognitive behavioural therapy in people with intellectual disability: A systematic review. *Research in Developmental Disabilities*. 60, 37-51, (referred to extensively throughout and given in appendix C).

**Stott, J.**, Sweeney, J. M., Koschalka, L., O'Connor, L., & Mwale, A. (2017). People with dementia as peer workers, challenges, and benefits: a thematic analysis and nominal groups study. *International Psychogeriatrics*. 29(7), 1185-1192, (referred to in chapter 6).

**Stott, J.** Spector, A., Orrell, M., Scior, K., Sweeney, J., & Charlesworth, G. (2017). Limited validity of the Hospital Anxiety and Depression Scale (HADS) in dementia: evidence from a confirmatory factor analysis. *International Journal of Geriatric Psychiatry*, 32(7), 805-813, (published version of chapter 3).

## Table of Contents

<b>Abstract</b> .....	<b>3</b>
<b>Impact Statement</b> .....	<b>5</b>
<b>Acknowledgments</b> .....	<b>7</b>
<b>Contribution</b> .....	<b>9</b>
<b>Dissemination</b> .....	<b>10</b>
<b>Table of Contents</b> .....	<b>11</b>
<b>List of Figures</b> .....	<b>15</b>
<b>List of Tables</b> .....	<b>16</b>
<b>Chapter 1: Introduction - CBT for Anxiety and Depression in People Living with Dementia; the Importance of Cognitive Restructuring</b> .....	<b>18</b>
<b>Overview</b> .....	<b>19</b>
<b>Dementia</b> .....	<b>19</b>
<b>Depression and Anxiety in Dementia</b> .....	<b>22</b>
<b>CBT in PLWD</b> .....	<b>25</b>
<i>The need for more evidence to adapt CBT for PLWD.</i>	<i>27</i>
<i>A focus on core cognitive elements of CBT.</i>	<i>27</i>
<i>Cognitive restructuring.</i>	<i>30</i>
<b>Summary and Next Steps</b> .....	<b>30</b>
<b>Chapter 2: CBT Pre-Therapy Skills - Integrating the Dementia Literature with a Systematic Review of the Intellectual Disabilities Literature</b> .....	<b>32</b>
<b>Overview</b> .....	<b>33</b>
<b>CBT Readiness</b> .....	<b>33</b>
<b>Pre-Therapy Skills</b> .....	<b>34</b>
<i>Cognitive restructuring components as pre-therapy skills.</i>	<i>34</i>
<b>Summary and Aims of the Current Thesis</b> .....	<b>36</b>
<i>Aim 1: to develop tools to measure pre-therapy skills in PLWD.</i>	<i>37</i>
<i>Aim 2: to examine pre-therapy skill performance in PLWD, OA and YA.</i>	<i>40</i>
<i>Aim 3: to understand reasons for any pre-therapy skill differences.</i>	<i>41</i>
<i>Aim 4: to validate measures of confounding variables used in this thesis.</i>	<i>44</i>
<i>Aim 5: to examine neurocognitive correlates of pre-therapy skills in PLWD.</i>	<i>45</i>
<i>Summary of aims.</i>	<i>47</i>
<b>Chapter 3: Validity of the Hospital Anxiety and Depression Scale (HADS) in PLWD - Evidence from Confirmatory Factor Analysis</b> .....	<b>50</b>
<b>Abstract</b> .....	<b>51</b>
<b>Introduction</b> .....	<b>52</b>
<b>Method</b> .....	<b>54</b>
<i>Participants.</i>	<i>54</i>
<i>The HADS.</i>	<i>54</i>
<i>Statistical analyses.</i>	<i>55</i>
<b>Results</b> .....	<b>57</b>
<i>Data characteristics and initial analyses.</i>	<i>57</i>
<i>CFA results.</i>	<i>61</i>
<i>Measurement invariance.</i>	<i>64</i>

<b>Discussion</b> .....	<b>68</b>
<i>Conclusions.</i>	72
<b>Chapter 4: The Validity of the Test of Premorbid Function (TOPF)</b> .....	<b>73</b>
<b>Abstract</b> .....	<b>74</b>
<b>Introduction</b> .....	<b>75</b>
<i>Assumption 1: TOPF and ACE-III scores will correlate in an OA group.</i>	76
<i>Assumption 2: TOPF performance will be robust to dementia.</i>	77
<b>Method</b> .....	<b>78</b>
<i>Design.</i>	78
<i>Participants.</i>	78
<i>Recruitment procedures.</i>	81
<i>Data collection procedure.</i>	81
<i>Measures.</i>	82
<i>Statistical analyses.</i>	83
<b>Results</b> .....	<b>84</b>
<i>Participant flow.</i>	84
<i>Demographic and clinical characteristics.</i>	86
<i>Analysis of aims.</i>	88
<b>Discussion</b> .....	<b>90</b>
<i>The TOPF is associated with ACE-III performance.</i>	90
<i>The TOPF may not be robust to dementia.</i>	90
<i>Strengths and limitations.</i>	91
<i>Conclusions.</i>	92
<b>Chapter 5: Thought-Feeling Discrimination in PLWD - Development and Validation of a Clinical Tool</b> .....	<b>93</b>
<b>Abstract</b> .....	<b>94</b>
<b>Introduction</b> .....	<b>95</b>
<b>Methods</b> .....	<b>97</b>
<i>Participants.</i>	97
<i>Modification of the BTFQ.</i>	97
<i>Data collection and measures.</i>	100
<i>Statistical analyses.</i>	101
<b>Results</b> .....	<b>103</b>
<i>Sample characteristics.</i>	103
<i>Factor analysis.</i>	105
<i>The BTFQ- dementia (BTFQ-D) – validity and scoring.</i>	112
<b>Discussion</b> .....	<b>116</b>
<i>Lack of behaviour scale utility.</i>	116
<i>Structural and convergent validity.</i>	117
<i>Research and clinical implications.</i>	118
<i>Strengths and limitations.</i>	119
<i>Conclusions.</i>	120
<b>Chapter 6: Cognitive Mediation in PLWD - Development and Validation of a Clinical Tool</b> .....	<b>121</b>
<b>Abstract</b> .....	<b>122</b>
<b>Introduction</b> .....	<b>123</b>
<b>Methods</b> .....	<b>125</b>
<i>Participants.</i>	125
<i>Modification of the cognitive mediation measure.</i>	126

<i>Data collection and measures.</i>	135
<i>Statistical analyses.</i>	135
<b>Results .....</b>	<b>137</b>
<i>Sample characteristics.</i>	137
<i>Factor analysis.</i>	142
<i>The cognitive mediation - dementia Version (CM-DEM).</i>	145
<b>Discussion .....</b>	<b>148</b>
<i>A one factor structure.</i>	148
<i>Validity and reliability.</i>	148
<i>Research and clinical implications.</i>	149
<i>Strengths and limitations.</i>	150
<i>Conclusions.</i>	151
<b>Chapter 7: CBT Pre-Therapy Skills in PLWD, OA and YA - Investigating Group Differences and the Role of Neurocognition, Mood and Anxiety ...</b>	<b>152</b>
<b>Abstract.....</b>	<b>153</b>
<b>Introduction .....</b>	<b>155</b>
<b>Methods .....</b>	<b>158</b>
<i>Participants.</i>	158
<i>Data collection and measures.</i>	159
<i>Statistical analyses.</i>	160
<b>Results .....</b>	<b>162</b>
<i>Sample characteristics.</i>	162
<i>Aim 1: Do pre-therapy skills represent separable constructs?</i>	165
<i>Aim 2: OA, YA and PLWD group differences in pre-therapy skills.</i>	167
<i>Aim 3: Mediator and confounder analysis.</i>	170
<i>Aim 4: Correlates of pre-therapy skills in PLWD.</i>	173
<b>Discussion .....</b>	<b>175</b>
<i>CBT pre-therapy skills; separable and differentially affected in PLWD.</i>	175
<i>The centrality of neurocognition.</i>	176
<i>A deficit in CBT pre-therapy skill performance in OA vs. YA.</i>	177
<i>The role of anxiety and depression.</i>	177
<i>Research and clinical implications.</i>	178
<i>Strengths and limitations.</i>	179
<i>Conclusions.</i>	180
<b>Chapter 8: Discussion - Implications of Results for CBT Readiness, Research and Practice with PLWD and OA.....</b>	<b>182</b>
<b>Thesis Summary .....</b>	<b>183</b>
<b>The Wider Context of CBT Readiness.....</b>	<b>183</b>
<i>Common therapeutic processes.</i>	184
<i>Behavioural elements of CBT.</i>	185
<i>Other dimensions of CBT readiness.</i>	185
<b>Implications for Adaptation of CBT .....</b>	<b>187</b>
<b>Pre-therapy Skills Training.....</b>	<b>187</b>
<b>Evaluating Pre-therapy Skills in the Context of CBT Outcome.....</b>	<b>188</b>
<b>Implications for CBT with OA.....</b>	<b>189</b>
<b>Implications Outside of a CBT Context.....</b>	<b>190</b>
<b>Overall Strengths and Limitations.....</b>	<b>191</b>
<b>Conclusion.....</b>	<b>194</b>
<b>References .....</b>	<b>196</b>

<b>Appendices .....</b>	<b>210</b>
<b>Appendix A: Guidance on Undertaking a PhD While Supervising the     Research of Professional Doctorate Students .....</b>	<b>211</b>
<b>Appendix B Declarations of Overlap. ....</b>	<b>213</b>
<b>Appendix C: Measures of Readiness for Cognitive Behavioural Therapy     in People with Intellectual Disability: A Systematic Review.....</b>	<b>218</b>
<b>Appendix D: Tetrachoric Correlation Matrix Heatmap - BTFQ .....</b>	<b>260</b>
<b>Appendix E: The BTFQ-D .....</b>	<b>261</b>
<b>Appendix F: Qualtrix Prompt Event Valence Survey .....</b>	<b>264</b>
<b>Appendix G: Tetrachoric Correlation Heatmap – Cognitive Mediation</b>	<b>270</b>
<b>Appendix H: The CM-DEM .....</b>	<b>271</b>

## List of Figures

Figure 1. <i>The three most commonly proposed factor structures for the HADS...</i>	55
Figure 2. <i>Flow of participants through the study.....</i>	85
Figure 3. <i>Three-factor measurement model of CBT pre-therapy skills.....</i>	166
Figure 4. <i>Mediation of differences between PLWD and OA in pre-therapy skill performance .....</i>	172

## List of Tables

Table 1. <i>Datasets used in each chapter of this thesis</i> .....	49
Table 2. <i>Sample characteristics and descriptive statistics of HADS items</i> .....	59
Table 3. <i>Fit indices of original and respecified versions of one factor, two factor and three factor models</i> .....	63
Table 4. <i>Series of model comparisons to test measurement invariance of two and three factor models</i> .....	66
Table 5. <i>Demographic variables and TOPF and ACE-III scores for PLWD and OA groups</i> .....	87
Table 6. <i>Demographic and clinical characteristics of PLWD and OA groups</i> ...	104
Table 7. <i>BTFQ brief item statistics for all items</i> .....	106
Table 8: <i>Final factor loadings of BTFQ items in the full sample</i> .....	109
Table 9. <i>Final factor loadings of BTFQ items in the PLWD group</i> .....	110
Table 10. <i>Final factor loadings of BTFQ items in the OA group</i> .....	111
Table 11. <i>Median scores for BTFQ-D and convergent validity measures in all samples</i> .....	114
Table 12. <i>Correlations between ER-40 and BTFQ-D scales (convergent validity).</i> .....	115
Table 13. <i>Results of prompt event emotional valence and intensity survey</i> .....	130
Table 14. <i>Cognitive mediation brief item statistics for all items</i> .....	139
Table 15. <i>Final factor loadings of cognitive mediation items in the full sample</i> .	143
Table 16. <i>Final factor loadings of cognitive mediation items in the PLWD group</i> .....	144
Table 17. <i>Final factor loadings of cognitive mediation items in the PLWD group</i> .....	145

Table 18. <i>Correlations between CM-DEM, BTFQ-D and ER40 scales (convergent validity)</i> .....	147
Table 19. <i>Demographic and clinical characteristics of PLWD, OA and YA samples</i> .....	164
Table 20. <i>Model fit statistics for one- and three-factor models</i> .....	167
Table 21. <i>Between-group comparisons on pre-therapy skill measures</i> .....	169
Table 22. <i>Correlations between covariates and CBT pre-therapy skill measures in the combined OA and PLWD sample</i> .....	171
Table 23. <i>Correlations of CBT pre-therapy skill measures with neurocognitive and mood variables in the PLWD sample</i> .....	174

**Chapter 1: Introduction - CBT for Anxiety and Depression in People Living with Dementia; the Importance of Cognitive Restructuring**

## **Overview**

The purpose of the current chapter is to provide an introduction and, in conjunction with chapter 2, a rationale for, the empirical work presented in chapters 3 to 7 of this thesis. Key concepts will be defined and discussed including: dementia; anxiety and depression in dementia; and cognitive behavioural therapy (CBT) for anxiety and depression in dementia. The overarching aim of the thesis will be presented, namely, to investigate 'pre-therapy skills' required to be ready for the cognitive elements CBT in PLWD. The chapter will end by identifying cognitive restructuring as a central cognitive element of CBT and the basis for the pre-therapy skills examined in this thesis.

## **Dementia**

Dementia is a syndrome – usually of a chronic or progressive nature – in which there is deterioration in neurocognitive function (i.e. the ability to process thought) beyond what might be expected from normal ageing (Prince et al., 2016; Wortmann, 2012). It may affect memory, thinking, orientation, comprehension, calculation, learning capacity, language, or judgement in isolation or in combination. Consciousness is not affected. The impairment in neurocognitive function is commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation (Wortmann, 2012). Dementia is caused by a variety of diseases and injuries that primarily or secondarily affect the brain, such as Alzheimer's disease (the most common cause of dementia), or stroke (Wortmann, 2012). In 2015, around 850,000 people in the UK had dementia (Prince et al., 2016).

To avoid confusion, the term 'neurocognitive' (rather than cognitive) is used to describe impairments in *processing* thought - such as deficits in episodic memory, executive functions or language ability that characterize dementia. The

term 'cognitive' will be used to describe the thought modifying elements of CBT which are discussed in more detail below and in chapter 2. Additionally, the term 'PLWD' is used throughout this thesis in line with the current focus on rights of PLWD, which permeates current policy and service provision (Bartlett & O'Connor, 2007) and has led to a consequent emphasis on what PLWD want to be called (Bartlett & O'Connor, 2007).

There is no definitive diagnostic test for most types of dementia, including for the most common form; Alzheimer's disease, and for many types, including Alzheimer's disease, the cause is unclear (Livingston et al., 2017). National Institute for Health and Care Excellence (NICE) (NICE, 2018) guidance recommends that diagnosis is based on comprehensive clinical assessment involving history taking from the person with suspected dementia and (if possible) an informant. Diagnosis should include assessment of neurocognitive, behavioural and psychological symptoms and the impact of these on daily life as well as physical examination and tests to exclude reversible causes of dementia. In all cases, there should be formal neurocognitive testing with a structured standardised instrument (NICE, 2018). Further neurocognitive testing should be used to clarify diagnosis where level or cause of cognitive impairment is unclear (NICE, 2018). All PLWD recruited for the empirical studies in the current thesis were assessed for dementia according to these NICE recommended procedures.

Subtype of dementia should be determined by application of standardised criteria (NICE, 2018). NICE recommend National Institute on Aging (NIA) (McKhann et al., 2011) criteria for Alzheimer's disease, National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) criteria for vascular dementia (Román et al., 1993) and international consensus criteria for

dementia with Lewy Bodies (McKeith, 2006). These criteria generally require the use of imaging to distinguish dementia subtypes, with Cerebrospinal Fluid examination used in some cases. In the empirical work presented in this thesis, subtype of dementia was determined using these criteria. However, all types of dementia are included as this reflects CBT trials for PLWD where there were not exclusion on the basis of dementia subtype (Orgeta, Qazi, Spector, & Orrell, 2014).

The lack of definitive tests, and the facts that (i) PLWD may be living with more than one form of dementia simultaneously (Korczyn, 2002), and (ii) may approach services at different times in their dementia journey (Grimmer et al., 2015) mean that there is great heterogeneity in presentation (Livingston et al., 2017). This heterogeneity is particularly evident in the neurocognitive effects of dementia, which tend to start with specific deficits and become more global as the syndrome progresses (Salmon & Bondi, 2009). The specific neurocognitive effects of the initial stages of dementia typically vary between different diagnostic subtypes (Salmon & Bondi, 2009). For example, the pathology of the most common type of dementia, Alzheimer's disease, normally presents first in medial temporal lobe areas and thus initially affects the new learning served by those brain areas (Squire & Zola-Morgan, 1991). By contrast, pathology in behavioural variant frontotemporal dementia, for example, presents first in the prefrontal cortical areas with consequent impact on emotional and executive control processes served by those areas (Piguet, Kumfor, & Hodges, 2017). To complicate the picture further, there is substantial variation in impairment even within a diagnostic subtype. As an example, the pathology for one atypical form of Alzheimer's disease (Posterior Cortical Atrophy) initially presents in the

occipital region of the brain and affects spatial processing rather than memory in the first instance (Crutch et al., 2012).

This neurocognitive heterogeneity has relevance for the current work, because, as discussed in chapter 2, certain types of neurocognitive deficits may be more associated with CBT relevant skills than others, but this is under-investigated. Consequently, performance in specific neurocognitive domains will be measured in the current thesis.

Most types of dementia are chronic, with ongoing deterioration in neurocognition and there are no disease modifying treatments for the most common types of dementia (Livingston et al., 2017). For Alzheimer's disease, dementia with Lewy bodies and Parkinson's disease dementia, some pharmacological interventions are recommended in NICE guidelines (NICE, 2018) and also in the Lancet Commission on Dementia (Livingston et al., 2017). These include Donepezil, Galantamine, Rivastigmine and Memantine. These medications target biochemical abnormalities underlying neuronal loss (Lane, Potkin, & Enz, 2006) and have small clinically important effects on cognition and function (Lanctôt et al., 2003; Winblad, Jones, Wirth, Stöfler, & Möbius, 2007). With regards to non-pharmacological interventions for neurocognition, group (neuro)cognitive stimulation is recommended in NICE guidelines (NICE 2018), with recent reviews suggesting that (neuro)cognitive rehabilitation and exercise may also be effective in enhancing neurocognition or at least slowing decline (Livingston et al., 2017).

### **Depression and Anxiety in Dementia**

While the primary diagnostic criterion for dementia is the presence of neurocognitive decrements, 'neuropsychiatric symptoms' are common in dementia. These typically increase with dementia severity (Srikanth, Nagaraja, &

Ratnavalli, 2005), and there is evidence that they affect nearly everyone with dementia at some point (Srikanth et al., 2005). 'Neuropsychiatric symptoms' is an umbrella term and factor analysis suggests that the various symptoms cluster into psychosis, apathy, hyperactivity and affective sub-syndromes (Aalten et al., 2008). The affective cluster constitutes anxiety and depression and it is anxiety and depression in dementia that have been the focus of CBT interventions, and which are thus most relevant to, and measured in, the current thesis.

In people without dementia across all age ranges, anxiety and depression are the most common mental health problems (Steel et al., 2014), with mood disorders affecting 9.6 % of adults in the general population at some point in their life, and anxiety disorders affecting 12.9 % (Steel et al., 2014). There is, however, a lack of consensus as to the definition and aetiology of depression and anxiety in a dementia context (Banerjee et al., 2011; Seignourel, Kunik, Snow, Wilson, & Stanley, 2008). There has been more written about depression than anxiety (Seignourel et al., 2008) and it has been suggested that depression in PLWD probably differs from depression in people without dementia in psychological, biological and social terms (Enache, Winblad, & Aarsland, 2011). Specifically, a recent Lancet Commission (Livingston et al., 2017) suggested that there may be three types of depression in PLWD: (i) recurrent depression, earlier episodes of which predated the dementia; (ii) depression as a reaction to the dementia, and (iii) a 'syndrome' that looks like depression but is a direct consequence of the neurobiological changes inherent in dementia. It has been suggested that the type of depression might influence efficacy of pharmacological treatments (Enache et al., 2011) but there is limited evidence for this (Livingston et al., 2017).

Perhaps in part due to the lack of conceptual clarity, depression and anxiety in are difficult to diagnose in PLWD (Korczyn & Halperin, 2009; Seignourel et al., 2008). This is particularly the case in later stages of dementia given the reliance on self-report in traditional assessment methods (Enache et al., 2011). There is also a lack of brief, validated diagnostic self-report tools for both anxiety and depression in PLWD (Enache et al., 2011; Seignourel et al., 2008). This lack of tools is one of the reasons that an early aim of the current thesis is to examine the validity of a brief measure of anxiety and depression - the Hospital Anxiety and Depression Questionnaire (HADS) - in PLWD.

Despite the issues with conceptualisation and diagnosis, there is a consensus that both anxiety and depression are relatively common in PLWD (Enache et al., 2011; Seignourel et al., 2008) and can occur at any point in the disease trajectory. (Livingston et al., 2017). Prevalence estimates vary widely between studies depending on criteria used and population studied, but probably over 20% of PLWD have depression at any one time, with many more experiencing symptoms (Chi et al., 2015), and up to 21% of PLWD have a concurrent anxiety disorder (Seignourel et al., 2008). This means that, at current estimated dementia prevalence rates, around 178,000 PLWD in the UK would meet criteria for a diagnosis of depression with a similar number meeting criteria for an anxiety diagnosis (although given the high comorbidity of anxiety and depression (Kessler et al., 2005) these groups are likely to overlap).

Quite apart from the inherent subjective distress caused by anxiety and depression, they are likely to have a significant impact on PLWD and their carers. Depression in dementia is associated with (i) decreased quality of life (González-Salvador et al., 2000), (ii) earlier institutionalisation (Dorenlot, Harboun, Bige, Henrard, & Ankri, 2005), and (iii) greater caregiver burden (Ornstein & Gaugler,

2012). Anxiety, while less studied than depression in PLWD, is also associated with a number of deleterious outcomes (Gibbons et al., 2002; Seignourel et al., 2008). Consequently, while anxiety and depression are not inevitable, they are important issues to address for PLWD.

In PLWD, there are no recommended medications for anxiety disorders and anti-depressants may not work. The largest randomised controlled trial (RCT) concluded that side effects outweigh benefits (Banerjee et al., 2011) and a recent comprehensive review suggested that antidepressants should not be used as a first line treatment in PLWD (Livingston et al., 2017). Perhaps as a consequence of such findings, there has been interest in finding effective non-pharmacological interventions for depression and anxiety in PLWD. In particular, with the trend towards earlier diagnosis (Livingston et al., 2017) and consequent increased ability to engage in complex interventions, there has been an interest in adapting talking therapies for PLWD. The focus of this thesis is on a particular type of talking therapy for anxiety and depression – CBT (Roth & Pilling, 2008).

### **CBT in PLWD**

CBT is the form of talking therapy that has been most evaluated in dementia (Orgeta et al., 2014). It is an efficacious intervention for those with anxiety and depression (Cuijpers et al., 2013; S. G. Hofmann & Smits, 2008) in people who do not have dementia. The British Association of Cognitive and Behavioural Psychotherapists define CBT as a *'talking therapy that looks at how we think about a situation and how this affects the way we act and how, in turn, our actions can affect how we think and feel'* (BABCP, 2018). They further add that in CBT the *'therapist and client work together in noticing whether any thoughts or behaviours are unhelpful for the client and think about whether these could be changed.'* (BABCP, 2018)

The evidence for CBT for PLWD has focussed on anxiety and depression, and is reviewed below. However, in practice, it is also possible that CBT could be useful for PLWD who do not have anxiety and depression as it has been for people living with other long term conditions (for example, in adjusting to a diagnosis (Moss-Morris et al., 2013)). This is important to note because it means that PLWD receiving CBT are likely, but not a-priori, going to also be anxious or depressed. To reflect this, in this thesis, levels of anxiety and depression were measured as covariates but not used to determine inclusion.

Several case studies have examined adapted CBT interventions for anxiety and depression in PLWD e.g. (Kraus et al., 2008; Walker, 2004). However, the strongest evidence for efficacy is provided by two RCTs. The findings from these RCTs were also incorporated into a recent Cochrane review of talking therapies in PLWD (Orgeta et al., 2014). Importantly, because the comparison groups in both RCTs were 'treatment as usual', any intervention effect on anxiety or depression could be due to (i) CBT specific processes, (ii) factors which are common across most or all talking therapies, or (iii) a combination of both (Wampold, 2015).

For anxiety, the Cochrane review reports a small to medium effect of two different adapted CBT interventions (Spector et al., 2015; Stanley et al., 2013) on clinician rated anxiety. The studies were very small, potentially affecting statistical power and generalisability. The Spector et al. (2015) result reported in the Cochrane review is somewhat contradicted by the published trial, which reports a *non-significant* effect of the intervention on anxiety. The trial report is likely to be more accurate, because change in anxiety was adjusted for baseline scores unlike in the review. Consequently, evidence from two trials as to the impact of CBT on anxiety in PLWD is inconsistent and somewhat limited.

For depression, results are also inconsistent. Specifically, the Spector et al. (2015) trial found a large effect of the intervention on clinician rated depression, but the Stanley et al. (2013) trial found no such effect.

### **The need for more evidence to adapt CBT for PLWD.**

Overall then, the evidence for CBT for PLWD is potentially encouraging, but limited and inconsistent. One reason for this inconsistency in findings might be that, while these interventions both call themselves CBT, they are actually quite different from one another (Spector et al., 2015). A key reason for this difference might be that, while there is consensus that having dementia means that CBT needs to be adapted, there is not consensus on what adaptations to make and little evidence to inform this (Spector et al., 2015). An overarching aim of this thesis is to start to understand the extent to which PLWD have the 'pre-therapy skills' to be ready for CBT, so that future interventions can be adapted with this in mind. It also aims to develop measures of these pre-therapy skills to facilitate idiosyncratic adaptation of CBT in practice by providing clinicians with useful tools to assess who with dementia needs what adaptation. The precise pre-therapy skills measured in this thesis and the potential impact of dementia on these are a focus of chapter two. Before going on to this, it is necessary to be clear what is meant by 'CBT' in this thesis.

### **A focus on core cognitive elements of CBT.**

CBT is an umbrella term encompassing interventions that are common to all talking therapies as well as a wide range of behavioural and cognitively oriented interventions (Roth & Pilling, 2008). In this context 'cognitive interventions' refer to those aimed at reducing symptoms through identifying and changing unhelpful thoughts and appraisals and behavioural interventions relate to reducing symptoms through focussing on direct behaviour change activities

such as pleasant event scheduling (Roth & Pilling, 2008). Given the wide range of potential interventions, it would be impossible to focus on all pre-therapy skills required to be ready for all elements of CBT. Thus, the focus in this thesis is on the pre-therapy skills necessary to be ready for the cognitive elements of CBT specifically. This will probably have the most impact on CBT research and practice because (i) it is hypothesised that cognitive elements of CBT are particularly affected by neurocognitive impairments or limitations (Doherr, Reynolds, Wetherly, & Evans, 2005; Johnco, Wuthrich, & Rapee, 2014; Joyce, Globe, & Moody, 2006); (ii) it is cognitive elements that tend to be significantly adapted (Stanley et al., 2013) or removed (Teri, Logsdon, Uomoto, & McCurry, 1997) in CBT interventions for PLWD; and (iii) there is a lack of consensus as to the level of adaptation of cognitive elements required, with some authors (Stanley et al., 2013) proposing more adaptation than others (Spector et al., 2015) for the same mental health problem in PLWD.

To identify pre-therapy skills, it is necessary to define more precisely what is meant by cognitive elements of CBT (Doherr et al., 2005). A central issue with defining any aspect of CBT, which has been highlighted in the wider CBT literature (Roth & Pilling, 2008), is that, even for a specific mental health problem such as depression, CBT is not a unitary therapy, but encompasses a wide range of approaches (Roth & Pilling, 2008).

There have been two main ways of dealing with the fact that CBT is an 'umbrella term' to arrive at agreed definitions of CBT: the data driven approach and a focus on core elements of CBT.

### ***The data driven approach.***

This approach seeks agreement on what published interventions are evidence-based variants of CBT, and argues that these interventions together

constitute CBT. This data driven approach has been used in a number of influential competence frameworks (Hollon & Ponniah, 2010; Roth & Pilling, 2008) and has good content validity as it captures a diverse range of CBT interventions (Roth & Pilling, 2008). However, it may lack inter-rater reliability. For example, two influential attempts to delineate CBT for the same condition (depression) differed as to whether therapies such as the Cognitive Behavioural Analysis System of Psychotherapy and mindfulness constituted 'CBT' (Hollon & Ponniah, 2010; Roth & Pilling, 2008).

***The core elements approach.***

A second approach is to generate expert consensus on the 'core elements' one would expect in a CBT intervention and use this to define an intervention as CBT or not. This approach is used by those seeking to define CBT in reviews where well-defined evidence based variants of CBT do not exist (for example, a Cochrane review of CBT for people with schizophrenia (Jones, Hacker, Cormac, Meaden, & Irving, 2012). It is also used in the literature examining whether performance on 'CBT skills' improves over the course of therapy (Jarrett et al., 2013; Strunk, DeRubeis, Chiu, & Alvarez, 2007). While this core elements approach may improve reliability (through gaining agreement of several experts) it may lack content validity in that it does not capture all aspects of CBT (Roth & Pilling, 2008).

Despite this potential disadvantage, the second 'core elements' approach will be used here to define the cognitive elements of CBT. There are three reasons for this.

- (i) This approach is in line with the approach of authors who have measured CBT relevant skills (Dagnan, Chadwick, & Proudlove, 2000; Jarrett et al., 2013; Strunk et al., 2007) in other populations.

- (ii) Using the data driven approach would require listing all elements used in or needed for all interventions that are labelled CBT and be impractical.
- (iii) The core elements approach is conceptually more aligned with the aims of this thesis, which are to identify skills relevant to most, if not all, CBT approaches rather than peripheral aspects of particular CBT interventions.

### **Cognitive restructuring.**

Given the content validity issue highlighted in the preceding discussion, there is unlikely to be just one core cognitive element of CBT and a number could potentially have been focussed on. However, cognitive restructuring was chosen as the core cognitive element of CBT to focus on in the current thesis. Cognitive restructuring can be defined as 'strategies that focus on the exploration, evaluation, and substitution of the maladaptive thoughts, appraisals, and beliefs that maintain psychological disturbance' (D. A. Clark, 2013). It was chosen for four reasons: (i) It is central to early conceptualisations of CBT (Beck, 1979). (ii) It is frequently identified as a core element by other authors examining CBT relevant skills (Jarrett et al., 2013). (iii) It is frequently used by reviewers of the CBT literature (Vernooij - Dassen, Draskovic, McCleery, & Downs, 2011) to define interventions as CBT. (iv) The role of cognitive restructuring in CBT efficacy is supported by empirical evidence, with work indicating potential mediation of CBT outcome by developing cognitive restructuring skills in CBT for depression (Hundt, Mignogna, Underhill, & Cully, 2013).

### **Summary and Next Steps**

This chapter defined dementia, discussed its neuropsychiatric correlates and identified that the overarching aims of this thesis were to understand and

measure pre-therapy skills required to be ready for cognitive elements of CBT in PLWD. 'Cognitive restructuring' was identified as the element of CBT that will form the basis for defining CBT pre-therapy skills evaluated in empirical chapters. The next chapter will define the specific pre-therapy skills focussed on in this thesis and integrate the dementia literature with a systematic review of the intellectual disabilities literature in order to specify the aims of the empirical elements of this thesis.

**Chapter 2: CBT Pre-Therapy Skills - Integrating the Dementia Literature with  
a Systematic Review of the Intellectual Disabilities Literature.**

## **Overview**

A central proposition of the current thesis is that certain pre-therapy skills are necessary to be ready for CBT. Consequently, this chapter will first define what is meant by 'CBT readiness' and the precise pre-therapy skills necessary to be ready for the core CBT component of cognitive restructuring will then be outlined. In the main body of this chapter, findings from the literature related to neurocognitive and neuropsychiatric effects of dementia will be integrated with findings from a systematic review of pre-therapy skills in intellectual disabilities to delineate the specific aims of this thesis.

## **CBT Readiness**

CBT Readiness is a term originating in the intellectual disability literature and, in the current context, refers to whether, at the point of referral, PLWD are able to take part in CBT as offered by a healthcare system. The term 'readiness for CBT' (Willner, 2006) is preferred here to 'suitability', which is sometimes used in other populations (Dagnan et al., 2000; Safran, Segal, Vallis, Shaw, & Samstag, 1993), as unlike 'suitability', readiness implies potential for change (Rollnick, 1998). This implication of mutability is both ethically desirable in not withholding interventions due to presumed 'unsuitability', and empirically justified as there is evidence that aspects of 'readiness' can be trained in cognitively impaired populations (Bruce, Collins, Langdon, Powlitch, & Reynolds, 2010).

Willner and others (Safran et al., 1993; Willner, 2006) have suggested that in order to be ready for CBT (or indeed any talking therapy) certain skills are required. These 'pre-therapy skills' (pre-therapy skills required for cognitive restructuring in particular) are the focus of this thesis. However, to be 'ready' for CBT is more than just possessing sets of pre-therapy skills. Specifically, Willner (2006) has argued that it requires motivation to take part in CBT as well.

Furthermore, in line with models commonly used to understand intervention uptake outside of a dementia context (Michie, Van Stralen, & West, 2011), it is also suggested in this thesis that 'system readiness' – the availability of opportunities to engage with an intervention (Michie et al., 2011) - is important. Motivation and system readiness, will be returned to in chapter 8, but the rest of this chapter and chapters 3-7 will concentrate on pre-therapy skills.

### **Pre-Therapy Skills**

The pre-therapy skills required to be ready for CBT include skills necessary for (i) processes common to any or most talking therapies (for example, skills related to an ability to form a working alliance), and (ii) processes specific to CBT (e.g. cognitive restructuring). Common therapeutic processes are related to outcome (Krupnick et al., 1996) and pre-therapy skills necessary to benefit from these common therapy processes (e.g. ability to form a working alliance) might be affected by many types of dementia (Potkins et al., 2003). This is of import (particularly since it could be through common therapy processes that CBT for PLWD has its impact (Spector et al., 2015)) and will be returned to in chapter 8. However, the focus here is on CBT specific pre-therapy skills, and, in particular, cognitive restructuring pre-therapy skills, which are discussed in more detail below.

#### **Cognitive restructuring components as pre-therapy skills.**

By definition, pre-therapy skills necessary to be ready for CBT should be present before CBT commences. For reasons discussed in chapter 1, CBT pre-therapy skills focussed on in this thesis will be derived from the concept of cognitive restructuring. Given that cognitive restructuring as a whole is expected to be acquired during CBT (Hundt et al., 2013) it is unreasonable to expect CBT naïve participants with neurocognitive limitations to be able to demonstrate

cognitive restructuring skills in their entirety, prior to CBT. However, cognitive restructuring is a multicomponent construct and while there are several definitions in the literature (D. A. Clark, 2013; Johnco et al., 2014) all include the following skills:

1. Recognising and being aware of mental states such as emotions, thoughts, and behaviours;
2. Discriminating between emotions, thoughts and behaviours;
3. Understanding the connections between thoughts, emotions, behaviours and situations and recognise their relationship to mental health;
4. Regulating emotions, particularly using cognitive reappraisal, i.e., to change the meaning of a stimulus in order to change emotions or behaviours with the ultimate aim of improving mental health.

For the purpose of the current thesis, pre-therapy skills required to be ready for CBT are defined as cognitive restructuring components 1, 2 and 3 above. The reason for not including component 4 in this definition is that there is a hierarchy of component skills in terms of complexity, with skills generally following the order detailed above in the process of CBT (Padesky & Greenberger, 2012). The pre-therapy skills required to be ready for CBT (rather than skills developed within it) are likely to be the less complex skills in the earlier stages of this hierarchy (i.e. components 1,2 and 3) (Padesky & Greenberger, 2012).

Despite this, even for these proposed 'early stage' components (1,2 and 3 above), there is a lack of consensus as to whether they are pre-therapy skills required for a client to be ready for CBT as suggested by some (Willner, 2006) or are to a degree learnt during therapy even by those without a cognitive impairment (Padesky & Greenberger, 2012). One of the reasons that a non-

cognitively impaired control group is included in the current work is to account for this by examining not absolute but *relative* skill level i.e. whether these pre-therapy skills are *more* affected in PLWD than in people without dementia.

The focus of this thesis will be specifically on cognitive restructuring components 2 and 3 above, i.e. pre-therapy skills in discriminating between thoughts, feelings and behaviours and linking thoughts, feelings and events. The reason for this is that, unlike component 1, which has been extensively examined in research on emotion recognition (Bora, Velakoulis, & Walterfang, 2016; Kumfor & Piguet, 2013) in PLWD, components 2 and 3 have not been examined in PLWD in published work. Furthermore, components 2 and 3 mirror those skills identified in other commonly used models of pre-therapy CBT skills originating in the intellectual disability literature (Dagnan et al., 2000; Oathamshaw & Haddock, 2006)

### **Summary and Aims of the Current Thesis**

In chapter 1 and above, CBT for PLWD and why CBT pre-therapy skills might be important to understand were discussed. The CBT pre-therapy skills that will be the focus for the current thesis were defined as cognitive restructuring components 2 and 3. It is recognised that these skills do not encompass the totality of skills necessary to be ready for CBT (something which was alluded to earlier and will be returned to in the discussion). However, for simplicity and clarity, in what follows they will simply be referred to as ‘CBT pre-therapy skills’ or ‘pre-therapy skills’. The remainder of this chapter integrates findings from a published systematic review by the author with findings from the dementia literature to map out the thesis aims, their rationale, and the chapters they will be addressed in.

An initial scoping review revealed that there were no published papers examining discrimination and connection of behaviours thoughts, emotions and events in PLWD. Consequently, measures and findings as to these skills in the intellectual disability literature (where there they have been examined) were systematically reviewed with a view to identifying appropriate measures for adaptation and findings that might inform the current work.

The rationale and aims of the current thesis are informed by combining the findings from this review with the preceding discussion as to pre-therapy skills and the literature on neuropsychiatric and neurocognitive impacts of dementia alluded to in chapter 1. Thus, rather than including the intellectual disabilities review as a standalone chapter, the published version (Stott, Charlesworth, & Scior, 2017) is included as appendix C and, in the remainder of this chapter, the findings from it are integrated with the preceding discussion to inform each of the aims of current work. Of note, the term 'readiness skills' is used in the aforementioned published literature review - its meaning is exactly the same as the term 'pre-therapy skills' used here.

**Aim 1: to develop tools to measure pre-therapy skills in PLWD.**

The label of 'dementia' encompasses an extremely heterogeneous group of people (Livingston et al., 2017). There is, consequently, highly likely to be substantial within group variability in PLWD's readiness for CBT. One way of approaching this is to idiosyncratically tailor clinical interventions based on an individual's pre-therapy skill level (Padesky & Greenberger, 2012). However, to do this, measures of pre-therapy skills validated for PLWD are required and these are not currently available. Thus, the first aim of this thesis is to develop tools, which will help guide assessment of which PLWD might be ready to take part in which aspects of CBT and which PLWD will need some adaptation of intervention

or training in CBT pre-therapy skills. Such tools may be particularly useful for clinicians who are not used to working with PLWD but may encounter PLWD in their clinical practice such as Improving Access to Psychological Therapies (IAPT) therapists.

To address aim 1, measures of discriminating between thoughts, feelings and behaviours and linking thoughts, feelings and events were identified from the systematic review of the intellectual disabilities literature discussed above and three were selected for adaptation and validation in the current work. These were:

1. a measure of behaviour-thought-feeling discrimination - The Behaviour Thought Feeling Questionnaire (BTFQ) (Oathamshaw & Haddock, 2006)-
2. two measures that tap the ability to link thoughts, feelings and events:
  - a) a measure of event-emotion linkage (the Reed Clements Task) (Reed & Clements, 1989)
  - b) a measure of cognitive mediation (the ability to recognise the interceding role of a thought between an event and its emotional consequence). (Dagnan, Mellor, & Jefferson, 2009)

Chapters 5 and 6 of this thesis detail the modification and validation of the BTFQ and cognitive mediation measures respectively, as well as the reasons for selection of these specific tools to measure pre-therapy skills. The factor analyses reported therein are also pertinent to the wider pre-therapy skill literature, since one finding from the systematic review (appendix C) was that there was very little evidence for validity or reliability of any task measuring pre-therapy skills and no task had been subjected to factor analysis. The Reed Clements task is not the focus of a validation chapter for two reasons. Firstly, (as will be obvious in chapters 5 and 6) it had a ceiling effect in PLWD and the lack of

variance made further evaluation of its psychometric properties not possible.

Secondly, unlike the other two measures, the Reed Clements task has been used with PLWD previously in an unpublished thesis and has been found to be feasible and acceptable (Harter, 2003)

In addition to validating pre-therapy skills measures in PLWD, a sub-aim in chapters 5 and 6 is to validate these tools for older adults without a recognised neurocognitive impairment (OA). This is because there is some evidence that performance on some measures of pre-therapy skills (e.g. emotion recognition may be lower in current cohorts of OA relative to younger adults (YA) (Orgeta & Phillips, 2007). Such work has led to suggestions that CBT for current generations of OA should be adapted for this (Laidlaw, Thompson, Gallagher-Thompson, & Dick-Siskin, 2015)– for example, by extending the period of ‘socialisation to the CBT model’ (Laidlaw et al., 2015). However, to the author’s knowledge, there are no measures of behaviour-thought-feeling discrimination or cognitive mediation validated for OA, which can be used to examine performance of these pre-therapy skills in this group.

Convergent validity of the tools developed in this thesis will be assessed with a measure of facial emotion recognition. This construct was chosen as it has been proposed that it taps component 1 of cognitive restructuring above (awareness of emotions) (Dagnan et al., 2000; Dagnan & Proudlove, 1997) and is thus hypothetically related to the pre-therapy skills measured in the current thesis, a prediction which has been borne out in empirical work in people with intellectual disabilities (Joyce et al., 2006). The Emotion Recognition 40 (ER40; was used to measure emotion recognition, because (i) it is validated for use in people living with mild dementia, and (ii) in line with recommendations on measurement of facial emotion recognition more generally (Edwards, Jackson, &

Pattison, 2002), stimuli used in it are photographs representing emotions of fear, disgust, sadness, happiness, and no emotion in people who vary in age, gender and ethnicity. The use of the ER-40 builds significantly upon work examining CBT pre-therapy skills in other populations, which have either used line drawings of faces or non-evoked emotions as stimuli, both of which have significant flaws as measures of facial emotion recognition (Edwards et al., 2002).

**Aim 2: to examine pre-therapy skill performance in PLWD, OA and YA.**

A key aim of this thesis is to inform future intervention development by developing evidence as to which CBT pre-therapy skills PLWD may particularly struggle with and consequently which aspects of CBT may require modification, training or, potentially, omitting from CBT interventions for this population. Thus, having identified and developed tools to measure CBT pre-therapy skills, it will be of clinical utility to inform future CBT intervention adaptation through understanding which, if any, these skills are impaired in PLWD. This has not been investigated for behaviour-thought-feeling discrimination previously and only one unpublished study has investigated it in relation to event, thought, emotion linkage (Harter, 2003). This study found that PLWD performed worse than OA who, in turn, performed worse than YA on measures of this construct. However, this study used non-validated measures of pre-therapy skills (making it hard to understand the meaning of poor performance on those measures) in a poorly defined sample of PLWD (meaning it was hard to know to whom the research applied).

Thus, in chapter 7, this thesis will expand upon the existing unpublished work to compare performance of a well-defined group of PLWD with an OA group on each of the validated pre-therapy skills measures described above. Given

Harter (2003)'s findings that current cohorts of OA may score lower than current cohorts of YA on CBT pre-therapy skills measures, a YA group will be included in chapter 7 so that differences in pre-therapy skills between OA and YA can also be examined. Finally, the systematic review of the intellectual disabilities literature in appendix C (Stott, Charlesworth, et al., 2017) suggested that some pre-therapy skills may have higher rates of above chance performance than others, perhaps indicating they are 'easier' and may be less affected by having a neurocognitive impairment. This issue will also be examined in the current thesis (chapter 7). The CBT pre-therapy skills measures used have different response modes (coded free response vs forced choice) and thus scores are not directly comparable. Consequently, the relative impact of having dementia on tasks will be assessed through examination of the effect sizes of differences in performance between OA and PLWD for each task.

**Aim 3: to understand reasons for any pre-therapy skill differences.**

It is useful clinically to understand potential reasons for any observed differences in CBT pre-therapy skills in order that any proposed adaptations to CBT can be optimally focussed (Spector et al., 2015). So, if the anticipated between group differences are found, the nature of these differences will be explored in terms of confounders and mediators. In this context, confounding variables are variables in which a difference between those with and without dementia is associated with CBT pre-therapy skill performance and is associated with *but not caused by* having dementia (McNamee, 2003). If confounders are not accounted for, erroneous conclusions may be drawn, whereby dementia is proposed to affect pre-therapy skill performance when in fact a diagnosis of dementia is simply acting as a marker variable for the true association between a confounder and pre-therapy skill level (McNamee, 2003). A mediating variable

differs from a confounder conceptually in that *is proposed to be caused by* dementia and to have a direct effect on CBT pre-therapy skill performance and thus represents a causal step between having dementia and reduced CBT pre-therapy skill level (Preacher & Kelley, 2011). Due to this proposed causal role, identifying mediators is important to inform development of treatment strategies (Kraemer, Wilson, Fairburn, & Agras, 2002).

One key potential mediating variable is neurocognition. The systematic review of the intellectual disabilities literature pointed to greater neurocognitive impairment being associated with poorer pre-therapy skill performance (Stott, Charlesworth, et al., 2017). Additionally, an underlying assumption in all CBT interventions (Spector et al., 2015; Stanley et al., 2013) that have been adapted for PLWD is that the neurocognitive difficulties that are, by definition, present in dementia will lead to difficulties in engaging with and benefitting from CBT. Consequently, chapter 7 will examine the hypothesis that overall levels of neurocognitive impairment mediate any differences in performance between groups of PLWD and OA on measures CBT pre-therapy skills. It should be noted that it is not expected that OA groups will show neurocognitive impairment relative to YA groups and it may well be that the underlying causes of any differences in CBT pre-therapy skills in OA are not the same as in PLWD (perhaps being due to cohort effects) (Chand & Grossberg, 2013). So, whilst differences in CBT pre-therapy skill performance between OA and YA groups are expected, mediation of these differences by neurocognition is not predicted.

Whether or not neurocognition plays a mediating role, other factors may confound any group differences in pre-therapy skill levels between those with and without dementia. Four potential confounders are considered in this thesis: age;

anxiety; depression; and cognitive reserve. The rationale for each of these is briefly described below:

**Age.**

As noted above, there may be age related differences in pre-therapy skill performance (probably due to the impact of cohort rather than age per se). It is common (and it is the case in this volume) that OA samples are younger than PLWD samples. Consequently, given that there may be differences between people of different ages in CBT pre-therapy skill performance and age may differ between people with and without dementia, age will be considered as a potential confounding variable for any differences in CBT pre-therapy skills between PLWD and OA groups.

***Anxiety and depression.***

Higher levels of anxiety and depression are associated with a dementia diagnosis (prevalence estimates for anxiety and depression are higher in PLWD relative to OA (Andreas et al., 2017; Chi et al., 2015; Enache et al., 2011; Seignourel et al., 2008)). Anxiety and depression are also associated with poorer performance on measures of other pre-therapy skills (emotion recognition) (Gur et al., 1992; Surcinelli, Codispoti, Montebanocci, Rossi, & Baldaro, 2006). Thus, by extension, they may affect the related pre-therapy skills studied here. Consequently, anxiety and depression will be measured as potential confounding variables for any observed differences between PLWD and OA. Understanding the role of anxiety and depression in CBT pre-therapy skill performance is of particular import, given that, as detailed in chapter 1, PLWD accessing CBT are a priori likely to *both* have dementia *and* be anxious or depressed. This will also be of utility in the general CBT readiness literature, where the role of anxiety and depression has been little examined (no studies examining this were found in the

systematic review of the intellectual disabilities literature (Stott, Charlesworth, et al., 2017)).

***Cognitive reserve.***

Cognitive reserve is the theory that individuals compensate for brain changes through application of pre-existing neurocognitive strategies (Stern, 2006). In PLWD with equivalent neurobiological changes, higher, rather than lower, cognitive reserve may lead to better task performance (Stern, 2012). It is therefore possible that pre-therapy skill performance (like that of other tasks) will be influenced by level of cognitive reserve. Since cognitive reserve is something that is built up prior to a diagnosis of dementia (Stern, 2012) it can only be measured by proxy variables (Stern, 2009). Two widely used proxies for cognitive reserve are measured in the current thesis:

- (i) Premorbid IQ, which can be seen as an index of an individual's prior general cognitive abilities and associated cognitive reserve (Stern, 2009). This is measured in the current thesis by the Test of Premorbid Function (TOPF) (Wechsler, 2011) a measure of irregular word reading
- (ii) Self-reported years of education (Stern, 2006).

**Aim 4: to validate measures of confounding variables used in this thesis.**

Addressing the potential effects of anxiety and depression on CBT pre-therapy skill performance as described above requires measurement of these constructs. The pragmatics of the current thesis require a brief self-report measure that captures both anxiety and depression. There are few, if any, such measures validated for use in PLWD (Enache et al., 2011). Thus, as a preliminary to examining CBT pre-therapy skills, the current work will examine the

validity of a brief measure of anxiety and depression (the HADS) (Zigmond & Snaith, 1983) in a sample of PLWD (chapter 3 of this volume), to inform its use in later chapters. As noted in chapter 1, validating the HADS will also have a wider utility for clinicians and researchers wishing to assess anxiety and depression in PLWD more generally.

Similarly, before deciding whether to use the TOPF (Wechsler, 2011) or years of education as the primary measure of cognitive reserve in the current study, it is of use to examine whether the TOPF validly taps premorbid IQ in PLWD and this is the aim of a brief chapter (chapter 4 of this volume).

**Aim 5: to examine neurocognitive correlates of pre-therapy skills in PLWD.**

As mentioned, the main reason that CBT has been adapted for PLWD is to account for the neurocognitive impairments inherent in the diagnosis (Spector et al., 2015; Stanley et al., 2013). In particular, perhaps due to the prominence of memory deficits in Alzheimer's disease, many of the strategies used in CBT for PLWD thus far have focussed on supporting memory – for example, frequent summaries and checking in with clients, use of external memory aids, and use of cognitive rehabilitation techniques such as spaced retrieval (Spector et al., 2015). Other adaptations have focused on overall reduction of cognitive complexity –for example, making use of the inherent structure implied within the CBT approach, proceeding at a slow pace, and presenting material as simply as possible (Spector et al., 2015). However, despite making these adaptations there is a lack of empirical work examining the extent and scope of the association of neurocognitive impairments with CBT relevant skills and consequently it is unclear which neurocognitive impairments may particularly affect which CBT relevant skills. Thus, the final aim of this thesis is to explore whether pre-therapy

skill performance in PLWD is particularly associated with specific aspects of neurocognitive function. This is theoretically important as it helps us to understand what might underlie differences in specific pre-therapy skill levels, and, could have clinical impact (as knowing which elements of neurocognition are particularly important may help with devising strategies to support or compensate). Results pertaining to this aim are preliminary because, pragmatic considerations meant that the measure of neurocognition used (the Addenbrooke's Cognitive Examination-III (ACE-III) (Hsieh, Schubert, Hoon, Mioshi, & Hodges, 2013), was relatively brief, with only limited construct validity for measurement of specific neurocognitive functions (Hsieh, Schubert, et al., 2013). Two aspects of neurocognitive functioning will be particularly focussed on:

***Language.***

CBT is a language-based therapy; thus, it is theoretically likely that scores on measures of language abilities will be associated with pre-therapy skill performance. This hypothesis is borne out in the intellectual disabilities literature, where the systematic review found relatively consistent associations between language scores and scores on CBT pre-therapy skills measures (Stott, Charlesworth, et al., 2017). It is hypothesised here that performance on the ACE-III language subscale will be positively correlated with CBT pre-therapy skill performance (this hypothesis is investigated in chapter 7).

***Executive functioning.***

To the author's knowledge, the role of executive functioning has not been assessed in relation to the pre-therapy skills measured in this thesis. However, it is associated with cognitive restructuring as a whole in OA samples (Johnco et al., 2014) with adaptations to CBT interventions proposed to account for this (Mohlman, 2008, 2013). Furthermore, other pre-therapy skills such as emotion

recognition are more compromised in types of dementia where pathology is focussed around prefrontal brain areas (e.g. Behavioural Variant Fronto-Temporal dementia) (Bora et al., 2016). Executive functioning difficulties are strongly related to these brain areas (Piguet et al., 2017) and also (perhaps because they are a proxy for other difficulties underpinned by these brain areas) to difficulties in emotion recognition (Bora et al., 2016). In the light of this, it is hypothesised that performance on the pre-therapy skills measured in the current study will be associated with performance on the ACE-III verbal fluency subscale (which taps executive function) (Strauss, Sherman, & Spreen, 2006).

### ***Other elements of neurocognitive functioning.***

As noted in above, CBT has frequently been adapted to account for memory difficulties, and the hypothesis that CBT pre-therapy skills are associated with ACE-III memory subscale scores will be tested. While there are no specific hypotheses relating to other areas of neurocognitive function tapped by the ACE-III subscales (attention and visuospatial abilities), correlations of performance with these areas will be provisionally explored for the purposes of hypothesis generation about the potential impact of these functions on CBT pre-therapy skill performance.

### **Summary of aims.**

Thus, the aims of this thesis (in order of the chapters they will be addressed in) are (i) to assess the validity of the HADS (chapter 3), and TOPF (chapter 4) in PLWD (these aims are examined first as they are a preliminary to using these tools to measure potential confounders in chapter 7); (ii) to develop measures of thought-feeling discrimination (chapter 5) and cognitive mediation (chapter 6) validated for use with PLWD; (iii) to compare performance of PLWD, OA and YA on these validated pre-therapy skill measures (chapter 7); (iv) to

examine whether neurocognition mediates, or cognitive reserve, age, anxiety of depression confound, any observed differences between PLWD and OA in pre-therapy skill performance (chapter 7); (v) to examine neurocognitive correlates of pre-therapy skill measures in PLWD with a focus on of language and executive function and memory (chapter 7); and (vi) to discuss the clinical and research implications of findings as a whole (chapter 8).

In order to address these aims, several overlapping and non-overlapping datasets were used. Some datasets also overlapped to differing degrees with the aforementioned DClinPsy theses. To provide clarity, a brief overview of the dataset used in each chapter and overlap with DClinPsys is given in table 1 below. More details on eligibility, recruitment and demographic breakdown of samples are given in the relevant chapters.

Table 1. *Datasets used in each chapter of this thesis*

Chapter	Dataset
3	Baseline HADS Data from two trials examining interventions for PLWD; Orrell et al. (2017) and Spector et al. (2015); N = 268 <sup>†</sup> .
4	Data collected from OA aged > 65 and PLWD for the purposes of this thesis; total N = 179, PLWD n =102, OA n = 77 <sup>§</sup>
5	The same sample (but different measures) as in chapter 4*
6	The same sample (but different measures) as in chapter 4*
7	OA and PLWD samples were the same as chapter 4, 5 and 6. The YA sample (n=56) was collected for this chapter. Overall N= 230 <sup>¶</sup>

---

Note: <sup>†</sup>no data or participant overlap with trainee theses; <sup>§</sup> 121 (68%) of participants in this chapter also contributed TOPF and ACE-III data to DClinPsys; \* no data overlap but 121(68%) participants were shared between these chapters and DClinPsys; <sup>¶</sup>no data overlap for main analyses, but 176 (76%) of participants were shared with DClinPsys and data used in sub-analyses (ACE-III, HADS, TOPF) was shared with DClinPsys.

## **Chapter 3: Validity of the Hospital Anxiety and Depression Scale (HADS) in PLWD - Evidence from Confirmatory Factor Analysis<sup>1</sup>**

---

<sup>1</sup> A version of this chapter has been published in a peer reviewed journal (Stott, Spector, et al., 2017)

## **Abstract**

*Objectives:* The HADS is a well-validated, self-report measure of both anxiety and depression. It is frequently used with PLWD and was used in the current thesis. However, its structural validity has never been examined in this population. The current study used confirmatory factor analysis (CFA) to assess this. *Methods:* Baseline data from two intervention studies for people living with mild to moderate dementia were combined (N = 268). CFA was used to test whether a one, two or three factor structure best fit the data. Indices of model misspecification were examined to test for poor quality items, and models re-specified accordingly. Finally, measurement invariance across gender and different levels of cognitive impairment was assessed. *Results:* A one-factor structure did not fit the data. Two and three factor structures fitted the data equally well. Model fit was improved by removal of two items. Measurement invariance was adequate across gender, but poor across groups with differing levels of cognitive impairment. *Conclusion:* The HADS is acceptable and feasible although somewhat difficult to interpret in PLWD. We suggest that it should be interpreted as measuring two separate factors of anxiety and depression and not one 'distress' factor. However, two items may need to be removed, affecting cut-off scores. Poor measurement invariance means the HADS may not be a good tool for measuring differences in anxiety and depression between those with mild and those with moderate cognitive impairment. Recommendations for use of the HADS with PLWD in clinical and research contexts (including this thesis) are given.

## **Introduction**

As detailed in chapters 1 and 2, depression and anxiety are common comorbidities in PLWD (Enache et al., 2011; Wolitzky-Taylor, Castriotta, Lenze, Stanley, & Craske, 2010) and are associated with negative outcomes (Gibbons et al., 2002). PLWD who are offered CBT will, by definition, also be presenting with anxiety and depression. Consequently, measuring anxiety and depression is important when assessing pre-therapy skills in PLWD. Doing so requires a valid self-report measure of anxiety and depression for use in PLWD.

Self-rating of mood in those with cognitive impairment is complex as self-report measures typically require memory for mood over a period of time and also self-awareness, both of which may be affected by a dementia diagnosis (Feher, Larrabee, & Crook, 1992). However, PLWD referred for CBT and the participants in the current thesis tend to have mild dementia with associated increased self-awareness and less impaired memory (Grimmer et al., 2015). Self-report measures are therefore probably relevant for the participants in this thesis and are increasingly relevant for PLWD more generally, particularly given the trend towards diagnosis at an earlier stage in disease progression (Grimmer et al., 2015). One particular advantage of such measures is that they can be used to measure mood in PLWD who have no available informant (Alzheimer's Association, 2012).

The HADS (Zigmond & Snaith, 1983), a 14-item self-report measure of anxiety and depression, is used in the current thesis. It is appealing for use in this project and more generally in PLWD as it is relatively brief, measures both anxiety and depression and can be used in those with comorbid physical health problems. It is also widely used in clinical practice and dementia research (e.g.

(Clare et al. (2012)). Therefore, evaluation of the measurement properties of the HADS in PLWD is important for the aims of this thesis and also more generally.

The utility of any measure stands or falls on its reliability and validity, both of which are multi-faceted constructs (Mokkink et al., 2010). The HADS performs well on some aspects of validity and reliability across different populations, for example, it shows good internal consistency, test-retest reliability and diagnostic accuracy in those with physical health problems and psychiatric inpatients (Bjelland, Dahl, Haug, & Neckelmann, 2002).

Structural validity, the degree to which item scores are an adequate reflection of dimensional structure is an important aspect of validity without which measurements cannot be adequately interpreted (Mokkink et al., 2010). Evidence for structural validity of the HADS in populations without dementia is mixed, with studies suggesting that the HADS measures one single 'distress' factor (Razavi, Delvaux, Farvacques, & Robaye, 1990), separate 'anxiety' and 'depression' factors (Zigmond & Snaith, 1983) or even three factors, following L. A. Clark and Watson (1991)'s tripartite model of anxiety, depression and negative affectivity (Cosco, Doyle, Ward, & McGee, 2012). While the structural validity of the HADS has not been specifically examined in PLWD, its evaluation with a medically-hospitalised OA sample (Helvik, Engedal, Skancke, & Selbæk, 2011) and a cognitively-intact nursing home sample (Haugan & Drageset, 2014) favours a two-factor structure. Such findings are not generalisable to PLWD (Cosco et al., 2012), as anxiety and depression present differently (Banerjee et al., 2011), and some items (e.g. the fourth item on the depression subscale (I feel slowed down)) may be confounded by cognitive functioning (Haugan & Drageset, 2014).

Hence, in the current chapter the usefulness and interpretability of the HADS in a sample of PLWD is assessed, evaluating its factor structure to

determine interpretation as an outcome measure or clinical tool. Implications for the use of the HADS in this thesis and more generally for research and clinical practice will be considered.

## **Method**

### **Participants.**

Data analysed in this study were the combined baseline data for participants living with mild to moderate dementia (diagnosed according to Diagnostic and Statistical Manual-IV (DSM-IV) (APA, 1994) taken from two clinical trials, one examining home based support for PLWD (Orrell et al., 2017), and the other, CBT for anxiety in dementia (Spector et al., 2015). The recruitment procedures and samples for these trials have been described in detail elsewhere (Orrell et al., 2017; Spector et al., 2015). All participants in both trials gave written informed consent. Ethical approval was obtained from 'East London 3 Research Ethics Committee' (reference number 10/H0701/124) for use of the Spector et al. (2015) data and from the Outer North East London Research Ethics Committee (reference number: 09/H0701/54) for use of the Orrell et al. (2017) data.

Age, gender, Mini Mental State (MMSE) (Folstein, Folstein, & McHugh, 1975) scores and dementia diagnosis (only recorded in the Orrell et al. (2017) study) of the combined sample are presented in Table 2.

### **The HADS.**

The HADS comprises 14 items each rated from 0-3, with higher scores indicating greater anxiety or depression. The anxiety and depression subscales each have seven items and a maximum score of 21 (Zigmond & Snaith, 1983) and cut-offs of 11 for caseness. Although structural validity and measurement invariance are not clear, reliability and other forms of validity are well established in non-dementia populations (Bjelland et al., 2002).

## Statistical analyses.

Confirmatory factor analysis (CFA) was used to test the fit of the HADS data from the combined dataset with the three most commonly proposed factor structures: the two-factor model of Zigmond and Snaith (1983), the one-factor model of Razavi et al. (1990), and the three-factor non-hierarchical model of Dunbar, Ford, Hunt, and Der (2000). Diagrams illustrating these models are shown in Figure 1.

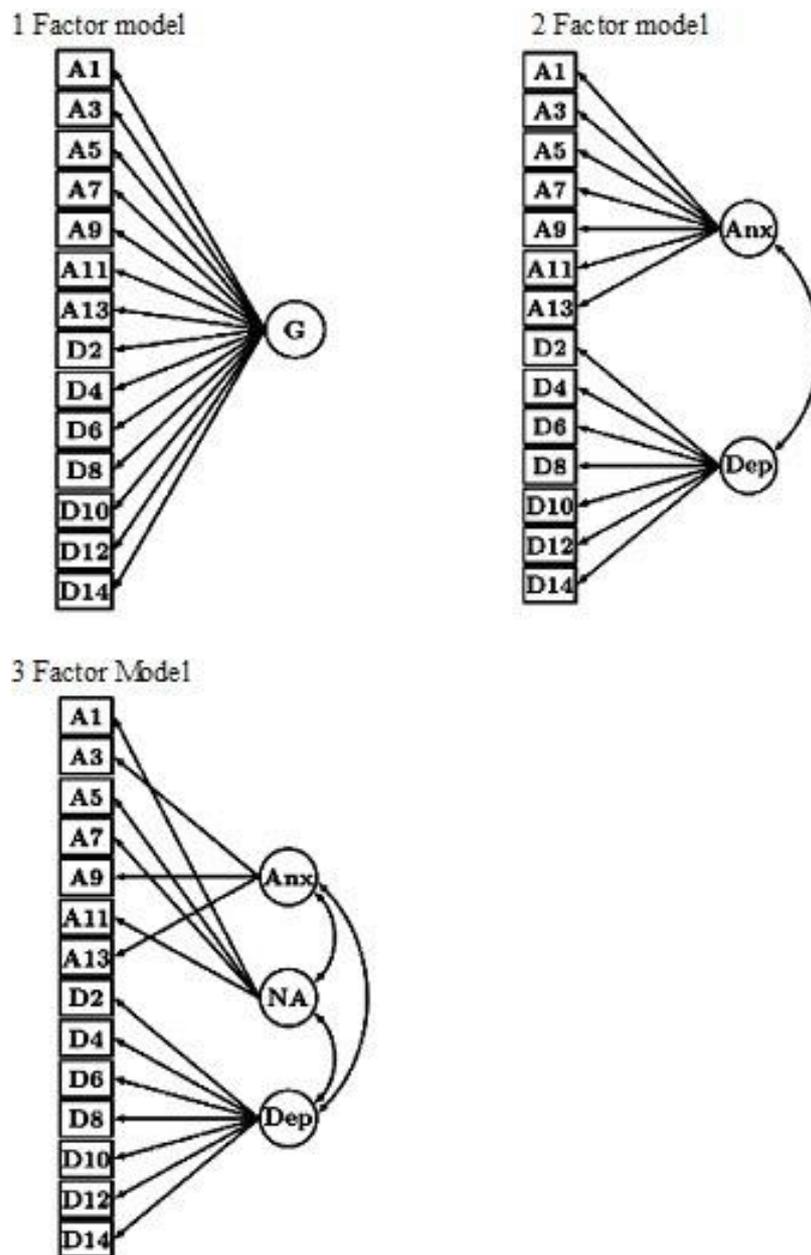


Figure 1. *The three most commonly proposed factor structures for the HADS*

CFA was performed in R (version 3.2.2) (R Core Team, 2013) statistical software using Lavaan (Rosseel, 2012) semplots and semtools packages (SemTools Contributors, 2015). For all models, independence of error terms was specified, and factors were allowed to correlate. The metric of latent variables was set by fixing the loading of one of the indicators for each variable at 1 (Byrne, 2013). Where assumptions of univariate normality (assessed by Shapiro Wilks' test) and multivariate normality (assessed by Mardia's test) were not met, Satorra Bentler corrected (robust) indices were used to examine fit of models (Hu, Bentler, & Kano, 1992).

### ***Indices of model fit.***

In line with the literature (Dunbar et al., 2000; Hu et al., 1992), model fit was assessed by several indices with cut-off scores used to determine good, adequate or poor fit. Two of the indices used here - the Standardized Root Mean Square Residual (SRMR) and the Root Mean Square Error of Approximation (RMSEA) - have cut-off scores of good fit,  $<0.05$ ; adequate fit,  $<0.08$ ; and poor fit  $> 0.08$ . Two of the others - the Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI) - have cut-off scores of good fit,  $>0.95$ ; adequate fit,  $>0.9$  and poor fit,  $<0.9$ . The final index, the Bayesian Information Criterion (BIC) is a comparative fit index, with smaller values indicating better fitting models, but no cut-off.

### ***Specification searching.***

Fit indices should be supplemented with information about how well individual items fit within a CFA model (Byrne, 2013). Here, items with standardised residuals with values in excess of 2.58 and high modification indices were classified as misspecified (Byrne, 2013). Both standardised residuals and modification indices were used to adapt and improve, or 're-specify', models through specification searching (Byrne, 2013). Once a good-

fitting model was obtained, parameters were examined for interpretability, size and statistical significance, and the presence of out of range values (Byrne, 2013). To avoid undue influence being given to the idiosyncrasies of a particular data set, *a priori* concerns were used to drive specification searching (Byrne, 2013).

### ***Measurement invariance.***

Measurement invariance of the HADS is necessary if it is to be used to test for differences in anxiety and depression across particular subgroups of PLWD. Measurement invariance is assumed if individuals in different groups with the same levels of the latent construct have the same expected raw-score on the measure (Drasgow & Kanfer, 1985). To test for measurement invariance of the HADS, the data were split into subgroups according to gender and cognitive impairment. In line with evidence relating MMSE to stage of dementia (Pernecky et al., 2006), mild impairment was defined as  $MMSE \geq 21$  and compared to a moderate impairment subgroup ( $MMSE \leq 20$ ). Following this, the models were examined for the different types of measurement invariance (configural, metric, strong and strict) through comparison of progressively more constrained models, with a change in CFI greater than 0.01 taken to indicate change in model fit across constraints and therefore lack of invariance between groups (Chen, 2007).

## **Results**

### **Data characteristics and initial analyses.**

Of the combined dataset (N = 339), 65 participants did not attempt the HADS. Of those who attempted the HADS, six were 'non-completers' (missing data for one or more items).

The data were examined for differences in gender, age and MMSE scores between those who completed, attempted and did not attempt the HADS. Chi

square was used to test for differences in gender and ANOVA for differences in MMSE and age. Groups did not differ for gender or age but did differ in MMSE, ( $F_{2,304} = 25.97$   $p < .001$ ) with planned comparisons revealing that non-attempters had lower MMSE scores than completers (Games Howell MD = -7.25,  $p < .001$ ).

For those who attempted the HADS (N= 274), Little's MCAR test revealed that data were missing completely at random ( $\chi^2 = 50.48$  (36),  $p = 0.06$ ). Where missing data are MCAR and under 5%, listwise deletion of cases with missing data is acceptable (Graham, 2009). Consequently, only those with full data available (N=268) were subject to CFA.

The final CFA sample consisted of 125 males (47%), 142 (52.9%) females and one unstated, mean age was 69 years (Standard Deviation (SD) of 12.3) and mean MMSE score was 19.8 (SD, 5.4). Descriptive data for the HADS items are shown in Table 2. Graphical inspection and significant Shapiro Wilk tests for all HADS items indicated significant univariate non-normality, with the sample generally reporting low levels of depression and anxiety and consequent positive skew. Mardia's test indicated significant multivariate non-normality ( $\chi^2$  skew = 1784.7,  $p < .001$ , Z Kurtosis = 26.2,  $p < .001$ ). Given this non-normal ordinal data, the CFA approach of Maximum Likelihood (ML) methods with Satorra Bentler corrected (robust) chi square was used to examine fit of all models (Finney & DiStefano, 2006).

Table 2. *Sample characteristics and descriptive statistics of HADS items*

	N (%)	Mean	SD	Skew	Kurtosis	Shapiro Wilk
<b>SAMPLE CHARACTERISTICS</b>						
Male	125(48)			NA	NA	NA
Age		69.1	12.3	NA	NA	NA
MMSE		19.8	5.4	NA	NA	NA
Dementia Diagnosis				NA	NA	NA
Not recorded (all participants in (Spector et al., 2015) trial and those with missing data in Orrell et al. (2017) trial.)	106 (39.3)			NA	NA	NA
Alzheimer's disease	109 (40.3)			NA	NA	NA
Vascular dementia	27(10)			NA	NA	NA
Frontotemporal dementia	3(1.1)			NA	NA	NA
Lewy Body Dementia	3(1.1)			NA	NA	NA
Any other type of dementia	23(8.2)			NA	NA	NA
<b>HADS ITEMS</b>						
1. I feel tense or 'wound up'		.74	0.78	1.07	1.029	0.769
2. I still enjoy the things I used to enjoy		.81	0.93	0.95	-.011	0.787
3. I get a sort of frightened feeling		.72	0.92	0.98	-.174	0.754
4. I can laugh and see the funny side of things		.51	.781	1.49	1.520	0.675

	N (%)	Mean	SD	Skew	Kurtosis	Shapiro Wilk
5. Worrying thoughts go through my mind		.75	.893	1.06	.308	0.769
6. I feel cheerful		.56	.750	1.35	1.569	0.715
7. I can sit at ease and feel relaxed		.78	.799	0.87	.320	0.796
8. I feel as if I have slowed down		1.19	.905	0.69	-.175	0.816
9. I get a sort of frightened feeling		.54	.751	1.40	1.656	0.707
10. I have lost interest in my appearance		.58	.829	1.30	.706	0.706
11. I feel restless as if I have to be on the move		.87	.888	0.74	-.287	0.816
12. I look forward with enjoyment to things		.68	.961	1.22	.296	0.709
13. I get sudden feelings of panic		.68	.793	1.04	.575	0.768
14. I can enjoy a good book or radio or TV program		.53	.901	1.70	1.825	0.625

Note: n was 268 for all items; all Shapiro Wilk statistics were significant  $p < .001$ ; NA = Not applicable.

### **CFA results.**

*Parameter estimates:* Parameter estimates were deemed adequate using Byrne (2013)'s three criteria: consistency with underlying theory, values falling inside admissible ranges, and parameters being statistically significant.

*Initial model fit:* Table 3 shows the fit indices for all three models. None of the models show a good fit with the data, with the one factor model performing particularly poorly across all indices. The two and three factor models both performed similarly and adequately on three indices (SRMR, CFI, RMSEA) but were poor on one (TLI).

### ***Specification searching.***

Given the mixed evidence as to the adequacy of fit of the two and three factor models but lack of difference in fit between them, specification searching (examination of modification indices and standardized residuals) of the two and three factor models was conducted to understand sources of model misspecification. This was first done in relation to item four on the depression subscale 'I feel slowed down', which had been specified *a priori*, as potentially problematic. The highest modification indices in the two factor (31.7) and three factor models (31.9) were associated with cross-loadings of this item onto latent variables other than depression (anxiety and negative affectivity). Additionally, in both models, this item was associated with the largest standardized residual covariance values (5.16, in the three-factor model and 5.01 in the two-factor model) and also the highest number of these in excess of 2.58 (five in both the two and three-factor models). Given clear evidence that this item was a source of misspecification, it was removed from the analyses, which were then re-run with results and fit indices detailed in Table 3. The fit of both two and three-factor models was improved, such that indices of fit were now 'adequate' for some fit

indices (TLI, CFI) and 'good' for others (RMSEA, SRMR). There was, however, still no discrimination between the models, with both models having almost identical fit indices

Table 3. *Fit indices of original and respecified versions of one factor, two factor and three factor models*

Model	$\chi^2$ (robust)	DF	SRMR	CFI	TLI	RMSEA (90% CI)	BIC
One factor	203.70	77	0.076	0.799	0.763	0.078 (0.067-0.089)	8831.78
Two factor original	135.90	76	0.066	0.905	0.886	0.054 (0.04-0.067)	8738.06
Two factor without HADD4	99.20	64	0.055	0.939	0.926	0.045 (0.030-0.060)	8047.85
Two factor without HADD4 or HADA4	70.10	53	0.042	0.967	0.959	0.035	7452.36
Three factor original	132.40	74	0.065	0.907	0.886	0.054 (0.041-0.067)	8746.52
Three factor without HADD4	96.40	62	0.054	0.941	0.925	0.045 (0.030-0.060)	8056.73
Three factor without HADD4 or HADA4	67.69	51	0.043	0.968	0.959	0.035 (0.009-0.053)	7460.95

Note: SRMR \_ standardized root mean residual (<0.05 suggests good fit, <0.08 suggests adequate fit, >0.08 suggests poor fit); CFI \_ comparative fit index (>0.95 suggests good fit, >0.9 suggests adequate fit, <0.9 suggests poor fit); TLI\_tucker Lewis Index (> 0.95 indicates good fit, >0.9 suggests adequate fit, <0.9 suggests poor fit ); RMSEA \_ root mean square error of approximation (< 0.05 is good fit, <0.08 is adequate fit, >0.08 is poor fit); CI\_confidence interval; BIC\_ Bayesian information criterion; HADD4, fourth item on the depression subscale; HADA4, fourth item on the anxiety subscale.

On inspection, the fourth item on the anxiety subscale ('I can sit at ease and feel relaxed') was associated with the next highest modification indices (22.1 in the three-factor model, 23.1 in the two-factor model) and next highest number of standardized residuals above 2.58 (three) in both the two and three factor models. While there was not an *a priori* reason for removing this item, given some evidence of misspecification, it was removed (along with the fourth item on the depression subscale) for an exploratory analysis (with results and fit indices detailed in Table 3). Removing item four on the anxiety scale improved model fit in both two and three factor models with all indices now suggestive of 'good fit'. However, once more there was no difference between models.

#### **Measurement invariance.**

In order to assess whether the HADS can be validly used to measure differences in depression and anxiety across groups who differ in cognitive functioning or gender, measurement invariance of the HADS was assessed across these groups. Measurement invariance assessment was conducted on both the two and three factor models with the fourth item of the depression subscale removed (the models with the fourth item on the anxiety subscale also removed were not subjected to this analysis as the removal of this item was exploratory). The data were first divided into subgroups according to gender (male  $n = 142$ ) and separately by MMSE score. For five participants, MMSEs were missing so the sample size for this analysis was 263, with  $n = 142$  falling into the low MMSE group and  $n = 121$  into the high MMSE group.

The results of the analysis of the different types of invariance (configural, metric, strong and strict) are shown in Table 4. Measurement invariance of the HADS was adequate across groups who differ in gender but inadequate across MMSE categories. Specifically, the data indicate that for groups differing

according to gender, configural invariance criteria were adequate for CFI (three-factor 0.91, two-factor 0.91) and RMSEA (three-factor 0.071, two-factor 0.069) and criteria for all other invariance types were met with CFI change always more than 0.01. For groups differing in cognition, the high cognition group for the three-factor model had a non-positive definite covariance matrix, meaning that it was difficult to interpret invariance for this model, and measurement invariance could not be assumed. For the two-factor model, the configural invariance assumption was not met (CFI of 0.88 and a RMSEA of 0.081) although measurement invariance was demonstrated across all other levels ( $CFI\Delta < 0.01$ ) aside from strict invariance where  $CFI\Delta = 0.013$

Table 4. Series of model comparisons to test measurement invariance of two and three factor models

Subgroup	Model	Invariance type	$\chi^2$ ( $\Delta\chi^2$ )	DF ( $\Delta$ DF)	$\Delta p$	CFI( $\Delta$ CFI)	RMSEA
Gender	Two factor model without HADD4	Configural	210.77	128	N/A	0.910	0.069
		Metric	(13.25)	11	(0.277)	(0.002)	NA
		Strong	(17.71)	11	(0.088)	(<0.001)	NA
		Strict	(7.83)	2	(0.020)	(0.002)	NA
	Three factor model without HADD4	Configural	207.05	124	N/A	0.909	0.071
		Metric	(12.69)	10	(0.241)	(0.003)	NA
		Strong	(17.81)	10	(0.058)	(0.009)	NA
		Strict	(8.00)	3	(0.046)	(0.005)	NA
MMSE	Two factor model without HADD4	Configural	238.63	128	N/A	0.884	0.081
		Metric	(8.88)	11	(0.632)	(0.002)	NA
		Strong	(25.07)	11	(0.009)	(0.015)	NA
		Strict	(36.67)	2	(0.102)	0.003	NA
	Three factor model without HADD4	Configural	229.80	124	N/A	0.889	0.081
		Metric	(9.346)	10	(0.499)	(0.001)	NA
		Strong	(22.026)	10	(0.015)*	(0.013)*	NA

Subgroup	Model	Invariance type	$\chi^2 (\Delta\chi^2)$	DF ( $\Delta$ DF)	$\Delta p$	CFI( $\Delta$ CFI)	RMSEA
		Strict	(7.843)	3	(0.049)	(0.005)	NA

Note: For configural invariance fit indices, CFI \_ comparative fit index (>0.95 suggests good fit, >0.9 suggests adequate fit, <0.9 suggests poor fit), RMSEA \_ root mean square error of approximation (< 0.05 is good fit, <0.08 is adequate fit, >0.08 is poor fit); for all other invariance types,  $\Delta$ CFI < 0.01 implies that the invariance assumption still holds;\* indicates that invariance assumption is not met according to these criteria.

## **Discussion**

This study suggests that people living with mild to moderate dementia can complete the HADS, but does raise some concerns about the structural validity and consequent interpretation of HADS scores in PLWD.

### ***A single distress factor?***

In line with previous CFA studies in other populations (Cosco et al., 2012; Haugan & Drageset, 2014) there was no evidence that the HADS measures a single distress factor. It is suggested, that the HADS should not be interpreted in this way in PLWD in clinical or research contexts.

### ***Structural ambiguity and pragmatic use of a two-factor model.***

It was not possible to distinguish between two interpretations of the HADS; that it measures two factors of 'anxiety' and 'depression' or that it measures three factors of anxiety, depression and negative affectivity. The inability to distinguish between different interpretations has been termed 'structural ambiguity' (Wang, Lopez, & Martin, 2006) and makes understanding HADS scores in PLWD difficult.

Structural ambiguity has been found with the HADS in other populations (Wang et al., 2006) and is in line with the general lack of clarity over HADS structure (Cosco et al., 2012). This has led some authors to advocate abandoning it altogether (Coyne & van Sonderen, 2012). One strategy for deciding between structural models is to favour the most parsimonious structure.

However, fit indices used in the current study (e.g. BIC) take model parsimony into account (Neath & Cavanaugh, 2012) and did not indicate that the two factor structure should be preferred. A strategy to disambiguate in future research would be to test the indices derived from the two and three factor structures for other forms of validity (e.g. convergent or criterion validity) in PLWD

and to see which performs best. Until this research has been done, it is suggested that, if the HADS is to be used with PLWD, a two-factor interpretation might be preferred due to its greater simplicity of scoring. Consequently, a two-factor interpretation will be used in the current thesis.

Given this and that fit indices are so similar for two and three factor models, the rest of this discussion will focus on the two-factor interpretation.

### ***Removal of items.***

Model fit was improved by removal of two items (the fourth items on the anxiety and depression subscales). The prediction that the fourth item on the depression subscale (I feel slowed down...) would be confounded by cognitive impairment was supported by the poor fit of this item coupled with its relatively high mean score, which could reflect individuals endorsing it due to cognitive impairment regardless of depression. It was more surprising that the fourth item on the anxiety scale (I can sit at ease and feel relaxed) did not relate to the underlying construct of anxiety. This may be a data idiosyncrasy but has been found before (Haugan & Drageset, 2014) and warrants further consideration.

Given their poor fit to the data, it is suggested that HADS users should definitely remove the fourth item on the depression scale and consider removing the fourth item on the anxiety subscale in scoring the HADS for PLWD. The removal of one or both of these items will affect the ability to use HADS cut-offs for anxiety and depression caseness, so it is suggested that future work with the HADS in PLWD could also focus on developing cut-offs for shortened HADS subscales excluding these items. Consequently, in the current thesis, when the HADS scales are used as continuous variables (e.g. in measuring associations with measures of CBT pre-therapy skills) the 4<sup>th</sup> items on both depression and anxiety scales will be removed. This 12 item HADS will be referred to as 'HADS

(dementia modified) version' in the following chapters. However, when HADS scales are used for assessing caseness, the 4<sup>th</sup> items will be included in the measure, with caveats as to interpretation noted.

***Measurement invariance.***

The measurement invariance data suggest that, in PLWD, differences in mean HADS scores between moderate and milder impairment groups may be un-interpretable. This is because such differences may be due either to between group variation in the relationship of raw HADS scores to the latent constructs of anxiety and depression or to between group differences in anxiety and depression themselves (Hirschfeld & Von Brachel, 2014). The implication of this is that research using the HADS to examine differences in anxiety and depression between mild and moderate impairment groups will be hard to interpret. Similarly, in clinical work, where normative reference groups differ in the degree of cognitive impairment to a person living with dementia, HADS scores will be difficult to meaningfully understand. As measurement invariance is better across gender, comparisons in HADS scores between men and women living with dementia can be performed.

In writing this chapter on the structural validity of the HADS in PLWD it became clear that this issue had not been examined in carers of PLWD despite the fact that such investigation would have significant utility in clinical and research practice (Stott, Orrell, & Charlesworth, 2017). This is of some relevance to the current work, since CBT for PLWD frequently also includes carers in the intervention and it is also important to measure how their anxiety and depression levels change over the course of CBT. Consequently, as an adjunct to this thesis, the utility of the HADS in a large sample of carers of PLWD was evaluated by the author of the current work and the results published (Stott, Orrell, et al., 2017).

### ***Strengths and limitations.***

This chapter reports on the first study to examine the structural validity of the HADS in PLWD. A strength is the use of a CFA approach to test hypotheses as to which structure proposed in the literature best fits for PLWD. Some limitations require noting. The mean score on HADS items was low. Future work should examine this measure in samples where there is more variance and higher levels of depression and anxiety. Dementia is an umbrella term and factor structure may differ across specific dementia diagnoses, which have different patterns of impairment. For example, those with behavioural variant Frontotemporal Dementia may lack insight (Rosen et al., 2014) and under-report anxiety or depression. There were not the data available to examine this, but future research is recommended. A number of individuals in the Orrell et al. (2017)'s study did not attempt the HADS, and were excluded from the analysis. Consequently, these results are only representative of those PLWD, who attempt the HADS, not PLWD as a whole. The higher MMSE scores of attempters compared to non-attempters suggests that this population may have higher cognitive functioning, although there were attempters with very low MMSE scores, indicating low MMSE scores should not be used to rule out use of the HADS.

The relatively small sample size (for a CFA study) may result in the structural ambiguity found here (Wang et al., 2006). Replication with a larger sample is recommended. Finally, although the most frequently proposed structures in the literature were evaluated, not all potential HADS structures were considered. Future research should examine the bi-factor structure (Norton, Cosco, Doyle, Done, & Sacker, 2013) and the impact of measurement artefacts (Straat, van der Ark, & Sijtsma, 2013). Item Response Theory studies may be

useful to conduct in PLWD as these provide strong evidence of latent variable structure, with the particular advantage of being generalizable beyond a sample to a population (Cosco et al., 2012).

### **Conclusions.**

This study suggests that the HADS is feasible for use with PLWD, but is somewhat difficult to interpret. The HADS should not be used to measure one factor of 'distress' in this population. While two and three factor structures are equally supported here, it is suggested that the HADS is used to measure two factors of anxiety and depression for simplicity of scoring and this is how it will be used in the current thesis. Two HADS items may not be useful in PLWD and further work is needed to develop cut-off scores for a reduced item version. Consequently, in the current thesis, while the reduced item (HADS (dementia modified)) version will be used for measuring anxiety and depression as continuous constructs, caseness will be assessed with the original version. Lack of measurement invariance means that the HADS may not be suitable to measure differences in anxiety and depression where groups differ in level of cognitive impairment.

## Chapter 4: The Validity of the Test of Premorbid Function (TOPF)<sup>2</sup>

---

<sup>2</sup> A version of this chapter has been published in a peer reviewed journal (Stott, Scior, Mandy, & Charlesworth, 2017)

## **Abstract**

*Objectives:* The TOPF is a widely-used measure of premorbid IQ. However, its validity and clinical utility in PLWD is relatively under-investigated. This study investigates validity of the TOPF in relation to two key assumptions (associations with neurocognitive measures in people without dementia and robustness to dementia). *Methods:* 179 UK based adults (102 memory service attendees diagnosed with dementia and 77 typically functioning volunteers over the age of 65 without subjective memory impairments) completed the ACE-III and the TOPF. Validity was assessed using hierarchical multiple linear regression to examine (i) the association between TOPF and ACE-III scores in typically functioning over 65s, and (ii) the association of dementia diagnosis with TOPF scores in the whole sample, independent of potential confounding variables of gender, education and age. *Results:* Although ACE-III scores were associated with TOPF scores independently of demographic indices, suggesting that it may validly tap premorbid ACE-III performance, TOPF performance was also associated with dementia status and thus may not be robust to dementia. *Conclusion:* The TOPF meets one validity criterion (association with neurocognitive measures in those without dementia). However, cross sectional findings did not support robustness to dementia and, consequently, the TOPF will not be used as the main measure of cognitive reserve in future chapters of this thesis.

## Introduction

As discussed in chapter 2, measures of premorbid IQ are widely used as a way of tapping cognitive reserve - the ability to use pre-existing neurocognitive strategies to compensate for neurological damage. A diagnosis of dementia is associated with lower cognitive reserve (Stern, 2012) and it is hypothesised here that lower cognitive reserve may also be associated with lower pre-therapy skill levels. As such, cognitive reserve is an important variable to measure in understanding the potential reasons for any differences in CBT pre-therapy skills between people with and without dementia. In light of this, the current chapter will examine evidence for validity of the measure of premorbid IQ used in the current thesis, the TOPF (Wechsler, 2011), in PLWD.

The TOPF is a test of irregular word reading and the reading of irregular English words is the most established methodology for premorbid IQ measurement in PLWD (Crawford, Stewart, Parker, Besson, & Cochrane, 1989). There are a number of English irregular word reading tests (Nelson & Willison, 1991; Wechsler, 2001, 2011). Most research has been conducted using the National Adult Reading Test (NART) (Nelson & Willison, 1991) However, this test is now over 30 years old and the NART has not been co-normed with most recent versions of IQ tests, so its association with up to date measures of IQ is unknown (Wechsler, 2001, 2011).

In light of this, a number of newer English irregular word reading tests have been developed (Wechsler, 2001, 2011). The TOPF (Wechsler, 2011) is the most recent of these. It was developed and co-normed with the most widely used current measure of IQ, the Wechsler Adult Intelligence Scale IV (WAIS-IV) (Wechsler, Coalson, & Raiford, 2008) and is thus used in the current study.

Premorbid IQ in PLWD is hard to measure as it requires a test that meets two assumptions: (i) high association with current IQ (and other measures of neurocognitive function) in individuals without dementia, and (ii) robustness to the effects of dementia (Crawford et al., 1989). This brief chapter will examine evidence for the validity of the TOPF in relation to each of these assumptions:

**Assumption 1: TOPF and ACE-III scores will correlate in an OA group.**

To some extent, there is good evidence that the TOPF meets the first assumption above as it has a high correlation with a gold standard measure of IQ (WAIS-IV) (Wechsler et al., 2008) in non-cognitively impaired individuals (Wechsler, 2011). Indeed, this was the main reason for its use in the current study. However, the cognitive reserve hypothesis (Stern, 2012), and the use of the TOPF to assess neurocognitive deterioration in dementia (Salmon & Bondi, 2009) is based on the assumption that performance on TOPF is associated with performance on a wider range of neurocognitive measures than just IQ (Salmon & Bondi, 2009) and this includes dementia screening tools (Stern, 2012) Such associations have not been investigated for the TOPF, but associations between performance on other measures of premorbid IQ function (NART) and dementia screening tools have been found (Alves, Simões, Martins, Freitas, & Santana, 2013). Consequently, the first aim of this study is to examine whether the TOPF validly taps a broader range of neurocognitive measures than simply IQ by examining the association of TOPF scores with performance on a widely used dementia screening tool, the ACE-III (Hsieh, Schubert, et al., 2013) in people without dementia.

**Assumption 2: TOPF performance will be robust to dementia.**

The TOPF's Robustness to dementia is less well evidenced than its association with IQ. To the author's knowledge, the only published study to evaluate the impact of dementia on the TOPF is one detailed in the TOPF manual (Wechsler, 2011). In this study, performance on the TOPF was compared to performance on the WAIS – IV (Wechsler et al., 2008) and the Wechsler Memory Scale IV (WMS-IV) (Wechsler, 2009) in a sample of individuals with mild Alzheimer's Disease. In line with expectations TOPF *predicted* WAIS-IV IQ scores and, most markedly, TOPF predicted WMS-IV memory scores were higher than the *obtained* WMS and WAIS-IV scores. The conclusion drawn from these findings was, that, in comparison to the WAIS-IV and WMS-IV the TOPF is relatively robust to the effects of dementia.

While this study is useful, it only demonstrates that the TOPF is robust to Alzheimer's disease relative to the WAIS-IV and WMS-IV. It could be that despite being robust to Alzheimer's disease *relative* to these measures, the TOPF is still significantly affected by having dementia (just not as much as the WAIS and WMS-IV). Furthermore, this study only included individuals with Alzheimer's Disease and it may be that results do not apply in a sample of PLWD with varying subtypes and more heterogeneous neurocognitive impairments (Salmon & Bondi, 2009).

Consequently, the second aim of the current chapter is to further assess the robustness of the TOPF in a heterogeneous sample of PLWD using methodologies that have been used to validate other premorbid IQ tests (Cockburn, Keene, Hope, & Smith, 2000; Crawford, Deary, Starr, & Whalley, 2001; Dykiert & Deary, 2013). Specifically, the chapter will investigate whether performance on the TOPF is associated with having a dementia diagnosis, with

the implication that, if replicated in longitudinal work, such an association could indicate an impact of dementia on TOPF performance (Cockburn et al., 2000).

In summary, this chapter has two novel aims and associated hypotheses:

1. To examine the association of ACE-III with TOPF performance in OA and thus the potential validity of TOPF as a premorbid estimate of neurocognitive functions beyond IQ in line with the cognitive reserve hypothesis. It is predicted that there will be a significant association of TOPF with ACE-III performance in OA independent of age, gender and years of education.
2. To evaluate the association of having a diagnosis of dementia with TOPF scores, providing cross sectional evidence as to the robustness of TOPF performance to brain pathologies underlying dementia. The hypothesis that having a diagnosis of dementia is significantly associated with poorer TOPF performance independently of gender, age or years of education will be tested

## **Method**

### **Design.**

A cross sectional correlational design was employed to examine the association of the TOPF with the ACE-III in an OA group and association with dementia diagnosis in the whole sample.

### **Participants.**

As will be apparent from the datasets table (Table 1 above), the participants reported on in this chapter overlapped with the participants reported on in subsequent chapters. The eligibility criteria, recruitment and data collection procedures for the OA and PLWD groups are the same across chapters.

Consequently, these are discussed in detail below and referred back to in subsequent chapters.

***Eligibility criteria.***

All participants were fluent in English, had no self-reported literacy issues and had capacity to consent. Exclusion criteria included a DSM-IV (APA, 1994) Axis 1 diagnosis of bipolar disorder or schizophrenia, diagnosed intellectual disability, and significant uncorrected sensory deficits. This sample was the same as the sample used in the following chapters relating to pre-therapy skills. Thus, as past CBT experience may influence performance on pre-therapy skills measures, participants reporting current or previous experience of CBT were excluded. There were two groups; PLWD and OA (over 65s). These groups are detailed below. All participants from both PLWD and OA groups gave written informed consent to participate in the study. Ethical approval was given by NRES Committee London – City Road & Hampstead (REC Reference 14/LO/0554).

Required Sample size was calculated using G\*Power for the main regression analysis which would have the smallest number of participants in it (the analysis for aim 1- evaluating association of the ACE-III and TOPF in the OA group). The calculation was based on using a hierarchical multiple linear regression and was powered to detect an expected medium effect size ( $f^2 = 0.15$ ) for the  $R^2$  change when adding ACE-III as a predictor of TOPF over and above potential confounding variables of gender, age and years of education. Power (1-B) was set at 0.8 and alpha at 0.05. This calculation indicated that 55 OA would be needed for the analysis. The achieved sample sizes (77) for this and the other analysis, which included the PLWD group, were much larger than this because this larger sample size was necessitated for the analyses used in later chapters (5, 6 and 7).

### ***The PLWD group.***

The PLWD group consisted of 102 people living with mild dementia (last MMSE (Folstein et al., 1975) score  $\geq 24$  or equivalent on other cognitive screen (Law et al., 2013)) who were clients at four memory clinics in north London. Dementia (including subtype) was diagnosed according to NICE (2018) recommended consensus criteria (Emre et al., 2007; McKhann et al., 2011; Neary et al., 1998; Román et al., 1993) by a psychiatrist led, multi-disciplinary memory clinic. All clients had a cognitive assessment, with extent of assessment driven by client need as per NICE and British Psychological Society guidelines (Guss et al., 2014). Specifically, results of cognitive screening tests (ACE-III, Addenbrooke's Cognitive Examination Revised (ACE-R), MMSE or, Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) from an initial diagnostic interview were discussed in the multidisciplinary team with interpretation guided by experienced clinical psychologists specializing in neuropsychology. Where diagnosis was unclear, a more extensive neuropsychological assessment was conducted by a clinical psychologist. In determining client functioning, client and informant report were used. Where informant report was unavailable and functional status was unclear, occupational therapy assessment was used to clarify. Diagnostic subtypes included Alzheimer's disease (McKhann et al., 2011) (n=65), vascular dementia (Román et al., 1993) (n=8), mixed dementia (n=16), (McKhann et al., 2011) dementia in Parkinson's disease (Emre et al., 2007) (n=2) and frontotemporal dementia (Neary et al., 1998) (n=1) When subtype criteria were not met, but criteria for dementia were, a diagnosis of dementia not otherwise specified was made according to ICD-10 criteria (WHO, 1993) (n=10). Intermediate diagnoses such as possible Alzheimer's disease were not included.

### ***The OA group.***

The OA group constituted a convenience sample of 77 healthy volunteers over the age of 65 without a diagnosis of dementia (determined through self-report) and not reporting subjective memory problems. They were recruited by advertisement distributed to: members of 'The University of the Third Age', which was developed for retired and semi-retired people who wish to learn new skills; age UK, a charity for OA; the Join Dementia Research database (Join Dementia Research, 2016) and London based community groups.

### **Recruitment procedures.**

Participants first gave consent to be contacted. This was either by responding to an advert (OA group), by agreeing to be on a research register, or being contacted by an involved clinician (PLWD group). After participants had consented to be contacted they were screened for eligibility by phone or e mail or in person. Those who met the inclusion criteria were invited to meet the researchers in person either at UCL, in clinics or in the participant's home. Information was sent to participants 48 hours prior to meeting with them to give them time to understand the study and was discussed again when meeting with them. Informed consent was then sought.

### **Data collection procedure.**

Where informed consent was given, demographic information was obtained verbally and for the PLWD group was verified via electronic healthcare records. Measures were then administered. Measures included two neurocognitive measures; the TOPF and ACE-III, which are described below and are used in analyses in this chapter. They also included four 'CBT pre-therapy skills measures': (i) the Reed Clements task (Reed & Clements, 1989); (ii) the BTFQ- Dementia (BTFQ-D, see chapter 5); (iii) the cognitive mediation in

dementia questionnaire (CM-DEM, see chapter 6); and (iv) the Emotion Recognition-40 (ER-40) (Kohler, Bilker, Hagendoorn, Gur, & Gur, 2000), as well as the HADS (described in chapter 3). These measures are described in detail where they are used in the analyses in later chapters. Three other measures (the Mindfulness Breath Attention Task (Frewen, Lundberg, MacKinley, & Wrath, 2011), a measure of mindfulness; the Four Mountains Task (Bird et al., 2010), a measure of allocentric spatial memory; and the Trail Making Test (Tombaugh, 2004), a measure of cognitive flexibility and switching, were given to subsets of participants. These measures were used in the aforementioned three DClinPsys supervised by the author of the current work, and are not discussed further in this thesis.

Measures were presented in two blocks 1) neurocognitive measures and 2) Pre-therapy skills, mindfulness and mood measures. The order of blocks was counterbalanced and the order of presentation of measures within blocks was randomised to avoid order effects. Researchers administering measures included the author of this thesis, the three DClinPsys, an MSc student, and three psychology graduates. All researchers were trained in administration and scoring of the TOPF and ACE-III and supervised by the author of this thesis who is a clinical psychologist with a postgraduate diploma in clinical neuropsychology.

### **Measures.**

#### ***TOPF.***

The TOPF (Wechsler, 2011) involves reading up to 70 irregular English words. The raw score (total number correct) can be converted into two estimates of premorbid IQ. The unadjusted premorbid IQ is based on published tables developed through regression with TOPF alone as a predictor of IQ. The adjusted premorbid IQ is obtained through entering TOPF score along with age, gender

and years of education into a regression equation. In support of construct validity, both unadjusted and adjusted TOPF premorbid IQ show strong associations with current measures of IQ and are reliable over time (Wechsler, 2011).

### ***ACE-III.***

The ACE-III is a validated 'pen and paper' neurocognitive screening tool for dementia (Hsieh, Schubert, et al., 2013) covering five cognitive domains including memory, language, orientation and attention, verbal fluency, and visuospatial abilities, with a maximum score of 100. The recommended cut-off score for screening neurocognitive impairments related to dementia is 87/88 (Hsieh, Schubert, et al., 2013).

### **Statistical analyses.**

In examining between group demographic differences, all continuous variables were assessed for parametric assumptions. Where these were met, t - tests were conducted, where not; Mann Whitney U tests were used. For categorical variables, Chi Squared or Fisher's exact test were used, depending on minimum cell counts (Field, 2009). Where necessary to quantify uncertainty, 95% confidence intervals (CIs) were calculated.

The association of ACE-III with TOPF in the OA group (aim 1) and the association of dementia diagnostic status with TOPF performance (aim 2) were both assessed using hierarchical multiple linear regression to statistically control for potential confounding variables of age, gender, and years of education, all of which can influence TOPF performance (Wechsler, 2011). For both aims, these potential confounders were entered in a first block with variables of interest added in a second block. For all regression models, outliers with undue influence on coefficients were investigated, and where necessary, removed. Where assumptions of regression (Field, 2009) were not met, bootstrapped bias-

corrected accelerated CIs were generated to increase model robustness (Field, 2009).

The unadjusted TOPF premorbid IQ was used in all analyses due to potential multicollinearity (Field, 2009) in regression equations containing the adjusted premorbid score. Multicollinearity was expected as the adjusted premorbid IQ score incorporates gender, age and years of education (all of which were included in the main analysis regression equations as potential confounders). Sensitivity analyses were performed with the adjusted score to ensure that its use did not alter results.

Data were analysed using R (R Core Team, 2013) specifically the QuantPsych (Fletcher, 2012) and BaylorEdPsych (Beaujean, 2012) packages were used.

## **Results**

### **Participant flow.**

345 people were initially approached to take part in the study. Of the 285 potentially eligible, 179 participants (102 PLWD, 77 OA) took part. In the OA and PLWD groups those potentially eligible did not differ from participants in gender ( $X^2 = 0.006$  and  $0.002$  respectively,  $p > 0.05$ ) or age ( $t = 0.77$ , and  $-0.06$  respectively,  $p > 0.05$ ). Those in the PLWD sample did not differ in diagnostic subtype (whether they had Alzheimer's disease or another dementia type) either ( $X^2 = 0.2$ ,  $p > 0.05$ ). Figure 1 shows a flow diagram detailing flow of participants through the study and reasons for exclusion.

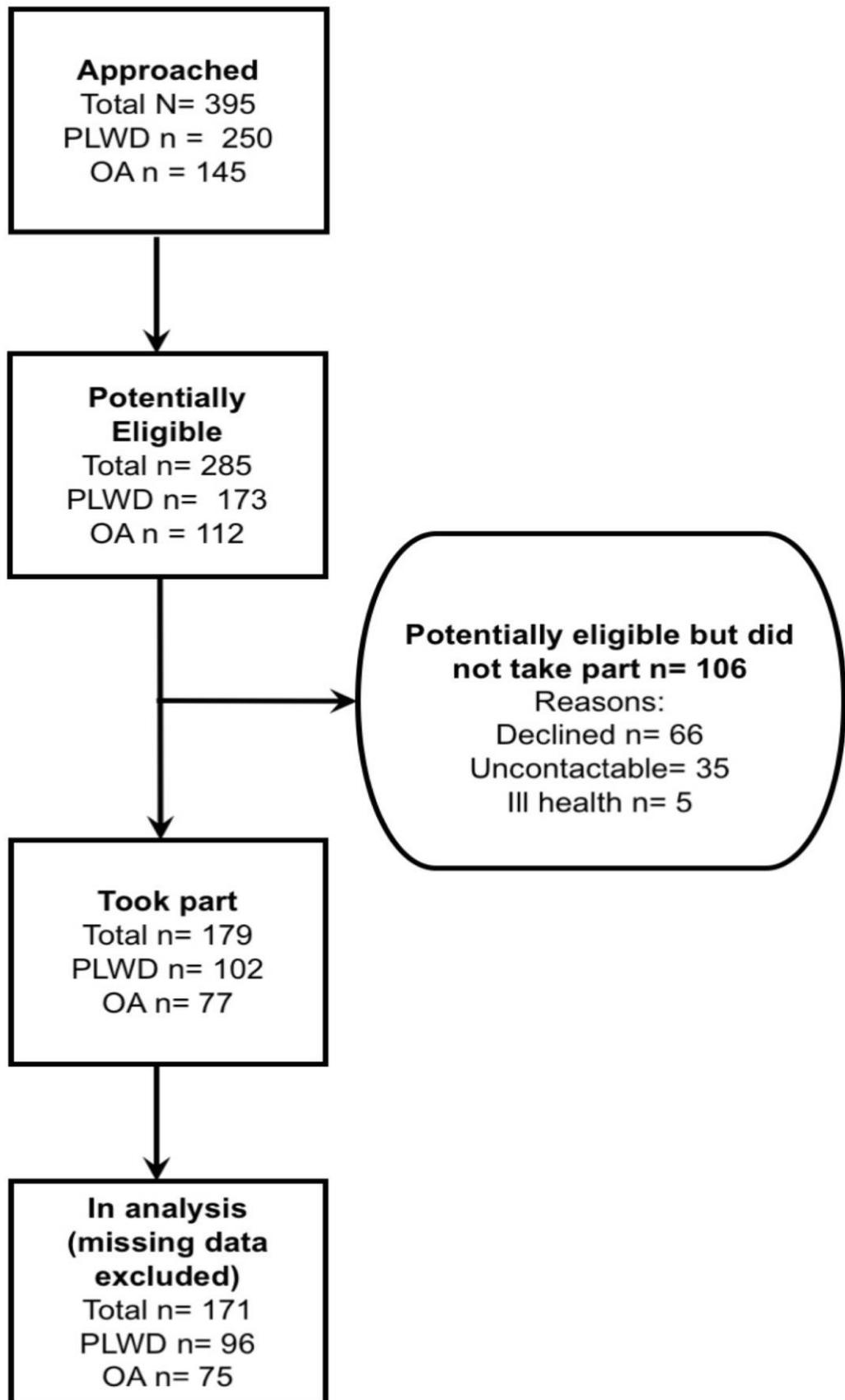


Figure 2. *Flow of participants through the study*

### **Demographic and clinical characteristics.**

Table 5 shows summary statistics for demographics as well as ACE-III (which in the PLWD group indicates level of cognitive impairment) and TOPF scores for both groups. The PLWD group were significantly older, had significantly fewer years' education and lower ACE-III and TOPF scores than the OA group. Gender and ethnicity did not significantly differ between groups.

Table 5. Demographic variables and TOPF and ACE-III scores for PLWD and OA groups

Variable	PLWD (n= 102)		OA group (n=77)		Significant contrast <sup>†</sup>
	Median (min-max)	% (n)	Median (min-max)	% (n)	
Age	81 (58-97)		72 (65-92)		PLWD > OA
Sex (M)		43 (44)		36 (28)	N/S
Ethnicity (White)		90 (92)		100 (77)	N/S
Education (years)	12 (5-25)		16 (7-25)		OA > PLWD
Cognitive impairment (ACE-III score)	74 (43-98)		95 (67-100)		OA > PLWD
TOPF	104.5(56-120)		119 (85-140)		OA > PLWD

Note: <sup>†</sup>Significant at p < .05, adjusted for Benjamini-Hochberg (B-H) (Benjamini & Hochberg, 1995) false discovery rate; for continuous variables, medians and ranges reported due to non-normally distributed data; N/S, no significant contrasts (Mann Whitney U test was used to compare continuous variables and  $\chi^2$  to compare categorical variables across groups).

### **Analysis of aims.**

Analyses were hierarchical multiple linear regressions. Standardized betas ( $\beta$ s) for all significant (at  $p < .05$ ) predictors among the variables of interest and potential confounders (gender, age, education) after adjustment for all other predictors are reported. Bootstrap bias corrected accelerated 95% CIs were used to quantify uncertainty. Eight participants had missing data on one or more measures. Data were missing completely at random (Little's MCAR test  $p < .05$ ) and  $< 5\%$ , so were removed list-wise as recommended by Graham (2009). Thus, the final analysis for aim 1 included 75 OA and for aim 2 there were 171 participants (96 PLWD and 75 OA).

#### ***Aim 1 association between TOPF and ACE-III scores in the OA group***

In the OA group, higher TOPF premorbid IQ scores independently predicted higher ACE-III scores ( $\beta=0.65$ , CI 0.38 to 0.81). This was the case even when controlling for gender, age and education, with older age ( $\beta = -0.24$ , CI -0.48 to -0.05) the only other significant independent predictor of higher ACE-III scores. Consequently, the results suggest that in those without dementia, ACE-III scores are highly and independently associated with premorbid IQ scores: This provides support for hypothesis 1, that performance on the TOPF is associated with performance on the ACE-III in people without dementia and thus could provide an indicator of premorbid functioning on the ACE-III.

#### ***Aim 2 robustness of the TOPF to dementia.***

In an analysis conducted on the whole sample, having a dementia diagnosis was a significant independent predictor of poorer TOPF performance ( $\beta=-0.45$ , CI -0.56 to -0.33). This was the case even when controlling for the significant prediction of TOPF performance by years of education ( $\beta=0.38$ , CI 0.25 to 0.48) and age ( $\beta=0.17$ , CI 0.06 to 0.29). The  $\beta$  coefficient of -0.45

indicates that having a dementia diagnosis is associated with a 0.45 standard deviation reduction in TOPF score. Consequently, results suggest that TOPF performance is associated with dementia and that potentially the measure may not be robust to dementia.

To further investigate the relationship between current cognitive functioning and TOPF performance in PLWD, the impact of degree of cognitive impairment (ACE-III score) and dementia subtype on TOPF performance were assessed in the PLWD sample. Lower ACE-III score was predictive of lower TOPF score ( $\beta=0.49$ , CI 0.31 -0.66) even when controlling for the fact that lower age ( $\beta=0.19$ , CI 0.04 -0.33) and fewer years of education ( $\beta=0.34$ , CI 0.18 -0.48) were also significant predictors of lower TOPF scores. Type of dementia (Alzheimer's disease vs. any other dementia subtype) was not associated with TOPF performance. Consequently, results indicate that TOPF performance in PLWD may be lower in those with increased cognitive impairment (as measured by ACE-III scores) independent of any differences in age and education, but there is no indication that TOPF performance is affected by subtype of dementia.

It is also of note that the  $\beta$  coefficient (0.49) for association of ACE-III and TOPF in PLWD is smaller than that reported in the OA group in aim 1 (0.64) This smaller association between TOPF and current functioning in the PLWD than in the OA group could indicate that, while it is not robust to dementia, TOPF is less affected by dementia than ACE-III.

For both aims 1 and 2, planned sensitivity analyses were run with the adjusted TOPF score used instead of the unadjusted TOPF score. All results above were unchanged.

## **Discussion**

This brief chapter examined the validity of the TOPF in relation to two criteria. The results suggested that while the TOPF may be a useful proxy for screening performance (the ACE-III) in those without dementia, performance may not be robust to the neuropathological impairments inherent in dementia.

### **The TOPF is associated with ACE-III performance.**

The high association of TOPF with ACE-III score in the OA group suggests that the first criterion for construct validity i.e. the ability to provide an indication of ACE-III scores in a sample without dementia, is met. This is in line with findings relating TOPF scores to IQ and memory measures. (Wechsler, 2011) It also suggests that where TOPF is robust to brain injury (as it may be to some types) (Wechsler, 2011) it could be a useful proxy for premorbid ACE-III scores.

### **The TOPF may not be robust to dementia.**

TOPF performance was associated with having dementia independently of potential confounding variables of age, years of education and gender. This could suggest that the TOPF does not meet the second validity criterion of robustness to dementia (and raises questions as to its utility as a measure of premorbid IQ and, consequently, cognitive reserve, in PLWD). There are, however, some caveats to this interpretation. The design used here does not allow inference of causality. In particular, reverse causality is possible (Stern, 2012), with the implication that the lower TOPF scores in the PLWD sample may, in fact, reflect a difference in underlying IQ between those with and without dementia.

Additionally, findings that the TOPF is associated with degree of cognitive impairment in dementia echo those for other irregular word reading tests and perhaps indicate that impact of dementia may be mitigated by less severe cognitive impairment (Cockburn et al., 2000) albeit within the limited range

afforded by the mild PLWD sample. This is important clinically, as it is in mild cases and in the clarification of potential false negatives that premorbid IQ measurement is particularly clinically useful (Crawford et al., 1989). Finally, the finding that TOPF scores in PLWD appeared less associated with the ACE-III than in the OA group perhaps suggests that the TOPF is, if not unaffected, less affected by dementia than the ACE-III (Wechsler, 2011).

### **Strengths and limitations.**

The study reported in this chapter had some strengths. In particular, generalizability of the results (of this chapter and the others that follow) to clinical practice is enhanced by the fact that, unlike many studies, having a carer was not an inclusion criterion. Generalisability was also augmented by the fact that the dementia diagnostic subtype breakdown in the PLWD group was very similar to the UK national picture (Prince et al., 2014). Additionally, the OA and PLWD groups were similar in gender and age to eligible non-participants. Participants living with dementia were also similar in diagnostic breakdown to eligible non-participants, suggesting no selection bias on these domains.

Aside from the aforementioned limitations on conclusions that can be drawn from cross sectional analyses, a key limitation to conclusions about TOPF robustness to dementia was the fact that the OA group were, on average, younger with more years of education than the PLWD group. Although differences were controlled for statistically and findings were unchanged when age and years of education were added to regression models, this limitation should be addressed by sample matching in future. Another limitation was that the sample in general was high functioning (as indicated by the relatively high TOPF median score in both groups, limiting the applicability of the findings to people with lower levels of intellectual functioning.

## **Conclusions.**

The TOPF may be useful in providing an indication of ACE-III performance in those without a cognitive impairment. However, the robustness of the TOPF to dementia was not supported, particularly where dementia it is more severe. While interpretation is somewhat equivocal, the TOPF will not be used as the main measure of cognitive reserve in this thesis and it is recommended it should be used with caution in clinical practice in PLWD.

**Chapter 5: Thought-Feeling Discrimination in PLWD - Development and  
Validation of a Clinical Tool**

## **Abstract**

*Objectives:* There is emerging evidence that CBT can be effective for treating anxiety and depression in PLWD. Discriminating between thoughts and feelings is a critical CBT pre-therapy skill. The aim of the present chapter was to modify and validate an existing measure of thought-feeling discrimination for use in PLWD. *Methods:* The Behaviour Thought Feeling questionnaire (BTFQ) was modified via expert and service user consultation for use in PLWD. 102 PLWD and 77 people aged over 65 years (OA group) who did not have measurable cognitive impairments completed the modified measure along with two measures of emotional recognition and reasoning (the ER-40 and the Reed Clements tasks). The factor structure of this measure was examined and the number of items reduced. *Results:* Factor analysis suggested a two-factor solution with thought and feeling items loading on separate factors. The behaviour items were dropped due to high cross-loading and ceiling effects, leaving a 14-item measure with two subscales. Thus, a new measure was created (named the BTFQ-D) which showed moderate convergent validity in the PLWD, but not the OA group. Both thought and feeling subscales showed adequate to good internal consistency. *Conclusions:* The BTFQ-D showed preliminary validity as a measure of thought-feeling discrimination in PLWD. It may have some clinical utility in measuring CBT pre-therapy skill levels. However, further validation is required before it is used to assess suitability for CBT.

## Introduction

As detailed in chapter 2, a CBT pre-therapy skill focussed on in this thesis is the skill of discriminating between behaviours, thoughts and feelings. To the author's knowledge, there are no measures of this construct that have been validated for use in PLWD, and a key aim of this thesis is to develop such measures for use as tools in CBT practice. Consequently, as discussed in chapter 2, a systematic review of the literature in the intellectual disabilities population (Stott, Charlesworth, et al., 2017) (where this skill has been measured) was conducted and two tools were identified (Oathamshaw & Haddock, 2006; Quakley, Reynolds, & Coker, 2004). The most widely used of these is the BTFQ (Oathamshaw & Haddock, 2006). Unlike the other identified tool (Thought Feeling Behaviour task (TFB; (Quakley et al., 2004)), the BTFQ was developed specifically for use as a practical tool within CBT sessions (Oathamshaw & Haddock, 2006; Padesky & Greenberger, 2012). Consequently, given the aim of developing useable tools for clinical practitioners, the BTFQ rather than the TFB was selected for development in the current thesis.

The BTFQ has 23 items. For each item, a participant is asked to identify whether a prompt is either a behaviour (e.g. 'having a bath'), a feeling (e.g. 'frightened') or a thought (e.g. 'this is hard'). Responses are scored as correct or incorrect and summed to give behaviour (range 0-8), feelings (range 0-8) and thought scores (range 0-7). A score of 6 or more on any sub-scale indicates above chance responding (at  $p < .05$ ).

The BTFQ cannot simply be adopted in an unmodified form for a dementia context without some investigation of its suitability. There are clear differences between an intellectual disability population and PLWD in relation to developmental trajectories as well as the nature and onset of cognitive

impairment. Such differences can affect measures in a number of ways. For example, concept(s) or dimension(s) may be missing, the meaning or appropriateness of concepts may differ, or PLWD may interpret items differently or use different styles of responding (Stewart, Thrasher, Goldberg, & Shea, 2012).

It is not always the case that a modified measure will need revalidation in the target population (Stewart et al., 2012). However, despite reasonably widespread use in CBT research (Oathamshaw & Haddock, 2006; Sams, Collins, & Reynolds, 2006; Vereenoghe, Gega, Reynolds, & Langdon, 2016) and use in clinical contexts (Dagnan, Chadwick, Stenfert Kroese, Dagnan, & Loumidis, 1997; Oathamshaw, Barrowcliff, & Haddock, 2012) there has been limited assessment of the validity or reliability of the BTFQ (Stott, Charlesworth, et al., 2017). In particular, factor structure of the BTFQ, which is critical for scoring and interpretation (Stott, Spector, et al., 2017), has not previously been evaluated in any population. This is especially pertinent as ambiguity as to scoring and interpretation is apparent in the literature. Specifically, the BTFQ has been interpreted as measuring three factors (the separate skills in discriminating thoughts, feelings and behaviours from one another) (Oathamshaw & Haddock, 2006) or one (discriminating thought, feelings and behaviours as a single skill) (Lickel, MacLean, Blakeley-Smith, & Hepburn, 2012).

Consequently, the first purpose of this chapter is to assess the suitability of, and, where necessary, modify the BTFQ for use in PLWD using an established framework for cross-population modification of measures as a guide. The second aim is to establish the factor structure and validate the modified BTFQ in PLWD. As discussed in chapter 2, the BTFQ may also be useful in working with OA, given potential variability in skill level even in this group

(Stanley et al., 2013). Thus, a sub aim is to examine psychometric properties of the modified measure in a non-cognitively impaired older population.

## **Methods**

### **Participants.**

The sample for validity analyses was from the same cohort as that in a previously published paper (Stott, Scior, et al., 2017) and the preceding chapter, where characteristics, recruitment procedures and inclusion and exclusion criteria are reported in detail. In brief, it consisted of two groups: (i) 102 people with mild dementia (PLWD group) and (ii) 77 people aged over 65 without dementia (OA group). The PLWD group was a consecutive referrals sample from a memory clinic and the OA group was recruited from community groups.

All participants from both groups gave written informed consent to participate in the study. Ethical approval was given by NRES Committee London – City Road & Hampstead (REC Reference 14/LO/0554).

While the sample was from the same cohort as reported in the preceding chapter, slightly different clinical characteristics were relevant to this chapter (i.e. HADS cases). Consequently, sample demographics are reported again in table 6 below.

### **Modification of the BTFQ.**

Prior to validity analyses, the version of the BTFQ used in the intellectual disabilities literature was assessed for suitability for use with PLWD, with modification conducted where indicated. Procedures followed recommendations of Stewart et al. (2012) Assessment for modification was conducted contemporaneously with assessment for modification of the cognitive mediation measure reported on in the next chapter.

### ***Expert consultation.***

Following initial review by the author and supervisors of this thesis (JS, GC, KS), the original BTFQ was circulated to five CBT experts (MFA, VH, NR, HC, DD) asking for comments on the measure in relation to suitability for measuring the construct of behaviour-thought-feeling discrimination and suitability for use with PLWD. Feedback from this consultation was noted down under headings relating to the different potential types of modifications identified within Stewart et al. (2012)'s framework i.e. modifications to item content, response options, dimensionality, format, or presentation. Information from expert review was supplemented with evidence from the author's recent review of the intellectual disabilities literature (Stott, Charlesworth, et al., 2017). The only changes arising from this were the necessity for the addition of an item to the thought scale (to ensure it was of the same length as others) and a suggestion of pre-testing the measure within a group of PLWD to assess for issues with presentation, item content and/or response options.

### ***Item generation.***

To generate a new thought item, a brief focus group was conducted with five PLWD. This was structured using a modified nominal groups methodology, a procedure to reduce the impact of specific group processes on decision making (Van de Ven & Delbecq, 1972) which the author of this thesis has recently adapted and used with PLWD in a consensus making context (Stott, Sweeney, Koschalka, O'Connor, & Mwale, 2017). In brief, this consisted of silent generation of ideas where participants were asked to write down potential thought items (or, if this was difficult, to describe any potential items to the facilitator who then wrote them down). The thoughts generated in this process were then discussed in the

whole group for relevance and acceptability, with two (*'I want to see my friend'*, *'I'm going to miss my train'*) taken forward.

#### ***Expert validity assessment of two new items.***

The two new thought items were subjected to 'expert validity assessment' to determine which should be added to the final measure (Oathamshaw & Haddock, 2006). This involved presentation of items to a group of 20 CBT professionals (trainee psychologists who had completed a year of CBT teaching and had been on year-long CBT placements) with the expectation that all items would be 100% correctly answered (because CBT pre-therapy skills would be expected to be universally present in this group). Only one of the two newly generated items (*I want to see my friend*) met this criterion and this was added to the thought scale.

#### ***Pre-test of measure with PLWD.***

The BTFQ used in pre-test consisted of the original version with the new thought item added. The purpose of pre-testing was to assess item content, instructions, response options and presentation format (Smith et al., 2005). Five individual interviews were conducted with PLWD covering a range of severity as assessed informally by the interviewer (JS) and who were not involved in the main validation study. A double interview technique was used as has been recommended for pre-test in PLWD (Smith et al., 2005). Interviews continued until the BTFQ and cognitive mediation measure (reported in the next chapter) had been finished or the person living with dementia wished to stop. Interviews lasted around 30 minutes and responses as well as any other issues arising during pre-test were noted down in field notes. Two modifications arose from pre-testing: (i) the need to supplement the verbal presentation of the measure with large written prompts ((20pt) black Arial font) to support memory, and (ii) the

need to develop standardised administration instructions about what to do if a participant failed to provide a response or asked for clarification. Instructions were developed based on other standardised measure instructions used in a dementia population (Smith et al., 2005). This version of the BTFQ was then administered to the OA and PLWD groups for the validity analyses.

### **Data collection and measures.**

Data collection procedures were the same as reported in the previous chapter. Convergent validity was assessed by examining inter-correlations with measures of two other constructs that, as detailed in chapter 2, have been identified as pre-therapy skills (Dagnan et al., 2000; Dagnan et al., 2009; Reed & Clements, 1989) and are empirically related with behaviour-thought-feeling discrimination in other populations (Oathamshaw & Haddock, 2006). Depression, anxiety and neurocognition were also measured for the purposes of sample characterisation.

### ***CBT pre-therapy skills measures.***

1. Emotion recognition – ER- 40 (Kohler et al., 2000). The ER-40 examines the ability to categorically identify facial expressions of emotion according to emotional valence, and has been validated in populations with mild Alzheimer’s disease (Kohler et al., 2005). It is a computer-based test consisting of 40 randomly presented colour photographs of people of varying age, gender and ethnicity with felt or evoked, sad, happy, angry, fearful or neutral facial expressions of varying intensity. An overall recognition index is calculated (0-40).
2. Event-emotion linkage (Reed Clements’ task) (Reed & Clements, 1989). Six simple first person scenarios are described and also presented in written format. Participants are asked to identify whether they would feel

happy or sad in that particular situation. A total score between 0-6 represents the number of scenarios answered correctly. This measure has been used previously in a dementia context and found to be acceptable and feasible (Harter, 2003).

### ***Anxiety and depression (The HADS).***

The HADS comprises 14 items each rated from 0 to 3, with higher scores indicating greater anxiety/depression. The anxiety and depression subscales each have seven items and a maximum score of 21. Caseness was established using a score >8 on either scale. As discussed in chapter 3 above, the original rather than dementia specific version was used for case ascertainment as the dementia version does not have a cut-off score.

### ***Cognition (ACE-III).***

A validated measure of neurocognitive functioning, developed for dementia screening purposes with a score range 0-100 (described in more detail in the preceding chapter).

### **Statistical analyses.**

#### ***Sample size.***

was calculated for the factor analysis element of the study since this requires larger sample sizes than the other reliability and validity analyses (Mokkink et al., 2010). A sample size estimate for factor analysis of binary response data (Pearson & Mundform, 2010) was used. Assumptions included, high communalities between items, an 80:20 ratio of correct: incorrect responses and, based on theory, that the number of factors extracted would be three at maximum. These criteria suggest that a minimum sample size of 175 is needed for a structure that is likely to be stable across samples (Pearson & Mundform,

2010). Thus, analysis was first performed in the entire sample before cross-checking the fit of factor structure in OA and PLWD subsamples.

### ***Data screening.***

Data were initially examined for floor or ceiling effects, with any item having more than 90% or less than 10% correct response in the PLWD sample removed prior to factor analysis (Pearson & Mundform, 2010). To determine the factor structure and reduce items, factor analysis was then conducted.

### ***Factor analysis.***

As data were binary, the tetrachoric correlation matrix was used because standard correlations may attenuate parameter estimates (Field, 2013). Tetrachoric correlation matrices are frequently non-positive definite (Revelle, 2017). This was addressed using a smoothing algorithm (Debelak & Tran, 2013). However, smoothing is a recent procedure and there is not yet an accumulation of evidence to support it (Revelle, 2017). Consequently, as a sensitivity analysis, the final factor structure was rerun on a conventional correlation matrix to check whether the structure was replicated across methodologies. Maximum likelihood factor analysis was used with an oblique rotation (Oblimin) given the theoretical supposition that any obtained factors will be highly inter-correlated (Fabrigar, Wegener, MacCallum, & Strahan, 1999). The number of factors to extract from the initial item set was based on Kaiser's criterion of eigenvalues greater than one, scree plot analysis and theoretical considerations (Fabrigar et al., 1999). Once factors had been extracted from the initial item set, item reduction was conducted whereby items which loaded in ways not expected by theory, items with high cross loadings (Field, 2013) (identified by item complexity factor (R. J. Hofmann, 1978; Pettersson & Turkheimer, 2010)) and/or low loadings on the primary factor were considered for removal (Field, 2013). Following item removal,

factor analysis was rerun on the final item set, with Kaiser criterion and scree plot analysis used to determine the final number of factors to extract (Fabrigar et al., 1999).

### ***Other analyses.***

Internal consistency of the final measure was assessed with Cronbach's alpha and convergent validity was assessed through correlations between BTFQ and the other measures detailed above. Spearman's rank correlations were used due to non-parametric data distribution. Significance of correlations was adjusted for type 1 error using B-H false discovery rate adjustment (Benjamini & Hochberg, 1995), a method that minimises type II error inflation (Verhoeven, Simonsen, & McIntyre, 2005). All data were analysed in the R environment using the Psych package (Revelle, 2017).

## **Results**

### **Sample characteristics.**

Table 6 shows clinical and demographic characteristics for all groups. PLWD had significantly lower ACE-III scores, were significantly older and had fewer years of education than the OA group. They also had significantly higher levels of anxiety and depression as indicated by HADS caseness.

Table 6. *Demographic and clinical characteristics of PLWD and OA groups*

Variable	PLWD (n=102)		OA (n=77)		Significant contrast <sup>†</sup>
	Median (min-max)	% (n)	Median (min-max)	% (n)	
Age	81 (58-97)		72 (65-92)		PLWD > OA
Sex (M)		43 (44)		36 (28)	N/S
Ethnicity (White)		90 (92)		100 (77)	N/S
Education (years)	12 (5-25)		16 (7-25)		OA > PLWD
Cognitive impairment (ACE-III score)	74 (43-98)		95 (67-100)		OA > PLWD
Anxiety and/or depression cases (HADS)		44 (44)		14 (11)	PLWD > OA

Note: <sup>†</sup>Significant at  $p < .05$ , adjusted for Benjamini-Hochberg (B-H) (Benjamini & Hochberg, 1995) false discovery rate; for continuous variables, medians and ranges reported due to non-normally distributed data; N/S, no significant contrasts (Mann Whitney U test was used to compare continuous variables and  $\chi^2$  to compare categorical variables across groups).

### **Factor analysis.**

*Assumptions.* There were no missing data on the BTFQ. Two behaviour items (6 and 19) were removed as over 90% of PLWD responded correctly to them, leaving 22 items in the factor analysis. Several other items, particularly items relating to identification of behaviours in the OA group, tended towards ceiling (Table 7 gives % correct for all items). The tetrachoric correlation matrix (appendix D) supported data factorability (Field, 2013), with nearly all correlations between items of at least moderate (Revelle, 2017) effect size (0.3 or above) and in the expected direction

Table 7. *BTFQ brief item statistics for all items*

Item no	Item	Correct answer	%(n) correct – PLWD	%(n) correct – OA	%(n) correct – all
6	Answering the phone	Behaviour	91(93)	99(76)	94(168)
19	Washing up	Behaviour	91(93)	100(77)	94(168)
7	Happy	Feeling	88(90)	95(73)	91(163)
9	Making a cup of tea	Behaviour	88(90)	97(75)	92(165)
1	Sad	Feeling	85(87)	96(74)	90(161)
15	Having a bath	Behaviour	85(87)	100(77)	92(165)
16	Frightened	Feeling	85(87)	97(75)	91(163)
18	Frustrated	Feeling	84(86)	96(74)	89(159)
24	Miserable	Feeling	83(85)	96(74)	89(159)
4	Working	Behaviour	80(82)	99(76)	88(158)
3	Angry	Feeling	78(80)	94(72)	85(152)
22	Upset	Feeling	78(80)	99(76)	87(156)
13	Gardening	Behaviour	77(79)	100(77)	87(156)
2	Talking to a friend	Behaviour	74(75)	96(74)	83(149)
21	Playing darts	Behaviour	71(72)	99(76)	83(149)
11	Worried	Feeling	66(67)	87(67)	75(134)
12	I don't know what to do for the best	Thought	64(65)	71(55)	67(120)
14	I hope this works out	Thought	56(56)	70(54)	62(111)

Item no	Item	Correct answer	%(n) correct – PLWD	%(n) correct – OA	%(n) correct – all
20	I'm a good person	Thought	50(51)	74(57)	60(107)
23	I'm looking forward to my holiday	Thought	43(44)	74(57)	56(100)
5	This is hard	Thought	42(43)	70(54)	54(97)
10	I'm good at things	Thought	42(43)	70(54)	54(97)
8	I want to see my friend	Thought	36(37)	64(49)	48(86)
17	I've achieved something	Thought	24(24)	57(44)	38(68)

### ***Factor extraction.***

Both Kaiser's criterion of eigenvalues greater than one (Field, 2013) and scree plot analysis suggested two factors should be extracted. All thought items loaded onto one unique factor; however, behaviour items and feeling items clustered together onto the same factor. Behaviour items showed other psychometric weaknesses: two items had already been removed due to ceiling effects and the four items with highest complexity scores were behaviour items (items 2, 4, 13 and 15). Therefore, and in order to maintain the theoretical coherence of the measure, items related to behaviour were removed from the analysis to create a measure of thought-feeling discrimination.

Factor analysis was rerun with just the thought and feelings items. Two items were removed as they showed high cross-loading on opposing factors in the PLWD sample (items 1 and 8). Kaiser criterion and scree plot analysis suggested that two factors should be extracted, with all items loading onto the expected factor in each of the full, OA, and PLWD samples. This left a two-subscale measure with seven items in each scale. Planned sensitivity analyses examining factor structure in conventional correlation matrices supported the above findings with factor structure replicated albeit with loadings attenuated as expected (Field, 2013).

### ***Model fit.***

The model showed good statistical properties for the full (Table 8) and PLWD samples (table 9) (accounting for 52% and 46% of the variance respectively). Loadings were above 0.4 with communalities 0.35 or above (indicating that an acceptable proportion of the variance in each item was accounted for by the factors). Complexity factors generally indicated low cross loadings. By contrast, the model in the OA group (table 10) showed lower

communalities and higher complexity scores. The thoughts, feelings and total scales showed adequate to good internal consistency across all samples (alphas = 0.72-0.81).

Table 8: *Final factor loadings of BTFQ items in the full sample*

Items	'Correct answer'	Feeling Loading	Thought loading	Communalities	Complexity score <sup>†</sup>
22	Feeling	<b>0.9</b>	0.09	0.87	1
3	Feeling	<b>0.84</b>	-0.18	0.64	1.1
24	Feeling	<b>0.83</b>	0.09	0.74	1
11	Feeling	<b>0.74</b>	-0.11	0.51	1
18	Feeling	<b>0.74</b>	-0.04	0.53	1
7	Feeling	<b>0.71</b>	0.13	0.58	1.1
16	Feeling	<b>0.6</b>	0.14	0.43	1.1
20	Thought	-0.1	<b>0.73</b>	0.5	1
10	Thought	-0.01	<b>0.72</b>	0.51	1
17	Thought	0.12	<b>0.63</b>	0.47	1.1
12	Thought	0.07	<b>0.62</b>	0.42	1
23	Thought	-0.03	<b>0.62</b>	0.38	1
14	Thought	0.03	<b>0.57</b>	0.34	1
5	Thought	0.16	<b>0.51</b>	0.34	1.2
Eigenvalues	-	4.33	1.84	-	-
% variance	-	33	28	-	-

Note: <sup>†</sup>Complexity score is an indicator of how much cross-loading there is on a particular item (R. J. Hofmann, 1978). There is no cut-off for complexity scores but they represent the number of latent variables needed to account for a manifest variable and at a maximum can equal the number of factors in a model (i.e. two in this case). Highest loadings for each item are indicated in bold.

Table 9. *Final factor loadings of BTFQ items in the PLWD group*

Items	'Correct answer'	Feeling Loading	Thought loading	Communalities	Complexity score <sup>†</sup>
22	Feeling	<b>0.84</b>	0.07	0.74	1
11	Feeling	<b>0.78</b>	-0.25	0.56	1.2
3	Feeling	<b>0.73</b>	0	0.53	1
18	Feeling	<b>0.7</b>	-0.04	0.48	1
24	Feeling	<b>0.68</b>	0.19	0.57	1.2
7	Feeling	<b>0.66</b>	0.28	0.62	1.4
16	Feeling	<b>0.55</b>	-0.04	0.28	1
12	Thought	0.06	<b>0.92</b>	0.89	1
23	Thought	0.19	<b>0.7</b>	0.44	1.2
14	Thought	0.07	<b>0.61</b>	0.4	1
10	Thought	0.02	<b>0.46</b>	0.22	1
5	Thought	0.23	<b>0.43</b>	0.4	1.5
17	Thought	0.2	<b>0.42</b>	0.27	1.4
20	Thought	0.06	<b>0.4</b>	0.15	1
Eigenvalues		4.46	1.8		
% variance		27	19	-	-

Note: <sup>†</sup>Complexity score is an indicator of how much cross-loading there is on a particular item (R. J. Hofmann, 1978). There is no cut-off for complexity scores but they represent the number of latent variables needed to account for a manifest variable and at a maximum can equal the number of factors in a model (i.e. two in this case). Highest loadings for each item are indicated in bold.

Table 10. *Final factor loadings of BTFQ items in the OA group*

Items	'Correct answer'	Feeling Loading	Thought loading	Communalities	Complexity score <sup>†</sup>
22	Feeling	<b>0.91</b>	0.09	0.82	1.0
16	Feeling	<b>0.89</b>	0.24	0.81	1.1
24	Feeling	<b>0.86</b>	-0.13	0.77	1.0
18	Feeling	<b>0.84</b>	-0.15	0.74	1.1
7	Feeling	<b>0.83</b>	0.15	0.69	1.1
3	Feeling	<b>0.79</b>	-0.49	0.92	1.7
11	Feeling	<b>0.53</b>	0.41	0.41	1.9
17	Thought	-0.12	<b>0.71</b>	0.54	1.1
5	Thought	0.07	<b>0.66</b>	0.44	1.0
20	Thought	0.03	<b>0.61</b>	0.38	1.0
10	Thought	0.1	<b>0.53</b>	0.28	1.1
23	Thought	0.23	<b>0.5</b>	0.28	1.4
12	Thought	0.02	<b>0.41</b>	0.17	1.0
14	Thought	-0.08	<b>0.36</b>	0.14	1.1
Eigenvalue		3.02	1.42		
% variance		34	19	-	-

Note: <sup>†</sup>Complexity score is an indicator of how much cross-loading there is on a particular item (R. J. Hofmann, 1978). There is no cut-off for complexity scores but they represent the number of latent variables needed to account for a manifest variable and at a maximum can equal the number of factors in a model (i.e. two in this case). Highest loadings for each item are indicated in bold.

As a final check on whether removing behaviour items was warranted, the Spearman's rank correlations between the final feeling factor and the combined behaviour-feeling factor found in the original analysis was assessed. These were significant across samples at the  $p < .001$  level and of very large effect size; (.9 in the full sample and PLWD subsample and .94 in the OA subsample). This provides further evidence that the behaviour items add little to the measure.

Given the consistent findings across samples and methodologies, this two factor structure with seven items tapping thoughts and seven tapping feelings was adopted in all further analyses.

### **The BTFQ- dementia (BTFQ-D) – validity and scoring.**

#### ***Measure name and scoring.***

As the factor analysis indicated that behaviour items were not of utility in this group, and 1 item from the original BTFQ feeling scale was removed, the resulting measure was renamed the Behaviour-Thoughts- Feelings Dementia Questionnaire (BTFQ-D) to distinguish from the original (Appendix E).

BTFQ-D Thoughts and Feelings scores were calculated by summing responses to the seven items in each respective scale. A total score (ranging from 0-14) was calculated to capture thought-feeling discrimination ability as a whole. New cut-off scores to signify above-chance responding at the  $p < .05$  level were calculated using binomial probabilities (the same methodology used in the original measure). These calculations indicated cut-off scores of 5 or above for subscales and 9 or above for the total score. These scores are relatively high, as due to the limited range of possible responses, scores that are not near the maximum may indicate chance or inconsistent responding.

#### ***Administering behaviour items.***

Despite their lack of use in scoring, it is recommended that behaviour items should be retained when using the measure for face validity purposes. They are quick to administer (the whole measure takes approx. five minutes, once explained) and response to the scale requires identification of an item as 'thought', a 'feeling' or a 'behaviour' which could create confusion if a behaviour was never a correct response.

***Descriptive statistics.***

Table 11 gives the medians and range for all subscale scores. The feeling subscale indicated a marked ceiling effect in the OA group and possible ceiling effect for the PLWD sample. The thought subscale did not appear to exhibit ceiling effects. The Reed Clements score was also at ceiling across all samples.

Table 11. Median scores for BTFQ-D and convergent validity measures in all samples

Measure	PLWD group Median (min-max)	OA group Median (min-max)	Full Sample Median (min-max)	Cut-off scores for above- chance responding
BTFQ-D - Total	9 (0-14)	12 (7-14)	11 (0-14)	≥9
BTFQ-D-Feelings	6 (0-7)	7 (0-7)	7 (0-7)	≥5
BTFQ-D-Thoughts	3 (0-7)	5 (0-7)	4 (0-7)	≥5
Reed Clements	6 (0-6)	6 (5-6)	6 (0-6)	-
ER-40	27 (17-34)	30 (19-37)	29 (17-37)	-

Note: Medians and ranges reported due to non-normally distributed data; n for PLWD = 102; OA= 77, full sample = 179.

### **Convergent validity.**

In the combined sample, all BTFQ-D scales were significantly correlated with the ER-40. In the subsamples, the only significant correlation with the ER-40 in the PLWD sample was with the BTFQ-D total scale (Table 12). The Reed-Clements measure was not included in the convergent validity analysis due to respondents scoring at ceiling.

Table 12. *Correlations between ER-40 and BTFQ-D scales (convergent validity).*

		ER-40	BTFQ-D Scale	
			Feeling	Thought
Full sample <sup>†</sup>	Feeling	<b>.22</b>	-	-
	Thought	<b>.26</b>	<b>.22</b>	-
	Total	<b>.32</b>	<b>.58</b>	<b>.89</b>
PLWD <sup>§</sup>	Feeling	.19	-	-
	Thought	.19	.16	-
	Total	<b>.27</b>	<b>.62</b>	<b>.84</b>
OA group <sup>*</sup>	Feeling	-.03	-	-
	Thought	-.003	.05	-
	Total	.04	<b>.38</b>	<b>.90</b>

Note: N= <sup>†</sup>163, <sup>§</sup>87, <sup>\*</sup>76 due to missing data on ER40.

Correlations in bold were significant at  $p < .001$ . All correlations were Spearman's rank due to assumptions of normality of distribution not being met. All p values were corrected for type 1 error using B-H false discovery rate.

### **BTFQ-D performance and ACE-III reading item score.**

Finally, although those with self-reported literacy difficulties were excluded from the study, there is a possibility that performance was affected by poor reading ability, given that the measure was presented in combined written and verbal form. An exploratory analysis was performed to assess group differences

in BTFQ-D score between those who scored 0 (incorrect) and those who scored 1 (correct) on the final language item of the ACE-III, which assesses reading ability. This analysis was only conducted within the PLWD group (20 participants scoring 0, 78 scoring 1) as there was no variability in ACE-III reading score in the OA group (all participants scored 1). A Mann Whitney U test was used as assumptions for parametric tests were not met, and no significant difference was found between groups ( $p < .05$ ). Thus, there was no evidence from this exploratory analysis that inability to do the ACE-III reading item affects BTFQ-D performance.

## **Discussion**

Discriminating between thoughts and feelings has been identified as a CBT pre-therapy skill. This study is the first to explore the use of a measure of this ability in OA and PLWD samples. Given the substantial changes that were required, this new measure was named the BTFQ-D. Findings are of clinical relevance as they could be used in assessment to inform how best to conduct CBT (Oathamshaw & Haddock, 2006), a promising intervention for anxiety and depression in PLWD and OA (Orgeta et al., 2014).

### **Lack of behaviour scale utility.**

Factor analysis supported a two-factor structure of 'thoughts' and 'feelings'. Behaviour items from the BTFQ showed a number of psychometric weaknesses, including ceiling effects, cross-loading and not clustering on a separate dimension to feeling items. Consequently, in the populations studied here, it is argued that behaviour items should not be scored although they might be maintained for face validity purposes. Future research in other populations should carefully consider whether it is necessary to include behaviour items in the BTFQ-D.

### **Structural and convergent validity.**

Structural validity was good in all samples especially in the full sample and in PLWD. In the full sample, where analyses had greater power and scores had more variance, there is good evidence for association of all BTFQ-D scales with another measure of CBT relevant abilities (ER-40). In PLWD, there is better evidence of convergent validity for the total score than the thoughts and feelings scores. In the OA group, there was limited evidence for convergent validity of any of the scales, though this may have been caused by reduced variance due to ceiling effects especially in the feeling scale where the median was the maximum possible score. The structural validity findings provide a basis for measure utility in all samples assessed here.

Convergent validity evidence is preliminary and derives from correlations with measures of related constructs rather than exactly the same construct. This means that the low correlations between some BTFQ-D scales and ER-40 scales across samples do not in and of themselves contraindicate use. Additionally, the use of a measure of a related rather than the same construct, while consistent with work in other populations (Dagnan et al., 2000; Oathamshaw & Haddock, 2006), means it is unclear as to whether the BTFQ-D specifically measures thought-feeling discrimination or rather measures a general 'CBT readiness' or some other construct common to it and emotion recognition. Future work should seek to specify the construct measured more precisely through developing other thought/feeling discrimination measures and measuring divergent as well as convergent validity (Mokkink et al., 2010). Perhaps, most importantly, future work should examine the relationships between scores on the BTFQ-D and CBT outcomes.

### **Research and clinical implications.**

The BTFQ-D could be used in research to examine thought-feeling discrimination in PLWD samples. This is supported by the findings of internal consistency of 0.7 or above which are sufficient for group comparison purposes of the sort reported in chapter 7 (Wells & Wollack, 2003). The potential ceiling effect and lack of convergent validity in the OA group will, however, limit the ability to interpret associations with other variables in this group.

It has been recommended in other populations that the BTFQ could be used as a clinical tool in the context of CBT (Oathamshaw et al., 2012; Oathamshaw & Haddock, 2006). It is suggested here that this measure should *not* be used to determine suitability for CBT. The psychometric findings are preliminary and need validation. In particular, excellent internal consistency (0.9 or above) is required for use of a measure as a single high-stakes assessment tool (Wells & Wollack, 2003). Additionally, and most importantly, there is no evidence currently of the association between this measure and CBT outcomes, a necessary criterion for making such a decision (Hebblethwaite, Jahoda, & Dagnan, 2011).

It would perhaps be more warranted to use cut-off scores to provide an indication of potential areas of difficulty and allow CBT to be tailored appropriately as discussed in chapter 2. Tailoring might involve the provision of thought/feeling discrimination training (e.g. as recommended in a widely used CBT manual) (Padesky & Greenberger, 2012), although current evidence in intellectual disabilities suggests this may not be effective (Vereenooghe et al., 2016). It might also involve use of an intervention that gives less emphasis on thought/feeling discrimination (for example, pleasant event scheduling, which has emerging evidence of utility in a dementia population (Livingston et al., 2017)).

### **Strengths and limitations.**

Recommended procedures for measure modification were used to develop a measure with relevance to the population in which the measure is to be applied (Mokkink et al., 2010). The theoretical coherence of the BTFQ-D was bolstered through consultation with experts (Smith et al., 2005), by basing the measure on a review of existing measures from other populations (Stott, Charlesworth, et al., 2017) (Smith et al., 2005) and a particular model of CBT pre-therapy skills elucidated in chapter 2. The relevance of the measure to PLWD was increased by engaging with them to develop initial item sets, formats and instructions using pre-testing and focus group methodologies adapted (including by the author of this volume) for this population (Dening, Jones, & Sampson, 2012; Smith et al., 2005; Stott, Charlesworth, et al., 2017).

The current study has limitations: sample size limitations meant that the initial factor analysis was conducted on a heterogeneous population with factor structure checked in individual subpopulations. While this approach suggested that factor structure was replicated within groups, future research should take a CFA approach in a large heterogeneous sample and examine measurement invariance formally (Stott, Orrell, et al., 2017; Stott, Spector, et al., 2017) as was done in chapter 3 of this volume. It would also be useful to replicate findings in a clinical sample selected on the basis of anxious and depressive symptomatology as it would be this group normally seen for CBT. However, generalisability of the results is enhanced by the fact that a significant proportion of participants met caseness for anxiety or depression, particularly in the PLWD group (44 %)

Assessment of convergent validity, while in line with the literature in other populations (Dagnan et al., 2009; Doherr et al., 2005; Oathamshaw & Haddock, 2006) was preliminary, and limited. Finally, it is a conceptual limitation that the

BTFQ-D measures the ability to comprehend thought/feeling distinctions in relation to abstract verbal statements rather than in relation to an individual's own mental states, as is expected in CBT. However, the ability to differentiate between verbal statements as to thoughts and feelings might function as a proxy for self-awareness of those states, as has been shown in other areas of emotional understanding (Shimokawa et al., 2001) and, even if not, the measure reported here is still useful in relation to CBT practice where such distinctions are frequently used in conversation (Greenberger & Padesky, 1995).

### **Conclusions.**

This chapter reports on the modification of the BTFQ for use in PLWD. Factor analysis suggests that the resulting measure is best interpreted as assessing only thought-feeling discrimination. Convergent validity evidence is presented but is somewhat limited in scope and further work should be done on this. Results support the use of this measure in the group analyses reported in chapter 7 of this thesis. The measure may have some utility in measuring pre-therapy CBT skills at a group level in PLWD. It may also be of clinical use in aiding decisions about how to adapt therapy in PLWD. It should not, however, be used to assess for suitability for therapy.

## **Chapter 6: Cognitive Mediation in PLWD - Development and Validation of a Clinical Tool**

## **Abstract**

*Objectives:* One of the key pre-therapy skills required for 'CBT readiness' is cognitive mediation - understanding the mediating role of cognitions between an antecedent event and its emotional consequences. The aim of the present study is to modify and validate an existing measure of cognitive mediation for use in PLWD. *Methods:* A measure of cognitive mediation was modified via expert and service user consultation for use in PLWD. The factor structure of this measure was examined and the measure reduced. The same PLWD and OA groups as reported in chapter 5 completed the modified measure along with three other measures of CBT pre-therapy skills; the ER-40, the Reed-Clements and the BTFQ-D. *Results.* A final measure of 10 items (named the CM-DEM) was subject to factor analysis yielding a single factor solution. The measure showed good psychometric properties in the PLWD group, including good model fit, high internal consistency and inter-rater reliability, and moderate convergent validity with related constructs. By contrast, although psychometric properties were adequate in the OA group, there was a lack of convergent validity. *Conclusions.* The CM-DEM showed preliminary validity as a measure of cognitive mediation in PLWD, but validity in OA was more mixed. The CM-DEM has some clinical utility in measuring CBT pre-therapy skills. However further validation is required before using to assess suitability for CBT.

## Introduction

As discussed in detail in chapter 2, one pre-therapy skill that has been identified as important for CBT (Dagnan et al., 1997) in the current thesis - as well as by authors in the adult mental health (Greenberger & Padesky, 1995), child (Quakley, Coker, Palmer, & Reynolds, 2003; Quakley et al., 2004), autism (Lickel et al., 2012) and intellectual disability (Oathamshaw & Haddock, 2006) literatures - is the ability to recognise the interceding role of a cognition between a triggering event and emotional response (cognitive mediation) (Dagnan et al., 2000; Dagnan et al., 1997; Doherr et al., 2005) .

To the author's knowledge, there are no measures of this construct validated for use in PLWD. Consequently, as described in chapter 2, the intellectual disabilities literature where this has been measured (Dagnan et al., 2000; Dagnan et al., 1997; Dagnan et al., 2009; Oathamshaw & Haddock, 2006), was reviewed to search for a measure that might be modified for use in PLWD. Using this approach, two measures of cognitive mediation were found, both developed by Dave Dagnan (Dagnan et al., 2000; Dagnan et al., 1997)

Of these, Dagnan et al. (1997)'s cognitive mediation measure was used in the current thesis because (i) an unpublished thesis suggested that Dagnan et al. (2000)'s cognitive mediation measure has floor effects in PLWD (Harter, 2003), (ii) Dagnan et al. (1997)'s measure has greater face validity for clinical use than Dagnan et al. (2000)'s measure, due to its free as opposed to forced choice response mode (Dagnan et al., 2009; Hebblethwaite et al., 2011) and, is thus recommended for CBT practice in clinical texts (Dagnan et al., 1997) .

Dagnan et al. (1997)'s cognitive mediation measure contains six items that describe a hypothetical event in the first person and an associated feeling of happiness or sadness. For each item, the participant is asked to identify a

thought congruent with the presented emotion. For example, one item is 'You see a group of friends but they do not say hello, and you feel sad'. What would you be thinking or saying to yourself?' An example of an accurate response would be 'They don't like me.' (Dagnan et al., 2009). Responses are coded on a 1-7 scale and thoughts deemed congruent with the valence of the presented emotion are scored as correct with the other six coding options detailing different types of 'error'. A binary correct/incorrect score is also obtained by classifying all error types as 'incorrect'. While the six prompt events were originally repeated once (Dagnan et al., 1997; Dagnan et al., 2009), a version with the prompts repeated twice, once associated with happy and once associated with sad emotions, has also been developed by Hebblethwaite et al. (2011). This latter version is used in the current thesis as correct responses to a different emotion presented with the same event clarifies that an individual is responding to the emotion presented and not the prompt event itself (Dagnan et al., 2009; Hebblethwaite et al., 2011).

As with the BTFQ, the cognitive mediation measure used here cannot simply be adopted from people with intellectual disabilities to PLWD in an unmodified form without first checking for suitability due to the potential impact of population differences on responses to items and score interpretation (Stewart et al., 2012). Consequently, the first aim of the current chapter is to assess the suitability of, and, if necessary, modify Hebblethwaite et al. (2011)'s version of Dagnan et al. (1997)'s measure of cognitive mediation for use in PLWD.

While not all modified measures will need revalidation in their target population (Stewart et al., 2012), it was notable in reviewing the intellectual disabilities literature, that despite their reasonably widespread use in CBT research and clinical texts (Dagnan et al., 1997; Oathamshaw et al., 2012), there has been very limited assessment of validity or reliability for any cognitive

mediation measure (Stott, Charlesworth, et al., 2017). In particular, some literature suggests that cognitive mediation measures might be divided into two dimensions (Dagnan et al., 2000): (i) the ability to perform the task when the emotion presented is *congruent* with the 'emotional valence' of the prompt event (e.g. 'you are sitting in the sunshine and you feel happy') and, (ii) the ability to do so when the emotion presented is *incongruent* with the emotional valence of the prompt event (e.g. 'you are sitting in the sunshine and you feel sad'). Within this two factor conceptualisation, the incongruent score is seen as a 'strong test' of cognitive mediation (Dagnan et al., 2000) since it allows a rater to dissociate whether an individual is responding to the presented prompt event or the presented emotion. However, this two factor structure is not universally adhered to and some authors interpret the measure as a single dimension (Dagnan et al., 1997; Dagnan et al., 2009). To the author's knowledge, no study has investigated which of these interpretations holds true, or, indeed assessed factor structure of any cognitive mediation measure at all.

Consequently, this chapter has two aims. The first is to assess suitability, and, if indicated, modify the described measure of cognitive mediation for use in PLWD. The second is to establish the factor structure and validate this measure for use with PLWD. As alluded to in chapters 1 and 2, this measure may also be useful in OA given likely variability in ability even in this group (Chand & Grossberg, 2013). Consequently, a sub-aim is to examine the psychometric properties of the measure in an OA group too.

## **Methods**

### **Participants.**

The sample is the same as that reported on in chapters 4 and 5. In brief, it consisted of two groups: (i) PLWD (n=102) and (ii) OA (n=77). The PLWD group

was a consecutive referrals sample from a memory clinic and the OA group was recruited from community groups. Demographics and clinical characteristics are reported in Table 6 in chapter 5.

### **Modification of the cognitive mediation measure.**

Prior to validity analyses, the Hebblethwaite et al. (2011) version of the cognitive mediation measure was assessed for suitability for use with PLWD, with modification conducted where indicated. Procedures followed recommendations of Stewart et al. (2012).

The sources of information were very similar to those reported for the BTFQ in the previous chapter. The process of expert consultation/literature review, prompt generation, expert validity and pre-testing were done for both measures contemporaneously. Only results relating to cognitive mediation are reported here.

### ***Expert consultation/literature review.***

As discussed in chapter 5, feedback from expert review was noted down under headings relating to the different types of modifications (Stewart et al., 2012) and was supplemented with information arising from the author's review of the intellectual disabilities literature (Stott, Charlesworth, et al., 2017). Expert review procedures were as reported in chapter 5 and four recommendations followed from this process:

1. To generate more prompt events; there are only six prompt events (each repeated twice) in the original measure and some of the original prompts may not be suitable for PLWD (e.g. 'you have been asked to go and see the centre manager'). This was done using a focus group as discussed below.

2. To test the 'emotional valence' (the immediate emotion evoked by the prompt event) of each prompt event, to allow the categorisation of emotions as 'congruent' or 'incongruent' to prompt events. To date prompt valence has been merely stated by the authors of the measure, rather than tested.
3. To measure the perceived emotional intensity of prompt events; intensely emotionally evocative events might lack acceptability when paired with an opposing emotion (for example, being asked to produce a 'happy thought' to an item such as 'your partner dies and you feel happy.....'). The test of points 2 and 3 is reported under 'Item validity checks' below.
4. To test acceptability and response format issues. In particular, it was suggested that the response options (free response) and presentation of the items (verbally only and with a Makaton face to represent happy or sad emotions) might not fit for a dementia population. The test of point 4 is reported under 'measure pre-test' below.

***Prompt event generation.***

To generate new prompt events with relevance to PLWD, a brief focus group structured using a modified nominal groups methodology (Stott, Sweeney, et al., 2017; Van de Ven & Delbecq, 1972) was conducted with five PLWD. Prompt events generated in the individual generation phase of the nominal group along with the original prompt events from the cognitive mediation measure were discussed in the group discussion phase of the nominal group for relevance and acceptability. Twenty new prompts in addition to the six original prompt events (shown in table 13) were generated and all 26 were taken forward for further validity checks

### ***Item validity checks.***

Two checks were conducted to select the optimal items for inclusion in the final measure, one to determine emotional valence and intensity of the prompt events and the other, an expert validity check of the items (Oathamshaw & Haddock, 2006).

#### ***Validity check 1: prompt event valence and intensity.***

A survey methodology was used to test which group of prompt events generated from the consultation had the desired characteristics for inclusion in the final measure (i.e. a set of prompts, some associated with positive and some with negative emotional valences of moderate intensity). The 26 prompt events were administered to a convenience sample of 55 over 65s. All participants identified as White British; 32 were female; the median (range) age was 70(65-85); and the mean (SD) years of education was 12 (5)). This sample was separate from the main sample for analysis and the survey was administered online using Qualtrics software (see appendix F for full survey). Order of presentation of events was randomized across participants. For each of the 26 prompt events, participants were presented with a list of emotions adapted from Izard's (Izard, Libero, Putnam, & Haynes, 1993) emotional taxonomy and asked which two emotions were most associated with the prompt. Responses were coded into three categories of 'positive emotional valence', 'negative emotional valence' or 'neutral'. Emotional intensity was measured in line with recommendations on verbal rating scales for intensity (Hjermstad et al., 2011) on six-point verbal rating scales with anchors ranging from 'slightly (I would hardly feel this at all)' – 'As strongly as I have ever felt this'. (Hjermstad et al., 2011) Survey results for each prompt given in Table 13. Prompts were selected on the basis of two criteria: (i) prompts where at least 60% of participants indicated a

positive or a negative emotional valence, and (ii) prompts with moderate intensity scores (i.e. scoring in the bottom 60% of intensity for the sample – a median score of 3 or less). Ten prompts met these criteria - prompt numbers: 4; 5; 7; 11; 15 (negative emotional valence), and prompt numbers: 2; 3; 17; 18; and 25 (positive emotional valence) - and were taken forward to the expert validity test.

Table 13. Results of prompt event emotional valence and intensity survey

Prompt number	Prompt event	Positive valence %	Neutral valence %	Negative valence %	N	Intensity Median
1	You are going on a trip to the seaside and the sun is out	94.3	2.8	2.8	106	4
2	<b>You are listening to the radio and hear a song that you used to dance to</b>	<b>82.7</b>	<b>9.2</b>	<b>8.2</b>	<b>96</b>	<b>3</b>
3	<b>You notice the flowers blooming in the park</b>	<b>100</b>	<b>0</b>	<b>0</b>	<b>96</b>	<b>3</b>
<u>4</u>	<u>You are eating a meal at home on your own</u>	<u>61</u>	<u>31.8</u>	<u>7.2</u>	<u>97</u>	<u>3</u>
<u>5</u>	<u>A friend calls to cancel a trip you had planned</u>	<u>2</u>	<u>14.3</u>	<u>83.7</u>	<u>98</u>	<u>3</u>
6	The leaves on the trees are going brown. It is the end of summer	42.7	25	32.3	96	3
<u>7</u>	<u>Your GP who has treated you for years tells you she is retiring</u>	<u>19.4</u>	<u>15.3</u>	<u>65.3</u>	<u>98</u>	<u>3</u>
8	You see a group of your friends and they do not say hello	3.1	38.5	58.3	96	4
9	You are about to go to an event where you don't know anyone	53.1	8.3	38.5	96	3

Prompt number	Prompt event	Positive valence	Neutral valence	Negative valence	N	Intensity
		%	%	%		Median
10	You want to go on a special trip but there is only one place and your friend is chosen to go instead	12.5	12.5	75	96	3.5
11	<i>You walk into a room where there are a group of your friends and as you walk in they start to laugh</i>	60	24	16	100	3
12	You are in bed one night and you hear a loud noise downstairs	8.2	8.2	83.7	98	4
13	You look at your diary and see that you have no plans for the next week	41.7	41.7	16.7	96	3.5
14	You are talking and laughing with a group of your friends	99	1	0	96	4
<u>15</u>	<u><i>Your daughter calls you to tell you her relationship has broken down*</i></u>	<u>8</u>	<u>16</u>	<u>76</u>	<u>100</u>	<u>3</u>
16	You look at your medication. You see the large number of tablets you have to take	11.8	29.4	58.8	102	3.5

Prompt number	Prompt event	Positive valence	Neutral valence	Negative valence	N	Intensity
		%	%	%		Median
<b>17</b>	<b><u>You are shopping and you see a friend you have not seen for ages</u></b>	<b><u>79.4</u></b>	<b><u>17.6</u></b>	<b><u>2.9</u></b>	<b><u>102</u></b>	<b><u>3</u></b>
<b>18</b>	<b><u>You are watching television when one of your favourite films comes on</u></b>	<b><u>95</u></b>	<b><u>5</u></b>	<b><u>0</u></b>	<b><u>100</u></b>	<b><u>3</u></b>
19	You are at a centre sitting with people your own age	57	18	25	100	3
20	You lose at a game of cards	25.3	27.3	47.5	99	2
21	You are given a job to do and you do it quicker than everyone else	84.4	15.6	0	96	4
22	You tell a neighbour to stop getting on your nerves	15.3	5.1	79.6	98	4
23	You win a prize in a competition you have entered	78	22	0	100	4
24	You are in a line and someone pushes in front of you and you tell them not to push in	10.8	9.8	79.4	102	3.5

Prompt number	Prompt event	Positive valence	Neutral valence	Negative valence	N	Intensity
		%	%	%		Median
<b>25</b>	<b><u>You are sitting in the park and the sun is out</u></b>	<b>96.9</b>	<b>3.1</b>	<b>0</b>	<b>98</b>	<b>3</b>
26	The cafe you used to go to all the time has been turned into a posh bar	25.6	15.3	59.1	97	3

Note: Prompt events in bold or italics were the **positively** or *negatively* valenced prompts respectively that were included in expert validity testing; prompts underlined were included in the measure that went through to pre-test and factor analysis;

\* This prompt was changed to begin 'A very close friend calls you....', following pre-test.

### ***Validity check 2 – expert validity.***

A preliminary cognitive mediation measure with the instructions, response format and presentation described by Hebblethwaite et al. (2011) was constructed. It consisted of 20 items (the 10 prompt events retained following the prompt event valence survey presented twice, once paired with the emotion 'happy' and once with the emotion 'sad'). This measure was presented to the same group of 20 CBT professionals as described in the previous chapter. Fourteen items - seven prompt events presented twice with opposing emotions (4, 5, 7, 15, 17, 18, and 25) - were 100% correctly answered and included in the final measure for pretesting.

### ***Measure pre-test.***

The 14 item measure was presented to five PLWD, who were not involved in the main validation study. Each prompt event was presented twice, once with the emotion 'happy' and once with the emotion 'sad'. On the basis of feedback in pre-testing, one prompt ('your daughter calls you to tell you that a relationship has broken down') was changed to 'a very close friend calls you.....' since it was deemed not relevant if you did not have a daughter. Three other issues were addressed: (i) Verbal presentation of the measure was supplemented with large written prompts to support memory. (ii) Makaton faces were kept rather than removed. (iii) Standardised administration instructions about what to do if a participant failed to provide a response or asked for clarification were developed based on other standardised measure instructions used in a dementia population (Smith et al., 2005). The seven prompt (14 item) measure with these adaptations made was taken forward for psychometric analyses reported on below.

### **Data collection and measures.**

Data collection procedures were the same as those reported in the preceding two chapters. As in the previous chapter, ACE-III and HADS were used to characterise the levels of neurocognitive performance and anxiety and depression caseness of the samples respectively. Convergent validity was assessed by examining inter-correlations with measures of three other constructs that have been identified as CBT pre-therapy skills (Dagnan et al., 2000; Dagnan et al., 2009; Reed & Clements, 1989) and are empirically related with cognitive mediation in other populations. These are detailed in chapter 5, but in brief, they were:

1. Emotion recognition – ER- 40 (Kohler et al., 2000) - a measure of facial emotion recognition, validated in people living with mild Alzheimer’s disease (score range 0-40).
2. Event-emotion linkage -Reed Clements’ task (Reed & Clements, 1989) (score range of 0-6).
3. Thought-feeling discrimination - The BTFQ-D and its subscales (score range 0-7 for each of the two subscales, with a score of 5 or more indicating above chance responding).

### **Statistical analyses.**

#### ***Sample size.***

As in chapter 5, sample size was calculated for the factor analysis element of this study since this requires larger sample sizes than the other analyses. Assumptions included: binary response data; high communalities between items; a variable to factor ratio of seven; an 80:20 ratio of correct: incorrect responses and a maximum of two factors extracted (based on theory). This suggests that a minimum sample of 120 is needed for a structure that is likely to be stable

(Pearson & Mundform, 2010) (this is different to the requirement for the BTFQ because the number of expected factors is two rather than three.) Consequently, factor analysis was performed first in the entire sample (n=179) before cross checking the fit of factor structure in subsamples with and without dementia as described below (Pearson & Mundform, 2010).

***Item screening.***

Items were initially examined for floor or ceiling effects, and any item with more than 90% or fewer than 10% correct responses in the PLWD sample was removed prior to factor analysis. For a subset of individuals (22% of the sample) two raters (a master's student (EC) and JS) independently coded items and Cohen's Kappa for inter-rater reliability of each item was calculated. Only items with good interrater reliability (Kappa >0.8) were included in the factor analysis.

***Factor analysis.***

Factor analysis procedures are as detailed in chapter 5. In brief, due to binary data, tetrachoric correlation matrices were analysed (Field, 2013). Non-positive definite matrices were addressed using a smoothing algorithm (Debelak & Tran, 2013). Maximum likelihood factor analysis with an oblique rotation (Oblimin) was used given the theoretical supposition that any obtained factors will be highly inter-correlated (Fabrigar et al., 1999). The number of factors to extract was based on Kaiser criterion, scree plot analysis and theoretical considerations (Fabrigar et al., 1999). Once factors had been extracted from the initial item set, item reduction was conducted, whereby items with high cross loadings (determined by item complexity factor) (Field, 2013) and/or low loadings on their primary factor were considered for removal (Field, 2013). Following item removal, factor analysis was rerun on the final item set, with Kaiser Criterion and scree plot analysis used to determine final number of factors to extract (Fabrigar et al.,

1999). Final factor structure was determined by fit with theory, pragmatic considerations as to the number of items per factor as well as statistical issues such as factor loadings.

### ***Other validity checks.***

Internal consistency of factors was assessed with Cronbach's Alpha (Zinbarg, Revelle, Yovel, & Li, 2005). Interrater reliability of the total score between two predetermined independent raters (EC and JS) was assessed on a subset of 54 participants (22% of the sample) using a mixed model intraclass correlation coefficient. The sample size for this was based on an a priori criterion set by the COSMIN international consensus framework on measure quality (Mokkink et al., 2010), which suggests that 50+ participants is an 'excellent' sample size for such analyses. Finally, convergent validity was assessed through correlations of the cognitive mediation measure and the other measures detailed above. Spearman's rank correlations were used due to non-parametric distribution of data. Multiple testing was accounted for using B-H false discovery rate adjustment (Verhoeven et al., 2005).

All data were analysed in the R environment using the Psych package (Revelle, 2017)

## **Results**

### **Sample characteristics.**

The sample was exactly the same as that use in chapter 5 and thus clinical and demographic characteristics for both groups are reported there. As reported in that chapter, PLWD had significantly lower ACE-III scores, were significantly older and had fewer years of education than the OA group and significantly more met HADS caseness for anxiety or depression.

*Assumptions:* There were small amounts of missing data (4%) on the cognitive mediation measure in the dementia group. Data were missing completely at random (Little's MCAR test  $p < .05$ ) and  $< 5\%$ , so were removed list-wise as recommended by Graham (2009). Thus, there were 178 participants (98 PLWD and 77 OA) in the factor analysis.

All items met pre-specified criteria for inter-rater reliability, floor and ceiling effects (see table 14) and were included in factor analysis. The tetrachoric correlation matrix (heatmap given as appendix G) supported data factorability (Field, 2013) with nearly all correlations between items of at least moderate (Revelle, 2017) effect size (.3 or above) in the expected direction.

Table 14. *Cognitive mediation brief item statistics for all items*

Item	Paired item <sup>†</sup>	Item: Prompt (presented emotion)	Congruent?	All <sup>§</sup> %(n)	PLWD* %(n)	OA <sup>¶</sup> %(n)	Interrater reliability <sup>‡</sup> Kappa
1	8	<b>You are sitting in the park and the sun is out (happy)</b>	<b>Congruent</b>	<b>71(124)</b>	<b>59(58)</b>	<b>86(66)</b>	<b>.94</b>
2	9	<b>You are eating a meal at home on your own (sad)</b>	<b>Congruent</b>	<b>61(107)</b>	<b>48(47)</b>	<b>77(59)</b>	<b>.96</b>
3	10	A friend calls to cancel a trip you had planned (happy)	Incongruent	58(102)	38(37)	83(64)	.92
4	11	<b>You are shopping and you see a friend you have not seen for ages (happy)</b>	<b>Congruent</b>	<b>69(120)</b>	<b>59(58)</b>	<b>82(62)</b>	<b>.89</b>
5	12	<b>Your GP who has treated you for years tells you she is retiring (happy)</b>	<b>Incongruent</b>	<b>56(98)</b>	<b>37(36)</b>	<b>81(62)</b>	<b>.96</b>
6	13	A very close friend calls to tell you their relationship has broken down (sad)	Congruent	54(95)	42(41)	70(54)	.81

Item	Paired item <sup>†</sup>	Item: Prompt (presented emotion)	Congruent?	All <sup>§</sup>	PLWD*	OA <sup>¶</sup>	Interrater reliability <sup>‡</sup>
				%(n)	%(n)	%(n)	Kappa
7	14	<b>You are watching television when one of your favourite films comes on (sad)</b>	<b>Incongruent</b>	<b>50(88)</b>	<b>38(37)</b>	<b>66(51)</b>	<b>.88</b>
8	1	<b>You are sitting in the park and the sun is out (sad)</b>	<b>Incongruent</b>	<b>47(82)</b>	<b>24(23)</b>	<b>77(59)</b>	<b>.88</b>
9	2	<b>You are eating a meal at home on your own (happy)</b>	<b>Incongruent</b>	<b>63(110)</b>	<b>45(44)</b>	<b>87(67)</b>	<b>.87</b>
10	3	A friend calls to cancel a trip you had planned (sad)	Congruent	54(95)	39(38)	73(56)	.96
11	4	<b>You are shopping and you see a friend you have not seen for ages (sad)</b>	<b>Incongruent</b>	<b>49(86)</b>	<b>29(28)</b>	<b>75(58)</b>	<b>.93</b>
12	5	<b>Your GP who has treated you for years tells you she is retiring (sad)</b>	<b>Congruent</b>	<b>56(98)</b>	<b>41(40)</b>	<b>75(58)</b>	<b>.81</b>
13	6	A very close friend calls to tell you their relationship has broken down (happy)	Incongruent	61(107)	44(43)	83(64)	.88

Item	Paired item <sup>†</sup>	Item: Prompt (presented emotion)	Congruent?	All <sup>§</sup>	PLWD*	OA <sup>¶</sup>	Interrater reliability <sup>‡</sup>
				%(n)	%(n)	%(n)	Kappa
<b>14</b>	<b>7</b>	<b>You are watching television when one of your favourite films comes on (happy)</b>	<b>Congruent</b>	<b>70(123)</b>	<b>62(61)</b>	<b>79(61)</b>	<b>.91</b>

Note: <sup>†</sup> Item with same prompt but different emotion; <sup>§</sup>n = 175; <sup>¶</sup>n=98; n=77; <sup>‡</sup>n=54; Kappa cut-offs were <.7 = poor, .7-.8 =good >.8 =excellent (Landis & Koch, 1977); items included in the final CM-DEM measure are shown in bold.

## **Factor analysis.**

### ***Factor extraction.***

Kaiser criterion and scree plot analysis of the tetrachoric matrix in the full sample indicated that one factor should be extracted. All items loaded significantly onto the single factor with large magnitude (range 0.5- 0.8). This factor structure was replicated in the PLWD and OA groups. In the PLWD sample, all loadings were large and significant (Range 0.4-0.9). However, in the OA group, items 3, 10 and 6 had very low loadings (0.17, 0.22 and 0.22 respectively) and were therefore removed. To maintain a consistent structure of presenting each prompt twice (once paired with the emotion 'happy' and the other with the emotion 'sad', item 13 (the item with the same prompt event as item 6 but presented with the emotion sad rather than happy)) was also removed before re-running analyses, leaving 10 items. Factor analysis was conducted again in all samples. The one-factor solution was confirmed in this final 10-item set.

### ***Model fit/loadings.***

The model showed good statistical properties for the full (Table 15) and PLWD samples (Table 16), accounting for 44% and 36% of the variance respectively, with loadings above 0.45 and communalities 0.20 or above. The model fit was reasonable (Revelle, 2017) although less good in the OA group (table 17) with the model accounting for 30% of the variance, and factor loadings at 0.3 or above with some communalities below 0.1, indicating that for items 2 and 8, little variance in the items was accounted for by the factor. Internal consistency was 'adequate' to 'good' in all samples ( $\alpha = 0.70-0.82$ ). Planned sensitivity analyses examining factor structure in conventional correlation matrices supported the above findings, with factor structure replicated albeit with loadings attenuated as expected (Field, 2013). Correlations between the overall

summed score for the reduced 10 item version and the summed score for the original full 14 item set were very high across the full, PLWD and OA groups (.96, .95 and .92 respectively). This suggests little was added to scoring by retaining the four removed items. Consequently, the 10-item single factor measure was adopted as the final version and used in all future analyses.

Table 15. *Final factor loadings of cognitive mediation items in the full sample*

Items	Rotated factor loadings	Communalities
11	0.83	0.68
7	0.81	0.66
5	0.75	0.56
12	0.69	0.48
9	0.67	0.45
1	0.61	0.37
14	0.61	0.37
8	0.57	0.33
2	0.52	0.27
4	0.52	0.27
Eigenvalues	4.45	-
%of variance accounted for	46	-
Internal consistency	.82	-
Inter-rater ICC (SEM)	.964(0.855)	-

Note: ICC=Intraclass correlation coefficient; SEM=Standard error of measurement; n=54 for ICC calculation.

Table 16. *Final factor loadings of cognitive mediation items in the PLWD group*

Items	Rotated factor loadings	Communalities
7	0.87	0.75
11	0.69	0.48
5	0.64	0.41
12	0.63	0.4
1	0.6	0.35
8	0.5	0.25
9	0.5	0.25
2	0.49	0.24
4	0.49	0.24
14	0.46	0.21
Eigenvalues	3.6	-
%of variance accounted for	36	-
Internal consistency	.80	-
Inter-rater ICC (SEM)	.954(0.932)	-

Note: ICC=Intraclass correlation coefficient; SEM=Standard error of measurement; n=32 for ICC calculation.

Table 17. *Final factor loadings of cognitive mediation items in the PLWD group*

Items	Rotated factor loadings	Communalities
14	1	0.995
11	0.69	0.475
4	0.56	0.317
9	0.5	0.251
12	0.5	0.251
1	0.45	0.204
7	0.42	0.178
5	0.41	0.169
8	0.31	0.093
2	0.3	0.091
Eigenvalues	2.91	-
%of variance accounted for	30	-
Internal consistency	.7	-
Inter-rater ICC (SEM)	.902(1.06)	-

Note: ICC=Intraclass correlation coefficient; SEM=Standard error of measurement; n=22 for ICC calculation.

### **The cognitive mediation - dementia Version (CM-DEM).**

#### ***Measure name and scoring.***

The final version of the revised measure was altered considerably from the original cognitive mediation measure (the new measure had different prompt content and a different number of items). To reflect this, it was named the Cognitive Mediation – Dementia Version (CM-DEM) and is given as appendix H.

The CM-DEM total score (ranging from 0-10) was calculated by summing correct responses. A preliminary cut-off of 4 was chosen, as this score was

obtained by fewer than 5% of the OA group. Scores below this level therefore could represent significantly greater difficulty (at the  $p < .05$  level) than is generally found in OA.

***Descriptive statistics/inter-rater reliability.***

Table 14 shows the number and percentage correct for each item in the CM-DEM. A Wilcoxon test showed that participants had significantly more correct responses on congruent items compared to incongruent (median congruent=4 vs median incongruent=3;  $V=7990.5$ ;  $p < .001$ ) in line with the hypothesis that such items would be less challenging. For all samples, the total scores were non-normally distributed and the median score (range) of the measure was 6(0-10) in the full sample, 4(0-10) in the PLWD sample and 8(3-10) in the OA group. Inter-rater reliability was high in all samples with intraclass correlation coefficients ranging from .90 (OA) .96 (full sample).

***Convergent validity.***

Table 18 details Spearman's rank correlations between the CM-DEM and the BTFQ-D (total score and subscales) and ER-40. As per the previous chapter the Reed Clements task was not included in this analysis due to a ceiling effect. Coefficients were all significant and mostly of moderate effect size (0.3 or above) when measured in the entire sample. In the dementia sample, findings were similar to the full sample. The OA group showed very little evidence of convergent validity with no significant correlations between the CM-DEM and other related measures.

Table 18. *Correlations between CM-DEM, BTFQ-D and ER40 scales (convergent validity)*

	Full sample <sup>†</sup>	PLWD <sup>§</sup>	OA <sup>*</sup>
BTFQ-D -Feelings	<b>.38</b>	<b>.32</b>	.06
BTFQ-D -Thoughts	<b>.42</b>	<b>.38</b>	.17
BTFQ-D -Total	<b>.47</b>	<b>.42</b>	.15
ER40 –score	<b>.35</b>	<b>.29</b>	.12

Note: <sup>†</sup>n=160, <sup>§</sup>n=84, <sup>\*</sup>n=76 (due to missing data on CM-DEM and ER-40);

correlations in bold were significant at  $p < .001$ ; all correlations were Spearman's rank due to assumptions of normality of distribution not being met; all P values were corrected for type 1 error using B-H method.

#### ***CM-DEM performance and ACE-III reading item score***

Finally, although those with self-reported literacy difficulties were excluded from the study, there is a possibility that performance was affected by poor reading ability, given that the measure was presented in combined written and verbal form. An exploratory analysis was performed to assess group differences in CM-DEM score between those who scored 0 (incorrect) and those who scored 1 (correct) on the final language item of the ACE-III, which assesses reading ability. This analysis was only conducted within the PLWD group (20 participants scoring 0, 78 scoring 1) as there was no variability in ACE-III reading score in the OA group (all participants scored 1). A Mann Whitney U test was used as assumptions for parametric tests were not met, and no significant difference was found between groups ( $p < .05$ ). Thus, there was no evidence from this exploratory analysis that reading affects CM-DEM performance.

## **Discussion**

As discussed in chapter 2 and above, cognitive mediation is an important CBT pre-therapy skill. This chapter is the first to report the modification and validation of a measure of this construct for use with PLWD (named CM-DEM).

### **A one factor structure.**

Factor analysis indicated that the measure had a one-factor structure in PLWD and OA groups suggesting that the splitting up of scores into congruent or incongruent cognitive mediation as has been suggested (Dagnan et al., 2000) is not warranted, at least for the CM-DEM. However, incongruent questions were answered incorrectly more frequently, perhaps in line with the idea that cognitive mediation is a skill that can be measured on a single dimension running from the easier 'weak test' of congruent cognitive mediation to a more difficult 'strong test' of incongruent cognitive mediation.

### **Validity and reliability.**

Structural validity was good in the full sample and in PLWD and adequate in the OA group. The measure also showed adequate to good internal consistency and good inter-rater reliability across samples.

The measure showed good convergent validity in PLWD indicated by correlations with measures of emotion recognition and thought-feeling identification. The significant correlations between BTFQ-D scales and the CM-DEM also provide further evidence for the convergent validity of the BTFQ-D (discussed in chapter 5) in PLWD, since, as CBT pre-therapy skills, thought-feeling discrimination and cognitive mediation are hypothetically related (Oathamshaw & Haddock, 2006).

By contrast, the measure showed poor convergent validity with measures of related constructs in the OA group. For the feelings subscale of the BTFQ-D

this could be explained by a ceiling effect in the OA group, with consequent lack of variability. For the thoughts subscale of the BTFQ-D and the ER-40, results might be partially explained by the relatively smaller sample size in the OA group leading to lack of power, but the small effect sizes of obtained correlations (0.12-0.15) do not point to large associations even if sample size were larger.

Lack of convergent validity evidence perhaps should not be used as a contraindication for use of the CM-DEM (or, indeed, for the BTFQ-D) as it derives from correlations with measures of related, rather than the same, construct. In this thesis, the generally positive structural validity findings will be taken as a basis for measure utility in all samples assessed here, but the need for further research on convergent validity in OA samples in particular is recognised.

As with the BTFQ-D, this use of measures of related rather than the same construct, while consistent with work in other populations (Dagnan et al., 2000; Oathamshaw & Haddock, 2006), means it is unclear as to whether the CM-DEM specifically measures cognitive mediation or some other construct common to it the BTFQ-D and emotion recognition.

### **Research and clinical implications.**

The current study suggests that CM-DEM can be reasonably interpreted as a single factor of 'cognitive mediation' for PLWD, and while evidence for convergent validity in OA samples is lacking, structural validity findings provide some evidence for use in this population too. The adequate to good internal consistency and high inter-rater reliability of the measure indicate that it could be used in research for group level analyses (e.g. between-group differences and within-group associations) of the sort described in chapter 7 (Wells & Wollack, 2003).

Like the BTFQ-D, the CM-DEM should not be used to determine suitability for CBT as there is no evidence yet that performance on it relates to CBT outcome. The preliminary normative cut-off scores on the measure could be used in clinical practice to provide an indication of when a person living with dementia might need more support than that offered to someone without dementia in terms of developing this skill. Such support might take the form of pre-therapy skills training (which has been shown to be effective in people with intellectual disabilities) (Bruce et al., 2010). A limitation to this approach is that the preliminary cut-off score is based on performance of an OA sample that is, on average, younger, with more years of education and higher average premorbid IQ scores than many PLWD samples. However, with further validation, the CM-DEM could perhaps be used as part of a battery of tests to help inform clinical decisions about which intervention within the CBT umbrella for a given client might be most appropriate (e.g. less cognitively demanding pleasant event scheduling vs potentially more demanding cognitive restructuring).

### **Strengths and limitations.**

The current study has many of the strengths detailed in the previous chapter. In addition, the use of a survey to determine typical valence of emotional response to prompt scenarios is, to the author's knowledge, a novel way of developing such prompts and may be of interest to others in the CBT literature. However, a limitation in the current work is the use of an OA sample rather than PLWD as a proxy group to do this survey.

The current study also has limitations. While tetrachoric correlation matrices were used appropriately for factor analysis and sample size was calculated a-priori, actual sample size needed may have been larger given that one assumption of the calculation was high communalities which, in practice,

were not obtained. As discussed in the previous chapter, further research using CFA and formal measurement invariance approaches would be useful.

Conceptually, the use of only happy and sad as emotional responses to the prompt events provides simplicity. However, this approach does limit the content validity of the measure and its applicability to cognitive mediation in the context of other emotional consequences of thoughts (e.g. fear). Finally, and critically, although there is some limited evidence that the construct of cognitive mediation may change over the course of CBT in people with intellectual disabilities (Hartley et al., 2015), the relationship of this measure to CBT outcome remains unknown for PLWD.

### **Conclusions.**

This chapter reports on the development of a measure of cognitive mediation for use in PLWD – the CM-DEM. It is proposed that the CM-DEM can be used in the group level analyses of the sort reported in the following chapter (e.g. comparison of performance between PLWD and those without). Clinically, this measure should not be used in clinical practice to determine suitability for therapy but might be used to indicate specific areas of difficulty that could be addressed in the course of CBT.

**Chapter 7: CBT Pre-Therapy Skills in PLWD, OA and YA - Investigating  
Group Differences and the Role of Neurocognition, Mood and Anxiety**

## **Abstract**

*Objectives:* The main aim of this chapter was to inform CBT practice and research through (i) understanding whether CBT pre-therapy skills represent separable constructs, (ii) comparing CBT pre-therapy skills in PLWD to non-cognitively impaired YA and OA controls, (iii) examining potential confounders and mediators, and (iv) exploring the associations of particular neurocognitive abilities with pre-therapy skill performance in PLWD. *Methods:* CFA in the entire sample (N=230) was used to assess the validity of measuring discrimination of thoughts (BTFQ-D-Thoughts), feelings (BTFQ-D-Feelings), and cognitive mediation (CM-DEM) as separate factors. Performance on each of these pre-therapy skills measures was compared between PLWD (n=102), OA (n=77) and YA (n=56) groups using Kruskal Wallance and Dunn Post Hoc tests. Mediators and confounders of differences in pre-therapy skill performance between OA and PLWD groups were assessed using structural equation modelling. Spearman's rank correlations were used to examine the relationship of pre-therapy skills with neurocognition (ACE-III subscales) in PLWD. *Main results:* The measurement of pre-therapy skills as separate constructs was supported. Pre-therapy skill performance followed the pattern YA > OA > PLWD, with effect sizes of differences bigger for discrimination of thoughts and cognitive mediation than discrimination of feelings. The difference between OA and PLWD was mediated by neurocognition for all skills. This was not the case for differences between OA and YA groups. In PLWD, language was associated with performance on all skills. *Conclusions:* PLWD may have a relative difficulty in CBT pre-therapy skills that require identification and use of thoughts. There is, however, substantial variability in this and mild dementia does not necessarily preclude CBT readiness. In PLWD, the role of neurocognition may be important and strategies

to adapt CBT for PLWD should take this into account. Age is also associated with CBT pre-therapy skill deficits, but this does not appear to be due to neurocognitive factors.

## **Introduction**

As detailed in chapters 1 and 2, there is evidence that CBT can be effective with PLWD. However, questions have been raised about the ability of PLWD to access the cognitive elements of CBT due to their neurocognitive difficulties (Spector et al., 2015; Stanley et al., 2013). It has also been argued that (to a lesser extent) current cohorts of OA may have deficits in their understanding of CBT, with consequent adjustments to CBT also proposed for this group (Mohlman, 2008, 2013).

As noted in chapter 2, the question of what the cognitive elements of CBT are – and the pre-therapy skills required to access them – is a complex one given the umbrella nature of the concept (Doherr et al., 2005; Roth & Pilling, 2008). The model adopted here is based on the core components of cognitive restructuring and mirrors the most widely used model in the intellectual disabilities literature (Oathamshaw & Haddock, 2006). It proposes that CBT pre-therapy skills include the ability to discriminate (i) thoughts and, (ii) feelings, (iii) the ability to link events to emotions, and (iv) and the ability to recognise the interceding role of a cognition between an event and its emotional consequence (cognitive mediation) (Greenberger & Padesky, 1995; Lickel et al., 2012; Oathamshaw & Haddock, 2006; Quakley et al., 2003; Quakley et al., 2004).

In chapters 5 and 6, there were inter-correlations in some samples between measures of (i), (ii) and (iv) above. This might indicate that, rather than representing separable entities, they could reflect a single underlying dimension of 'CBT readiness'. This single factor approach has been used to interpret performance on pre-therapy skills measures in the past (Lickel et al., 2012), but has not been empirically tested. Consequently, the first aim of the current chapter

is to test the structural validity of conceptualising pre-therapy skills as separable entities rather than one underlying dimension.

As discussed in chapters 1 and 2, understanding whether there is a deficit in CBT pre-therapy skill performance in PLWD (and potentially in OA too) is a clinically important question. Thus, if measured skills do represent separable constructs, it will also be useful to understand which are most affected in PLWD and OA groups. This could inform the adaptation of CBT interventions to emphasise skills that PLWD (and potentially OA too) generally find easier, and provide training on those skills that are more challenging (Dagnan et al., 2000; Joyce et al., 2006; Oathamshaw & Haddock, 2006). For example, understanding the link between events and emotions, but not cognitive mediation may suggest a person living with dementia is ready to engage with pleasant event scheduling (which may entail comprehension of links between events and feelings) but may need more support to engage in cognitive restructuring (which is likely to require comprehension of cognitive mediation)(Oathamshaw & Haddock, 2006).

Consequently, the second aim of this chapter is to assess how PLWD and OA samples perform on validated measures of CBT-pre-therapy skills both in comparison to each other and a YA control group. The hypothesis is that for all skills, the YA group will score highest, followed by the OA and then the PLWD group. Where differences are found, this study will examine whether some skills are more affected than others through exploration of effect sizes. The hypothesis here is that between group differences for identifying cognitive mediation (theoretically the most complex skill) will have a larger effect size than differences in discriminating thoughts, which, in turn will have a larger effect size than differences in discriminating emotions or linking feelings to events (since

discriminating thoughts directly taps more 'cognitive' elements of CBT – which it is suggested PLWD struggle with).

As discussed in chapter 2, if the anticipated group differences are found, the nature of these differences will be explored in terms of confounders and mediators, with a mediating variable differing from a confounder conceptually in that *is proposed* to represent a causal step between having dementia and poorer CBT pre-therapy skill performance (Preacher & Kelley, 2011). As discussed in chapter 2, neurocognition will be measured as a mediator, with anxiety and depression, age and cognitive reserve measured as potential confounders. Thus, the third aim of this chapter is to examine the hypothesis that neurocognition will mediate any relationship between having dementia and CBT pre-therapy skill performance independent of potential confounders.

While the preceding aims of this chapter relate to the impact of dementia as a whole on CBT pre-therapy skill performance, there is substantial heterogeneity within the diagnostic category of dementia, particularly in terms of degree and type of neurocognitive deficit (Salmon & Bondi, 2009). Consequently, the final aim of this chapter and of this thesis is to start to explore whether within group neurocognitive heterogeneity as well as within group differences in anxiety and depression might have implications for CBT pre-therapy skills. In particular, it is predicted that specific aspects of neurocognition, which are empirically related to CBT skills in other populations or routinely adapted for in CBT interventions for PLWD (language, executive function and memory) will be associated with pre-therapy skill performance in a group of PLWD.

In summary, the overarching aim of the study presented in this chapter is to compare CBT pre-therapy skills in PLWD to non-cognitively impaired YA and OA controls, examining potential confounders and mediators and exploring within

PLWD sample heterogeneity. The main hypotheses are that (i) A CBT-pre-therapy skills model of separable skills as opposed to a model of one underlying dimension of 'CBT readiness' will be supported by the data; (ii) PLWD will show lower scores on all pre-therapy skill measures than the OA group, who, in turn, will perform worse than the YA group; (iii) any difference between OA and PLWD groups in pre-therapy skill performance will be mediated by neurocognition and will be independent of potential confounders (anxiety, depression, cognitive reserve, and age); and (iv) variability in pre-therapy skill performance within PLWD will be associated with measures of executive functioning (measured here by verbal fluency), language and memory.

## **Methods**

### **Participants.**

The OA and PLWD groups are from the same source as reported in chapters 5 and 6 and recruitment and sample descriptions are reported there. In brief, they consisted of (i) a PLWD group (N=102), (ii) an OA group (N=77). For the purposes of the current chapter, 56 young adults (YA group) were also recruited. This YA group were a convenience sample of university students and other YA aged 18 to 25, not reporting any subjective cognitive problems who had registered their interest in participating in research with the university's psychology subject-pool. Like the OA and PLWD groups, they were fluent in English, had no self-reported literacy problems and had capacity to consent. Exclusion criteria included a DSM-IV (APA, 1994) Axis 1 diagnosis of bipolar disorder or schizophrenia, diagnosed intellectual disability, and significant uncorrected sensory deficits. As past CBT experience may influence performance on measures, participants reporting current or previous experience of CBT were excluded. All participants gave written informed consent to

participate in the study. Ethical approval was given by NRES Committee London – City Road & Hampstead (REC Reference 14/LO/0554). Demographics and clinical characteristics of the three groups are presented in Table 19 in results below.

### **Data collection and measures.**

Data collection procedures were the same as those reported in the preceding three chapters. Measures used were:

#### ***Pre-therapy skill measures.***

1. Event-emotion linkage - Reed Clements' task (Reed & Clements, 1989). This is described in more detail in chapter 5 but measures the ability to link events to emotions and has a score range of 0-6.
2. Thought/feeling discrimination - The BTFQ-D and its subscales is described in chapter 5. In brief, the score range is 0-7 for each of the two 'thought' and 'behaviour' subscales with a score of 5 or more indicating above chance responding.
3. Cognitive mediation - *The CM-DEM* as described in detail in the previous chapter. Scoring is from 1-10.

#### ***Current cognition.***

The ACE-III (Hsieh, Schubert, et al., 2013). The ACE-III is described in preceding chapters. However, for the purpose of this chapter, subscale scores (attention (range 0-18), memory (range 0-26), fluency (range 0-14), language (range 0-26), visuospatial functioning (range 0-16)) were used to tap different neurocognitive functions. There is some evidence for convergent and divergent validity of subscales derived from correlations with established neuropsychological measures of the respective constructs (Hsieh, Schubert, et al., 2013)

### ***Cognitive reserve.***

This was measured by self-reported number of years of education (Stern, 2009). The TOPF was not used as the main measure due to its potential lack of robustness to dementia (as discussed in chapter 4), but was used in a sensitivity analysis as described in 'statistical analyses' below.

### ***Anxiety and depression.***

The HADS (dementia-modified) version (Stott, Spector, et al., 2017). This is described in chapter 3. It has 12 items each rated from 0 to 3, with higher scores indicating greater anxiety or depression. The anxiety and depression subscales each have six items and a maximum score of 18. This differs from the usual seven item version used in previous chapters, because, as discussed in chapter 3, the psychometric properties in PLWD were improved by removing one item from each of the anxiety and depression scales. As discussed previously, where caseness was assessed the original 14-item HADS was used.

### ***Statistical analyses.***

#### ***Sample size.***

This was estimated for the structural equation modelling analysis used to examine confounding and mediating effects as this was anticipated to have the largest sample size requirements of analyses used in this paper (Wolf, Harrington, Clark, & Miller, 2013). Bentler and Chou (1987) 's heuristic of five observations per estimated parameter was used. In the current study, 21 parameters were expected to be estimated (seven parameters for each of the three pre-therapy skills measures). Estimated parameters for each skill included: the direct effect of group on neurocognition; the direct effect of neurocognition on pre-therapy skill performance; the indirect effect of group through neurocognition on pre-therapy skill performance; and the effects of each of the four confounders

on this relationship. Thus, the minimum sample size was estimated at 105. The obtained sample was larger than this, since a larger sample size had been necessary for factor analyses reported in earlier chapters.

***Missing data.***

Missing data ranged from 0 - 3% across the different measures. Visual inspection and Little's MCAR test revealed that all data were missing at random. As less than 5% of data were missing this was handled through listwise deletion (Graham, 2009), ns for specific analyses are given in results below.

***Aim 1: testing whether CBT-pre-therapy skills represent separable constructs.***

CFA was performed to test the latent structure of the BTFQ-D (Thoughts), BTFQ-D (Feelings) and CM-DEM. Two models were tested: (i) a one-factor model with all items loading on a 'CBT-readiness' factor, and (ii) a three-factor model corresponding to thought discrimination, feeling discrimination and cognitive mediation. Analysis was conducted on the entire sample to improve power and model stability (Wolf et al., 2013). Robust versions of multiple fit indices were used to determine fit for reasons discussed in chapter 3. It should be noted that event-emotion linkage (which is part of the pre-therapy skills model described in the introduction) was not included in the CFA. This was due to the marked ceiling effect, and consequent lack of variability on the Reed-Clements noted in chapters 5 and 6, which made this measure unsuitable for inclusion in a CFA (Byrne, 2013)

***Aim 2: between-group performance on pre-therapy skills measures (and clinical and demographic indices).***

For categorical variables, Chi-Square tests with post-hoc comparisons were used. Continuous data were found to be non-normally distributed (through

visual inspection and Shapiro Wilkes' tests) so non-parametric Kruskal-Wallis with Dunn Post hoc tests were used. Epsilon<sup>2</sup> was used to measure of effect size in line with recommendations for non-parametric tests (Tomczak & Tomczak, 2014). To minimise type-II error inflation, the Benjamini-Hochberg (B-H) method was used to adjust for false discovery rate (Benjamini & Hochberg, 1995).

***Aim 3: confounder and mediator analysis.***

The influence of potential confounders and mediators on any observed difference was tested using structural equation modelling. Neurocognition was measured as a potential mediator. Age, years of education, HADS anxiety and HADS-depression were included as covariates in the model to account for any potentially confounding effects. The model was first tested for fit to data using robust versions of multiple indices as described in more detail in chapter 3. Indices included the: RMSEA; TLI; CFI and SRMR. Standardized Beta coefficients of indirect and direct paths were used to examine mediating and confounding effects.

***Aim 4: investigating correlations of neurocognitive variables, depression and anxiety with pre-therapy skills in PLWD.***

Spearman's rank correlation coefficients adjusted for false discovery rate using B-H method were used to examine correlations between pre-therapy skills, ACE-III subscale scores and HADS scores.

**Results**

**Sample characteristics.**

Table 19 shows clinical and demographic characteristics for all groups. On average, PLWD had significantly lower ACE-III scores, were significantly older and had fewer years of education than both the YA and OA groups. They had higher levels of anxiety than the OA group, and higher levels of depression than

both the YA and OA groups. By contrast, the YA group had higher levels of anxiety than the OA and PLWD groups. There were also differences in terms of ethnicity, with the PLWD and OA groups containing a significantly smaller proportion of people from a BME background than the YA group.

Table 19. *Demographic and clinical characteristics of PLWD, OA and YA samples*

Variable	PLWD (n=97)		OA (n=77)		YA (n=56)		Significant contrast <sup>†</sup>
	Median (min-max)	% (N)	Median (min-max)	% (N)	Median (min-max)	% (N)	
Age	81 (58-97)		72 (65-92)		21 (18-26)		PLWD>OA>YA
Sex (M)		43 (44)		36 (28)		27 (15)	N/S
Ethnicity (White)		90 (92)		100 (77)		63 (35)	PLWD, OA>YA
Education (years)	12 (5-25)		16 (7-25)		15 (12-19)		OA, YA>PLWD
ACE-III	74 (43-98)		95 (67-100)		96 (69-100)		- OA, YA>PLWD
HADS-A score	5 (0-15)		3 (0-14)		6 (0-14)		- YA>PLWD>OA
HADS-D score	3 (0-14)		1 (0-8)		1 (0-11)		- PLWD>OA, YA
HADS A or D cases <sup>§</sup>		44(44)		14(11)		24(44)	

Note: <sup>†</sup>Significant at  $p < .05$ , adjusted for false discovery rate; Kruskal Wallace and post hoc tests used to examine differences in continuous and  $\chi^2$  in categorical variables across groups; median (range) reported due to non-normal distributions; <sup>§</sup> Caseness was ascertained using the 14-item HADS with continuous scores calculated using the 12-item dementia modified version.

**Aim 1: Do pre-therapy skills represent separable constructs?**

Two models were tested: (i) a three-factor model in which the items from the BTFQ-D (thoughts), BTFQ-D (feelings) and CM-DEM loaded on separate correlated factors (Figure 3), and (ii) a one-factor “CBT-readiness” model with all items loading on one factor.

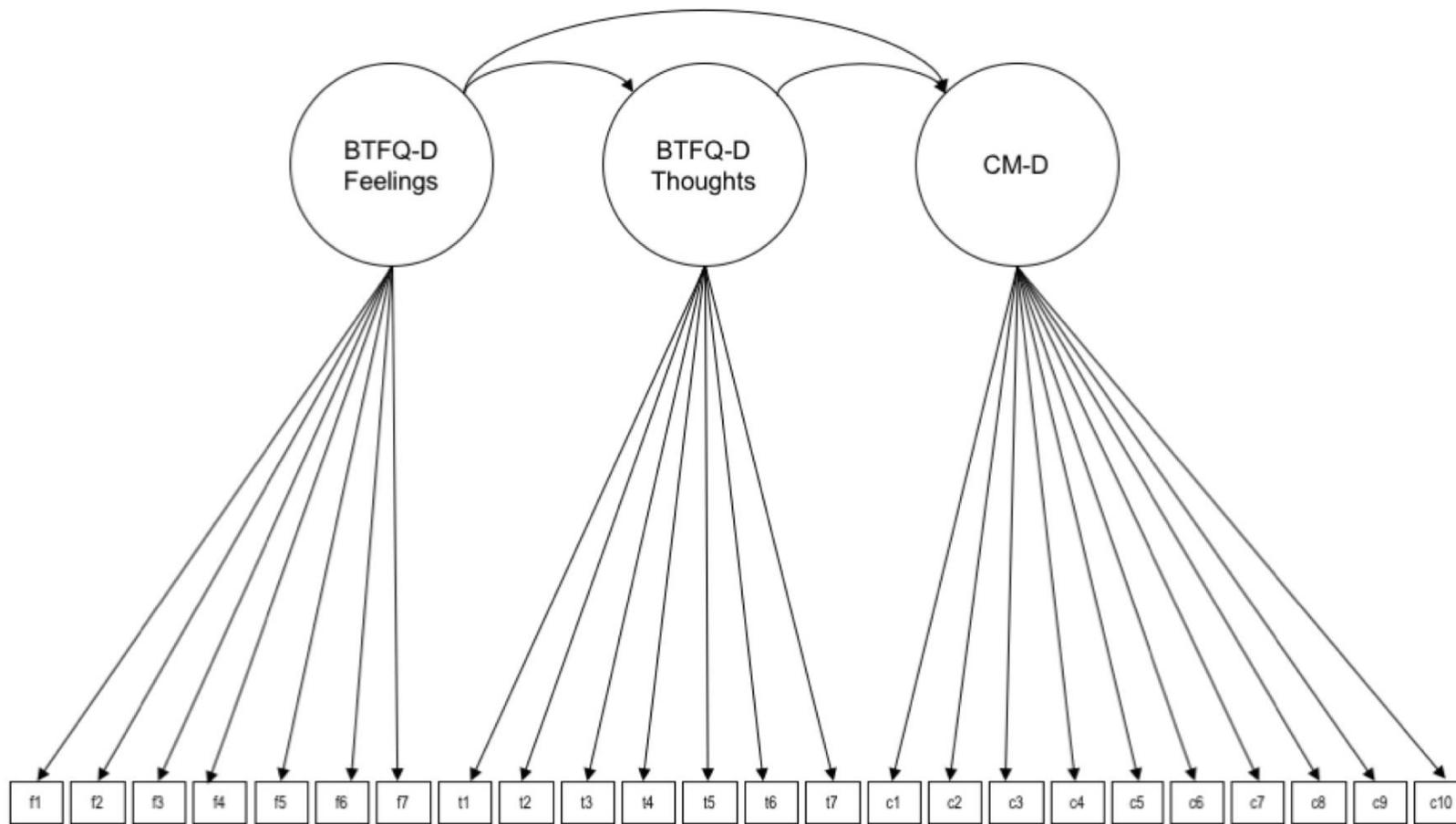


Figure 3. *Three-factor measurement model of CBT pre-therapy skills*

The three-factor model (Figure 3) showed good fit on all statistics except SRMR (Table 20). By contrast, the one-factor model showed poorer fit on all statistics except SRMR. Consequently, in subsequent analyses, pre-therapy skills measures were examined separately, rather than as one combined “CBT-readiness” construct.

Table 20. *Model fit statistics for one- and three-factor models*

Fit statistic	one Factor model	three factor model
Chi-squared	411.86	264.00
CFI	0.92	0.99
TLI	0.91	0.99
RMSEA	0.05	0.02
SRMR	0.14	0.09

Note: CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; n=226.

**Aim 2: OA, YA and PLWD group differences in pre-therapy skills.**

Table 21 shows between-group differences on CBT pre-therapy skills measures. Consistent with hypotheses, for thought-feeling discrimination (BTFQ-D) scales and CM-DEM, the YA group scored higher than the OA group, who in turn scored higher than the PLWD group. As mentioned previously, the event-emotion linkage (Reed-Clements) measure showed marked ceiling effects, with the median score at maximum for all groups. The only between group difference on this measure was that the YA group performed better than the PLWD group. Visual inspection suggested that, partly in line with hypotheses, effect sizes for differences in thought discrimination (BTFQ-D-Thoughts) and cognitive mediation

(CM-DEM) were larger than for either feeling discrimination (BTFQ-D Feelings) or event-emotion linkage (Reed Clements).

Table 21. *Between-group comparisons on pre-therapy skill measures*

Measure	PWD (n=102)		OA (n=77)		YA (n=56)		Significant Contrast <sup>†</sup>	Effect size <sup>§</sup>
	Median (min-max)	% (n) above chance	Median (min-max)	% (n) above chance	Median (min-max)	% (n) above chance		
BTFQ-D-Feelings	6 (0-7)	80.4 (82)	7 (0-7)	97.4(75)	7 (4-7)	98.2 (55)	YA>OA>PLWD	0.23
BTFQ-D-Thoughts	3 (0-7)	30.4 (31)	5 (0-7)	58 (45)	7 (4-7)	98.2 (55)	YA>OA>PLWD	0.46
CM-DEM*	4 (0-10)	-	8 (3-10)	-	9 (6-10)	-	YA>OA>PLWD	0.46
Reed Clements	6 (0-6)	89.1 (90)	6 (5-6)	88.3(68)	6 (6)	91 (51)	YA>PLWD	0.042

Note: <sup>†</sup>Dunn post-hoc test significant at  $p < .05$ , adjusted for false discovery rate, <sup>§</sup>Epsilon<sup>2</sup> used to measure effect size. Cut-offs for above-chance responding: BTFQ-D-Thoughts and feelings,  $\geq 6$ ; Reed-Clements,  $>6$ ; \*n for the dementia group in this comparison was 98 due to missing data.

### **Aim 3: Mediator and confounder analysis.**

To further explore differences between PLWD and OA groups on pre-therapy skills measures, neurocognition was assessed as a potential mediator using structural equation modelling. Potential mediators should be correlated with both the predictor and the outcome (Preacher & Kelley, 2011) and were assessed for this. For the difference between PLWD and OA groups, ACE-III clearly met this criterion, differing significantly between groups (see Table 19) and correlating with pre-therapy skills measures in the combined OA-PLWD sample (see Table 22). However, there was no difference between the YA group and OA group in ACE-III scores and the YA group were thus not included in the mediation analysis. Potential confounders should meet similar conditions to mediators (Preacher & Kelley, 2011) and were also assessed against these criteria. All confounders differed between OA and PLWD groups (see Table 19). Age and education were also correlated with performance on all pre-therapy skills in the combined OA and PLWD sample in the expected directions (Lower age and more education were associated with better performance; see Table 22). Levels of anxiety and depression correlated with performance on some pre-therapy skills measures (BTFQ-D Feelings and CM-DEM correlated with depression and BTFQ-D-Thoughts correlated with anxiety; see Table 22). Thus, all potential confounders met criteria to at least some extent and were included in the mediation analysis.

Table 22. *Correlations between covariates and CBT pre-therapy skill measures in the combined OA and PLWD sample*

Measure	ACE-III	Age	Education	HADS-A	HADS-D
BTFQ-D-Thoughts	<b>.52</b>	<b>-.24</b>	<b>.23</b>	<b>-.14</b>	-.14
BTFQ-D-Feelings	<b>.46</b>	<b>-.28</b>	<b>.35</b>	.08	<b>-.20</b>
CM-DEM	<b>.66</b>	<b>-.4</b>	<b>.29</b>	-.07	<b>-.27</b>

Note: n= 171-175; correlations in bold are significant at  $p < .05$  adjusting for B-H false discovery rate; all correlations were Spearman's rank due to assumptions of normality of distribution not being met.

Mediation was tested using a structural equation modelling approach. A binary variable was used to indicate membership of the OA or PLWD group. The three-factor measurement model (Figure 3) was combined with direct paths and indirect paths via ACE-III from the group variable to scores on each pre-therapy skill measure (BTFQ-D-Feelings, BTFQ-D-Thoughts and CM-DEM; three paths). Age, years of education, anxiety and depression were included as control variables to account for potential confounding effects. The Reed-Clements was excluded from this analysis as there were no between-group differences.

Figure 4 shows the mediation models. For clarity, the mediation paths for each outcome are shown separately and the measurement model is excluded; however, they were tested as part of one model. Model fit was good on most indices ( $X^2=378.83$  (N/S); CFI=0.94; TLI=0.93; RMSEA=0.03; SRMR=0.13). Examination of regression coefficients showed that performance on 'The indirect effects of group via ACE-III were all significant (BTFQ-D-Thoughts = -0.34; BTFQ-D-Feelings = -0.2; CM-DEM=-0.33).

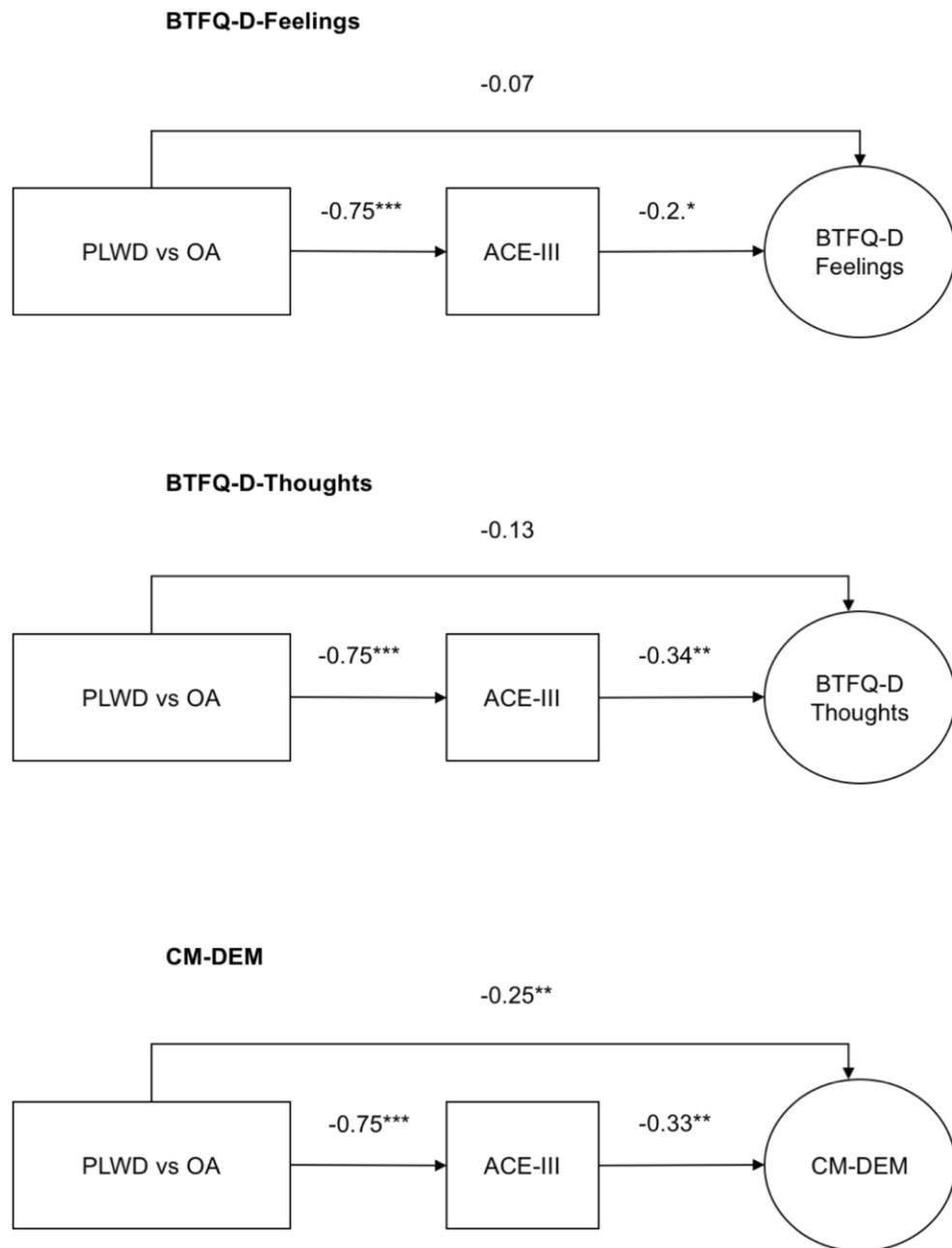


Figure 4. Mediation of differences between PLWD and OA in pre-therapy skill performance<sup>3</sup>

---

*n=167; Figures are standardised regression coefficients ( $\beta$ ), adjusted for covariates (age, years of education, anxiety and depression); measurement model not shown*

### ***Sensitivity analyses.***

As discussed above, the TOPF was not used as the primary measure of cognitive reserve due to questions about robustness to dementia. Rerunning the mediation analysis with TOPF included as the measure of cognitive reserve, confirmed the main result that ACE-III fully mediated the association of dementia with the BTFQ-D Thoughts scale, and partially mediated the association with the CM-DEM. However, the association of dementia with feeling discrimination (BTFQ-D-Feelings) performance was no longer mediated by ACE-III in this analysis.

While the YA group were not included in the mediation analysis for the reasons discussed above, it was investigated whether potential confounders might account for differences in pre-therapy skills between YA and OA groups. However, no variable met criteria for confounding of being correlated with predictor and dependent variables and no further analysis of this difference was undertaken.

### **Aim 4: Correlates of pre-therapy skills in PLWD.**

Finally, the associations between particular aspects of neurocognition with CBT pre-therapy skills (Table 23) were explored. Largely, in line with hypotheses, all pre-therapy skills measures were positively correlated with ACE-III language scores and all except BTFQ-D-Thoughts were associated with ACE-III fluency scores. CM-DEM was also associated with memory scores. The other significant un-hypothesised correlations were: BTFQ-D-Feelings and CM-DEM with visuospatial functioning scores, BTFQ-D-Thoughts with attention scores, and BTFQ-D-Feelings with HADS-anxiety scores.

Table 23. Correlations of CBT pre-therapy skill measures with neurocognitive and mood variables in the PLWD sample

Measure	ACE language	ACE Fluency	ACE Memory	ACE Attention	ACE Visuospatial	HADS-A	HADS-D
BTFQ-D-'Feelings	<b>.32</b>	<b>.32</b>	.10	.07	<b>.29</b>	<b>.26</b>	.001
BTFQ-D-Thoughts	<b>.27</b>	.20	.22	<b>.31</b>	.17	.05	-.05
CM-DEM	<b>.41</b>	<b>.30</b>	<b>.38</b>	.08	<b>.26</b>	.13	-.08

Note: Bold represents significant correlation at  $p < .05$ , adjusted for False Discovery rate; n=97-101.

## Discussion

The main aim of this chapter was to compare the performance of a group of PLWD on measures of CBT pre-therapy skills with both OA and YA control groups. As hypothesised, PLWD scored significantly lower on these measures compared to the OA group. This effect was mediated by overall neurocognitive level and this was the case even when the differences in age, education, anxiety and depression levels between the samples were accounted for. Performance on pre-therapy skills was also poorer in the OA than the YA group. However, this effect did not appear to be mediated by cognition or confounded by other measured variables. Within the PLWD group, language function was associated with performance on all pre-therapy skills, with other aspects of neurocognitive functioning associated with some but not all pre-therapy skills measured here.

### **CBT pre-therapy skills; separable and differentially affected in PLWD.**

The measurement model results supported the idea that CBT pre-therapy skills should be measured as separate components rather than as a single readiness factor. Whilst there were between-group differences on all pre-therapy skills measures, the largest effect sizes for differences in performance were on the 'thought-specific' measures (CM-DEM; BTFQ-D-Thoughts) as opposed to the more feeling oriented measures (the BTFQ-D-Feelings and the Reed Clements). Event-emotion linkage, in particular, appeared to be largely unaffected with the performance only differing between PLWD and YA groups. Thus, the results perhaps suggest that, in general, PLWD struggle most with measures that require the identification or manipulation of thoughts.

One caveat to this interpretation is that task demands rather than skill difficulty per se may have contributed to the relative size of group differences. In particular, the CM-DEM required free as opposed to the forced-choice response

used in other tasks and free response may be particularly hard for PLWD (Smith et al., 2005). The thought identification task requires comprehension of longer item stems than the feeling identification task and although this is partly inherent in the constructs themselves (thoughts but not feelings are invariably more than one word) (Padesky & Greenberger, 2012), it may have affected results. Consequently, findings require replication using other measures, for example the cognitive mediation measure devised by Dagnan et al. (2000) which may reduce the difference in task demands although perhaps at the cost of ecological validity (Dagnan et al., 2000; Dagnan et al., 2009).

### **The centrality of neurocognition.**

Although without longitudinal data causation cannot be established, the results support the idea that the reason PLWD struggle on CBT pre-therapy skills measures is because of neurocognitive deficits. However, the partial (rather than full) mediation of CM-DEM differences suggest, for this more complex skill at least, other unmeasured differences between groups may play a role.

Subject to confirmation of causality using appropriate methods, findings also suggested that, in PLWD, language impairment is important in all pre-therapy skills and also that executive function (represented by ACE-III fluency scores) may play an important role in two of three skills. This is in line with previous work (Dagnan et al., 2000; Johnco et al., 2014; Joyce et al., 2006). The role of memory was only supported for the CM-DEM, which is of interest given that many adaptations to CBT for PLWD have focussed on memory. The association of visuospatial functioning with CM-DEM and BTFQ-D-Feelings was not predicted. This could reflect the fact that visuo-constructional tasks making up the 'visuospatial' subscale of the ACE-III have substantial executive components (Freeman et al., 2000) with CBT pre-therapy skill associations being due to this.

The visuospatial finding may also be due to task demands as CBT pre-therapy tasks were presented visually as well as verbally. However, neither of these explanations account for the lack of association with BTFQ-D-Thoughts performance and further investigation is required.

### **A deficit in CBT pre-therapy skill performance in OA vs. YA.**

The OA group had deficits in all pre-therapy skills (aside from event-emotion linkage) relative to the YA group. However, in contrast to the above finding, there was no evidence for mediation of difference by neurocognition and neither was there evidence of a role for any measured potential confounder. As discussed previously, a possible explanation for the findings is that the observed effect of age was due to cohort effects with the current generation of OA having less of a culture of talking about thoughts, emotions and their linkage than the current generation of YA (Chand & Grossberg, 2013). Whilst the finding that age did not confound the effect of dementia (vs OA group) on outcomes may seem to contradict the importance of age, this could be because of the small differences in age between the two groups. Possibly, cohort effects are more pronounced when comparing 18-25 to 65+ samples as opposed to when comparing an OA group to a slightly older PLWD group. In order to further elucidate mechanisms, it may be useful in future work to measure cohort beliefs across age groups and examine their association with CBT pre-therapy skill performance.

### **The role of anxiety and depression.**

In the current study there was little evidence that anxiety and depression contributed to the variability in skill performance within PLWD or confounded the difference between PLWD and the OA group (despite the PLWD group having higher levels of anxiety and depression). This suggests that the results here might be applicable to PLWD samples who are universally anxious and

depressed (such as those attending CBT) and thus enhances the generalisability and clinical utility of results. Indeed, the PLWD sample in particular (44% meeting caseness for anxiety and depression) probably had considerable overlap with such CBT samples.

### **Research and clinical implications.**

#### ***PLWD.***

There are a number of clinical implications of the results that pertain to PLWD. Firstly, the seemingly smaller impact of dementia on feeling identification and event-emotion linkage and high above chance performance on these skills (80 and 89% respectively) would support the use of pleasant event scheduling (Jacobson et al., 1996), as feeling identification and event-emotion linkage are likely to be core skills in that approach (Jacobson et al., 1996).

The results do not, however, necessarily imply across-the-board removal of cognitive elements of CBT for PLWD. There was substantial variability in performance within the PLWD group, even on thought-related measures. Indeed, 30% of PLWD scored above chance on the BTFQ-D-Thoughts. Consequently, dementia might be better viewed as a risk factor for poor CBT pre-therapy skills. Within this conceptualisation, levels of pre-therapy skills necessary for cognitive aspects of CBT could be established through idiosyncratic assessment using the current measures in the context of clinical judgement. Changes to CBT practice could then be individually applied on a case by case basis.

Were cross-sectional findings as to neurocognition and the roles of language in particular to represent causal relationships, it would support the use of strategies that focus on simplifying language. These might include very frequent capsule summaries, support of verbal with written material and regular checking of understanding, all of which are already incorporated within standard

CBT. Given the verbal fluency associations and the possible executive explanation for visuospatial findings, reducing executive demands (for example, imposition of a clear session structure) may also be useful. The moderately large association of memory with CM-DEM, but not other measures, suggests that, potentially, for more complex skills where working memory may be required to process and manipulate information, memory deficits are important. Consequently, mini formulations which focus on reducing general neurocognitive load of cognitive elements of CBT (Spector et al., 2015) may be a useful adaptation. Further longitudinal research to elucidate the causal role, or otherwise, of neurocognition in CBT pre-therapy skill performance is warranted.

#### **OA.**

The differences between OA and YA groups were not mediated by neurocognition and different strategies for amelioration may be useful. It may well be that this is a cohort -based difference in emotional understanding and possibly the best way of supporting OA to develop pre-therapy skills will be through training and explanation of constructs. Again, within group performance was highly variable with high above chance performance suggesting that idiosyncratic assessment may be a sensible strategy.

#### **Strengths and limitations.**

This study was the first to examine CBT pre-therapy skills in PLWD using measures designed for that population. It has strengths in using measures which were guided by a theoretical model and have been validated (albeit in an overlapping sample). It was the first study to test the structural validity of conceptualising CBT pre-therapy skills as separable constructs, and it improved methodologically on previous work in intellectual disabilities through comparison to a control group using a sample size with power to detect subtle differences.

Measurement of potentially important dementia-relevant neurocognitive variables allowed hypotheses as to the reasons for pre-therapy skill deficits in PLWD to be generated, something that was not possible in previous research.

Several limitations should be noted. While structural equation modelling sample size was calculated a priori using a recognised heuristic, it may be more appropriate to use Monte-Carlo simulation techniques (Wolf et al., 2013) and future work should do this. The design was cross-sectional with consequent limitations on the ability to ascertain cause and effect. As discussed in earlier chapters, the measures used only reflect a subset of the potential pre-therapy skills required to be ready for CBT, and there are other factors required of a client and health and social care systems for a client to make use of CBT (Stott, Charlesworth, et al., 2017). The Reed Clements task had a ceiling effect in all groups and other measures had ceiling effects in the YA group in particular, meaning that the upper end of traits measured here might not be fully examined and lead to consequent lack of ability to detect important differences between groups. However, the lower end of the trait range is arguably of the greatest clinical significance. Perhaps the most important limitation is that the measures used here have not been used in the context of actual CBT for PLWD, consequently their relationship with CBT outcomes is not known and needs to be established.

### **Conclusions.**

While PLWD may have a relative difficulty in CBT pre-therapy skills that require identification and use of thoughts, there is substantial variability in this and this study does not suggest that mild dementia in itself precludes readiness for cognitive elements of CBT. In PLWD, the role of neurocognition may be important and strategies to adapt CBT for PLWD should take this into account.

Older age is also associated with poorer performance on CBT pre-therapy skills (although not due to neurocognition, but perhaps through cohort effects). Future research should use a longitudinal design to examine the role of cohort beliefs, with clinical strategies for OA perhaps focussed on idiosyncratic adaptation of socialisation to the model based on assessment of CBT pre-therapy skill levels.

**Chapter 8: Discussion - Implications of Results for CBT Readiness,  
Research and Practice with PLWD and OA**

## **Thesis Summary**

The focus of this thesis was on the 'pre-therapy skills' necessary to be ready for CBT in PLWD and, to a lesser extent, OA. CBT is not a unitary construct and this thesis argued that cognitive elements of CBT are particularly important to consider. It was further suggested that cognitive restructuring is a core cognitive element of CBT and, as a consequence, the CBT pre-therapy skills focussed on this thesis were two components of cognitive restructuring; the skills of (i) discriminating between emotions, thoughts and behaviours and (ii) understanding the connections between thoughts, emotions, behaviours and situations. The development of measures of these skills and implications of findings related to the impact of dementia on them are detailed in the preceding chapters. The focus in the discussion is on setting findings within the wider context of CBT readiness in particular, and on implications for CBT for PLWD and OA more generally. Research and clinical implications will be discussed as will the strengths and limitations of the thesis as a whole.

## **The Wider Context of CBT Readiness**

This thesis was concerned with whether PLWD are 'ready' to take part in CBT at the point at which it is offered (Willner, 2006). However, as mentioned in chapter 2, readiness for CBT is more than just proficiency in skills related to specific components of cognitive restructuring. Thus, the pre-therapy skills measured in the current thesis are probably only a small subset of the pre-therapy skills necessary to be ready for CBT as a whole. In particular, Safran et al. (1993) and Willner (2006) have suggested that pre-therapy skills required to be ready for CBT include skills necessary for (i) 'common factors' that are recognised to be important across all talking therapies (for example, skills

involved in forming a working alliance or engaging in a meaningful conversation), and (ii) CBT specific factors (for example, cognitive restructuring).

### **Common therapeutic processes.**

This thesis only focussed on factors specific to CBT. It is, however, well established that common therapeutic factors are important in CBT outcome in people without dementia (Wampold, 2015). While the role of common factors has not been directly investigated in PLWD, no trial of CBT for PLWD has included the 'active' control necessary to draw the conclusion that CBT specific processes (rather than other common factors) are important for outcome. The potential role of common factors is also supported by evidence that other psychotherapeutic approaches have benefits in improving depression and anxiety (Cheston, Jones, & Gilliard, 2003) suggesting that positive outcomes are not unique to CBT. Consequently, it may well be that common factors are actually the active therapeutic ingredients in CBT interventions for PLWD. Given their potential import for therapy outcome, future work could meaningfully focus on readiness for these common therapeutic processes in PLWD. This work might proceed in a similar way to the work presented in the current thesis i.e. by (i) specifying the pre-therapy skills necessary to be ready for these common therapeutic factors, (ii) identifying which of these skills have not been studied in dementia previously, (iii) developing tools to measure them and, (iv) investigating relationships with hypothetically important correlates. For example, it may be that specific types of dementia (e.g. Behavioural variant Frontotemporal dementia) are associated with specific problems (e.g. lack of empathic concern (Hsieh, Irish, Daveson, Hodges, & Piguet, 2013)), which may particularly effect common therapeutic factors (e.g. ability to participate in a working alliance).

### **Behavioural elements of CBT.**

Even putting aside common therapeutic factors, not all pre-therapy skills required to be ready for *CBT specific* processes were measured in this thesis. In particular, CBT comprises behavioural elements (e.g. behaviour change activities such as pleasant event scheduling (Roth & Pilling, 2008)) as well as the cognitive elements focussed on here. It is assumed by many authors that PLWD need less adaptation of the behavioural elements of CBT than the cognitive ones (Teri et al., 1997). As mentioned in previous chapters, the work in this thesis (particularly the ceiling effect on the Reed Clements task) does provide preliminary support for this assertion. However, the planning and organisational skills related to behavioural elements (e.g. the ability to plan and organise an activity schedule (Greene, Hodges, & Baddeley, 1995)) of CBT may be affected by many types of dementia, including Alzheimer's disease. Investigating the relationship of these potential pre-therapy skills to readiness for behavioural elements of CBT might be a fruitful area for future research.

### **Other dimensions of CBT readiness.**

There are also other aspects to CBT readiness beyond pre-therapy skills. In particular, widely used models which pertain to uptake of interventions outside of a dementia context (e.g. the COM-B model (Michie et al., 2011)) suggest that skills (or capabilities – the 'C' in the COM-B model) are only one aspect of readiness for interventions. Such models suggest that 'motivation' and opportunity, the 'M' and 'O' in the COM-B model are also of import.

#### ***Motivation.***

The importance of motivation in CBT readiness has been outlined in the intellectual disabilities literature (Willner, 2006). Furthermore, in the general psychological therapies literature, poorer motivation is associated with poorer

outcome (Wampold, 2015). Motivation includes, at a minimum, expectancies around the potential for change (Wampold, 2015) and trust in professionals (Willner, 2006). When an individual has a cognitive impairment, the motivation of carers or family members to be involved in CBT is also likely to be important (Willner, 2006). As such, motivational readiness may be affected by psychological and other social aspects of dementia. For example, it could be that due to the hopelessness which PLWD and their carers sometimes experience (Lopez et al., 2003), expectancies around the potential for change may be low, leading to low uptake of psychological interventions and potentially reducing effectiveness (Wampold, 2015). Motivation may also be directly affected by neurocognitive processes, with some dementia subtypes (for example, Behavioural variant Frontotemporal Dementia) directly affecting neurobiological systems related to volition (Eslinger, Moore, Antani, Anderson, & Grossman, 2012). While not the focus here, motivation to engage with CBT (and indeed any psychosocial intervention) is under-investigated in PLWD and will be useful to consider in future work.

### ***Opportunity.***

Healthcare systems may also not be 'ready' to offer CBT to PLWD. In particular, a diagnosis of dementia is likely to affect the opportunities available within a healthcare system (for example, services may exclude PLWD directly or indirectly, PLWD may need carers to attend sessions, professionals may need training, or measures may not be validated for this group). This 'system readiness' of psychological therapy services for PLWD has not been evaluated and would be useful to examine going forward. A clinical psychology doctoral student is currently being supervised by the author of this thesis to undertake this

work using a recent study evaluating accessibility of IAPT for people with intellectual disabilities (Marwood, Chinn, Gannon, & Scior, 2018) as a template.

In summary, while findings in this thesis are useful in understanding pre-therapy skills that are important for the cognitive elements of CBT, the empirical work presented does not encompass the totality of readiness for CBT and future work should evaluate: readiness for: common therapy factors; behavioural aspects of CBT; motivation to take part in CBT; and system readiness. It is of note that much of this work would also have benefits for thinking about access of PLWD to psychosocial interventions other than CBT too.

### **Implications for Adaptation of CBT**

It is hoped that by developing tools, which can be used by clinicians to identify issues with specific pre-therapy skills, this research will facilitate idiosyncratic adaptation of CBT in practice that is tailored to individual needs of PLWD and OA. It is also hoped that by identifying the particular pre-therapy skills that might be most affected in PLWD, as well as the correlates of those pre-therapy skills, this work will help with developing future CBT interventions that emphasise the elements PLWD find easy and potentially omit or provide training in the areas that they might struggle with.

### **Pre-therapy Skills Training**

The focus in the current thesis was on pre-therapy skills at a single point in time and understanding which skills might be particularly affected in PLWD has important implications for design of future interventions. However, as has been alluded to earlier, it is also useful to know whether the specific CBT pre-therapy skills examined here (and indeed any others) can be trained. There has been some encouraging work relevant to this in PLWD, which has shown that a pre-therapy skill related to those focussed on here (facial emotion recognition) can be

trained over the course of a group intervention of a few weeks in duration (García-Casal et al., 2017). However, in the context of a relatively brief CBT intervention, a shorter training period would probably be needed, and the ability to train the specific pre-therapy skills discussed in this thesis is unknown. There are some pragmatic examples of brief training packages for cognitive mediation and behaviour-thought-feeling discrimination in the intellectual disabilities literature (e.g. (Tsimopoulou, Kroese, Unwin, Azmi, & Jones, 2018)) and some papers detailing these are included in the literature review in appendix C. It may be that adapting these training packages for PLWD is a next step to the research reported in the current thesis. This is particularly pertinent, because it may be that the principal effect of dementia and, in particular, Alzheimer's disease (given its typical impact on new learning (Salmon & Bondi, 2009)) is less on pre-therapy skill performance per se, but more on the ability to develop pre-therapy skills with training. Once training packages have been developed, the tools developed in this thesis could provide indicators of who needs what training, when.

### **Evaluating Pre-therapy Skills in the Context of CBT Outcome**

As discussed in preceding chapters, the importance of the skills measured here in a CBT context is, to a large degree, dependent on whether improving skill performance improves outcome. This is typically examined through mediation, which entails establishment of four logical conditions: (i) CBT is efficacious. (ii) CBT improves pre-therapy skill performance. (iii) Change in pre-therapy skills effect the outcome variable (e.g. anxiety, depression). (iv) Any CBT effects can be attributed to this causal pathway (Weersing & Weisz, 2002).

Given that evidence even as to (i) is currently preliminary (Livingston et al., 2017). it would be useful to conduct well planned RCTs of CBT for PLWD with thoughtfully spaced, repeated measures of pre-therapy skills, sufficient power,

and an active control group (Kazdin, 2007). Such trials would allow examination of outcome and of any mediators of that outcome. Conducting such trials is highly resource intensive, and multiple potential mediators should be examined in one trial, both for efficiency and because pathways to impact are likely to be highly complex (Kazdin, 2007). Given the high cost of large scale RCTs, it may also be useful to follow updated MRC guidance in developing complex interventions (Moore et al., 2015) and conduct smaller RCTs with clearly specified mixed-method process evaluation components to understand mechanisms of change in CBT for PLWD. If larger RCTs are conducted it will also be important to explore heterogeneity of dementia presentation in relation to outcome. The work in this thesis suggests that, potentially, PLWD with overall lower neurocognitive function, and those with language and executive functioning deficits, may have particular difficulties in CBT pre-therapy skill performance. Consequently, work looking at these variables as predictors of CBT outcome in PLWD would be useful.

### **Implications for CBT with OA.**

The main focus of the current thesis was on PLWD, but results also pertain to CBT for OA (defined here as people aged over 65). In particular, results suggested that OA groups may perform less well on the pre-therapy skills measures than YA groups. The reasons for this were not clear, but, unlike for PLWD, did not appear to be due to neurocognitive factors. As was briefly alluded to earlier, an intuitively plausible explanation for obtained differences is the impact of birth cohort effects. In particular, a number of authors have suggested that current older cohorts may be less socialised than younger cohorts into discussing and thinking about constructs related to psychological therapy (Chand & Grossberg, 2013; Laidlaw et al., 2015) and thus may need a

more extended period of socialisation to the CBT model (Chand & Grossberg, 2013; Laidlaw et al., 2015). The work in this thesis would support this idea. Furthermore, given that OA are a highly heterogeneous group and amount of socialisation to the model needed is likely to vary across individuals (Chand & Grossberg, 2013), measures developed here might also be useful tools to measure which older individuals struggle with which pre-therapy skills, and hence, who will actually need extended socialisation periods. The caveat to this individualised case by case approach is that the pre-therapy skills measures had less evidence for validity in OA than in PLWD.

It is interesting that despite the evidence here that OA may be 'less ready' for CBT than YA, evidence from the national IAPT dataset (NHS Digital, 2016) and recent national surveys of psychological services (Chaplin, Farquharson, Clapp, & Crawford, 2015) suggests that the current cohort of OA may actually have better psychological therapy outcomes than the current cohort of YA. Future research might usefully examine the reasons for this apparent disconnect and, as with the dementia literature, might also include the pre-therapy skills measured here as potential predictors or mediators of CBT outcomes in OA.

### **Implications Outside of a CBT Context**

Finally, while the main focus of this thesis was on CBT, the results might also have relevance outside of this context and it could be argued that the skills measured here are important in managing one's emotions in everyday life. In particular, emotion regulation strategies such as changing appraisals of a situation to upregulate positive, or downregulate negative, emotions (cognitive reappraisal) would seem to necessitate understanding of event, thought, feeling, discrimination and linkage. Given that difficulties with emotion regulation are associated with multiple psychological health problems and social problems

(Aldao, Nolen-Hoeksema, & Schweizer, 2010) in adults without dementia, using the measures here to understand emotion regulation skills in PLWD might be useful clinically and in research. Additionally, to this author's knowledge there has been no research directly examining cognitive reappraisal in PLWD and it might be useful to do this in future, perhaps adapting some of the methods that have been used to examine this in OA groups e.g. (Shiota & Levenson, 2009).

In addition to measures of pre-therapy skills having potential use outside of a CBT context, the chapters on the TOPF and HADS reported in this volume also have implications beyond CBT. In particular, the lack of validity of the TOPF raises questions as to its use as a measure of premorbid IQ in diagnosis. Indeed, the published version of chapter 4 (Stott, Scior, et al., 2017) includes an analysis which suggests that the TOPF does not increase dementia screening accuracy over and above existing screening tools, which counts against its routine use in assessment contexts. The HADS is a widely-used tool in depression and anxiety measurement both clinically and in research and results in chapter 3 of the present volume will enable more accurate interpretation of findings from it.

### **Overall Strengths and Limitations**

There were some strengths related to particular aspects of this thesis and these are presented in their respective chapters. There were, however, some strengths relating to the thesis as a whole, which are discussed below:

The recruitment of PLWD from memory clinic settings allowed for clear characterisation of dementia and subtype and aided readers in knowing to whom the results are generalizable, something which previous research in PLWD (Harter, 2003) did not allow. Furthermore, the recruitment of OA and YA control groups is an advance on work in intellectual disabilities where a comparison group has not generally been used. This is particularly pertinent in examining pre-

therapy skills since a key purpose of examining performance on these skills is to inform adaptation from therapy used with typically functioning groups and thus it is difficulties *relative* to these typically functioning groups rather than absolute rates of ‘below chance performance’ that are probably most clinically relevant.

The validation of pre-therapy skills measures, and the use of factor analysis in particular, was an important element of this thesis and was an advance on previous work in any population as this has not been previously done, despite lack of clarity about the utility of items and the structural interpretation of measures.

The measurement of mood and anxiety and conceptualisation of them as covariates was an advance on previous work, which has not normally measured these variables despite their importance in CBT and the fact that mood and anxiety levels may be associated with performance on measures of other pre-therapy skills (Gur et al., 1992; Surcinelli et al., 2006).

The absence of a carer related inclusion criteria was a strength as the study thus provided some evidence as to what aspects of CBT PLWD might be able to do on their own. This has potential clinical utility since all CBT intervention studies for PLWD thus far have required a ‘supportive other’ to be involved, and, given that about a third of PLWD live alone (Miranda-Castillo, Woods, & Orrell, 2010), the necessity for this involvement could be a barrier to access.

As with strengths, limitations as to the specific empirical questions evaluated in this thesis were discussed in the respective chapters. However, there are some design and conceptual issues that relate to this thesis as a whole:

While it was a strength that the sample size was larger than in previous work, sample size restrictions precluded detailed analyses of between dementia subtype differences in CBT relevant skills. Given the potential impact that

subtype and the underlying neuropsychological deficits that tend to be associated with subtype, might have on CBT pre-therapy skill performance, this may be a fruitful area for future research. In that research, it may also be useful to use better validated, more extensive, measures of specific neurocognitive domains than the subscales of the ACE-III used here.

Furthermore, the measures of pre-therapy skills in the current thesis were pencil and paper measures of fractionated CBT abilities. Such measures might have limited ecological validity in representing the nuanced and contextualised conversations about event-emotion-thought linkage and discrimination that happen in CBT practice (Hebblethwaite et al., 2011). One approach to managing this is to design semi-structured interview measures to 'mock up' 'CBT like conversations' as has been done in the intellectual disabilities literature (Hebblethwaite et al., 2011). The author of this thesis has supervised an MSc student in developing an appropriate contextualised measure and a paper on differences between PLWD and OA on this measure as well as correlations with pencil and paper measures is being prepared for publication.

The second, and perhaps most important limitation, is that, while this thesis built on previous work by making clear arguments for defining (i) cognitive restructuring as a critical element of CBT, and (ii) linkage and discrimination of thoughts, feelings, situations and behaviours as core components of cognitive restructuring, there is still not certainty over what the core elements of CBT are and what pre-therapy skills are necessary for them. In particular, an argument could be made that the evidence for cognitive restructuring as a mediating process in CBT outcome is not conclusive (Hundt et al., 2013) and that the components identified here might not be specific to cognitive restructuring, but also be important in processes common to all therapies (e.g. an ability to self-

reflect on mental states). These issues of construct validity are not unique to cognitive restructuring and it is likely that similar criticisms could be levelled at any attempt to 'define CBT'. A suggestion for future research that the author of this thesis makes in the published literature review in appendix C is that '*Given the lack of a priori definition of what constitutes a CBT intervention (Doherr et al., 2005) and the multiplicity of interventions falling under the CBT umbrella (Roth & Pilling, 2008) it might be useful to move away from conceptualising readiness for CBT overall and to start with specific CBT interventions themselves, developing and evaluating measures and training procedures specifically based on the skills judged important within particular interventions, for example, behavioural experiments in CBT for social phobia (Roth & Pilling, 2008). This would automatically improve face validity and feasibility, allowing refinement of the concept of readiness and preliminary assessment of specific skills associated with therapy outcomes. More rigorous psychometrically sophisticated measures of relevant constructs could then be developed.*' (Stott, Charlesworth et al., 2017 p50)

Perhaps future work might adopt this strategy in addition to the other suggestions for future research mentioned above.

## **Conclusion**

In conclusion, the measures of CBT pre-therapy skills and findings related to them presented in this thesis could be of use in adapting CBT practice and also in intervention development for PLWD as well as for OA. The validation of the HADS and TOPF also have implications for working with PLWD outside of a CBT context and will hopefully be of use to clinicians in assessing depression and anxiety and premorbid functioning respectively. Future research might usefully focus on motivation and system readiness for CBT as well as evaluating

mediation of CBT outcome by performance on the pre -therapy skills measures developed here.

## References

- Aalten, P., Verhey, F. R., Boziki, M., Brugnolo, A., Bullock, R., Byrne, E. J., . . . De Deyn, P. P. (2008). Consistency of neuropsychiatric syndromes across dementias: results from the European Alzheimer Disease Consortium. *Dementia and Geriatric Cognitive Disorders*, *25*(1), 1-8.
- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: a meta-analytic review. *Clinical Psychology Review*, *30*(2), 217-237.
- Alves, L., Simões, M. R., Martins, C., Freitas, S., & Santana, I. (2013). Premorbid IQ influence on screening tests' scores in healthy patients and patients with cognitive impairment. *Journal of Geriatric Psychiatry and Neurology*, *26*(2), 117-126.
- Alzheimer's Association. (2012). 2012 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *8*(2), 131-168.
- Andreas, S., Schulz, H., Volkert, J., Dehoust, M., Sehner, S., Suling, A., . . . Da Ronch, C. (2017). Prevalence of mental disorders in elderly people: the European MentDis\_ICF65+ study. *The British Journal of Psychiatry*, *210*(2), 125-131.
- APA. (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. Washington, DC: American Psychiatric Association.
- BABCP. (2018). What is CBT. Retrieved from <http://www.babcp.com/Public/What-is-CBT.aspx>
- Banerjee, S., Hellier, J., Dewey, M., Romeo, R., Ballard, C., Baldwin, R., . . . Katona, C. (2011). Sertraline or mirtazapine for depression in dementia (HTA-SADD): a randomised, multicentre, double-blind, placebo-controlled trial. *The Lancet*, *378*(9789), 403-411.
- Bartlett, R., & O'Connor, D. (2007). From personhood to citizenship: broadening the lens for dementia practice and research. *Journal of Aging Studies*, *21*(2), 107-118.
- Beaujean, A. A. (2012). R Package for Baylor University Educational Psychology Quantitative Course (Version 0.5). Retrieved from <https://cran.r-project.org/package=BaylorEdPsych>
- Beck, A. T. (1979). *Cognitive Therapy of Depression*. New York: Guilford press.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *289*-300.
- Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*(1), 78-117.
- Bird, C. M., Chan, D., Hartley, T., Pijnenburg, Y. A., Rossor, M. N., & Burgess, N. (2010). Topographical short-term memory differentiates Alzheimer's disease from frontotemporal lobar degeneration. *Hippocampus*, *20*(10), 1154-1169.
- Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *Journal of Psychosomatic Research*, *52*(2), 69-77.
- Bora, E., Velakoulis, D., & Walterfang, M. (2016). Meta-analysis of facial emotion recognition in behavioral variant frontotemporal dementia: comparison with Alzheimer's disease and healthy controls. *Journal of Geriatric Psychiatry and Neurology*, *29*(4), 205-211.

- Bruce, M., Collins, S., Langdon, P., Powlitch, S., & Reynolds, S. (2010). Does training improve understanding of core concepts in cognitive behaviour therapy by people with intellectual disabilities? A randomized experiment. *British Journal of Clinical Psychology, 49*(1), 1-13.
- Byrne, B. M. (2013). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Hove, UK: Routledge.
- Chand, S. P., & Grossberg, G. T. (2013). How to adapt cognitive-behavioral therapy for older adults. *Current Psychiatry, 12*, 10-15.
- Chaplin, R., Farquharson, L., Clapp, M., & Crawford, M. (2015). Comparison of access, outcomes and experiences of older adults and working age adults in psychological therapy. *International Journal of Geriatric Psychiatry, 30*(2), 178-184.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504.
- Cheston, R., Jones, K., & Gilliard, J. (2003). Group psychotherapy and people with dementia. *Aging & Mental Health, 7*(6), 452-461.
- Chi, S., Wang, C., Jiang, T., Zhu, X.-C., Yu, J.-T., & Tan, L. (2015). The prevalence of depression in Alzheimer's disease: a systematic review and meta-analysis. *Current Alzheimer Research, 12*(2), 189-198.
- Clare, L., Nelis, S. M., Martyr, A., Roberts, J., Whitaker, C. J., Markova, I. S., . . . Morris, R. G. (2012). The influence of psychological, social and contextual factors on the expression and measurement of awareness in early-stage dementia: testing a biopsychosocial model. *International Journal of Geriatric Psychiatry, 27*(2), 167-177.
- Clark, D. A. (2013). Cognitive restructuring. In S. Hofman (Ed.), *The Wiley Handbook of Cognitive Behavioral Therapy* (Vol. 1, pp. 1-22). New York: Wiley.
- Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316-336.
- Cockburn, J., Keene, J., Hope, T., & Smith, P. (2000). Progressive decline in NART score with increasing dementia severity. *Journal of Clinical and Experimental Neuropsychology, 22*(4), 508-517.
- Cosco, T. D., Doyle, F., Ward, M., & McGee, H. (2012). Latent structure of the Hospital Anxiety And Depression Scale: a 10-year systematic review. *Journal of Psychosomatic Research, 72*(3), 180-184.
- Coyne, J. C., & van Sonderen, E. (2012). The Hospital Anxiety and Depression Scale (HADS) is dead, but like Elvis, there will still be citings. *Journal of Psychosomatic Research, 73*(1), 77-78.
- Crawford, J. R., Deary, I. J., Starr, J., & Whalley, L. J. (2001). The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval. *Psychological Medicine, 31*(3), 451-458.
- Crawford, J. R., Stewart, L., Parker, D., Besson, J., & Cochrane, R. (1989). Estimation of premorbid intelligence: combining psychometric and demographic approaches improves predictive accuracy. *Personality and Individual Differences, 10*(7), 793-796.
- Crutch, S. J., Lehmann, M., Schott, J. M., Rabinovici, G. D., Rossor, M. N., & Fox, N. C. (2012). Posterior cortical atrophy. *The Lancet Neurology, 11*(2), 170-178.
- Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult

- depression, alone and in comparison with other treatments. *The Canadian Journal of Psychiatry*, 58(7), 376-385.
- Dagnan, D., Chadwick, P., & Proudlove, J. (2000). Toward an assessment of suitability of people with mental retardation for cognitive therapy. *Cognitive Therapy and Research*, 24(6), 627-636.
- Dagnan, D., Chadwick, P., Stenfert Kroese, B., Dagnan, D., & Loumidis, K. (1997). Components of cognitive therapy with people with learning disabilities. In B. Stenfert Kroese, D. Dagnan, & K. Loumidis (Eds.), *Cognitive Therapy for People With Learning Disabilities* (pp. 110-123). London: Routledge.
- Dagnan, D., Mellor, K., & Jefferson, C. (2009). Assessment of cognitive therapy skills for people with learning disabilities. *Advances in Mental Health and Learning Disabilities*, 3(4), 25-30.
- Dagnan, D., & Proudlove, J. (1997). Using Makaton drawings to assess the ability to recognize facial expression of emotion in people with learning disabilities. *Clinical Psychology Forum*, 3-5.
- Debelak, R., & Tran, U. S. (2013). Principal component analysis of smoothed tetrachoric correlation matrices as a measure of dimensionality. *Educational and Psychological Measurement*, 73(1), 63-77.
- Dening, K. H., Jones, L., & Sampson, E. L. (2012). Preferences for end-of-life care: a nominal group study of people with dementia and their family carers. *Palliative Medicine*, 27(5), 409-417.
- Doherr, L., Reynolds, S., Wetherly, J., & Evans, E. H. (2005). Young children's ability to engage in cognitive therapy tasks: associations with age and educational experience. *Behavioural and Cognitive Psychotherapy*, 33(2), 201-215. doi:10.1017/S1352465804001894
- Dorenlot, P., Harboun, M., Bige, V., Henrard, J. C., & Ankri, J. (2005). Major depression as a risk factor for early institutionalization of dementia patients living in the community. *International Journal of Geriatric Psychiatry*, 20(5), 471-478.
- Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology*, 70(4), 662-680.
- Dunbar, M., Ford, G., Hunt, K., & Der, G. (2000). A confirmatory factor analysis of the Hospital Anxiety and Depression scale: comparing empirically and theoretically derived structures. *British Journal of Clinical Psychology*, 39(1), 79-94.
- Dykiert, D., & Deary, I. J. (2013). Retrospective validation of WTAR and NART scores as estimators of prior cognitive ability using the Lothian Birth Cohort 1936. *Psychological Assessment*, 25(4), 1361-1366.
- Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: a methodological review. *Clinical Psychology Review*, 22(6), 789-832.
- Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., . . . Gauthier, S. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*, 22(12), 1689-1707.
- Enache, D., Winblad, B., & Aarsland, D. (2011). Depression in dementia: epidemiology, mechanisms, and treatment. *Current Opinion in Psychiatry*, 24(6), 461-472.

- Eslinger, P. J., Moore, P., Antani, S., Anderson, C., & Grossman, M. (2012). Apathy in frontotemporal dementia: behavioral and neuroimaging correlates. *Behavioural Neurology*, *25*(2), 127-136.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299.
- Feher, E. P., Larrabee, G. J., & Crook, T. H. (1992). Factors attenuating the validity of the Geriatric Depression Scale in a dementia population. *Journal of the American Geriatrics Society*, *40*(9), 906-909.
- Field, A. (2009). *Discovering Statistics Using SPSS*. UK: Sage
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. UK: Sage.
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In R. M. Gregory Hancock (Ed.), *Structural Equation Modeling: A Second Course* (pp. 269-314).
- Fletcher, T. D. (2012). QuantPsyc: Quantitative Psychology Tools (Version 1.5). Retrieved from <https://cran.r-project.org/package=QuantPsyc>
- 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.
- Freeman, R. Q., Giovannetti, T., Lamar, M., Cloud, B. S., Stern, R. A., Kaplan, E., & Libon, D. J. (2000). Visuoconstructional problems in dementia: contribution of executive systems functions. *Neuropsychology*, *14*(3), 415-426.
- Frewen, P., Lundberg, E., MacKinley, J., & Wrath, A. (2011). Assessment of response to mindfulness meditation: meditation breath attention scores in association with subjective measures of state and trait mindfulness and difficulty letting go of depressive cognition. *Mindfulness*, *2*(4), 254-269.
- García-Casal, J. A., Goñi-Imizcoz, M., Perea-Bartolomé, M., Soto-Pérez, F., Smith, S. J., Calvo-Simal, S., & Franco-Martín, M. (2017). The efficacy of emotion recognition rehabilitation for people with Alzheimer's disease. *Journal of Alzheimer's Disease*, *57*(3), 937-951.
- Gibbons, L. E., Teri, L., Logsdon, R., McCurry, S. M., Kukull, W., Bowen, J., . . . Larson, E. (2002). Anxiety symptoms as predictors of nursing home placement in patients with Alzheimer's disease. *Journal of Clinical Geropsychology*, *8*(4), 335-342.
- González-Salvador, T., Lyketsos, C. G., Baker, A., Hovanec, L., Roques, C., Brandt, J., & Steele, C. (2000). Quality of life in dementia patients in long-term care. *International Journal of Geriatric Psychiatry*, *15*(2), 181-189.
- Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annual Review of Psychology*, *60*, 549-576.
- Greenberger, D., & Padesky, C. A. (1995). *Mind over mood: A Cognitive Therapy Treatment Manual for Clients*. New York, USA: Guilford press.
- Greene, J. D., Hodges, J. R., & Baddeley, A. D. (1995). Autobiographical memory and executive function in early dementia of Alzheimer type. *Neuropsychologia*, *33*(12), 1647-1670.
- Grimmer, T., Beringer, S., Kehl, V., Alexopoulos, P., Busche, A., Förstl, H., . . . Peters, H. (2015). Trends of patient referral to a memory clinic and towards earlier diagnosis from 1985–2009. *International Psychogeriatrics*, *27*(12), 1939-1944.

- Gur, R. C., Erwin, R. J., Gur, R. E., Zvil, A. S., Heimberg, C., & Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Research, 42*(3), 241-251.
- Guss, R., Middleton, J., Beanland, T., Slade, L., Moniz-Cook, E., Watts, S., & Bone, A. (2014). *Clinical Psychology in the Early Stage Dementia Care Pathway*. London: The British Psychological Society.
- Harter, L. (2003). *Assessment of the Suitability of Older Adults With Dementia for Cognitive Therapy*. (D. Clin. Psy. Clinical Psychology Doctoral Thesis), University of Lancaster.
- Hartley, S. L., Esbensen, A. J., Shalev, R., Vincent, L. B., Mihaila, I., & Bussanich, P. (2015). Cognitive behavioral therapy for depressed adults with mild intellectual disability: a pilot study. *Journal of Mental Health Research in Intellectual Disabilities, 8*(2), 72-97. doi:10.1080/19315864.2015.1033573
- Haugan, G., & Drageset, J. (2014). The Hospital Anxiety and Depression Scale—dimensionality, reliability and construct validity among cognitively intact nursing home patients. *Journal of Affective Disorders, 165*, 8-15.
- Hebblethwaite, A., Jahoda, A., & Dagnan, D. (2011). Talking about real-life events: an investigation into the ability of people with intellectual disabilities to make links between their beliefs and emotions within dialogue. *Journal of Applied Research in Intellectual Disabilities, 24*(6), 543-553. doi:10.1111/j.1468-3148.2011.00637.x
- Helvik, A.-S., Engedal, K., Skancke, R. H., & Selbæk, G. (2011). A psychometric evaluation of the Hospital Anxiety and Depression Scale for the medically hospitalized elderly. *Nordic Journal of Psychiatry, 65*(5), 338-344.
- Hirschfeld, G., & Von Brachel, R. (2014). Multiple-Group confirmatory factor analysis in R-A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation, 19*(7), 1-12.
- Hjermstad, M. J., Fayers, P. M., Haugen, D. F., Caraceni, A., Hanks, G. W., Loge, J. H., . . . Collaborative, E. P. C. R. (2011). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *Journal of Pain and Symptom Management, 41*(6), 1073-1093.
- Hofmann, R. J. (1978). Complexity and simplicity as objective indices descriptive of factor solutions. *Multivariate Behavioral Research, 13*(2), 247-250.
- Hofmann, S. G., & Smits, J. A. (2008). Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry, 69*(4), 621-632.
- Hollon, S. D., & Ponniah, K. (2010). A review of empirically supported psychological therapies for mood disorders in adults. *Depression and Anxiety, 27*(10), 891-932.
- Hsieh, S., Irish, M., Daveson, N., Hodges, J. R., & Piguet, O. (2013). When one loses empathy: its effect on carers of patients with dementia. *Journal of Geriatric Psychiatry and Neurology, 26*(3), 174-184.
- Hsieh, S., Schubert, S., Hoon, C., Mioshi, E., & Hodges, J. R. (2013). Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 36*(3), 242-250.
- Hu, L. T., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*(2), 351-362.

- Hundt, N. E., Mignogna, J., Underhill, C., & Cully, J. A. (2013). The relationship between use of cbt skills and depression treatment outcome: a theoretical and methodological review of the literature. *Behavior Therapy, 44*(1), 12-26. doi:10.1016/j.beth.2012.10.001
- Izard, C. E., Libero, D. Z., Putnam, P., & Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology, 64*(5), 847-860.
- Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., . . . Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*(2), 295-304.
- Jarrett, R. B., Minhajuddin, A., Kangas, J. L., Friedman, E. S., Callan, J. A., & Thase, M. E. (2013). Acute phase cognitive therapy for recurrent major depressive disorder: who drops out and how much do patient skills influence response? *Behaviour Research and Therapy, 51*(4-5), 221-230. doi:10.1016/j.brat.2013.01.006
- Johnco, C., Wuthrich, V. M., & Rapee, R. M. (2014). The influence of cognitive flexibility on treatment outcome and cognitive restructuring skill acquisition during cognitive behavioural treatment for anxiety and depression in older adults: results of a pilot study. *Behaviour Research and Therapy, 57*(1), 55-64. doi:10.1016/j.brat.2014.04.005
- Join Dementia Research. (2016). Retrieved from <https://www.joindementiaresearch.nihr.ac.uk/>
- Jones, C., Hacker, D., Cormac, I., Meaden, A., & Irving, C. B. (2012). *Cognitive Behavioural therapy Versus Otther Psychosocial Treatments for Schizophrenia (1465-1858)*. Retrieved from <http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD008712/full>
- Joyce, T., Globe, A., & Moody, C. (2006). Assessment of the component skills for cognitive therapy in adults with intellectual disability. *Journal of Applied Research in Intellectual Disabilities, 19*(1), 17-23. doi:10.1111/j.1468-3148.2005.00287.x
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593-602.
- Kohler, C. G., Anselmo-Gallagher, G., Bilker, W., Karlawish, J., Gur, R. E., & Clark, C. M. (2005). Emotion-discrimination deficits in mild Alzheimer disease. *The American Journal of Geriatric Psychiatry, 13*(11), 926-933.
- Kohler, C. G., Bilker, W., Hagendoorn, M., Gur, R. E., & Gur, R. C. (2000). Emotion recognition deficit in schizophrenia: association with symptomatology and cognition. *Biological Psychiatry, 48*(2), 127-136.
- Korczyn, A. D. (2002). Mixed dementia—the most common cause of dementia. *Annals of the New York Academy of Sciences, 977*(1), 129-134.
- Korczyn, A. D., & Halperin, I. (2009). Depression and dementia. *Journal of the Neurological Sciences, 283*(1), 139-142.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877-883.

- Kraus, C. A., Seignourel, P., Balasubramanyam, V., Snow, A. L., Wilson, N. L., Kunik, M. E., . . . Stanley, M. A. (2008). Cognitive-behavioral treatment for anxiety in patients with dementia: two case studies. *Journal of Psychiatric Practice, 14*(3), 186-192.
- Krupnick, J. L., Sotsky, S. M., Simmens, S., Moyer, J., Elkin, I., Watkins, J., & Pilkonis, P. A. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 64*(3), 532-539.
- Kumfor, F., & Piguet, O. (2013). Emotion recognition in the dementias: brain correlates and patient implications. *Neurodegenerative Disease Management, 3*(3), 277-288.
- Laidlaw, K., Thompson, L. W., Gallagher-Thompson, D., & Dick-Siskin, L. (2015). *Cbt for Older People*. London UK: SAGE.
- Lanctôt, K. L., Herrmann, N., Yau, K. K., Khan, L. R., Liu, B. A., LouLou, M. M., & Einarson, T. R. (2003). Efficacy and safety of cholinesterase inhibitors in Alzheimer's disease: a meta-analysis. *Canadian Medical Association Journal, 169*(6), 557-564.
- Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics, 33*, 363-374.
- Lane, R. M., Potkin, S. G., & Enz, A. (2006). Targeting acetylcholinesterase and butyrylcholinesterase in dementia. *International Journal of Neuropsychopharmacology, 9*(1), 101-124.
- Law, E., Connelly, P. J., Randall, E., McNeill, C., Fox, H. C., Parra, M. A., . . . Gray, S. (2013). Does the Addenbrooke's Cognitive Examination-Revised add to the Mini-Mental State Examination in established Alzheimer disease? Results from a national dementia research register. *International Journal of Geriatric Psychiatry, 28*(4), 351-355.
- Lickel, A., MacLean, W. E., Jr., Blakeley-Smith, A., & Hepburn, S. (2012). Assessment of the prerequisite skills for cognitive behavioral therapy in children with and without autism spectrum disorders. *Journal of Autism & Developmental Disorders, 42*(6), 992-1000. doi:10.1007/s10803-011-1330-x
- Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Mukadam, N. (2017). Dementia prevention intervention and care. *The Lancet, 390*(10113), 2673-2734.
- Lopez, O. L., Becker, J. T., Sweet, R. A., Klunk, W., Kaufer, D. I., Saxton, J., . . . DeKosky, S. T. (2003). Psychiatric symptoms vary with the severity of dementia in probable Alzheimer's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences, 15*(3), 346-353.
- Marwood, H., Chinn, D., Gannon, K., & Scior, K. (2018). The experiences of high intensity therapists delivering cognitive behavioural therapy to people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 31*(1), 76-86.
- McKeith, I. G. (2006). Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the Consortium on DLB International Workshop. *Journal of Alzheimer's Disease, 9*(s3), 417-423.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., . . . Mayeux, R. (2011). The diagnosis of dementia due to

- Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263-269.
- McNamee, R. (2003). Confounding and confounders. *Occupational and Environmental Medicine*, 60(3), 227-234.
- Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1), 1-42.
- Miranda-Castillo, C., Woods, B., & Orrell, M. (2010). People with dementia living alone: what are their needs and what kind of support are they receiving? *International Psychogeriatrics*, 22(4), 607-617.
- Mohlman, J. (2008). More power to the executive? a preliminary test of CBT plus executive skills training for treatment of late-life GAD. *Cognitive and Behavioral Practice*, 15(3), 306-316. doi:10.1016/j.cbpra.2007.07.002
- Mohlman, J. (2013). Executive skills in older adults with GAD: relations with clinical variables and CBT outcome. *Journal of Anxiety Disorders*, 27(1), 131-139. doi:10.1016/j.janxdis.2012.12.001
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737-745.
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., . . . Wight, D. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *British Medical Journal*, 350, h1258.
- Moss-Morris, R., Dennison, L., Landau, S., Yardley, L., Silber, E., & Chalder, T. (2013). A randomized controlled trial of cognitive behavioral therapy (CBT) for adjusting to multiple sclerosis (the saMS trial): does CBT work and for whom does it work? *Journal of Consulting and Clinical Psychology*, 81(2), 251-262.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699.
- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S. a., . . . Albert, M. (1998). Frontotemporal lobar degeneration. A consensus on clinical diagnostic criteria. *Neurology*, 51(6), 1546-1554.
- Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2), 199-203.
- Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test (NART)*: Nfer-Nelson Windsor.
- NHS Digital. (2016). *Psychological Therapies: Annual Report on the Use of IAPT Services 2015-16, England*. Retrieved from <http://www.content.digital.nhs.uk/catalogue/PUB22110/psych-ther-ann-rep-2015-16.pdf>
- NICE. (2018). *Dementia: Assessment, Management and support for People Living With Dementia and Their Carers (NG97)*. Retrieved from <https://www.nice.org.uk/guidance/ng97>
- Norton, S., Cosco, T., Doyle, F., Done, J., & Sacker, A. (2013). The Hospital Anxiety and Depression Scale: a meta confirmatory factor analysis.

- Journal of Psychosomatic Research*, 74(1), 74-81.  
doi:10.1016/j.jpsychores.2012.10.010
- Oathamshaw, S. C., Barrowcliff, A. L., & Haddock, G. (2012). Cognitive-behavioral therapy for people with intellectual disabilities and psychosis. In J. Taylor, W. Lindsay, R. Hasings, & C. Hatton (Eds.), *Psychological Therapies for Adults with Intellectual Disabilities* (pp. 157-172): John Wiley and Sons.
- Oathamshaw, S. C., & Haddock, G. (2006). Do people with intellectual disabilities and psychosis have the cognitive skills required to undertake cognitive behavioural therapy? *Journal of Applied Research in Intellectual Disabilities*, 19(1), 35-46. doi:10.1111/j.1468-3148.2005.00284.x
- Orgeta, V., & Phillips, L. H. (2007). Effects of age and emotional intensity on the recognition of facial emotion. *Experimental Aging Research*, 34(1), 63-79.
- Orgeta, V., Qazi, A., Spector, A. E., & Orrell, M. (2014). *Psychological Treatments for Depression and Anxiety in Dementia and Mild Cognitive Impairment*. Retrieved from <http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD009125.pub2/pdf>
- Ornstein, K., & Gaugler, J. E. (2012). The problem with “problem behaviors”: a systematic review of the association between individual patient behavioral and psychological symptoms and caregiver depression and burden within the dementia patient–caregiver dyad. *International Psychogeriatrics*, 24(10), 1536-1552.
- Orrell, M., Hoe, J., Charlesworth, G., Russell, I., Challis, D., Moniz-Cook, E., . . . D'Amico, F. (2017). *Support at Home: Interventions to Enhance Life in Dementia (SHIELD)–Evidence, evelopment and Evaluation of Complex Interventions*. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK424174/>
- Padesky, C. A., & Greenberger, D. (2012). *Clinician's Guide to Mind Over Mood*. UK: Guilford Press.
- Pearson, R. H., & Mundform, D. J. (2010). Recommended sample size for conducting exploratory factor analysis on dichotomous data. *Journal of Modern Applied Statistical Methods*, 9(2), 359-368.
- Pernecky, R., Wagenpfeil, S., Komossa, K., Grimmer, T., Diehl, J., & Kurz, A. (2006). Mapping scores onto stages: mini-mental state examination and clinical dementia rating. *The American Journal of Geriatric Psychiatry*, 14(2), 139-144.
- Pettersson, E., & Turkheimer, E. (2010). Item selection, evaluation, and simple structure in personality data. *Journal of Research in Personality*, 44(4), 407-420.
- Piguet, O., Kumfor, F., & Hodges, J. (2017). Diagnosing, monitoring and managing behavioural variant frontotemporal dementia. *The Medical Journal of Australia*, 207(7), 303-308.
- Potkins, D., Myint, P., Bannister, C., Tadros, G., Chithramohan, R., Swann, A., . . . Ballard, C. (2003). Language impairment in dementia: impact on symptoms and care needs in residential homes. *International Journal of Geriatric Psychiatry*, 18(11), 1002-1006.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93-115.

- Prince, M., Ali, G.-C., Guerchet, M., Prina, A. M., Albanese, E., & Wu, Y.-T. (2016). Recent global trends in the prevalence and incidence of dementia, and survival with dementia. *Alzheimer's Research & Therapy*, 8(1), 8-23.
- Prince, M., Knapp, M., Guerchet, M., McCrone, P., Prina, M., Comas-Herrera, A., . . . King, D. (2014). *Dementia UK: Update*. Retrieved from [https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/dementia\\_uk\\_update.pdf](https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/dementia_uk_update.pdf)
- Quakley, S., Coker, S., Palmer, K., & Reynolds, S. (2003). Can children distinguish between thoughts and behaviours? *Behavioural and Cognitive Psychotherapy*, 31(2), 159-168. doi:10.1017/S1352465803002030
- Quakley, S., Reynolds, S., & Coker, S. (2004). The effect of cues on young children's abilities to discriminate among thoughts, feelings and behaviours. *Behaviour Research and Therapy*, 42(3), 343-356. doi:10.1016/S0005-7967(03)00145-1
- R Core Team. (2013). R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>.
- Razavi, D., Delvaux, N., Farvacques, C., & Robaye, E. (1990). Screening for adjustment disorders and major depressive disorders in cancer in-patients. *The British Journal of Psychiatry*, 156, 79-83.
- Reed, J., & Clements, J. (1989). Assessing the understanding of emotional states in a population of adolescents and young adults with mental handicaps. *Journal of Mental Deficiency Research*, 33(3), 229-233.
- Revelle, W. (2017). Psych: Procedures for Psychological, Psychometric, and Personality Research (Version 1.7.5). Retrieved from <https://cran.r-project.org/package=psych>
- Rollnick, S. (1998). Readiness, importance, and confidence: Critical conditions of change in treatment. In N. H. W. R. Miller (Ed.), *Applied Clinical Psychology. Treating Addictive Behaviours* (pp. 49-60). New York, US: Plenum Press.
- Román, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J., Garcia, J. a., . . . Hofman, A. (1993). Vascular dementia diagnostic criteria for research studies: report of the NINDS-AIREN International Workshop. *Neurology*, 43(2), 250-250.
- Rosen, H. J., Alcantar, O., Zakrzewski, J., Shimamura, A. P., Neuhaus, J., & Miller, B. L. (2014). Metacognition in the behavioral variant of frontotemporal dementia and Alzheimer's disease. *Neuropsychology*, 28(3), 436-447.
- Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36.
- Roth, A. D., & Pilling, S. (2008). Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy*, 36(2), 129-147. doi:10.1017/S1352465808004141
- Safran, J. D., Segal, Z. V., Vallis, T., Shaw, B. F., & Samstag, L. W. (1993). Assessing patient suitability for short-term cognitive therapy with an interpersonal focus. *Cognitive Therapy and Research*, 17(1), 23-38. doi:10.1007/BF01172738
- Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology*, 60, 257-282.

- Sams, K., Collins, S., & Reynolds, S. (2006). Cognitive therapy abilities in people with learning disabilities. *Journal of Applied Research in Intellectual Disabilities, 19*(1), 25-33. doi:10.1111/j.1468-3148.2006.00303.x
- Seignourel, P. J., Kunik, M. E., Snow, L., Wilson, N., & Stanley, M. (2008). Anxiety in dementia: a critical review. *Clinical Psychology Review, 28*(7), 1071-1082.
- SemTools Contributors. (2015). emTools: Useful Tools for Structural Equation Modeling. (Version 0.4-14). Retrieved from <http://cran.r-project.org/package=semTools>
- Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., Sugai, Y., & Kohno, M. (2001). Influence of deteriorating ability of emotional comprehension on interpersonal behavior in Alzheimer-type dementia. *Brain and Cognition, 47*(3), 423-433.
- Shiota, M. N., & Levenson, R. W. (2009). Effects of aging on experimentally instructed detached reappraisal, positive reappraisal, and emotional behavior suppression. *Psychology and Aging, 24*(4), 890-900.
- Smith, S., Lamping, D., Banerjee, S., Harwood, R., Foley, B., Smith, P., . . . Levin, E. (2005). *Measurement of Health-Related Quality of Life for People With Dementia: Development of a New Instrument (DEMQOL) and an Evaluation of Current Methodology* (1366-5278). Retrieved from <https://researchonline.lshtm.ac.uk/13685/1/FullReport-hta9100.pdf>
- Spector, A., Charlesworth, G., King, M., Lattimer, M., Sadek, S., Marston, L., . . . Knapp, M. (2015). Cognitive-behavioural therapy for anxiety in dementia: pilot randomised controlled trial. *The British Journal of Psychiatry, 206*(6), 509-516.
- Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science, 253*(5026), 1380-1386.
- Srikanth, S., Nagaraja, A., & Ratnavalli, E. (2005). Neuropsychiatric symptoms in dementia-frequency, relationship to dementia severity and comparison in Alzheimer's disease, vascular dementia and frontotemporal dementia. *Journal of the Neurological Sciences, 236*(1), 43-48.
- Stanley, M. A., Calleo, J., Bush, A. L., Wilson, N., Snow, A. L., Kraus-Schuman, C., . . . Schulz, P. E. (2013). The Peaceful Mind program: a pilot test of a cognitive-behavioral therapy-based intervention for anxious patients with dementia. *The American Journal of Geriatric Psychiatry, 21*(7), 696-708.
- Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology, 43*(2), 476-493.
- Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease & Associated Disorders, 20*, S69-S74.
- Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*(10), 2015-2028.
- Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology, 11*(11), 1006-1012.
- Stewart, A. L., Thrasher, A. D., Goldberg, J., & Shea, J. A. (2012). A framework for understanding modifications to measures for diverse populations. *Journal of Aging and Health, 24*(6), 992-1017.
- Stott, J., Charlesworth, G., & Scior, K. (2017). Measures of readiness for cognitive behavioural therapy in people with intellectual disability: A systematic review. *Research in Developmental Disabilities, 60*, 37-51.

- Stott, J., Orrell, M., & Charlesworth, G. (2017). Validity and usefulness of the Hospital Anxiety and Depression Scale in carers of people with dementia: evidence from confirmatory factor analysis, concurrent validity, and measurement invariance in a large sample. *The American Journal of Geriatric Psychiatry*, 25(11), 1270-1280.
- Stott, J., Scior, K., Mandy, W., & Charlesworth, G. (2017). Dementia screening accuracy is robust to premorbid IQ variation: evidence from the Addenbrooke's Cognitive Examination-III and the Test of Premorbid Function. *Journal of Alzheimer's Disease*, 57(4), 1293-1302.
- Stott, J., Spector, A., Orrell, M., Scior, K., Sweeney, J., & Charlesworth, G. (2017). Limited validity of the Hospital Anxiety and Depression Scale (HADS) in dementia: evidence from a confirmatory factor analysis. *International Journal of Geriatric Psychiatry*, 32(7), 805-813.
- Stott, J., Sweeney, J. M., Koschalka, L., O'Connor, L., & Mwale, A. (2017). People with dementia as peer workers, challenges, and benefits: a thematic analysis and nominal groups study. *International Psychogeriatrics*, 29(7), 1185-1192.
- Straat, J. H., van der Ark, L. A., & Sijtsma, K. (2013). Methodological artifacts in dimensionality assessment of the Hospital Anxiety and Depression scale (HADS). *Journal of Psychosomatic Research*, 74(2), 116-121. doi:10.1016/j.jpsychores.2012.11.012
- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York, USA: Oxford University Press.
- Strunk, D., R., DeRubeis, R., J., Chiu, A., W., & Alvarez, J. (2007). Patients' competence in and performance of cognitive therapy skills: relation to the reduction of relapse risk following treatment for depression. *Journal of Consulting and Clinical Psychology*, 75(4), 523-530. doi:10.1037/0022-006X.75.4.523
- Surcinelli, P., Codispoti, M., Montebanocci, O., Rossi, N., & Baldaro, B. (2006). Facial emotion recognition in trait anxiety. *Journal of Anxiety Disorders*, 20(1), 110-117.
- Teri, L., Logsdon, R. G., Uomoto, J., & McCurry, S. M. (1997). Behavioral treatment of depression in dementia patients: a controlled clinical trial. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 52(4), 159-166.
- Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology*, 19(2), 203-214.
- Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 21(1), 19-25.
- Tsimopoulou, I., Kroese, B. S., Unwin, G., Azmi, S., & Jones, C. (2018). A case series to examine whether people with learning disabilities can learn prerequisite skills for cognitive behavioural therapy. *The Cognitive Behaviour Therapist*, 11(e1), 1-22.
- Van de Ven, A. H., & Delbecq, A. L. (1972). The nominal group as a research instrument for exploratory health studies. *American Journal of Public Health*, 62(3), 337-342.
- Vereenoghe, L., Gega, L., Reynolds, S., & Langdon, P. E. (2016). Using computers to teach people with intellectual disabilities to perform some of the tasks used within cognitive behavioural therapy: a randomised

- experiment. *Behaviour Research and Therapy*, 76, 13-23.  
doi:10.1016/j.brat.2015.11.002
- Verhoeven, K. J., Simonsen, K. L., & McIntyre, L. M. (2005). Implementing false discovery rate control: increasing your power. *Oikos*, 108(3), 643-647.
- Vernooij-Dassen, M., Draskovic, I., McCleery, J., & Downs, M. (2011). *Cognitive Reframing for Carers of People With Dementia* (1465-1858). Retrieved from <http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD005318.pub2/full>
- Walker, D. A. (2004). Cognitive behavioural therapy for depression in a person with Alzheimer's dementia. *Behavioural and Cognitive Psychotherapy*, 32(4), 495-500.
- Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14(3), 270-277.
- Wang, W., Lopez, V., & Martin, C. R. (2006). Structural ambiguity of the Chinese version of the Hospital Anxiety and Depression Scale in patients with coronary heart disease. *Health and Quality of Life Outcomes*, 4(6), 1-5. doi:10.1186/1477-7525-4-6
- Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2009). *Wechsler Memory Scale: WMS-IV; Technical and Interpretive Manual*. San Antonio, TX: Pearson.
- Wechsler, D. (2011). *Test of Premorbid Functioning. UK Version (TOPF UK)*. UK: Pearson Corporation.
- Wechsler, D., Coalson, D. L., & Raiford, S. E. (2008). *WAIS-IV: Wechsler Adult Intelligence Scale*. San Antonio, TX: Pearson
- Weersing, V. R., & Weisz, J. R. (2002). Community clinic treatment of depressed youth: benchmarking usual care against CBT clinical trials. *Journal of Consulting and Clinical Psychology*, 70(2), 299-310.
- Wells, C. S., & Wollack, J. A. (2003). *An Instructor's Guide to Understanding Test Reliability*. Retrieved from <https://testing.wisc.edu/Reliability.pdf>
- WHO. (1993). *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organisation.
- Willner, P. (2006). Readiness for cognitive therapy in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 19(1), 5-16. doi:10.1111/j.1468-3148.2005.00280.x
- Winblad, B., Jones, R. W., Wirth, Y., Stöfler, A., & Möbius, H. J. (2007). Memantine in moderate to severe Alzheimer's disease: a meta-analysis of randomised clinical trials. *Dementia and Geriatric Cognitive Disorders*, 24(1), 20-27.
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913-934.
- Wolitzky-Taylor, K. B., Castriotta, N., Lenze, E. J., Stanley, M. A., & Craske, M. G. (2010). Anxiety disorders in older adults: a comprehensive review. *Depression and Anxiety*, 27(2), 190-211.
- Wortmann, M. (2012). Dementia: a global health priority-highlights from an ADI and World Health Organization report. *Alzheimer's Research & Therapy*, 4(5), 40.
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*, 67(6), 361-370.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's  $\alpha$ , Revelle's  $\beta$ , and McDonald's  $\omega$  H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1), 123-133.

## Appendices

## **Appendix A: Guidance on Undertaking a PhD While Supervising the Research of Professional Doctorate Students**

1. There are many advantages to undertaking doctoral research as part of a collaborative team and this is encouraged. However in these circumstances it is essential that the contribution of each party and the way in which the thesis meets the following criteria (which apply to all doctoral programmes) is explicitly stated in a declaration and submitted with the thesis.

*The thesis will make a distinct contribution to the knowledge of the subject and will afford evidence of originality as shown by the discovery of new facts and/or the exercise of independent critical power.*

In the case of Professional Doctorate students the declaration should be signed by each of the students involved in the project and their supervisor.

The same examiner will be appointed for these theses.

In the case of staff undertaking a PhD, the declaration should be signed by the staff member themselves and their supervisor and the declarations from all Professional Doctoral thesis based on data which overlaps at all with data reported in the PhD thesis should be submitted with the declaration. The examiners of the PhD thesis should be advised that these Professional Doctoral theses are available to them to consult at their request.

2. In planning their thesis work, team members should ensure that no studies are planned which involve completely overlapping data. For example Professional Doctorate student 1 might collect data on variables A, B and C in Year 1, Professional Doctorate student 2 might collect data on

variables A, B and D in Year 2 and the staff member might analyse longitudinal data on the variables A and B in a PhD thesis study.

- 3.** For the PhD upgrading the staff member should, in addition to the other documentation required, submit a draft of the declaration they envisage submitting with their thesis so that any questions that need to be resolved can be addressed at this stage and plans with the regard to use of shared data can be formally approved.

## **Appendix B Declarations of Overlap.**

My PhD thesis contains overlapping data with three DClinPsy theses that I supervised, as such it is part of a collaborative endeavour. In such instances we are required to follow the guidance created by Norah Frederikson, Professor of Educational Psychology at UCL, to cover cases of overlap between professional doctorates and PhDs.

This guidance requires that I make a declaration that my PhD thesis will 'make a distinct contribution to the knowledge of the subject and will afford evidence of originality as shown by the discovery of new facts and/or the exercise of independent critical power.'

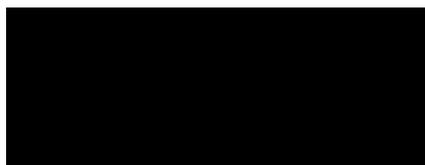
I confirm that this is the case. The questions asked in the DClinPsy theses relate to the diagnostic utility of the four mountains test (Lucy Gore), mindfulness abilities in people with dementia (Catherine Bousfield) and differences in mindfulness abilities between older and younger adults (Noor Habib). Consequently, I confirm that the questions asked in my PhD and those addressed in the three theses are completely separate questions

Furthermore, the guidance requires that I clarify that while my thesis does have some data in common with all three theses, there is not completely overlapping data. None of the three DClinPsy theses separately or combined contain all the participants or measures that are used in my PhD. Consequently, I confirm that my PhD thesis does not contain completely overlapping data with one or all of the DClinPsy theses.

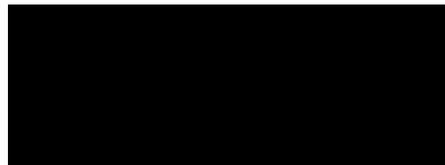
The guidance requires that the professional doctoral theses be made available on request, Lucy Gore, Catherine Bousfield and Noor Habib have all confirmed that they are happy for this to happen. As also required in the guidance the declarations of overlapping data from the three DClinPsy theses are enclosed

here and in each case the students confirm that they are happy for their theses to be made available upon request.

The guidance requires that my supervisor and I sign this declaration to confirm it is true and accurate to the best of our knowledge and we have done so below.



Joshua Stott  
PhD student/Research Fellow  
Date: 23/07/2018



Georgina Charlesworth  
Senior Lecturer  
Date: 25/07/2018

## **Lucy Gore: Joint Project Submission Declaration**

My thesis and my supervisor (Joshua Stott's) PhD thesis (which will be submitted in the future) contain some overlapping data and are part of a collaborative endeavour. In these cases we are required to follow the guidance created by Norah Frederikson, Professor of Educational Psychology at UCL, to cover cases of overlap between professional doctorates and PhDs.

This guidance requires that I make a declaration that my thesis will 'make a distinct contribution to the knowledge of the subject and will afford evidence of originality as shown by the discovery of new facts and/or the exercise of independent critical power.'

I confirm that this is the case and that the questions asked in my thesis and that which will be addressed in my supervisor's PhD are completely separate questions.

Furthermore the guidance requires that I clarify that while the two theses do have some data in common they do not contain completely overlapping datasets. I confirm that this is the case.

Finally it requires that I confirm that I am happy for my thesis to be made confirm that I am happy for this to happen.

The guidance requires that my supervisor and I sign this declaration to confirm it is true and accurate to the best of our knowledge and we have done so below.

Lucy Gore

Trainee Psychologist

19/06/2015

Joshua Stott

Senior Clinical Tutor

19/06/2015

## Catherine Bousfield: Joint project submission declaration

The D.Clin.Psy projects of Noor Habib and Catherine Bousfield and the PhD project of our supervisor Dr Joshua Stott contain overlapping data. This is a declaration that we have followed UCL guidance on overlapping research projects created by Norah Fredericson, Professor of Educational Psychology.

I am required to make a clear declaration that my thesis will 'make a distinct contribution to the knowledge of the subject and will afford evidence of originality as shown by the discovery of new facts and/or the exercise of the independent critical power.'

I confirm that my research questions are coherently different from those asked by my colleague Noor and my supervisor Joshua.

I also confirm that although the three projects have common data, they do not completely overlap.

As stated in my empirical paper, I collected all the data for my dementia sample, which is a subset of the data that Joshua Stott will be using for his dementia sample. The data I share with Noor Habib is that which she collected for the older adult comparison sample. This is used in my study for a clearly defined and limited purpose.

I am finally required to confirm that I agree for my thesis to be made available upon request to examiners of other theses and I confirm that I agree for this to happen.

The guidance states that my colleague, supervisor, and I sign this declaration to confirm the accuracy of the given information to our best knowledge.



Noor Habib  
Trainee Clinical Psychologist

20/6/16



Catherine Bousfield  
Trainee Clinical Psychologist

20/06/2016



Dr. Joshua Stott  
Senior Clinical Tutor

17/06/2016

## Noor Habib: Joint project submission declaration

Appendix : Joint Project Submission Declaration

Declaration of overlapping projects

The D.Clin.Psy projects of Noor Habib and Catherine B and the PhD project of our supervisor Dr Joshua Stott contain overlapping data. This is a declaration that we have followed UCL guidance of overlapping research projects created by Norah Fredericson, Professor of Educational Psychology.

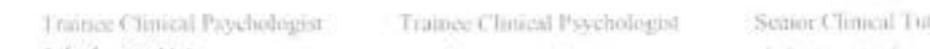
I am required to make a clear declaration that my thesis will "make a distinct contribution to the knowledge of the subject and will afford evidence of originality as shown by the discovery of new facts and/or the exercise of the independent critical power."

I confirm that my research questions are coherently different from those asked by my colleague Catherine and my supervisor Joshua.

I also confirm that although the three projects have common data, they do not completely overlap.

I am finally required to confirm that I agree for my thesis to be made available upon request to examiners of other theses and I confirm that I agree for this to happen.

The guidance states that my colleague, supervisor and myself sign this declaration to confirm the accuracy of the given information to our best knowledge.

		
Noor Habib	Catherine Housfield	Dr. Joshua Stott
Trainee Clinical Psychologist	Trainee Clinical Psychologist	Senior Clinical Tutor
16/06/16	16/6/16	16/06/2016

**Appendix C: Measures of Readiness for Cognitive Behavioural Therapy in People with Intellectual Disability: A Systematic Review<sup>4</sup>**

Joshua Stott<sup>a\*</sup>, Georgina Charlesworth<sup>a</sup>, Katrina Scior<sup>a</sup>.

<sup>a</sup>Research Department of Clinical, Educational and Health Psychology

University College London WC1E 7HB

\* Corresponding author address contact details as above, telephone 02076795950, email: [j.stott@ucl.ac.uk](mailto:j.stott@ucl.ac.uk)

Key Words, Cognitive Behavioural Therapy, CBT, Readiness, Suitability, Intellectual disabilities, measures.

Word count for total document except title page, abstract, acknowledgments and what this paper adds sections: 8000

---

<sup>4</sup> This is the UCL online repository version of a paper published in a peer reviewed journal (Stott, Charlesworth, et al., 2017)

## **Abstract**

Cognitive behavioural therapy (CBT) is a promising treatment for mental health problems in people with intellectual disabilities but some may not be suited or ready. This review critically evaluates the quality and utility of measures of CBT readiness in people with intellectual disabilities. Twelve studies of six measures based on three aspects of CBT readiness were identified through systematic review. Across measures, measurement quality was largely poor or un-assessed. Only one study evaluated measurement change over the course of CBT. Not all participants with intellectual disabilities could 'pass' readiness measures and performance may be affected by levels of language and cognitive functioning. There was some evidence that CBT readiness is trainable with brief interventions. Before using readiness measures in a clinical context, further work is needed to extend initial evidence on the recognising cognitive mediation as a CBT readiness ability. Given the lack of consensus as to the definition of CBT readiness and the heterogeneity of CBT interventions, future research could also focus on developing readiness measures using a bottom up approach, developing measures within the context of CBT interventions themselves, before further refining and establishing their psychometric properties.

## **What this paper adds?**

This paper is the first to systematically review measures of skills thought necessary to be ready for cognitive behavioural therapy in intellectual disabilities. The findings suggest that while readiness skills may be trainable with brief interventions, the available measures of these skills have not been fully evaluated for quality. Levels of functioning on these measures have yet to be established relative to those without intellectual disabilities and critically, there is very little evidence as to whether these skills are important in cognitive behavioural therapy

process and outcome. We suggest that future research could focus on those constructs where there is preliminary evidence for utility such as recognising cognitive mediation and also on developing the concept of readiness perhaps by developing measures within the context of specific CBT interventions.

Until this is done, clinicians should exercise caution in using these measures to assess readiness for cognitive behavioural therapy in people with intellectual disabilities.

## **Introduction**

Mental health problems are common in people with intellectual disabilities (Cooper, Smiley, Morrison, Williamson, & Allan, 2007) and cognitive behavioural therapy (CBT) is a promising treatment (Vereenoghe & Langdon, 2013). People with intellectual disabilities are a heterogeneous group and CBT is unlikely to be of benefit to all. For those who could use CBT, many may not be 'ready' to do so in an un-adapted form (Dagnan, Chadwick, & Proudlove, 2000). These individuals may need skills training using techniques such as errorless learning to engage in CBT or the therapy may need to be adapted perhaps with greater use of scaffolding (Vygotsky, 1978) to take account of their difficulties (Willner, 2006). In light of this, researchers have sought to measure 'readiness' to help make treatment decisions about how to best adapt CBT or support individuals in accessing it. (Dagnan, Chadwick, Stenfert Kroese, Dagnan, & Loumidis, 1997)

Readiness for CBT has motivational (e.g. expectations of therapy success) and skill components (Willner, 2006). The skill components include those that relate to any talking therapy (e.g. skills in holding a conversation) and aspects specific to CBT (Willner, 2006).

Furthermore, readiness may relate to either behavioural or cognitive elements of CBT (where cognitive refers to ability to change and reflect on thoughts rather than neurocognitive ability and behavioural refers to ability to make behavioural change) (Roth & Pilling, 2008). Readiness skills related to the cognitive elements of CBT are particularly important as these elements are the most cognitively complex elements of CBT and most affected by neurocognitive impairment (Stanley et al., 2013) Furthermore, engagement with cognitive elements is an important aspect of CBT efficacy in people with intellectual

disabilities (McGillivray & Kershaw, 2015). The focus of this review is thus on the cognitive elements of CBT.

CBT is not a unitary therapy, but an umbrella term encompassing interventions which have commonality in drawing on behavioural and cognitive models, but differ as to the precise theoretical framework underpinning them (Roth & Pilling, 2008) and thus have potentially different associated readiness skills (Doherr, Reynolds, Wetherly, & Evans, 2005). In the intellectual disabilities literature, measures of readiness (Dagnan et al., 1997) have focussed on an Antecedent Belief Consequence (ABC) model (Ellis, 1991). This model was originally outlined within Rational Emotive Behaviour Therapy (REBT) rather than CBT. Although REBT differs in application from traditional CBT, particularly in its use of disputation as a therapeutic technique (Ellis, 1980), its theoretical underpinnings have significant overlap with CBT (Ellis, 1980). In particular the ABC model, and its claim that beliefs mediate the relationship between antecedent events and their emotional or behavioural consequences can be seen as a central construct in CBT and REBT (Hyland & Boduszek, 2012). Consequently, it is the ABC model that has informed three skills (at a minimum) being identified as critical to being ready for CBT (Oathamshaw & Haddock, 2006).

These are:

1. Discriminating between emotions, thoughts and behaviours,
2. Making links between emotions and events; and
3. Understanding the mediating role of cognitions between an antecedent event and its consequences.

Consequently, while there may be other cognitive skills necessary for readiness and motivational components will be essential in accessing CBT, the

current review is a critical evaluation of measures of these three skills and findings related to them

As with any tools, useful measures of CBT readiness must have strong psychometric properties (Mokkink et al., 2010). Additionally, for any measure purporting to assess CBT readiness, measurement change should mediate CBT outcome (Hundt, Mignogna, Underhill, & Cully, 2013). It is also important to understand the performance of people with intellectual disabilities on these measures, as CBT should be adapted based on readiness skills that are affected in people with intellectual disabilities rather than those that are unaffected (Oathamshaw & Haddock, 2006). Finally, the trainability of CBT readiness skills is important as this determines adaptation; whether we remove elements from the therapy or train people to increase their skill level (Vereenooghe, Reynolds, Gega, & Langdon, 2015). Consequently, the questions addressed by this review are:

1. What are the measurement properties of tasks assessing the above CBT readiness skills?
2. What is the relationship between performance on measures of these skills and CBT outcome in people with intellectual disabilities?
3. What is the level of performance of people with intellectual disabilities on these measures?
4. What is the evidence for trainability of readiness skills in people with intellectual disabilities?

## **Methods**

### *Search Strategy*

Electronic searches of the following databases: PsycINFO, MEDLINE, SCOPUS were conducted. Search terms were identified based on previous

similar reviews in other populations (Muse & McManus, 2013) and recent reviews of the intellectual disability literature (Davies & Oliver, 2013). Search terms varied slightly according to databases due to differences in the keyword systems used, but were variants of developmental/learning/intellectual disabilities/mental handicap/retardation; Ability/readiness/suitability/preparedness/skills; CBT/Cognitive therapy/Cognitive behavioural therapy, combined using the Boolean terms 'OR' and 'AND'. An example of the full Medline search strategy is given in a supplementary file. Of the 311 papers identified in the initial search, duplicate or irrelevant articles were deleted, leaving 27 papers. After running citation searches in Web of Science and inspecting reference lists of remaining articles, four further papers were added. The full text of all potentially relevant (N=31) articles was reviewed against inclusion and exclusion criteria. See Figure 1 for details.

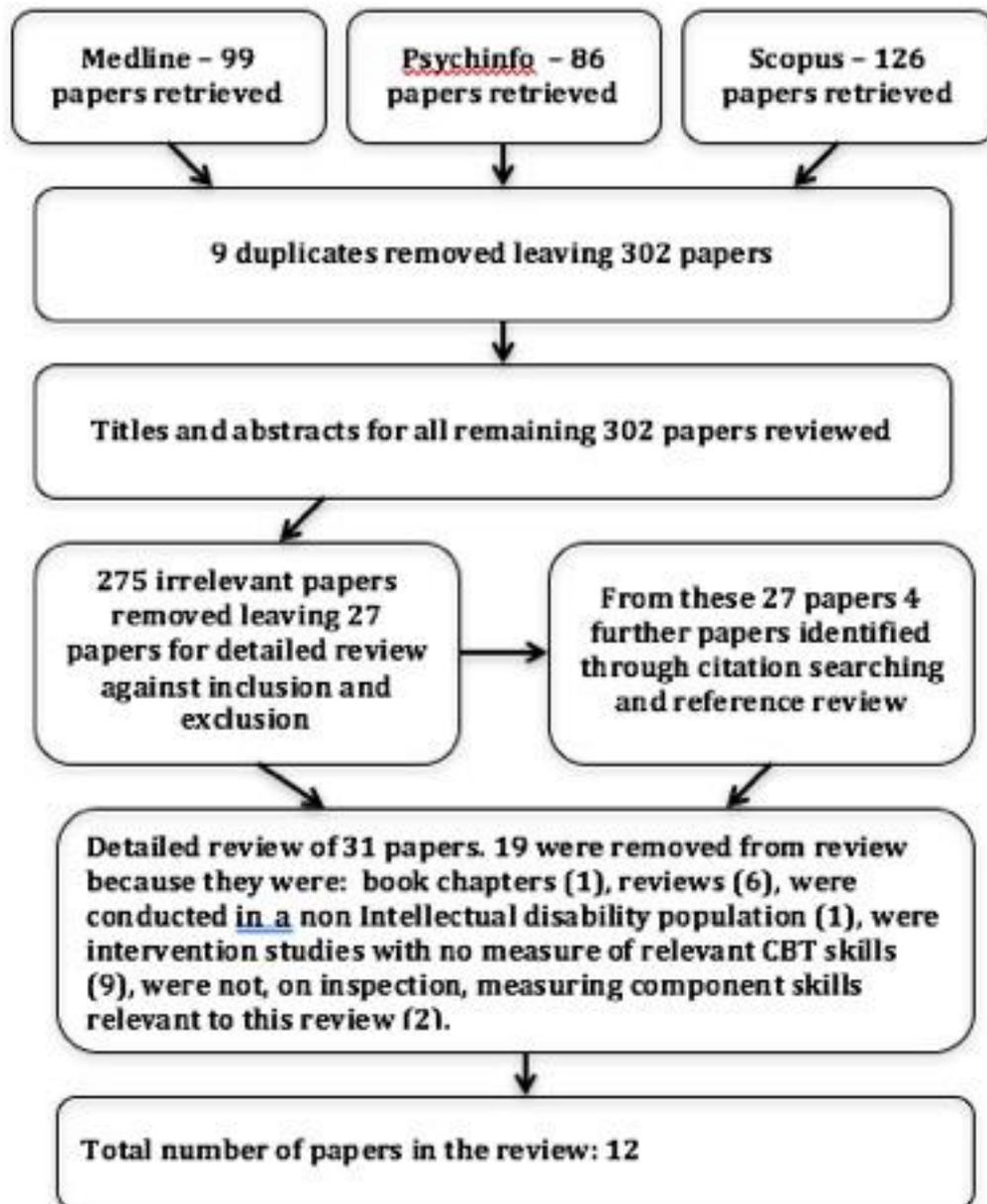


Figure 1 – Flow diagram of search strategy procedure

#### *Inclusion and exclusion criteria*

Studies were included if they were empirical studies with a population identified by authors as having an intellectual disability using quantitative measures of either discrimination between thoughts feelings and behaviours; linkage of emotions to events or the recognition of cognitive mediation. Only studies in English in peer-reviewed journals were considered.

### *Quality assessment*

The first aim was to assess measurement quality. This is a multi-faceted concept (Mokkink et al., 2010). To ensure consistent coverage of measurement properties, all measures were evaluated with reference to domains identified in an international consensus framework (Mokkink et al., 2010):

1. Reliability, subdivided into internal consistency, measurement error and consistency across raters/time.
2. Validity, subdivided into content (including face) and construct (including structural validity/hypothesis testing) and cross-cultural validity.
3. Responsiveness (no subdivision).

This review also evaluated evidence as to the relationship of skill level to CBT outcome, level of skill performance in an intellectual disability population, and the trainability of these skills. Given the heterogeneity of designs used to answer these questions, quality was assessed by the quantitative scale of the QualSyst (Kmet et al., 2004), a 14 item tool specifically designed for assessing quality of primary quantitative research articles of varying designs. QualSyst items are scored as having not been met (0), partially met (1), totally met (2), or not relevant to the article being rated (N/A). Inter-rater reliability for items varied from 40 percent to 100 percent (Kmet et al., 2004). Figure 2 gives details of all areas assessed by items. An overall quality score between 0 and 1 was generated for each article by summing the article score and dividing it by the total possible score (i.e.  $28 - (\text{number of 'n/a'} \times 2)$ ). As recommended by National Institute for Clinical Excellence (NICE, 2014), quality assessment was supplemented by critical appraisal and an overall rating of high (++), medium (+) or low (-) quality was assigned based on QualSyst rating and critical appraisal of how likely identified issues were to alter a study's main conclusion.

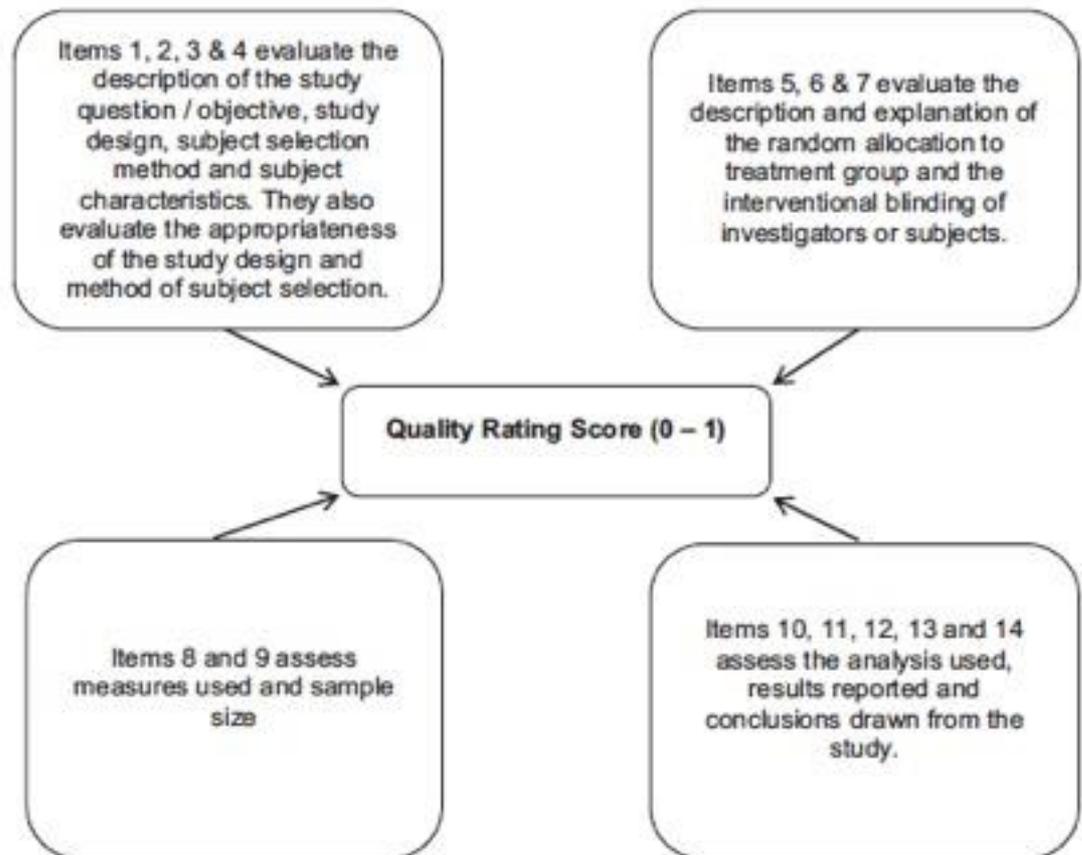


Figure 2 - Overview of the areas captured by QualSyst quantitative studies items

## Results

Search results with reasons for exclusion of studies are given in Figure 1.

Twelve articles reporting on results in relation to six measures were included in the review. Table 1 gives a description of all measures used. Table 2 provides an overview of all studies included and Table 3, their quality assessment scores. Most studies (eight) were judged to be of at least medium quality, with studies looking at trainability being of higher quality. Discussion of quality is integrated into the results presented below.

Table 1: Measures of CBT readiness used in studies included in the review

Instrument (Author/year)	Construct assessed	Dimensions (number of items)	Response options (range)	Pass criterion (Cut-off score)	Ease of scoring/administration	Sample items
Behaviour Thought Feeling Questionnaire (BTFQ) (Oathamshaw & Haddock, 2006)	Thought-feeling-behaviour discrimination	Total (23) Thoughts (7), Feelings (8) Behaviours (8)	Forced choice between, 'thought', 'feeling' or 'behaviour'	Overall (12) Thoughts (5) Feelings (6) Behaviours (6)	Easy to administer with clear instructions	Participant asked: <i>"sad" is that a thought, feeling or a behaviour</i>
Recognition of cognitive mediation 1 (RCM1) (Dagnan et al., 1997)	Recognition of cognitive mediation	One dimension (6 items) (Dagnan et al., 1997) or (12 items - prompt repeated with opposite emotion) (Hebblethwaite et al., 2011)	Free generation of thought given a prompt event and emotion	N/A – mean score	Medium – requires coding using guideline	Participant told: <i>'You walk into a room where there are a group of your friends; as you walk in they start to laugh and you feel happy'</i>  Then shown a happy Makaton face and asked: <i>'What would be thinking or saying to yourself?'</i>

Table 1: Measures of CBT readiness used in studies included in the review

Instrument (Author/year)	Construct assessed	Dimensions (number of items)	Response options (range)	Pass criterion (Cut-off score)	Ease of scoring/administration	Sample items
Recognition of cognitive mediation 2 (RCM2) (Dagnan et al., 2000)	Recognition of cognitive mediation	Six subscales, three each for thoughts as responses and emotions as responses. Scales are: Overall (10) Congruent (5) Incongruent (5)	Forced choice (from two thoughts or two emotions), responses are either congruent or incongruent with prompt event	Overall (8) Congruent (5) Incongruent (5)	Easy to administer with clear instructions	Participant told: ' <i>Your friend shouts at you and you feel sad</i> '  Then shown sad Makaton face and asked: ' <i>would you be thinking 'I'm a good person or I am a bad person</i> '
Reed Clements Task (Reed & Clements, 1989)	Event-emotion linkage	One dimension (6)	Forced choice (Happy or sad)	Errorless performance	Easy with clear instructions	Participant told ' <i>You take your dog for a walk. The dog breaks the lead. You have lost your dog.</i> '  Then shown happy/sad Makaton faces and asked: ' <i>Do you feel happy or sad?</i> '
Thought Feeling Behaviour task (TFB)	Thought-feeling-behaviour discrimination	Total score (18) Thoughts (6) Behaviours (6) Feelings (6)	Forced choice	N/A – mean score	Easy to score and administer	Participant read scenario: ' <i>Peter knew it was the last day of his holiday. Peter went to pack his suitcase.</i>

Table 1: Measures of CBT readiness used in studies included in the review

Instrument (Author/year)	Construct assessed	Dimensions (number of items)	Response options (range)	Pass criterion (Cut-off score)	Ease of scoring/administration	Sample items
(Quakley et al., 2004)						<p><i>Peter felt sad that he was going home today.</i></p> <p>Participant then given/read cards with each sentence element and asked to identify which card has a thought, which a feeling and which a behaviour</p>
Thought to feeling task (Doherr et al., 2005)	Recognition of cognitive mediation	One Dimension (6)	Free generation of feelings and thoughts	N/A – mean score	Difficult – requires coding with no identified coding system.	<p>Stick man/thought bubble and Makaton faces used</p> <p>Participant told: <i>Imagine that you are going away for a while. You say goodbye to your family</i>’ (indicate stick person)</p> <p><i>You think: ‘I can’t wait for my holiday</i>’. (Indicate thought bubble).</p> <p><i>How do you think you would feel if you thought, ‘I can’t</i></p>

Table 1: Measures of CBT readiness used in studies included in the review

Instrument (Author/year)	Construct assessed	Dimensions (number of items)	Response options (range)	Pass criterion (Cut-off score)	Ease of scoring/ administratio n	Sample items
						<i>wait for my holiday</i> ?(indicate Makaton faces)
						After answering, participants were asked ' <i>Why do you think you would feel</i> (insert answer)?'

### *Study setting and sample characteristics*

All studies were carried out in the UK. Most were of cross sectional design (N= 7). Four employed randomised experimental designs and one (Hartley et al., 2015) used a non-randomised pre-post design. The use of convenience samples of volunteers in 11 of the reviewed studies, and the lack of detail as to whether participants differed from non-participants limits generalisability to the wider intellectual disability population.

Sample sizes ranged from 19 to 59. The total number of participants with intellectual disability across studies was 462. Most studies only included participants over 18 years old (although one study had participants as young as 14 (Reed & Clements, 1989)). The percentage of female participants varied from 35.1 (Dagnan et al., 2000) to 79.3 percent (Vereenooghe, Gega, Reynolds, & Langdon, 2016). Where measured, mean full scale IQ varied from 50 (Vereenooghe et al., 2016) to 60.1 (Hebblethwaite, Jahoda, & Dagnan, 2011). Most studies explicitly excluded nonverbal participants and participants' mean receptive language score on the British Picture Vocabulary Scale (BPVS) (Dunn, Dunn, Whetton, & Pintille, 1982) ranged from 12.87 (Joyce, Globe, & Moody, 2006) to 88 (Oathamshaw & Haddock, 2006). Only three studies explicitly state that they excluded those with current (Vereenooghe et al., 2016; Vereenooghe et al., 2015) or prior CBT exposure (Sams, Collins, & Reynolds, 2006). This is important as CBT exposure may enhance performance, biasing results. Five studies (Dagnan et al., 2000; Dagnan, Mellor, & Jefferson, 2009; Joyce et al., 2006; McEvoy, Reid, & Guerin, 2002; Reed & Clements, 1989) did not check the intellectual disability status of their participants. This is important, as where cognitive functioning was checked, some participants were not in the intellectual

disability range and excluded (Hebblethwaite et al., 2011). Finally, sensory problems, which might affect tasks, were not routinely excluded.

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
Bruce et al. (2010)	UK PWID* Sample; day service/college attenders  N= 34; mean age = 40.5 (SD**, 13.8); 47% female; mean IQ = 55 (SD 3.3) in intervention group and 56 (SD, 4.9) in control group	Impact of training on CT*** skills	Pre post stratified Randomised experiment, ID participants identified by ‘local services’  Intervention: 1 hour on linkage and discrimination of behaviours, thoughts and feelings  Control: Relaxation training  Measures done up to a week before and a week after	Thought, feeling, behaviour task (TFB) (Quakley et al., 2004)  Thought to feeling task (Doherr et al., 2005)	IQ - Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999)  Language - British Picture Vocabulary Scale-II (BPVS-II) (Dunn et al. 1997)	Recognition of cognitive mediation but not discrimination of thoughts/feelings behaviours improved post training and generalized to novel task

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
Dagnan et al. (2000)	UK PWID sample in day services  N = 40; mean age 35.1 (SD, 9.5); female 52.5%; mean BPVS 64 (SD, 27)	Pass rates on CT measures, Associations with language and emotion recognition	Cross sectional design, participants identified by day centre staff	Reed Clements Task (Reed & Clements, 1989)  Recognition of Cognitive Mediation-2 (RCM2) (Dagnan et al., 2000)	Emotion recognition (Dagnan & Proudlove, 1997)  Language - BPVS (Dunn et al., 1982)	<p><i>Pass rates:</i></p> <p>1. Reed Clements Task - 75%</p> <p>2. RCM2 (Choose thought/choose emotion): Overall - 25%/10%; Congruent - 20%/37.5%; Incongruent - 12.5%/2.5%</p> <p><i>Associations:</i> BPVS differed across passers and failers for Reed Clements Task and RCM2 choose emotion overall and congruent subscales, and choose thought incongruent subscale. No correlations with emotion recognition</p>

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
Dagnan et al. (2009)	UK PWID N = 41; mean age = 39.2(SD,11.7); female = 34%; mean BPVS 61.48 (SD, 26.56)	Inter-rater reliability of Recognition of Cognitive Mediation 1 (RCM1) (Dagnan et al., 1997)  Associations with language and emotion recognition	Cross sectional design, not clear how sample recruited	Reed Clements Task  RCM1 - 6 item version	Recognition of emotions - (Dagnan & Proudlove, 1997)  Language - BPVS	Mean score on cognitive mediation was 2.16 (SD =2.1, Range 0 -6) Kappa for items = 1.  <i>Associations:</i> BPVS correlated with Reed Clements Task/RCM1  Reed Clements Task correlated with elements of RCM1 as hypothesised  No correlations between facial emotion recognition and CT measures
Hartley et al. (2015)	UK PWID and depression sample, living in a variety of settings.  Intervention N = 16; mean age	Improvement of CT skills in a CT intervention	Pre-post non-randomized study with 3 month follow up  Sample recruited via Fliers to case managers	RCM2  BTFQ	Depression: Self report depression questionnaire (Reynolds & Baker, 1988)	<i>Effect of intervention</i> Significant differences in depression and behaviour problems maintained on 3 month follow up.  <i>Effect of intervention on CT measures</i>

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	33.8 (SD, 10.92); female = 50%; IQ = 62.4 (SD, 7.4).  Control N = 8, Mean age; 40.3 (SD, 11.5); female = 47.5%; IQ = 61.1 (SD, 6.6)		Intervention: ‘Empower’ group CBT with caregivers also involved  Control: treatment as usual		Behaviour problems: Scales of Independent Behaviour-Revised (Bruininks, 1996)  Social skills the social performance survey (Matson & Hammer, 1996)	RCM1 but not BTFQ improved in CBT relative to control.
Hebblethwaite et al. (2011)	UK PWID sample – day centre attendees.  N=19; mean age = 42 (SD, 10.96); female	Ability on CT measures relative to controls without ID  Associations with IQ/’real	Cross sectional between groups design PWID participants recruited via keyworkers	RCM1 -12 item version	IQ -WASI  Cognitive emotive interview – measure designed for study to	ID sample mean score on RCM1 = 7.63 (SD= 1.8, range = 3-11). Kappa 0.86.  Significantly lower RCM1 for ID than controls.  <i>Associations</i>

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	= 63%. IQ 60.1 (SD, 6.22)  Controls: Staff/students at a college without ID.  N= 19; mean age = 43 (SD, 11.37); female; 57%; IQ = 95.44 (SD, 12.16)	life’ conversation			simulate real life ‘CBT like conversation	No correlations between RCM1 and IQ or cognitive emotive interview.
Joyce et al. (2006)	UK PWID sample in day services  N =52; mean age 40 (SD 11.6 ); female = 53% (range 21-81);	Pass rates on CT measures  Associations with language and Emotion recognition	Cross sectional design Participants ‘randomly selected’	Reed Clements Task  RCM2	Language - BPVS  Communication Assessment Skills Profile (Gaag, 1998)	<i>Pass rates</i> 1. Reed Clements Task – 50%  2. RCM2 (Choose thought/choose emotion)

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	mean BPVS =12.87 (SD, 6.9)				Emotion recognition 12 facial emotion recognition measure designed for study	Overall - 13%/11%; Congruent - 21%/19%; Incongruent - 6%/4%  <i>Associations</i> BPVS and CASP differed across passers and failers for Reed Clements Task and some RCM2 subtests  Reed Clements Task passers performed better than failers on identifying and labelling emotions.  RCM2 passers had better performance than failers on labelling but not identifying emotions.
McEvoy et al. (2002)	UK PWID Sample in one day service	Pass rates on CT measures  Associations with	Cross sectional study, sample identified by care staff.	Reed Clements Task	Language BPVS  Story task designed for study to look at	<i>Pass rates</i> Reed Clements Task – 41%  <i>Associations</i>

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	N= 41; mean age 36.8 (SD, 10.8); 39% female; BPVS mean 13.7 (SD, 5.9)	language/ concept of death			understanding of death.  Pictorial task designed for study to assess emotional response to bereavement	Reed Clements Task passers had higher BPVS and concept of death scores than failers
Oathamshaw and Haddock (2006)	UK  People with intellectual Disabilities (PWID) and psychosis community and hospital  N= 50; age =46 (SD, 11.3); 43% female; mean BPVS 88	Pass rates on CT measures  Associations with language	Cross sectional design, participants suggested by clinicians	Reed Clements Task  Behaviour Thought Feeling Questionnaire (BTFQ)  RCM2	Language - BPVS  Emotion Recognition (Dagnan & Proudlove, 1997)	<i>Pass rates:</i> 1. Reed Clements– 72%  2. BTFQ Feelings - 52%, Behaviours - 32%, Thoughts - 9%  3. RCM2 (Choose thought/choose emotion) Overall 10%/12%. Congruent 14%/30% Incongruent 4%./2%

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	(range 66-99.75),					<i>Associations:</i> BPVS score higher in passers than failers for Reed Clements Task, TFB Feelings and behaviours and some RCM2 subscales
Reed and Clements (1989)	UK ‘Mental Handicap’ sample  N=55; age range 14-25; female = 52%; mean BPVS 65.4 (Range 27-104)	Pass rates on CT measures  Association with language	Cross sectional design, not clear how sample recruited.	Reed Clements Task	Language - BPVS	<i>Pass rates:</i> 75% passed Reed Clements task  <i>Associations:</i> BPVS score higher in passers than failers for Reed Clements Task

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
Sams et al. (2006)	UK PWID sample in day centres and colleges  N= 59; age = 17-60 Female = 52.5%; mean IQ = 58 (range, 50-72); mean BPVS 85.4 (SD, 27.7)	Ability to 'do' TFB measure and enhancement of this by cues  Associations with IQ and language	Cross-sectional and experimental (for cue enhancement component), participants identified by key workers.	TFB	Language - BPVS-II  Recognition of emotion (Dagnan & Proudlove, 1997)  IQ - WASI	TFB mean (SD) scores: Thoughts - 3.9(1.6); behaviours -3.12 (2.1); Feelings - 2.76(1.89)  No enhancement of performance on TFB by cues.  <i>Associations</i> FSIQ/VIQ correlated with TFB total feelings (0.38, 0.4) and behaviours (0.51, 0.5).  BPVS II correlated with TFB total (0.53), behaviours (0.5).
Vereenooghe et al. (2015)	UK PWID Sample from day services,  Intervention N=32; mean	Impact of training on CT skills  Pass rates on CT tasks	Pre post stratified (on IQ) randomized experiment, staff at day centres identified participants.	Computerised version of RCM2	IQ – WASI	<i>Training effect:</i> RCM2 - choose emotion/ but not choose thought improved by training for congruent but not incongruent items

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	age 38.53 (SD, 12); female, 65%, IQ 53.3 (SD, 8.4)  Control N= 33, mean age = 38.2 (SD, 14.1); female 64%; IQ = 52.5 (SD, 8.5)	Associations with IQ	Intervention: one hour computerized training on event-emotion linkage.  Control group - attention control task using similar stimuli.  Pre and post measures immediately before and after training.			<i>Pass rates:</i> Reed Clements. – 65%  RCM2 subtest pass rates (Choose thought/choose emotion)– overall ,45%/59% (other rates not given)  <i>Associations:</i> IQ correlated with RCM2 choose emotion and choose thoughts scores
Vereenooghe et al. (2016)	UK PWID Sample from day services Intervention N=26; mean age 41 (SD, 14); female, 57.6%; IQ = 50 (40-69)	Impact of training on CT measures.  Associations between CT measures and with IQ.	Randomised pre post experimental design  Staff signposted users to study  Intervention: computerized version of TFB task	Computerised version of BTFQ  Computerised version of RCM2	IQ – WASI-II	<i>Training effect:</i> Effect on BTFQ overall but not other subscales  <i>Associations:</i> IQ correlated with BTFQ total, behaviours and feelings

Table 2 – Overview of studies included in the review

Study Author (date)	Location and sample characteristics	Relevant areas examined	Method	CBT readiness skills measures	Non CBT measures	Main relevant findings
	Control N= 29 mean age = 36 (SD, 13); female = 79.3%; IQ = 50 (40-67)		Control: Attention control using similar stimuli.			RCM2 choose thought was correlated with the aggregate BTFQ and feelings subtest. BTFQ participants identified 5 feelings, 5 behaviours, 2 thoughts on average

Table 3: Quality of included studies

Study	Qualsyst score - score/number of items, (ratio of score to items)	Main limitations:	Quality score
Bruce et al. (2010)	23/26 (0.88)	Small sample size and power not calculated	++
Dagnan et al. (2000)	14/22, (0.64)	Congruence and incongruence of presented scenario emotion/thought pairings central, but not clearly defined, previous/Current CBT not excluded, ID status not checked.	+
Dagnan et al. (2009)	12/20 (0.60)	No control group/task, Very limited information re sample, current/Previous CBT not excluded, ID status not checked.	-
Hartley et al. (2015)	16/26 (0.62)	No control group/task, Small sample size and power not calculated, non-randomized, non-blind design and involvement of assessors in treatment	-
Hebblethwaite et al. (2011)	20/22 (0.90)	Low power for correlational elements,	++
Joyce et al. (2006)	13/22 (0.59)	current/previous CBT not excluded Current CBT not excluded, ID status not checked, order of measures not counterbalanced.	+
McEvoy et al. (2002)	11/20(0.55)	Poorly defined sample, limited description of results, rationale for statistics used unclear	-
Oathamshaw and Haddock (2006)	15/22 (0.68)	No control group/task, order of measures not counterbalanced.	+
Reed and Clements (1989)	14/22 (0.64)	Previous/current CBT not excluded, No control group or task, no counterbalancing, current/Previous CBT not excluded, ID status not checked no control group/task	-
Sams et al. (2006)	18/26 (0.69)	Multiple correlations with no corrections for type 1 error	+
Vereenooghe et al. (2015)	22/28 (0.79)	N/A	++
Vereenooghe et al. (2016)	22/28 (0.79)	Power calculated but not achieved	++

Note: ++ = High quality; + = medium quality and - = low quality

### *Measurement quality of tasks used to assess readiness skills*

Measurement quality assessment was very limited. In terms of reliability only inter-rater reliability was assessed. As for validity, limited aspects of face, content and construct validity were assessed for some measures. Neither structural (factor analysis), nor cross-cultural validity, nor responsiveness were assessed for any measure.

There was a lack of clarity as to whether readiness skills are discontinuous or continuously distributed constructs, with some measures adopting a pass/fail criterion (Dagnan et al., 2000; Oathamshaw & Haddock, 2006) and others a mean score (Dagnan et al., 1997; Quakley, Reynolds, & Coker, 2004).

Two measures were used to assess thought-feeling-behaviour discrimination; the Behaviour Thought Feeling Questionnaire (BTFQ) (Oathamshaw & Haddock, 2006) asks respondents to identify if a prompt word or sentence is a thought, feeling or behaviour. It was used in three studies (Hartley et al., 2015; Oathamshaw & Haddock, 2006; Vereenoghe et al., 2016). The Thought Feeling Behaviour task (TFB) (Quakley et al., 2004) asks participants to identify the thought, behaviour and feeling elements of a set of standardised sentences about a person's reaction to scenarios. It was used in two studies (Bruce, Collins, Langdon, Powlitch, & Reynolds, 2010; Sams et al., 2006). Both measures have been scored as a single total, reflecting ability to recognise thoughts/feelings/behaviours as a whole (Bruce et al., 2010; Hartley et al., 2015; Vereenoghe et al., 2016) or as three separate subscales reflecting ability to recognise thoughts, feelings or behaviours separately (Oathamshaw & Haddock, 2006; Sams et al., 2006; Vereenoghe et al., 2016). A clear theoretical rationale was not given for either scoring method.

Given forced choice scoring, inter-rater reliability for both measures is likely to be high, particularly for computerised versions as there is no space for documentation or interpretation errors. Inter-rater reliability of the computerised versions may be further enhanced through the use of standardised recorded instructions and associated reduction in response bias and suggestive questioning. The content validity of the BTFQ was enhanced through involving people with intellectual disabilities in developing items and basing the BTFQ on a measure routinely used in CBT (Greenberger & Padesky, 1995). Forced choice responses affect the face validity of both the BTFQ and TFB as CBT requires free generation of response. The errorless performance of 20 CBT experts on the BTFQ provides some evidence of expert criterion validity (Oathamshaw & Haddock, 2006). There is mixed evidence of concurrent validity of the BTFQ through correlations with some subscales of Dagnan et al. (2000)'s recognition of cognitive mediation measure in a high quality study (Vereenooghe et al., 2016). The TFB has not been assessed for reliability or validity with an intellectual disability population.

One measure, the Reed Clements Task has been used to evaluate event-emotion-linkage (Reed & Clements, 1989). Participants are asked if they would feel happy or sad in a given scenario and it has been used in six studies (Dagnan et al., 2000; Dagnan et al., 2009; Joyce et al., 2006; McEvoy et al., 2002; Reed & Clements, 1989; Vereenooghe et al., 2015). The forced choice response between two emotions enhances inter-rater reliability but reduces face validity. There is some evidence of concurrent validity through correlations with measures of recognition of cognitive mediation described below.

Three measures have been used to examine recognition of cognitive mediation (Dagnan et al., 2000; Dagnan et al., 1997; Doherr et al., 2005):

The first measure (Dagnan et al., 1997) requires the participant to generate their thoughts given an emotion and a prompt event. Responses are subsequently coded and thoughts deemed congruent with the valence of the emotion scored as correct. This measure has been used in two studies of mixed quality with good evidence for inter-rater reliability of coding (Dagnan et al., 2009; Hebblethwaite et al., 2011). Free response generation and 'CBT like' prompt questions give good face validity. Evidence for concurrent validity is mixed; the measure correlates with the Reed Clements Task (Reed & Clements, 1989) as expected in a study assessed as of low quality (Dagnan et al., 2009) but not with an analogue of a CBT-like conversation in a high-quality study (Hebblethwaite et al., 2011).

The second measure (Dagnan et al., 2000) has two sections. The first requires participants to select which thought they would think from two choices given an event and an emotion. The second section requires the selection of the emotion that they would feel from two choices (happy/sad) on presentation of an event and a thought. For both sections, the 'correct' answer can be either congruent or incongruent with the valence of the prompt event. Correct incongruent answers are interpreted as reflecting strong evidence of cognitive mediation as they require ignoring event valence and making a choice based on the valence of the presented thought or emotion. Six subscales are generated and include overall scores for thought and emotion response modes and scores for subscales from each response mode based on response congruence or incongruence. Subscale validity has not been assessed through factor analysis.

This measure has been used in six studies in pencil and paper (Dagnan et al., 2000; Hartley et al., 2015; Joyce et al., 2006; Oathamshaw & Haddock, 2006) and computerised (Vereenoghe et al., 2016; Vereenoghe et al., 2015)

formats. The forced choice response format decreases face validity, but inter-rater reliability is likely to be high, particularly for computer-based versions for reasons described above. There is evidence of concurrent validity through correlations with some aspects of the BTFQ (Vereenooghe et al., 2016).

The thoughts-to-feeling task (Doherr et al. 2005) has been used in one study (Bruce et al. 2010), as the outcome measure in a randomised experiment. It has good face validity, but no other aspects of measurement quality have been assessed in this population

#### *Relationship of these skills to the process of CBT*

Whether scores on measures change in the process of CBT has only been examined in one study assessed as of low quality (Hartley et al., 2015).

Recognition of cognitive mediation (Dagnan et al., 2000), but not the BTFQ was found to change over the course of a CBT group intervention for people with intellectual disabilities, which also reduced depressive symptomatology. Critically, his study did not examine the relationship of readiness skill level to CBT outcome.

#### *Performance on readiness skills measures*

The most common goal of studies was to establish whether those with intellectual disabilities can 'do' aspects of CBT readiness. Seven studies provided information as to 'pass rates' in people with intellectual disabilities in terms of pre-defined cut-off scores (see Table 1 for scores) (Dagnan et al., 2000; Joyce et al., 2006; McEvoy et al., 2002; Oathamshaw & Haddock, 2006; Reed & Clements, 1989; Vereenooghe et al., 2016; Vereenooghe et al., 2015), Six studies (Bruce et al., 2010; Dagnan et al., 2009; Hartley et al., 2015; Hebblethwaite et al., 2011; Vereenooghe et al., 2016; Vereenooghe et al., 2015) provided mean scores.

One study (Hebblethwaite 2011) examined performance relative to a non-intellectual-disability control group. Consequently, it is unclear in most cases whether skill level is low relative to a general population. No study controlled for cognitive demands of tasks through use of a control task so it is unclear how much 'failure' is specific to the skills being measured and how much is a function of general task complexity.

Oathamshaw and Haddock (2006) hypothesised that event-emotion linkage is easier than thought-feeling-behaviour discrimination, which in turn is easier than recognition of cognitive mediation. This is supported here in terms of overall 'pass' rates and mean scores on measures.

Pass rates for thought-feeling-behaviour discrimination varied across studies. When the ability to identify thoughts, feelings or behaviours was examined separately (Oathamshaw & Haddock, 2006; Sams et al., 2006; Vereenooghe et al., 2016) each study showed different overall pass rates and different patterns of results emerged in relation to whether thoughts, behaviours or feelings were easier to identify.

Studies differed in population characteristics and mode of administration (computer vs. pencil and paper) and exact measure used, but given that two groups of people with intellectual disabilities matched for IQ reported highly discrepant pass rates for the total score on the BTFQ (71 percent vs. 48 percent) (Vereenooghe et al., 2016) measurement reliability or some unidentified factor may be influencing performance.

There is mixed evidence as to the relationship between thought-feeling-behaviour discrimination, language and cognitive difficulties. In two studies assessed as of medium and high quality, higher Verbal IQ (Sams et al., 2006) and total IQ (Sams et al., 2006; Vereenooghe et al., 2016) were related to higher

total, feelings and behaviour scores. In two medium quality studies, higher receptive language was related to higher behaviour and feelings scores on the BTFQ (Oathamshaw & Haddock, 2006) and with higher behaviour and total scores on the TFB measure (Sams et al., 2006). Thoughts subscales did not correlate with either IQ or language in either study, and provision of visual cues did not enhance performance (Sams et al., 2006)

Pass rates for event-emotion linkage were examined in six studies (Dagnan et al., 2000; Joyce et al., 2006; McEvoy et al., 2002; Oathamshaw & Haddock, 2006; Reed & Clements, 1989; Vereenooghe et al., 2015). They varied between 41 percent (McEvoy et al., 2002) and 75 percent (Dagnan et al., 2000; Reed & Clements, 1989). One study used random sampling (Joyce et al., 2006). Consequently, the 50 percent rate found in their study is perhaps most representative.

Pass rates vary with the mean receptive language ability across and within studies, with poor language skills associated with fewer passes. (Dagnan et al., 2000; Dagnan et al., 2009; Joyce et al., 2006; McEvoy et al., 2002; Oathamshaw & Haddock, 2006; Reed & Clements, 1989). To date, no study has examined the relationship of IQ to this measure, an omission, given that IQ may mediate the relationship with language.

Pass rates for the Dagnan et al. (2000) recognition of cognitive mediation task were examined in five studies (Dagnan et al., 2000; Joyce et al., 2006; Oathamshaw & Haddock, 2006; Vereenooghe et al., 2016; Vereenooghe et al., 2015). As predicted, pass rates were lower for incongruent (2.5–12.5 percent) than congruent (14-37.5 percent) scales. Pass rates did not systematically differ according to whether the response was a thought or an emotion. Task demands appeared to influence performance on this measure with overall pass rates much

higher for a computerised version of the task (45- 59 percent) (Vereenooghe et al., 2015) compared to a non-computerised version (10-25% percent) (Dagnan et al., 2000; Joyce et al., 2006; Oathamshaw & Haddock, 2006). This may be due to the greater use of pictorial cues or greater engagement with computerised versions (Vereenooghe et al., 2015).

When thought feeling behaviour linkage was conceptualised as a continuous construct using Dagnan et al. (1997)'s measure, all (Dagnan et al., 2009) or most participants (Hebblethwaite et al., 2011) gave some correct responses, indicating at least some level of better than chance ability given the open response format. A study assessed as of high quality found that performance on the Dagnan et al. (1997) measure was lower in people with than without intellectual disabilities who were matched for age and socioeconomic status (Hebblethwaite et al., 2011).

There is some evidence for association of recognition of cognitive mediation with language; BPVS scores differ between task passers and failers on some subtests of the Dagnan et al. (2000) task in three medium quality studies (Dagnan et al., 2000; Joyce et al., 2006; Oathamshaw & Haddock, 2006). They are also correlated with correct responses on the Dagnan et al. (1997) measure (Dagnan et al., 2009). There is mixed evidence as to associations with IQ in two studies assessed as of high quality. No correlations were found by Hebblethwaite et al. (2011), but the study's statistical power was low. In contrast, Vereenooghe et al. (2015) found correlations with IQ for some subscales but not others of the Dagnan et al. (2000) measure.

For both event-emotion linkage and recognition of cognitive mediation, facial emotion recognition is not correlated when measured using Dagnan and Proudlove (1997)'s measure (Dagnan et al., 2000; Dagnan et al., 2009). However

Event-emotion linkage performance and some subscales of Dagnan et al. (2000)'s measure are related when a more comprehensive measure of emotion recognition is used (Joyce et al., 2006).

#### *Trainability of readiness skills*

Three well-designed studies have examined whether it is possible to train people with intellectual disabilities to improve thought-feeling-behaviour discrimination and recognition of cognitive mediation. Interventions have focussed on developing event-feeling links (Vereenoghe et al., 2015), developing thought-feeling-behaviour discrimination (Vereenoghe et al., 2016) or both (Bruce et al., 2010). Studies training recognition of cognitive mediation (Bruce et al., 2010; Vereenoghe et al., 2015) found evidence of improved ability relative to a control group both immediately (Vereenoghe et al., 2015) and, on novel items, after a week (Bruce et al., 2010). Evidence for training in thought-feeling-behaviour discrimination was more mixed. Training improved performance on some measures but not others in one study (Vereenoghe et al., 2016) but not in another less well powered study (Bruce et al., 2010). There was no evidence that training in one skill (thought-feeling –behaviour discrimination) would have an effect on another (recognition of cognitive mediation) (Vereenoghe et al., 2016)

#### **Discussion**

This review critically examined measures of CBT readiness skills in relation to four areas; measurement quality, relationship to CBT outcome, level of functioning in people with intellectual disabilities, and trainability. It considered evidence from 12 studies using two measures of the ability to distinguish between thoughts, feelings and behaviours, one measure of the ability to link events to emotions, and three measures of the ability to recognise cognitive mediation. The use of convenience samples, and lack of checking of intellectual disability status

of participants, hinders the ability to generalise results to a wider intellectual disability population.

The measurement quality of tasks used to assess CBT readiness skills was largely indeterminate (Mokkink et al., 2010) with two measures having no assessment of measurement quality (Doherr et al., 2005; Quakley et al., 2004). Structural validity was unclear, as factor analysis has not been conducted on any of the tasks. The division of measures into subscales is thus hard to justify (Mokkink et al., 2010) and has been inconsistently applied across studies.

Relative levels of performance on the tasks are in line with conceptual models (Oathamshaw & Haddock, 2006) but pass rates vary widely across studies, possibly due to poor reliability of measurement, or differences in populations with mixed evidence that language and IQ may be related to performance. Only one study has examined performance in comparison to people without intellectual disability. As those without intellectual disability may also not 'pass' some measures (Harter, 2003) the degree to which having an intellectual disability effects performance as opposed to the measures being difficult for all adults, is unknown.

To be of clinical utility, strong psychometric properties are insufficient. CBT readiness measures must meaningfully tap underlying constructs. A key test of this is how measures relate to CBT outcome. The one study examining this (Hartley et al., 2015) offers provisional evidence that recognition of cognitive mediation changed over the course of an efficacious intervention. Before using measures of recognition of cognitive mediation clinically, however, more methodologically robust studies which assess whether change in skill level mediates change in CBT outcome are necessary, as change over an intervention does not, in itself, indicate a mechanism of action (Hundt et al., 2013)

Furthermore, only a small minority of those with intellectual disabilities were able to perform well or 'pass' on all tasks. This is at odds with the good efficacy of CBT for those with anger and depression (Vereenoghe & Langdon, 2013). This may be due to the measurement issues described above compromising validity, but another explanation is that 'pass' rates don't fully capture performance and that 'partial' ability is possible with partial ability 'enough' to be ready for CBT.

There is evidence that performance on measures is trainable using simple interventions offered over brief time periods, with stronger evidence for training recognition of cognitive mediation than thought-feeling-behaviour discrimination. This is interesting given the higher baseline pass rates in the latter and could suggest that thought/feeling/behaviour measures do not tap a component of the skill of recognising cognitive mediation as would be suggested by conceptual models (Oathamshaw & Haddock, 2006) but may tap a parallel skill, which is less difficult at baseline, but harder to train.

There were some limitations to the review. Thorough assessment of the identified articles against inclusion and exclusion criteria and quality assessment were carried out by the first author alone. Although the other authors were consulted in relation to queries and experts in the field were consulted to avoid missing articles, these are limitations as is the fact that further information was not requested from primary studies. Finally, grey literature was not searched so potentially pertinent unpublished literature was not included.

### **Conclusions and recommendations**

There is little evidence for the clinical utility of CBT readiness measures reviewed here, which are also of indeterminate measurement quality. However, given that recognising cognitive mediation changes over efficacious CBT and is

trainable, future research might focus on this ability. Given the lack of a priori definition of what constitutes a CBT intervention (Doherr et al., 2005) and the multiplicity of interventions falling under the CBT umbrella (Roth & Pilling, 2008) it might also be useful to move away from conceptualising readiness within an ABC model and to start with specific CBT interventions themselves, developing and evaluating measures and training procedures specifically based on the skills judged important within particular interventions, for example, behavioural experiments in CBT for social phobia (Roth & Pilling, 2008). This would automatically improve face validity and feasibility, allowing refinement of the concept of readiness and preliminary assessment of specific skills associated with therapy outcomes. More rigorous psychometrically sophisticated measures of relevant constructs could then be developed

**Acknowledgements:**

This systematic review was completed as part of a fellowship awarded to Author Joshua Stott by the Alzheimer's society. Grant number 236 (AS-CTF-14-005) and I would like to thank the society and acknowledge their on-going support

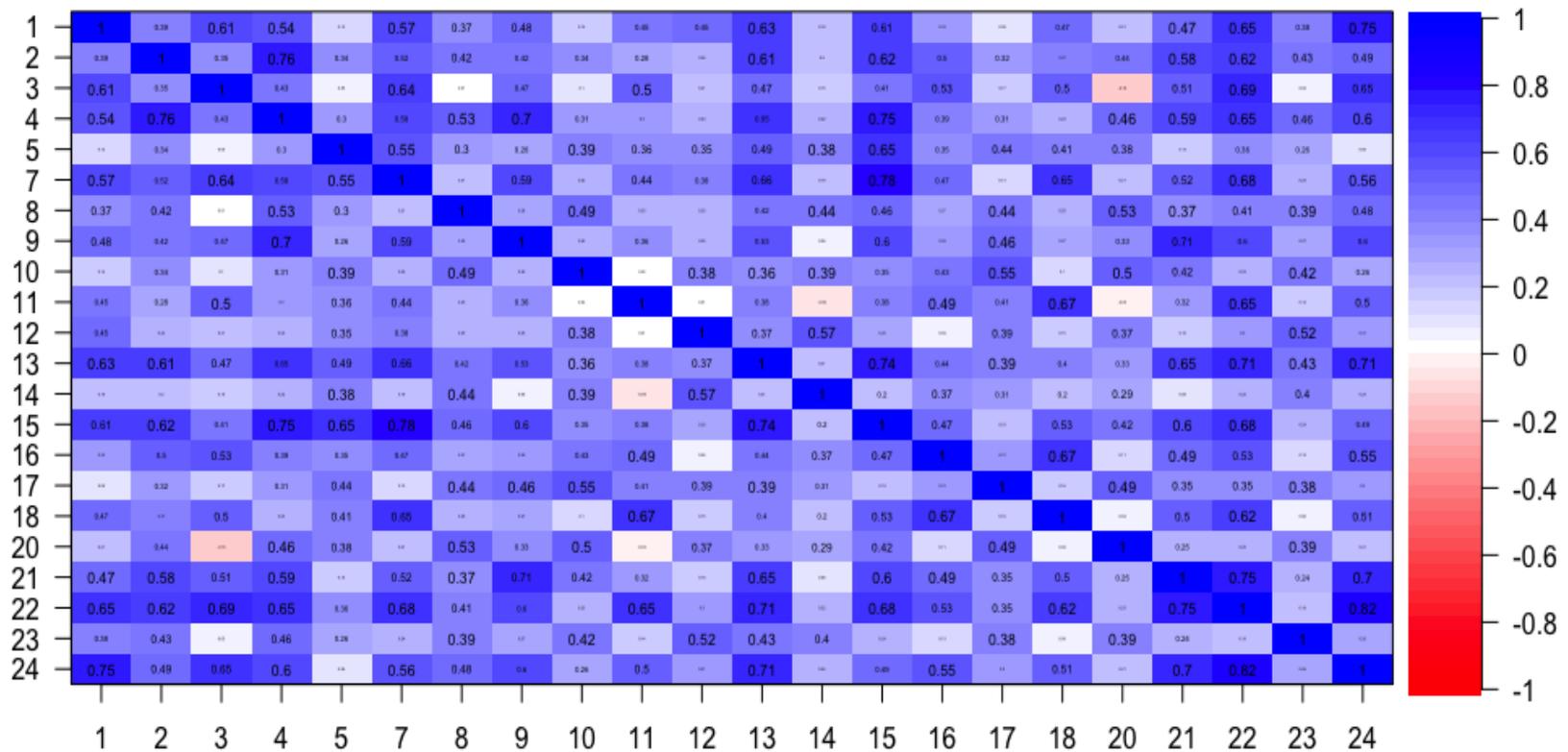
## References

- Bruce, M., Collins, S., Langdon, P., Powlitch, S., & Reynolds, S. (2010). Does training improve understanding of core concepts in cognitive behaviour therapy by people with intellectual disabilities? A randomized experiment. *British Journal of Clinical Psychology, 49*(1), 1-13.
- Bruininks, R. H. (1996). *SIB-R: Scales of independent behavior--revised*: Riverside Publishing Company.
- Cooper, S.-A., Smiley, E., Morrison, J., Williamson, A., & Allan, L. (2007). Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. *The British Journal of Psychiatry, 190*(1), 27-35.
- Dagnan, D., Chadwick, P., & Proudlove, J. (2000). Toward an assessment of suitability of people with mental retardation for cognitive therapy. *Cognitive Therapy and Research, 24*(6), 627-636.
- Dagnan, D., Chadwick, P., Stenfert Kroese, B., Dagnan, D., & Loumidis, K. (1997). Components of cognitive therapy with people with learning disabilities. *Cognitive therapy for people with learning disabilities, 110-123*.
- Dagnan, D., Mellor, K., & Jefferson, C. (2009). Assessment of cognitive therapy skills for people with learning disabilities. *Advances in Mental Health and Learning Disabilities, 3*(4), 25-30.
- Dagnan, D., & Proudlove, J. (1997). *Using Makaton drawings to assess the ability to recognize facial expression of emotion in people with learning disabilities*. Paper presented at the Clinical Psychology Forum.
- Davies, L., & Oliver, C. (2013). The age related prevalence of aggression and self-injury in persons with an intellectual disability: a review. *Research in Developmental Disabilities, 34*(2), 764-775.
- Doherr, L., Reynolds, S., Wetherly, J., & Evans, E. H. (2005). Young children's ability to engage in cognitive therapy tasks: Associations with age and educational experience. *Behavioural and Cognitive Psychotherapy, 33*(2), 201-215.
- Dunn, L. M., Dunn, D. M., Whetton, C., & Pintille, D. (1982). The British picture vocabulary scale. Adapted from the Peabody picture vocabulary test-revised edition: Windsor, Berks: NFER-Nelson.
- Ellis, A. (1980). Rational-emotive therapy and cognitive behavior therapy: Similarities and differences. *Cognitive Therapy and Research, 4*(4), 325-340.
- Ellis, A. (1991). The revised ABC's of rational-emotive therapy (RET). *Journal of Rational-Emotive and Cognitive-Behavior Therapy, 9*(3), 139-172.
- Gaag, A. (1998). Communication skills and adults with learning disabilities: eliminating professional myopia. *British Journal of Learning Disabilities, 26*(3), 88-93.
- Greenberger, D., & Padesky, C. A. (1995). *Mind over mood: A cognitive therapy treatment manual for clients*: Guilford press.
- Harter, L. (2003). Assessment of the suitability of older adults with dementia for cognitive therapy. *D. Clin. Psy. Thesis, Unniversity of Lancaster, Lancaster*.
- Hartley, S., Esbensen, A., Shalev, R., Vincent, L., B, Mihaila, I., & Bussanich, P. (2015). Cognitive Behavioral Therapy for Depressed Adults With Mild Intellectual Disability: A Pilot Study. *Journal of Mental Health Research in Intellectual Disabilities, 8*(2), 72-97.
- Hebblethwaite, A., Jahoda, A., & Dagnan, D. (2011). Talking about real-life events: An investigation into the ability of people with intellectual disabilities to make

- links between their beliefs and emotions within dialogue. *Journal of Applied Research in Intellectual Disabilities*, 24(6), 543-553.
- Hundt, N., E., Mignogna, J., Underhill, C., & Cully, J. A. (2013). The Relationship Between Use of CBT Skills and Depression Treatment Outcome: A Theoretical and Methodological Review of the Literature. *Behavior Therapy*, 44(1), 12-26.
- Hyland, P., & Boduszek, D. (2012). Resolving a difference between cognitive therapy and rational emotive behaviour therapy: Towards the development of an integrated CBT model of psychopathology. *Mental Health Review Journal*, 17(2), 104-116.
- Joyce, T., Globe, A., & Moody, C. (2006). Assessment of the Component Skills for Cognitive Therapy in Adults with Intellectual Disability. *Journal of Applied Research in Intellectual Disabilities*, 19(1), 17-23.
- Kmet, L. M., Lee, R. C., Cook, L. S., Lorenzetti, D., Godlovitch, G., & Einsiedel, E. (2004). *Systematic review of the social, ethical, and legal dimensions of genetic cancer risk assessment technologies*: Alberta Heritage Foundation for Medical Research Edmonton.
- Matson, J. L., & Hammer, D. (1996). Assessment of social functioning.
- McEvoy, J., Reid, Y., & Guerin, S. (2002). Emotion recognition and concept of death in people with learning disabilities. *British Journal of Developmental Disabilities*, 48(2), 83-89.
- McGillivray, J., A., & Kershaw, M. (2015). Do we need both cognitive and behavioural components in interventions for depressed mood in people with mild intellectual disability? *Journal of intellectual disability research: JIDR*, 59(2), 105-115.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19(4), 539-549.
- Muse, K., & McManus, F. (2013). A systematic review of methods for assessing competence in cognitive-behavioural therapy. *Clinical Psychology Review*, 33(3), 484-499.
- NICE. (2014). Developing NICE guidelines: the manual.
- Oathamshaw, S. C., & Haddock, G. (2006). Do people with intellectual disabilities and psychosis have the cognitive skills required to undertake cognitive behavioural therapy? *Journal of Applied Research in Intellectual Disabilities*, 19(1), 35-46.
- Quakley, S., Reynolds, S., & Coker, S. (2004). The effect of cues on young children's abilities to discriminate among thoughts, feelings and behaviours. *Behaviour Research and Therapy*, 42(3), 343-356.
- Reed, J., & Clements, J. (1989). Assessing the understanding of emotional states in a population of adolescents and young adults with mental handicaps. *Journal of Mental Deficiency Research*, 33(3), 229-233.
- Reynolds, W. M., & Baker, J. A. (1988). Assessment of depression in persons with mental retardation. *American Journal on Mental Retardation*.
- Roth, A. D., & Pilling, S. (2008). Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy*, 36(2), 129-147.

- Sams, K., Collins, S., & Reynolds, S. (2006). Cognitive Therapy Abilities in People with Learning Disabilities. *Journal of Applied Research in Intellectual Disabilities*, 19(1), 25-33.
- Stanley, M. A., Calleo, J., Bush, A. L., Wilson, N., Snow, A. L., Kraus-Schuman, C., . . . Schulz, P. E. (2013). The Peaceful Mind program: A pilot test of a cognitive-behavioral therapy-based intervention for anxious patients with Dementia. *The American journal of geriatric psychiatry*, 21(7), 696-708.
- Vereenoghe, L., Gega, L., Reynolds, S., & Langdon, P. E. (2016). Using computers to teach people with intellectual disabilities to perform some of the tasks used within cognitive behavioural therapy: a randomised experiment. *Behaviour Research and Therapy*, 76, 13-23.
- Vereenoghe, L., & Langdon, P. E. (2013). Psychological therapies for people with intellectual disabilities: A systematic review and meta-analysis. *Research in Developmental Disabilities*, 34(11), 4085-4102.
- Vereenoghe, L., Reynolds, S., Gega, L., & Langdon, P. E. (2015). Can a computerised training paradigm assist people with intellectual disabilities to learn cognitive mediation skills? A randomised experiment. *Behaviour Research and Therapy*, 71, 10-19.
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, 23(3), 34-41.
- Willner, P. (2006). Readiness for cognitive therapy in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 19(1), 5-16.

## Appendix D: Tetrachoric Correlation Matrix Heatmap - BTFQ



## Appendix E: The BTFQ-D

### **General Instructions for administration – The BTFQ-D**

Ensure that the person living with dementia (PLWD) is comfortable and happy to participate.

#### **1. Administering the questionnaire items:**

Read each question repeating the instructions: *'X is this something you do something you think or something you feel'*. You can stop repeating the instruction if the PLWD obviously understands and is becoming obviously irritated by repetition. When the PLWD has indicated his or her response, mark it on the questionnaire. Mark only one response for each question. If the PLWD does not or cannot answer an item (for any reason), record the response as missing.

#### **2. Possible queries and responses**

##### **Doesn't want to complete the questionnaire**

Tell the PLWD that participation is entirely voluntary. Say *'you are being asked to complete the questionnaire because it will help us to understand more about whether therapy is helpful for people with dementia'*.

##### **Stops completing the questionnaire because he/she does not understand**

If the PLWD does not understand a particular question, re-read it verbatim, but do not rephrase the question. The questions can be re-read as many times as is necessary and the PLWD can also be told that they have it written in front of them, but if it is clear that the PLWD does not understand then do not continue.

##### **Is concerned that someone will look at his/her answers**

Reassure the PLWD that all of his/her responses will be kept confidential to the research team. Explain that names will be replaced by a study number so that the questionnaires are completely confidential.

### **Asks you to interpret a question**

Re-read the item verbatim. Do not try to explain an item. Suggest that the PLWD base his/her answer on what *he/she* thinks the question means. If the PLWD is still querying or appears not to understand go on to the next question and reassure him/her that they're doing very well and it's fine to go on to the next question. Rephrasing or interpreting a question can bias results. It is very important that the questions are read verbatim and only the above strategy is used.

### **Answers 'don't know' or wants to miss out a question**

Acknowledge that it can be hard to respond but encourage the PLWD to respond. If a PLWD wants to miss out an item, explain to them that all the questions are very important. They should try to answer all of the questions. If the PLWD still does not want to answer a particular item, reassure them that it is alright, then go on to the next item.

### **Wants to know the meaning of his/her answers**

Tell the PLWD that all information is helpful and that there are no right or wrong answers. Remind the PLWD that all the information is kept confidential and that we will look at what everybody says together rather than anybody's questionnaire on its own.

### **If person misunderstands question (i.e. answering something else entirely):**

Say *is XXXXX – (the thing they have said) is that your answer?* If they say yes then mark it if they look doubtful/say no say *'remember we want to know X is this something you do, something you think or something you feel'*.

## Behaviour Thought Feeling Questionnaire Dementia version (BTFQ-D)

**Say:** 'I am now going to read a list of words and I'd like you to tell me if the word is something you do, something you think or something you feel.'

**Say:** 'So you can understand the task I will give you some examples for each example say 'X - is that something you do, something you think or something you feel' X = going shopping, Excited, I've got a lot to do.'

**If incorrect. Give feedback using the instructions below:**

*Say 'I can see how someone might think that that is a behaviour/ feeling/thought but we have it down as a behaviour because going shopping is something you do/a feeling because excited is something that you feel/a thought because I've got a lot to do is something that goes through your mind'*

**DO NOT GIVE FURTHER FEEDBACK ON THE OTHER ITEMS**

**Say** OK, That's great. Now let's try some more. Remember I want you to tell me if the word is something you do, something you think or something you feel.

Start by saying 'X -is that something you do, something you think or something you feel'  
If repeating the instruction becomes unnecessary (person understands/is irritated) just say the item without the instruction from then on

Question	Response	Correct response
1) Angry		F
2) Working		B
3) This is hard		T
4) Answering the phone		B
5) Happy		F
6) Making a cup of tea		B
7) I'm good at things		T
8) Worried		F
9) I don't know what to do for the best		T
10) Gardening		B
11) I hope this works out		T
12) Having a bath		B
13) Frightened		F
14) I've achieved something		T
15) Frustrated		F
16) Washing up		B
17) I'm a good person		T
18) Playing darts		B
19) Upset		F
20) I'm looking forward to my holiday		T
21) Miserable		F

Feeling /7 (above chance responding =5).

Thought/7(above chance responding = 5).

NB Behaviour items (in italics) are not included in the scoring but should be administered in the measure to maintain face validity

## **Appendix F: Qualtrix Prompt Event Valence Survey**

Below is the general form of the survey. Please note that for Q, 40 and 41 the questions were repeated for each potential scenario. Appendix F Table 1 gives the scenarios presented in the survey:

### **Survey Instructions**

In this study, we are looking at how specific situations make people feel. We hope that the results from this will help us to develop psychological tests to adapt psychological therapy for different groups of people. We would be very grateful if you were able to complete this survey which should take around 15 - 20 minutes.

After the brief questions about age, gender and education which follow, all questions will have the same format and be in three parts. Please don't think about your answers too carefully but give your immediate response.

For each question: Part 1 consists of a sentence detailing a short situation and a list of potential emotional responses to that situation. You are asked to rank the 2 emotions that you would be first and second most likely to feel in that situation. You might feel you need qualification of the situation in order to answer fully, but the situations are deliberately ambiguous and we are specifically looking for the first feelings that come to mind when you read the scenario.

The same set of emotions will be listed for every situation, but will be in a different order each time. Part 2 asks you to click to indicate how intensely you would feel the emotion you ranked as most likely in part 1. Part 2 is not asked if you indicate you would feel no emotion in part 1.

Part 3 examines how mixed the situation was emotionally by asking how likely it is you would feel an emotion other than the ones you identified in part 1.<sup>5</sup>

---

<sup>5</sup> NB part 3 was not used in deciding on which prompts to use as the response rate to this question was poor.

## Survey

Q1 Are you?

- Male (1)
- Female (2)

Q2 How old are you (in years)?

Q3 What is your highest qualification?

- Preschool certificate/O level (1)
- O level/school certificate or equivalent (2)
- A level (3)
- First degree (BA, BSc) (4)
- Master's degree (5)
- Doctoral degree (6)
- Other please specify (7) \_\_\_\_\_

Q4 What is your employment status?

- Employed for wages (1)
- Self-employed (2)
- Out of work but looking for work (3)
- Out of work but not looking for work (4)
- Homemaker (5)
- Student (6)
- Retired (7)
- Unable to work (8)
- Other please state (9) \_\_\_\_\_

Q5 What is your marital status?

- Now married (1)
- Never married (2)
- Widowed (3)
- Divorced (4)
- Cohabiting (5)
- Other - please state (6) \_\_\_\_\_

Q6 what is your ethnicity?

- White British (1)
- White Irish (2)
- Any other White background (5)
- Black/Black British - Caribbean (6)
- Black/Black British - African (24)
- Black/Black British - Other (25)
- Any other Black background (17)
- Mixed White and Black Caribbean (10)
- Mixed White and Black African (11)
- Any other mixed background (18)
- Asian/Asian British - Indian (13)
- Asian/Asian British Pakistani (15)
- Asian/Asian British Bangladeshi (19)
- Any other Asian background (20)

- Chinese/Chinese British (3)
- Any other ethnic background - please state (23) \_\_\_\_\_
- Don't wish to say (36)

Q7 Please type a 1 next to the emotion that best completes the sentence below and a 2 next to the emotion that you view as second best in completing the sentence. (Each of the scenarios are presented here in a randomised order (and are given in Appendix F: Table1 at the bottom of this survey)

- \_\_\_\_\_ happy (1)
- \_\_\_\_\_ neutral (no emotion) (2)
- \_\_\_\_\_ sad (3)
- \_\_\_\_\_ angry (4)
- \_\_\_\_\_ afraid (5)
- \_\_\_\_\_ disgusted (6)
- \_\_\_\_\_ surprised (7)
- \_\_\_\_\_ guilty (8)
- \_\_\_\_\_ ashamed (9)
- \_\_\_\_\_ irritated (10)
- \_\_\_\_\_ frustrated (11)
- \_\_\_\_\_ anxious (12)
- \_\_\_\_\_ content (13)
- \_\_\_\_\_ interested (14)
- \_\_\_\_\_ bored (15)
- \_\_\_\_\_ a sense of accomplishment (20)
- \_\_\_\_\_ excited (16)
- \_\_\_\_\_ proud (17)
- \_\_\_\_\_ amused (18)
- \_\_\_\_\_ embarrassed (19)

Q8 You have said that you would be most likely to feel (**the emotion you selected as "1" in Q7**). How intensely would you feel this?

- Slightly - I would hardly feel this at all (1)
- Mildly (2)
- Moderately (3)
- Strongly (4)
- Very Strongly (5)
- As strongly as I have ever felt this (6)

Q9 We can often feel different emotions in the same situation. How likely is it you would feel emotions other than (**the emotion you selected as "1" in Q7**) than or (**the emotion you selected as "2" in Q7**)?

- Very Unlikely - It is extremely hard to imagine feeling any other emotions (1)
- Unlikely (2)
- Somewhat Unlikely (3)
- Undecided (4)
- Somewhat Likely (5)
- Likely (6)
- Very Likely - It is very easy to imagine feeling other emotions (7)

Thank you for completing the survey, if you have any general comments or thoughts on specific scenarios, please do note them below:

---

Appendix F: Table 1. list of scenarios presented in the survey

---

You were going on a trip to the seaside and the sun was out

You are listening to the radio and hear a song that you used to dance to

You notice the flowers blooming in the park

You are eating a meal at home on your own

A friend calls to cancel a trip you had planned

The leaves on the trees are going brown. It is the end of summer

Your GP who has treated you for years tells you she is retiring

You see a group of your friends and they do not say hello

You are about to go to an event where you don't know anyone

You want to go on a special trip but there is only one place and your friend is chosen to go instead

You walk into a room where there are a group of your friends and as you walk in they start to laugh

You are in bed one night and you hear a loud noise downstairs

You look at your diary and see that you have no plans for the next week

You are talking and laughing with a group of your friends

Your daughter calls you to tell you her relationship has broken down

You look at your medication. You see the large number of tablets you have to take

You are shopping and you see a friend you have not seen for ages

You are watching television when one of your favourite films comes on

You are at a centre sitting with people your own age

You lose at a game of cards

You are given a job to do and you do it quicker than everyone else

You tell a neighbour to stop getting on your nerves

You win a prize in a competition you have entered

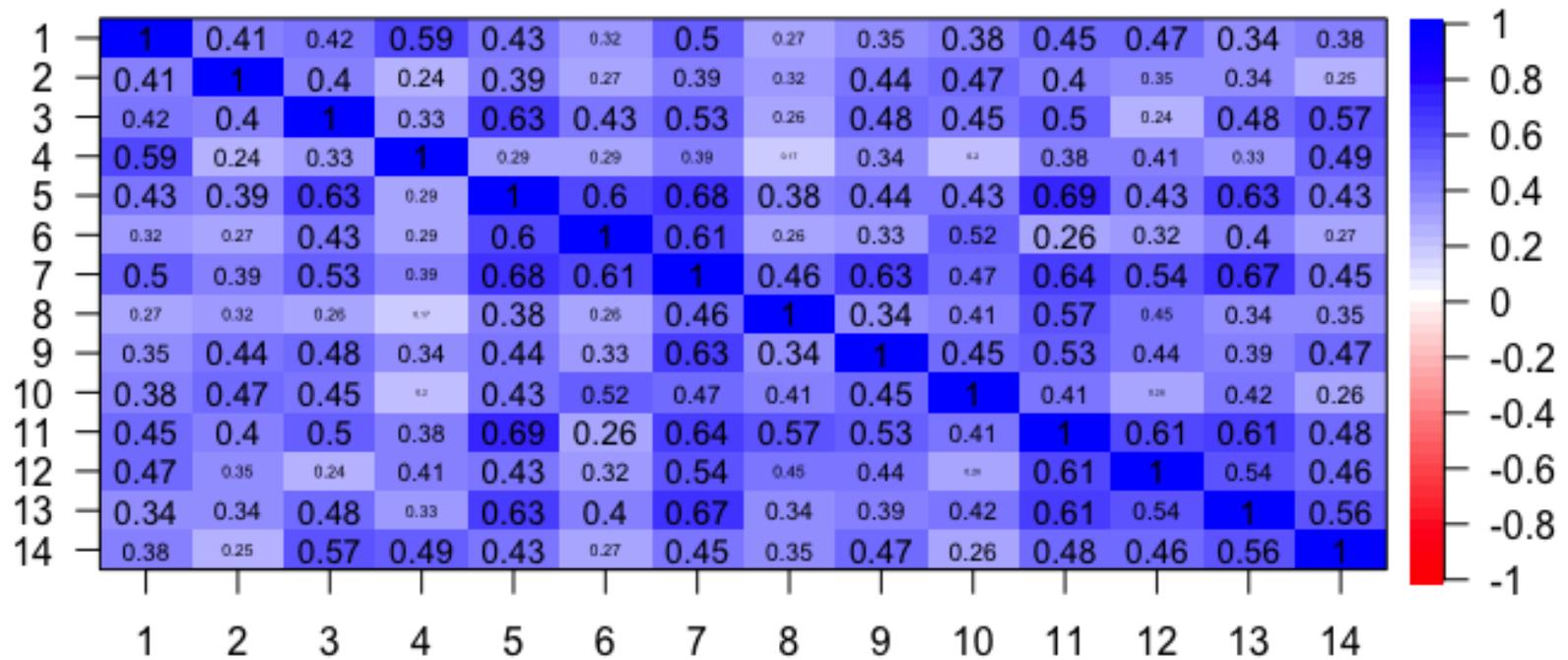
You are in a line and someone pushes in front of you and you tell them not to push in

You are sitting in the park and the sun is out

The cafe you used to go to all the time has been turned into a posh bar

---

Appendix G: Tetrachoric Correlation Heatmap – Cognitive Mediation



## **Appendix H: The CM-DEM**

Ensure that the person living with dementia (PLWD) is comfortable and happy to participate.

### **Instructions for administration**

Give the PLWD the card with the instructions on and say you are going to read the instructions out loud. If the PLWD indicates that they do not understand then repeat the instruction, particularly emphasising that you want to know their thoughts. If the PLWD still says they do not understand then move on saying ‘*That’s okay. Let’s try a few anyway*’ but note that have said they have not understood instructions.

#### **1. Administering the questionnaire items:**

Read each question repeating the instructions: “*What would you be thinking or saying to yourself?*” You can stop repeating the instruction if the PLWD obviously understands and is becoming obviously irritated by repetition. When the PLWD has indicated his or her response, write it on the questionnaire. Mark only one response for each question. If the PLWD does not or cannot answer an item (for any reason), record the response as missing.

#### **2. Possible queries and responses:**

##### **Doesn’t want to complete the questionnaire**

Tell the PLWD that participation is entirely voluntary. Say ‘*you are being asked to complete the questionnaire because it will help us to understand more about whether therapy is helpful for people with dementia*’.

##### **Stops completing the questionnaire because he/she does not understand**

If the PLWD does not understand a particular question, re-read it verbatim, but do not rephrase the question. The questions can be re-read as many times as

is necessary and the PLWD can also be told that they have it written in front of them, but if it is clear that the PLWD does not understand then do not continue.

### **Asks you to interpret a question**

Re-read the item verbatim. Do not try to explain an item. Suggest that the PLWD base his/her answer on what *he/she* thinks the question means. If the PLWD is still querying or appears not to understand go on to the next question and reassure him/her that they're doing very well and it's fine to go on to the next question. Rephrasing or interpreting a question can bias results. It is very important that the questions are read verbatim and only the above strategy is used.

### **Answers 'don't know' or wants to miss out a question**

Acknowledge that it can be hard to respond but encourage the PLWD to respond. If a PLWD wants to miss out an item, explain to them that all the questions are very important. They should try to answer all of the questions. If the PLWD still does not want to answer a particular item, reassure them that it is alright, then go on to the next item.

### **Wants to know the meaning of his/her answers**

Tell the PLWD that all information is helpful and that there are no right or wrong answers. Remind the PLWD that all the information is kept confidential and that we will look at what everybody says together rather than anybody's questionnaire on its own.

### **If person misunderstands question (i.e. answering something else entirely):**

Say *is XXXXX – (the thing they have said) is that your answer?* If they say yes then write it down if they look doubtful/say 'no' say '*remember we want to know*' and repeat the item

