Thought-feeling discrimination in people with dementia: adaptation & preliminary validation of the first dementia specific measure

Joshua Stott[1], Tim Cadman[2], Henry Potts[3], Katrina Scior[1], Janina Brede[1] and Georgina Charlesworth[1]

[1] Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom

[2] MRC Integrative Epidemiology Unit at University of Bristol, Population Health Science, Bristol Medical School, Bristol, United Kingdom

[3] UCL Institute of Health Informatics, University College London, London, United Kingdom

Correspondence to: Joshua Stott, Research Department of Clinical, Educational and Health Psychology, University College London, London, WC1E 6BT, UK. Tel.: +44 2076795950; Fax: +44 2079161989; E-mail: j.stott@ucl.ac.uk.

Abstract

Objective. There is emerging evidence that Cognitive Behavioural Therapy (CBT) can be effective for treating anxiety and depression in people living with dementia (PLWD). Discriminating between thoughts and feelings is a critical element of CBT and also of relevance to emotional understanding more generally. The aim of the present study was the structured adaptation and preliminary validation of an existing measure of thought-feeling discrimination for use in PLWD. Methods/Design. The Behaviour Thought Feeling questionnaire (BTFQ) was adapted via expert and service-user consultation for use in PLWD. 102 PLWD and 77 people aged over 65 years who did not have measurable cognitive impairments completed the adapted measure along with two measures of emotional recognition and reasoning. The factor structure of this measure was examined and the measure reduced. Results. Factor analysis suggested a two-factor solution with thought and feeling items loading on separate factors. The behaviour items were not included in scoring due to high cross-loading and ceiling effects, leaving a 14-item measure with two subscales. Thus an adapted measure was created (named the BTFQ-D), which showed moderate convergent validity in the PLWD, but not the older adult sample. Both thought and feeling subscales showed good internal consistency. Conclusions. The BTFQ-D showed preliminary validity as a measure of thought-feeling discrimination in PLWD. It may have utility in measuring readiness for CBT as part of clinical assessment. Further validation is required.

Keywords: Dementia, Cognitive Behavioural Therapy (CBT), Anxiety, Mood Disorders, Clinical assessment

Running title: Thought-feeling discrimination measure in dementia

Introduction

Anxiety and depression are common in people living with dementia (PLWD) (Enache et al., 2011, Seignourel et al., 2008) and are associated with a number of negative outcomes (Ballard et al., 2000, Gibbons et al., 2002). Antidepressants have limited evidence for efficacy in this group (Banerjee et al., 2011) and there is a growing literature on the use of cognitive behavioural therapy (CBT) (Orgeta et al., 2014, Spector et al., 2015, Stanley et al., 2013). CBT is an umbrella term encompassing a set of cognitive and behavioural change focussed interventions (Roth and Pilling, 2008). In CBT, 'cognitive' refers to managing and modifying unhelpful thought patterns (Roth and Pilling, 2008), rather than the neurocognitive difficulties in memory, attention and executive function that are a feature of all types of dementia (Salmon and Bondi, 2009).

It is often assumed that PLWD cannot comprehend the cognitive elements of CBT (Stanley et al., 2013) due to their neurocognitive difficulties. However, this assumption has not been empirically tested and there is emerging evidence that CBT (particularly when including the cognitive elements) can be effective in reducing depressive symptomatology in people with dementia with mild to moderate cognitive impairments (Spector et al., 2015).

This said, there will be significant variation in the ability of PLWD to understand the cognitive components of CBT. Research across a number of populations, including those whose ability to access CBT might be limited, such as children or individuals with intellectual disabilities, has suggested that a key skill required to engage in CBT is the ability to discriminate between thoughts, feelings and behaviours (Greenberger and Padesky, 1995, Quakley et al., 2004, Quakley et al., 2003, Lickel et al., 2012, Oathamshaw and Haddock, 2006). Assessing this ability in CBT clients is important to understand which elements of a CBT intervention an individual is likely to be able to engage in or might need support with (Oathamshaw et al., 2012). Measuring this ability in dementia also has utility outside of a CBT context as it is an important element of emotional understanding in its own right (Flavell, 1999, Quakley et al., 2003).

There are no measures of the ability to discriminate thoughts, feeling and behaviours that have been validated for use in PLWD. This ability has however been measured in adults with intellectual disabilities (Sams et al., 2006, Oathamshaw and Haddock, 2006) and our recent review of that literature (Stott et al., 2017a) indicates that the most widely used measure in this population is the Behaviour Thought Feeling questionnaire (BTFQ) (Oathamshaw and Haddock, 2006). The BTFQ was developed on general principles for CBT (Greenberger and Padesky, 1995), and had been used in children (Quakley et al., 2003), prior to its adaptation for use in adults with intellectual disability (Oathamshaw and Haddock, 2006).

The BTFQ as used by Oathamshaw and Haddock (2006) has 23 items. For each item, a participant is prompted to identify whether a word/phrase is either a behaviour (e.g. 'having a bath'), a feeling (e.g. 'frightened'), or a thought (e.g. 'this is hard'). The three response options are repeated by the administrator when introducing each item. Verbal responses are scored by the administrator 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 either sub-scale indicates above-chance responding (at $p<0.05$).

The BTFQ cannot simply be adopted in an unmodified form for a dementia context. 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 et al., 2012).

Further, the factor structure of the BTFQ has never been formally assessed and it is unclear how BTFQ performance should be interpreted. BTFQ performance could either be understood as measuring the single ability of discriminating between thoughts, feelings and behaviours (one factor) (Lickel et al., 2012), or as being indicative of three separate abilities to identify thoughts, feelings and behaviours (three factors) (Oathamshaw and Haddock, 2006). This has implications for the value of obtaining subscale vs total scores and for the use of BTFQ scores to inform adaptation of CBT delivery to match clients' ability.

Consequently, the first purpose of this paper is to adapt the BTFQ for use in PLWD using an established framework for cross-population adaptation of measures as a guide (Stewart et al., 2012). Our second aim is to establish the factor structure and preliminary validation of the adapted BTFQ in PLWD. The adapted BTFQ may also be useful in older individuals without a recognised neurocognitive impairment, given potential variability in ability even in this group (Stanley et al., 2013). Therefore, a secondary aim is to examine psychometric properties of the adapted measure in a non-cognitively impaired older adult population.

Materials and Methods

Sample

The sample is from the same cohort as that in a previously published paper, where eligibility criteria are outlined in detail (Stott et al., 2017c). In brief, the sample 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 were consecutive referrals from a memory clinic. Dementia was diagnosed according to 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 cognitive assessment with the extent of assessment driven by client need as per British Psychological Society guidelines (Guss et al., 2014). Client functioning was determined through client and informant report, with an additional occupational therapy assessment included where required. All PLWD participants included in this study had mild dementia characterized by a most recent MMSE score of 24 or higher, or equivalent on another cognitive screen (Law et al., 2013). Of the dementia group, 63 were diagnosed with Alzheimer's disease, 9 with vascular dementia, 17 with mixed Alzheimer's disease and Vascular dementia, two people with Parkinson's disease, one with Frontotemporal dementia and 10 with unspecified dementia.

The OA group was 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 in community groups and from the Join Dementia Research database (Join dementia research). Participants completed the Addenbrookes' Cognitive Examination-III (Hsieh et al., 2013), a screening instrument for dementia (maximum score 100). All healthy volunteer participants scored above the recommended cut-off of 82 for dementia.

All participants were fluent in English, had no self-reported literacy issues and had capacity to consent. Exclusion criteria included a DSM-IV Axis 1 diagnosis of bipolar disorder or schizophrenia, diagnosed intellectual disability, and significant uncorrected sensory deficits. As past CBT experience may influence performance on the measure, participants reporting current or previous experience of CBT were excluded. 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).

Adaptation of BTFQ

Prior to validity analyses, the version of the BTFQ used in the intellectual disabilities literature was assessed for adaptation, with modification conducted where indicated. Procedures followed recommendations by Stewart et al. (2012).

Expert consultation. Following initial review by the lead author and two co-authors (JS, GC, KS), the original BTFQ was circulated to five CBT and dementia experts asking for comments on the measure in relation to suitability for measuring the construct of behaviour/ thought/ feeling discrimination and for use with PLWD. The experts confirmed the suitability of the BTFQ and the potential value of using it in the PLWD population. The only change arising from this consultation was the recommendation for the addition of an item to the thought scale to ensure it was of the same length as others.

Item generation. To generate a new thought item, a focus group was conducted with five people with dementia. 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 and Delbecq, 1972, Stott et al., 2017d). Two thought items were generated from this focus group ('I want to see my friend', 'I'm going to miss my train').

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 and Haddock, 2006). This involved presentation of items to a group of 20 CBT professionals with the expectation that all items would be 100% correctly answered. 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. It also allowed us to keep the perspective of PLWD central to the measure development and provided an opportunity to address any additional queries raised by participants (Smith et al., 2005). Five individual interviews were conducted with PLWD who were not involved in the main validation study, using a double interview technique as has been recommended for pre-test in PLWD (Smith et al., 2005). This involved administering the BTFQ followed by two questions to probe the reasons behind each response choice, to ensure that respondents understood the question in the way that was intended. Two adaptations arose from pre-testing: (i) the supplementation of the verbal instructions with large written prompts to support memory; and (ii) the inclusion of standardised administration instructions to guide administrators' responses if a participant fails to provide a response or asks for clarification. These instructions were based on another validated measure used with PLWD, which in itself was developed using extensive pre-testing and consultation with PLWD (Smith et al., 2005). The resulting version of the BTFQ was then administered to the OA and PLWD samples and further revised to incorporate results from the factor analysis (see below).

Data collection

All participants were screened for eligibility in person, over the phone, or over email as appropriate. Eligible participants who gave informed consent were invited to take part in the study. Demographic information was gathered prior to the administration of measures. Measures were presented in a randomised order.

Measures

Convergent validity was assessed by examining inter-correlations with measures of two other constructs that have been identified as measures of readiness for CBT (Dagnan et al., 2000, Dagnan et al., 2009, Reed and Clements, 1989) and are empirically related with Behaviour/Thought/Feeling discrimination in other populations:

Emotion recognition – ER- 40 (Kohler et al., 2000). The Emotion Recognition Test (ER40) 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 felt or evoked, sad, happy, angry, fearful or neutral facial expressions. An overall recognition index is calculated (0-40).

Event emotion linkage – Reed and Clements' assessment (Reed and 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 Hospital Anxiety and Depression Scale (dementia-modified version) (Stott et al., 2016, Zigmond and Snaith, 1983). 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 is indicated by a score >8 in either scale.

Data analysis

Data were initially examined for floor or ceiling effects, with any item having more than 90% or less than 10% correct responses in the dementia sample removed prior to factor analysis (Pearson, 2008).

To determine the factor structure and reduce items, factor analysis was conducted. Recommendations indicated that a minimum sample of 175 was needed (Pearson, 2008), therefore analysis was first performed in the entire sample before cross-checking the fit of factor structure in subsamples with and without dementia. Factor analysis was based on the tetrachoric matrix using oblimin rotation. 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 (Hofmann, 1978, Pettersson and 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 to establish structural validity, with parallel analysis used to determine the final number of factors to extract (Fabrigar et al., 1999).

Internal consistency of the final measure was assessed with Cronbach's alpha. Convergent validity was assessed through correlations between BTFQ and the other measures detailed above. Significance of correlations was adjusted for type 1 error using False Discovery Rate adjustment, a method that minimises type II error inflation (Verhoeven et al., 2005). All data were analysed in the R environment using the Psych package (Revelle and Revelle, 2017).

Results

Sample characteristics

Table 1 shows clinical and demographic characteristics for both 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 mental health problems (above HADS caseness for either anxiety or depression).

Factor analysis

Assumptions. Two behaviour items (6 and 19) were removed as they scored over 90% correct in the dementia sample. The tetrachoric correlation matrix (not shown) supported data factorability (Field, 2013) with nearly all correlations between items of at least moderate (Revelle and Revelle, 2017) effect size (0.3 or above) in the expected direction.

Factor extraction. Both Kaiser's criterion of eigenvalues greater than 1 (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 dementia sample (items 1 and 8). As a final step, Kaiser's criterion and parallel analysis was run on the full, OA, and PLWD samples, and indicated that two factors should be extracted with all items loading onto the expected factor. This left a two subscale measure with 7 items in each scale.

Model fit. The model showed good statistical properties for the full (Table 2) and PLWD samples (accounting for 52% and 46% of the variance respectively). Loadings in the factor pattern matrix were above 0.4 with communalities 0.35 or above. Complexity factors generally indicated low cross loadings. By contrast, the model in the OA sample showed lower communalities and higher complexity scores. The thoughts, feelings and total scales showed good internal consistency across all samples (alphas = 0.72-0.81).

Validity analysis

Scoring. The revised version of the measure was named the BTFQ-D (Appendix A). Thought and Feeling subscale scores were calculated by summing responses to the 7 items in each respective scale. While factor structure indicated two factors, given inter-correlations between them (see convergent validity below) 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 $p<0.05$ 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.

Descriptive statistics. Table 3 gives the medians and range for convergent validity measures and BTFQ-D, including subscale scores. The feeling subscale indicated a marked ceiling effect in the OA sample and possible ceiling effect for the PWLD sample. The thought score did not appear to exhibit ceiling effects. The Reed Clements score was also at ceiling across both samples.

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 PWLD dementia sample was with the BTFQ-D total scale (Table 4). The Reed-Clements measure was not included in the convergent validity analysis due to respondents scoring at ceiling.

In the combined sample, there was a small, but significant negative correlation (-0.2, n=173) between depression and BTFQ-D feelings scores, indicating that more depression is associated with poorer performance on the BTFQ-D feelings subscale. There were no other correlations between BTFQ-D scales and anxiety or depression.

Discussion

Main findings
Discriminating between thoughts and feelings is a critical element of CBT and also of relevance to emotional understanding more generally. This study is the first to report on the structured adaptation and preliminary validation of a measure of this ability in older adults and PLWD, named the BTFQ-D. Findings are of clinical relevance as they could be used in assessment to inform how best to conduct CBT (Oathamshaw and Haddock, 2006), a promising intervention for anxiety and depression in dementia and older adult populations (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, given the presence of a 'behaviour' option in the responses, they might be included for face validity purposes. Future research in other populations should carefully consider whether it is necessary to include behaviour items in the BTFQ-D or whether an alternative behavioural subscale might be more appropriate.

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 older adults, there is 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 means it is unclear as to whether the BTFQ-D specifically measures thought-feeling discrimination or rather measures a general 'emotional literacy' or some other construct common to it and emotion recognition. Yet, it should be noted that the approach taken in this study is consistent with work in other populations (Dagnan et al., 2000, Oathamshaw and Haddock, 2006). 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). Most importantly, future work should examine the relationships between scores on this measure and CBT outcomes.

Research and clinical implications
The BTFQ-D could be used in research purposes to examine thought-feeling discrimination in dementia samples. This is supported by our findings of internal consistency of 0.7 or above which are sufficient for group comparison purposes (Wells and Wollack, 2003). The BTFQ-D is the only assessment of its kind specifically developed for use with PLWD. The potential ceiling effect and lack of convergent validity in older adults will, however, limit the ability to interpret associations with other variables in this group.

In other populations the BTFQ has been recommended to be used as a clinical tool in the context of CBT (Oathamshaw et al., 2012, Oathamshaw and Haddock, 2006). While the BTFQ-D might be of potential value to inform clinical decisions, we suggest that this measure should not yet be used as a stand-alone measure to determine suitability for CBT. Its

psychometric findings are preliminary and further validation is needed. In particular, excellent internal consistency (0.9 or above) is required for use of a measure as a single high-stakes assessment tool (Wells and 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 et al., 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. Tailoring might involve the provision of thought/feeling discrimination training (e.g. as recommended in a widely used CBT manual) (Greenberger and Padesky, 2015), although current evidence in other groups with cognitive impairment, such as individuals with 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 adaptation 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 by basing the measure on a review of existing measures from other populations (Stott et al., 2017a, Smith et al., 2005), a particular model of CBT readiness (Dagnan et al., 1997), and through consultation with experts and service users (Smith et al., 2005, Dagnan et al., 1997). Even though resources available limited the number of focus groups and pre-test participants, the relevance of this measure to PLWD was increased by engaging with them to develop the initial item set, administration and instructions using pre-testing and focus group methodologies specifically adapted for use in this population (Dening et al., 2012, Smith et al., 2005, Stott et al., 2017a).

The current study is a preliminary validation study. Sample size limitations meant that the initial factor analysis was conducted on a heterogeneous population, with factor structure being checked in individual subpopulations. While this approach suggested that factor structure was consistent and replicated within subgroups, future research should take a confirmatory factor analysis approach in a large heterogeneous sample and examine measurement invariance formally (Stott et al., 2017b, Stott et al., 2016). Generalisability of our results is enhanced by the fact that a significant proportion of the sample met caseness for anxiety or depression, which is the group that would normally be seen for CBT, particularly in the PLWD sample (44 %). Additionally, while there was a correlation between depression and poorer feeling identification, this was small, with no other correlations between thought or feeling identification and mood or anxiety. This further supports generalisability to a clinical population. However, it would also be useful to replicate findings in a clinical sample specifically selected on the basis of anxious and depressive symptomatology. The assessment of convergent validity so far is preliminary, but in line with the literature in other populations (Dagnan et al., 2009, Doherr et al., 2005, Oathamshaw and Haddock, 2006).

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. This is something that has been shown in other areas of emotional understanding (Shimokawa et al., 2001). In addition, the measure reported here is also generally useful in relation to CBT practice where such distinctions are frequently used in conversation (Greenberger and Padesky, 1995).

Conclusions

We report on the structured adaptation and preliminary validation 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. The measure may have utility in measuring readiness for cognitive therapy at a group level in PLWD for research purposes. It may also be of clinical use in aiding decisions about how to adapt therapy in PLWD. It should, however, not yet be used as a stand-alone measure to assess for suitability for therapy.

Conflict of interest declaration
None. No authors have any conflicts of interest to declare.

Description of authors' roles
JS led on conception, design, acquisition of, analysis of and interpretation of data as well as drafting the manuscript. TC, HP, KS, JB and GC all made substantial contributions to design, analysis and interpretation of data. TC and JB were involved in drafting the manuscript and GC and KS revised for critically important intellectual content. JB was also involved in data collection. All authors approve the final version for publication and agree to be accountable for all aspects of the work.

Acknowledgements
This study was completed as part of a fellowship awarded to Joshua Stott by the Alzheimer's society. Grant number 236 (AS-CTF-14-005). This funding source had no involvement in the study design, collection, analysis, and interpretation of data, writing the manuscript, and in the decision to submit the manuscript for publication.

Dr Cadman also receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206, LIFE-CYCLE project.

This work was supported by the UCLH NIHR BRC.

The authors would like to acknowledge the research participants who generously gave their time to take part in this research and also thank Noor Habib, Lucy Gore, Els Chadwick, Joanne Sweeney, and Catherine Bousfield for their role in data collection.

References

BALLARD, C., NEILL, D., O'BRIEN, J., MCKEITH, I., INCE, P. & PERRY, R. 2000. Anxiety, depression and psychosis in vascular dementia: prevalence and associations. *Journal of affective disorders,* 59**,** 97-106.

BANERJEE, S., HELLIER, J., DEWEY, M., ROMEO, R., BALLARD, C., BALDWIN, R., BENTHAM, P., FOX, C., HOLMES, C. & KATONA, C. 2011. Sertraline or mirtazapine for depression in dementia (HTA-SADD): a randomised, multicentre, double-blind, placebo-controlled trial. *The Lancet,* 378**,** 403-411.

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**,** 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**,** 25-30.

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**,** 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**,** 201-215.

EMRE, M., AARSLAND, D., BROWN, R., BURN, D. J., DUYCKAERTS, C., MIZUNO, Y., BROE, G. A., CUMMINGS, J., DICKSON, D. W. & GAUTHIER, S. 2007. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement disorders,* 22**,** 1689-1707.

ENACHE, D., WINBLAD, B. & AARSLAND, D. 2011. Depression in dementia: epidemiology, mechanisms, and treatment. *Current opinion in psychiatry,* 24**,** 461-472.

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**,** 272.

FIELD, A. 2013. *Discovering statistics using IBM SPSS statistics*, Sage.

FLAVELL, J. H. 1999. Cognitive development: Children's knowledge about the mind. *Annual review of psychology,* 50**,** 21-45.

GIBBONS, L. E., TERI, L., LOGSDON, R., MCCURRY, S. M., KUKULL, W., BOWEN, J., MCCORMICK, W. & LARSON, E. 2002. Anxiety symptoms as predictors of nursing home placement in patients with Alzheimer's disease. *Journal of Clinical Geropsychology,* 8**,** 335-342.

GREENBERGER, D. & PADESKY, C. A. 1995. *Mind over mood: A cognitive therapy treatment manual for clients*, Guilford press.

GREENBERGER, D. & PADESKY, C. A. 2015. *Mind over mood: Change how you feel by changing the way you think*, Guilford Publications.

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 Doctoral, Lancaster University.

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**,** 543-553.

HOFMANN, R. J. 1978. Complexity and simplicity as objective indices descriptive of factor solutions. *Multivariate Behavioral Research,* 13**,** 247-250.
HSIEH, S., SCHUBERT, S., HOON, C., MIOSHI, E. & HODGES, J. R. 2013. Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and geriatric cognitive disorders,* 36**,** 242-250.
JOIN DEMENTIA RESEARCH. Available: https://www.joindementiaresearch.nihr.ac.uk/ (date last accessed 7/11/2016) [Accessed].
KOHLER, C. G., ANSELMO-GALLAGHER, G., BILKER, W., KARLAWISH, J., GUR, R. E. & CLARK, C. M. 2005. Emotion-discrimination deficits in mild Alzheimer disease. *American Journal of Geriatric Psychiatry,* 13**,** 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**,** 127-136.
LAW, E., CONNELLY, P. J., RANDALL, E., MCNEILL, C., FOX, H. C., PARRA, M. A., HUDSON, J., WHYTE, L. A., JOHNSTONE, J., GRAY, S. & STARR, J. M. 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. *Int J Geriatr Psychiatry,* 28**,** 351-5.
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**,** 992-1000.
LIVINGSTON, G., SOMMERLAD, A., ORGETA, V., COSTAFREDA, S. G., HUNTLEY, J., AMES, D., BALLARD, C., BANERJEE, S., BURNS, A. & COHEN-MANSFIELD, J. 2017. Dementia prevention, intervention, and care. *The Lancet*.
MCKHANN, G. M., KNOPMAN, D. S., CHERTKOW, H., HYMAN, B. T., JACK, C. R., KAWAS, C. H., KLUNK, W. E., KOROSHETZ, W. J., MANLY, J. J. & 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**,** 263-269.
MOKKINK, L. B., TERWEE, C. B., PATRICK, D. L., ALONSO, J., STRATFORD, P. W., KNOL, D. L., BOUTER, L. M. & 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**,** 737-745.
NEARY, D., SNOWDEN, J. S., GUSTAFSON, L., PASSANT, U., STUSS, D., BLACK, S. A., FREEDMAN, M., KERTESZ, A., ROBERT, P. & ALBERT, M. 1998. Frontotemporal lobar degeneration A consensus on clinical diagnostic criteria. *Neurology,* 51**,** 1546-1554.
OATHAMSHAW, S. C., BARROWCLIFF, A. L. & HADDOCK, G. 2012. Cognitive-Behavioral Therapy for People with Intellectual Disabilities and Psychosis. *Psychological Therapies for Adults with Intellectual Disabilities.* 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**,** 35-46.
ORGETA, V., QAZI, A., SPECTOR, A. E. & ORRELL, M. 2014. Psychological treatments for depression and anxiety in dementia and mild cognitive impairment. *Cochrane Database Syst Rev,* 1.
PEARSON, R. H. 2008. *Recommended sample size for conducting exploratory factor analysis on dichotomous data*, University of Northern Colorado.
PETTERSSON, E. & TURKHEIMER, E. 2010. Item selection, evaluation, and simple structure in personality data. *Journal of research in personality,* 44**,** 407-420.
QUAKLEY, S., COKER, S., PALMER, K. & REYNOLDS, S. 2003. Can children distinguish between thoughts and behaviours? *Behavioural and Cognitive Psychotherapy,* 31**,** 159-168.

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**,** 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**,** 229-233.
REVELLE, W. & REVELLE, M. W. 2017. Package 'psych'.
ROMÁN, G. C., TATEMICHI, T. K., ERKINJUNTTI, T., CUMMINGS, J. L., MASDEU, J., GARCIA, J. A., AMADUCCI, L., ORGOGOZO, J.-M., BRUN, A. & HOFMAN, A. 1993. Vascular dementia Diagnostic criteria for research studies: Report of the NINDS-AIREN International Workshop. *Neurology,* 43**,** 250-250.
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**,** 129-147.
SALMON, D. P. & BONDI, M. W. 2009. Neuropsychological assessment of dementia. *Annual review of psychology,* 60**,** 257.
SAMS, K., COLLINS, S. & REYNOLDS, S. 2006. Cognitive Therapy Abilities in People with Learning Disabilities. *Journal of Applied Research in Intellectual Disabilities,* 19**,** 25-33.
SEIGNOUREL, P. J., KUNIK, M. E., SNOW, L., WILSON, N. & STANLEY, M. 2008. Anxiety in dementia: a critical review. *Clinical psychology review,* 28**,** 1071-1082.
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**,** 423-433.
SMITH, S., LAMPING, D., BANERJEE, S., HARWOOD, R., FOLEY, B., SMITH, P., COOK, J., MURRAY, J., PRINCE, M. & 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. *Health Technology Assessment (Winchester, England),* 9**,** 1-93.
SPECTOR, A., CHARLESWORTH, G., KING, M., LATTIMER, M., SADEK, S., MARSTON, L., REHILL, A., HOE, J., QAZI, A., KNAPP, M. & ORRELL, M. 2015. Cognitive-behavioural therapy for anxiety in dementia: Pilot randomised controlled trial. *British Journal of Psychiatry,* 206**,** 509-516.
STANLEY, M. A., CALLEO, J., BUSH, A. L., WILSON, N., SNOW, A. L., KRAUS-SCHUMAN, C., PAUKERT, A. L., PETERSEN, N. J., BRENES, G. A. & 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**,** 696-708.
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**,** 992-1017.
STOTT, J., CHARLESWORTH, G. & SCIOR, K. 2017a. 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. 2017b. The interpretability of the Hospital Anxiety and Depression Scale in carers of people with dementia: clinical and research implications. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association,* 13**,** P1238-P1239.
STOTT, J., SCIOR, K., MANDY, W. & CHARLESWORTH, G. 2017c. 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**,** 1293-1302.
STOTT, J., SPECTOR, A., ORRELL, M., SCIOR, K., SWEENEY, J. & CHARLESWORTH, G. 2016. Limited validity of the Hospital Anxiety and Depression Scale (HADS) in

dementia: evidence from a confirmatory factor analysis. *International Journal of Geriatric Psychiatry*.
STOTT, J., SWEENEY, J. M., KOSCHALKA, L., O'CONNOR, L. & MWALE, A. 2017d. People with dementia as peer workers, challenges, and benefits: a thematic analysis and nominal groups study. *International Psychogeriatrics,* 29**,** 1185-1192.
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**,** 337–342.
VEREENOOGHE, 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.
VERHOEVEN, K. J., SIMONSEN, K. L. & MCINTYRE, L. M. 2005. Implementing false discovery rate control: increasing your power. *Oikos,* 108**,** 643-647.
WELLS, C. S. & WOLLACK, J. A. 2003. An instructor's guide to understanding test reliability. *Testing & Evaluation Services. University of Wisconsin*.
ZIGMOND, A. S. & SNAITH, R. P. 1983. The hospital anxiety and depression scale. *Acta psychiatrica scandinavica,* 67**,** 361-370.

Supplementary material: Appendix A: final measure and instructions (separate file)

Table 1: Demographics of the sample

Variable	Older adults (n=77)		Dementia (n= 102)		Significant contrast[†]
	Median (min-max)	% (N)	Median (min-max)	% (N)	
Age	72 (65-92)		81 (58-97)		PLWD > OA
Sex (M)		36 (28)		43 (44)	No sig contrasts
Ethnicity (White)		100 (77)		90 (92)	No sig contrasts
Education (years)	16 (7-25)		12 (5-25)		OA > PLWD
Cognitive impairment (ACE-III score)	95 (67-100)		74 (43-98)		OA > PLWD
Above caseness for anxiety and/or depression (HADS)		14 (11)		44 (44)	PLWD > OA

Note:

[†]Significant at $p<0.05$, adjusted for false discovery rate

Due to the ordinal nature of the data and violation of assumptions for parametric analysis, Mann Whitney U tests were conducted to assess between group differences.

Medians and ranges are reported due to non-Normally distributed data.

Table 2: Factor pattern matrix of final factor loadings after oblique rotation for combined sample

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

Note:
+ Correlation between the factors was 0.4
†Complexity score is an indicator of how much cross-loading there is on a particular item(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. 2 in this case);
Highest loadings for each item are indicated in bold

Table 3: Median scores for BTFQ-D and convergent validity measures in all samples

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

Medians and ranges reported due to non-Normally distributed data.

Table 4: Spearman rank correlation coefficients between ER-40 and BTFQ-D scales (convergent validity) in the full and sub-samples

		ER-40	BTFQ-D Scale	
Sample*	BTFQ-D Scale		Feeling	Thought
Full sample	Feeling	**0.22**	-	-
	Thought	**0.26**	**0.22**	-
	Total	**0.32**	**0.58**	**0.89**
PLWD	Feeling	0.19	-	-
	Thought	0.19	0.16	-
	Total	**0.27**	**0.62**	**0.84**
Older adult	Feeling	-0.03	-	-
	Thought	-0.003	0.05	-
	Total	0.04	**0.38**	**0.90**

Bold represents significant correlation at $p<0.05$, adjusted for False Discovery rate
*this column indicates the sample or subsample in which the correlation was conducted