The development of the PPOM-C for family carers of people with

dementia: Psychometric properties and factor structure

Richard Pione

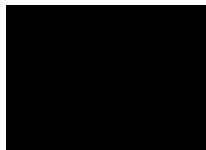D.Clin.Psy. thesis (Volume 1), 2020

University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

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

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

Signature:

Name: Richard Pione

Date: June 2020

Overview

This thesis focused on the quantitative measurement of positive psychology constructs in family carers of people living with dementia. The literature review (Part 1) had several aims. Firstly, it sought to identify positive psychology outcome measures currently in use with family carers. It identified 25 instruments across three positive psychology constructs. These instruments were then subjected to a psychometric appraisal using an established quality assessment tool. The most robust instrument measuring each of the three constructs were then explicitly recommended.

The empirical paper (Part 2) describes the process in which the Positive Psychology Outcome Measure (PPOM) was validated with family carers of people with dementia. The procedure involved recruiting carers using online methods to complete the PPOM amongst a battery of other measures. The PPOM possesses two subscales measuring hope and resilience. Whilst analysing the psychometric properties of the PPOM, two items were deleted to create the PPOM-C. The PPOM-C is a measure that is specific to family carers and possesses good psychometric properties. This empirical paper was completed jointly with another trainee clinical psychologist. This trainee validated an instrument measuring perceived social support and assisted in the design and recruitment stages of the study.

The critical appraisal (Part 3) reflected on the process of conducting the literature review and empirical paper. The nature of quantitative positive psychology measurement was discussed in addition to how challenges were negotiated during the completion of both papers.

Impact

There are approximately 700,000 family carers of people living with dementia in the UK. Research with this population is inundated with examples measuring the degree to which carers experience stress, burden and depression.

However, caring can also be associated with positive feelings and experiences. Family carers have spoken about how they have found meaning through the caring relationship or become closer to the person they provided care for. Others have experienced personal growth through caring.

Many research interventions with family carers seek to illustrate a reduction in instruments measuring constructs like depression or burden. When this occurs, such studies will designate the results as effective or successful. Interventions like this often disregard the pre-existing strengths and resources of family carers through not measuring them. By including robust instruments measuring positive psychology constructs, it is hoped interventions can develop which can augment them.

There are fewer robust positive psychology outcome measures compared to their negative counterparts. This thesis aimed to address this deficit using two approaches. Firstly, a systematic review was carried out to identify and evaluate positive psychology measures in use with family carers. Twenty-five positive measures were identified representing positive aspects to caregiving, resilience and self-efficacy. The measures were then subject to evaluation. Recommendations were then made with respect to which measures were the most psychometrically sound for each of the three categories. Secondly, the empirical paper validated the Positive Psychology Outcome Measure (PPOM) in a sample of 267 family carers. The PPOM was originally developed for people with dementia and measured levels of hope and resilience. After analysis of the results, the PPOM was modified to become the PPOM-C (Positive Psychology

Outcome Measure for Carers of People with Dementia). The PPOM-C is a 14-item

measure which is ready for further research in interventions with family carers. This

would test whether it is able to detect clinically significant changes in participants

following an intervention. Previous research suggests that similar measures are able to

do this, and thus, the PPOM-C is also expected to demonstrate this quality.

Table of Contents

List of Tables

Part 1: Systematic Review

Part 2: Empirical Paper

List of Figures

Part 1: Systematic Review

Part 2: Empirical Paper

Acknowledgements

I am extremely grateful to many people throughout this process of writing my doctoral thesis in clinical psychology. First, and foremost, I would like to thank my primary supervisor, Professor Aimee Spector. From the initial meeting to our last, her warmth, support and thoughtfulness were the perfect antidote to my thesis angst. My secondary supervisor, Dr Charlotte Stoner was also an invaluable support. Her expertise and encouragement buffered me through the trickier times writing this, for which I am grateful.

I would also like to express my appreciation to my research partner Anna Cartwright. It was great to share this journey with someone so resourceful, amiable and generous. I am also indebted to the Join Dementia Research organisation, without whom this project would have been far more difficult to complete. They were amazing in supporting us in contacting family carers.

Thank you to all the family carers who took the time to complete the battery of outcome measures. Without the time you spent completing this, this paper would not have been possible.

Finally, I would like to thank my family. To my wife, Hannah, thank you for your amazing support throughout the last three years and for continuously putting up with my mantra that 'the thesis is almost finished, and things can go back to normal soon'. They can now! Thank you for looking after our two wonderful children, Zach and Alex (especially during COVID-19!). Finally, I am grateful to my brother, Kevin who helped me stay sane playing computer games and my Mum who has encouraged me every step of the way.

Part 1: Literature Review

A psychometric appraisal of positive psychology outcome measures in use

with carers of people living with dementia: A systematic review

Abstract

Background: Positive psychology outcome measures developed for other populations are often adopted for use with family carers. However, the reliability and validity of many such measures are unknown. Therefore, the aim of this review was twofold; (i) To identify all positive psychology measures in use with family carers and (ii) to appraise their psychometric properties. **Method:** A systematic review of positive psychology outcome measures in use with family carers was completed. Searches were made in Ovid MEDLINE, PsychINFO and Web of Science. The development and validation papers for each instrument were then subject to a quality appraisal. **Results:** Twenty-five instruments representing the positive psychology constructs of positive aspects to caregiving, resilience and self-efficacy were found. The Gain in Alzheimer Care Instrument (Yap et al., 2010), the Resilience Scale (Wagnild & Young, 1993) and the Caregiver Efficacy Scale (Crellin et al., 2014) were found to be the highest scoring measures within their respective constructs. **Conclusions:** Although some robust instruments were identified, there were numerous examples of important psychometric properties being unevidenced in development or validation papers. Future researchers and clinicians should administer measures with robust psychometric properties representing positive and negative constructs to obtain a comprehensive picture of a person's wellbeing.

Introduction

Many people living with dementia reside in community settings and are often cared for by family members. In the UK, there are approximately 700,000 such carers (Lewis et al., 2014) caring for an estimated 850,000 people (Prince et al., 2014).

Research with family carers frequently employ instruments measuring concepts such as depression, stress and burden (Teahan et al., 2020). This approach is consistent with constructs described in the stress coping model (Lazarus & Folkman, 1984). This model conceptualises stress as a response to demands on an individual exceeding their perceived or actual resources. These frameworks have been critiqued for solely focusing on the negative aspects of caring and thus, providing an incomplete picture of a carer's wellbeing (Kramer, 1997).

Assessing carers' strengths or positive emotions are areas of research that some feel have been neglected (Tarlow et al., 2004). Qualitative research has evidenced that carers have described positive experiences such as personal growth and a closer relationship with the person they care for (Sanders, 2005). Further, family carers have recommended clinician's ask about positive aspects of caring in addition to the negative facets (Lloyd et al., 2016). This is an important and under-researched area as positive experiences can safeguard carers against burden (Koerner et al., 2009; Pope et al., 2018).

The positive psychology approach refers to how positive emotions, assets, strengths and capabilities can contribute to an individual flourishing (Seligman, 2002). Such strengths or capabilities include but are not limited to instances of growth, mastery, drive and building one's character (Seligman, 1998). There are numerous qualitative positive psychology accounts of the caring experience (Cohen et al., 2002), but fewer quantitative studies. Such quantitative efforts have been praised for their

intent but often critiqued for not using psychometrically robust outcome measures (Stoner, 2019).

The psychometric properties of such outcome measures should be an important consideration for researchers (Dow et al., 2018; Seligman et al., 2005). If a measure's psychometric properties lack validity or reliability, then the quality of data collected with it are uncertain. The authors of such measures typically conclude that their measure possesses adequate psychometric properties. Related literature reviews that evaluated these measures demonstrate they vary from low to medium quality however (Stansfeld et al., 2017; Stoner et al., 2015). Many development and validation papers failed to evidence the responsiveness or sensitivity to change of their respective measure. This is important as the aim of interventional research is to demonstrate significant differences using the measure across time. If responsiveness has not been suitably evidenced, then significant differences found using the measure could be due to other factors.

A recent literature review (Stansfeld et al., 2017) evaluated twelve positive psychology outcome measures that were developed or validated specifically with family carers. However, there are many such measures currently in use that were not developed or validated with family carers. As such, these measures were not included within the Stansfeld et al. (2017) review. This review also did not include information pertaining to the responsiveness of each measure.

The aim of this paper was to extend the Stansfeld et al. (2017) review by identifying all positive psychology outcome measures in use with family carers over the last twenty years. The intention was then to evaluate these measures using an established quality appraisal tool. In addition, data regarding measures' responsiveness was included. This extends the previous review as measures were not excluded if

17

originally developed or validated with a population other than family carers and data on responsiveness was purposefully sought.

Method

Design

The systematic search adhered to the principles set by the PRISMA group (Moher et al., 2009) with respect to searching, screening and appraising the texts. In line with PRISMA recommendations, consideration was made to the population, interventions, comparisons and outcomes (PICO). The population was family carers of people living with dementia. The intervention referred to finding research papers making use of positive psychology instruments. After these instruments were identified, searches were conducted to identify the associated development or validation paper. The comparison component utilised a quality appraisal criteria tool to evaluate all the measures. The outcome of which included recommending the most robust instruments for further use.

Search Strategy

Searches were conducted within the Ovid MEDLINE, PsychINFO and Web of Science databases on the 19th August 2019. Searches used the following headings from which search terms were derived. "Dementia" AND "carer" AND "positive psychology" AND "intervention". Positive psychology search terms were influenced by Seligman's (1998) definition of positive psychology and incorporated a variety of positive qualities which contribute to flourishing (Seligman et al., 2005). It was also influenced by related literature reviews (Stansfeld et al., 2017; Stoner et al., 2017). The heading 'intervention' was chosen with the intent to locate interventional research using outcome measures.

Synonyms and their associated truncations (where appropriate) of the below headings were used as follows.

Search terms for _dementia_ included: "Dementia" or "cognitive impairment" or "Alzheimer" or "senile". Search terms for _carer_ included: "Caregiver" or "family carer" or "relative" or "family" or "friend" or "spouse" or "informal carer" or "carer" or "supporter" or "supportive other". Search terms for _positive psychology_ included: "positive psychology" or "self-efficacy" or "gain" or "satisfaction" or "hope" or "resilience" or "wisdom" or "growth" or "development" or "outlook" or "coherence" or "autonomy" or "pleasure" or "uplift" or "self-realization" or "agency" or "gratitude" or "happiness" or "optimism" or "meaning" or "transcendence" or "affability" or "positivity" or "self-concept" or "humour" or "creativity" or "spirituality" or "love" or "compassion" or "mindfulness" or "acceptance" or "wellbeing" or "independence". Search terms for _intervention_ included: "intervention" or "therapy" or "treatment" or "group" or "group psychotherapy" or "support" or "support groups" or "education" or "psychoeducation" or "cognitive behaviour therapy or "psychotherapy" or "online therapy" or "computer assisted therapy".

Language was phrased to account for both British English and American English spelling. All texts were imported into Mendeley reference management software. Duplicates were removed and the remaining papers were vetted against the review's inclusion and exclusion criteria.

Inclusion Criteria.

1. A positive psychology outcome measure, as denoted by search terms, was employed.

2. The measure was used with family carers of people living with dementia.

3. The development or validation paper of the measure was published in a peer-reviewed journal.

4. The research paper citing the measure was published in a peer-reviewed journal.

5. The study using the measure was published between 1999-2019.

6. Available in English.

Exclusion Criteria.

1. The development or validation paper of the measure was not freely available.

2. Instruments related to external factors (e.g. social support) were excluded to limit the scope of this review to internal positive qualities.

Screening Process

Throughout screening, if the title, abstract or methodology were vague, they were included in the next phase of the process until certainty was provided. Research involving carers and those they cared for as a dyad were included throughout.

Firstly, 5109 titles were screened to determine if family carers were the target population. The abstracts of 907 papers were then assessed to determine whether a suitable methodology was employed. This led to a retrieval of 512 full texts where each method section was examined to identify positive psychology outcome measures. This yielded 130 measures suitable for appraisal.

Throughout screening, the most frequent reasons for exclusion were the omission of positive measures, the use of a qualitative methodology or a population other than family carers (Figure 1).

Refinement of Inclusion Criteria

Due to a large number of results and to ensure this review was manageable, two additional criteria were subsequently added. Firstly, constructs not deemed to be accurate reflections of positive psychology were excluded. Examples of this were 'wellbeing' (n=14 measures) and coping (n= 10 measures; Figure 1). Coping for instance, implies the presence of a negative or stressful event to be managed, rather than a positive emotion or trait as outlined in positive psychology approaches. Secondly, from the remaining constructs, the three most prevalent assessed by the frequency of associated outcome measures, were selected for appraisal.

This resulted in self-efficacy (n=16 measures), resilience (n=10 measures) and positive aspects to caregiving (n=10 measures) being selected. When the corresponding development texts for these measures were searched for, a proportion of them did not meet the inclusion or exclusion criteria. Therefore, a quality appraisal of positive aspects to caregiving (n=8; Table 1), resilience (n=6; Table 2) and self-efficacy (n=11; Table 3) was carried out.

Quality Appraisal Tool

Assessment of the psychometric properties of these measures were conducted using a quality appraisal tool designed to determine properties of health status questionnaires (Terwee et al., 2007). This has been used successfully in related research (Stansfeld et al., 2017; Stoner, et al., 2017) and evaluates measure development on nine criteria. Each criterion produces a score between zero and two indicating how well it had been evidenced within the text. A score of two indicates satisfactory reporting of the relevant statistics or an acceptable description of the design. A criterion scoring one indicated

some methodological deficits or design flaws. Zero was awarded if the criterion had serious methodological problems or no evidence was provided.

The assessed criteria include; content validity, internal consistency, criterion validity, construct validity, agreement, reliability, responsiveness, floor and ceiling effects and interpretability. Content validity assesses how well the construct is represented by items in the questionnaire. It also stipulates adequate consultation with the target population and experts. Internal consistency is appraised by satisfactory Cronbach's alpha and factor analyses. Criterion validity measures the extent to which the measure correlates with the gold standard. Construct validity is achieved through suitable correlations with theoretically related constructs. Agreement measures the absolute measurement error. Reliability measures temporal stability using an Intraclass Correlation Coefficient (ICC) or weighted kappa (≥ 0.70). Responsiveness evaluates the ability to identify clinically important changes over time. Floor and ceiling effects are considered absent if less than fifteen percent of respondents score the minimum or maximum score of the measure. Interpretability is assessed through the application of meaningful qualitative labels to quantitative scores.

The total possible score was 0-18. To increase the interpretability of the total score, descriptive labels were given. Scores from 0–4 were 'poor', 5–9 were 'moderate', 10–14 were 'good', and 15–18 were described as 'very good'. This quality appraisal was independently undertaken with my research partner. After, we met to discuss and resolve differences in scoring.

Identification

Records identified through database searching
(n = 5423)

Records after duplicates removed
(n = 5109)

Records excluded
(n = 4,202)

Most frequent exclusion reasons: qualitative research or with unidentified population

Screening

Records screened on title
(n = 907)

Abstracts deemed suitable for inclusion
(n = 512)

Records excluded
(n = 394)

Excluded outcome measures
(n = 93)

Analysed three most frequent constructs. Excluded: Wellbeing (14), Religiosity & Spirituality (12), Coping (10), Miscellaneous instruments with both positive and negative sub-scales (7), Satisfaction (7), Mindfulness (5), Mastery (4), Compassion (3), Meaning (3), Values (3), Gratitude (2), Growth (2), Happiness (2), Hope (2), Pleasant events (2), Self-esteem (2), Coherence (1), Empathy (1), Empowerment (1), Fortitude (1), Generativity (1), Mood questionnaire (1), Optimism (1), Preparedness (1), Psychological flexibility (1), Resourcefulness (1), self-care (1), Self-concept (1) & Self-transcendence (1)

Eligibility

Included

Total outcome measures identified
(n = 25)

Figure 1. *Systematic review screening process*

Interventional & Cross-Sectional Research using the Measures

The search strategy yielded 113 papers which collectively cited the selected twenty-five measures. This information was used to determine both the responsiveness and the frequency with which each measure was employed. Of the 113 texts, 30 were interventions and determined whether there were any statistically significant changes whilst using the measure. The remaining citations used the measures within a cross-sectional design. The number of citations each measure received was recorded (Table 4) to provide information regarding acceptability by researchers. This was not a factor in determining the overall score of a measure.

Results

After the screening process, 25 outcome measures representing three constructs were subjected to quality appraisal. These constructs included positive aspects to caregiving, resilience and self-efficacy. The Gain in Alzheimer Care Instrument (GAIN; Yap et al., 2010) was found to be the highest scoring measure across all constructs. The Resilience Scale (RS; Wagnild & Young, 1993) was the highest scoring measure for resilience. Finally, the Caregiver Efficacy Scale (CES; Crellin et al., 2014) was the highest scoring self-efficacy measure. The Revised Scale for Caregiving Self-Efficacy (RSSE; Steffen et al., 2002) possessed the highest number of citations evidencing significant differences following an intervention using the measure.

Quality Appraisal

Each construct incorporated measures not originally developed or validated for carers of people with dementia but nevertheless had been used in research. For clarity, measures developed for other populations were identified as such within the appraisal. Each

measure's score on the Terwee et al. (2007) quality appraisal criteria are presented in Tables 1-3. The relevant statistics contributing to each measure's score have been presented in Table 4.

(i)   *Positive Aspects to Caregiving*

For the purposes of this review, positive aspects to caregiving included positive appraisals, gains or rewards linked to being a family carer. Five measures were developed for family carers; Positive Aspects of Caregiving (PAC; Tarlow et al., 2004), Perceived Caregiver Rewards Scale (PCRS; Picot et al., 1997), GAIN (Yap et al., 2010), Gains Associated with Caregiving Scale (GAC; Faba et al., 2017) and the Positive Aspects of Caregiving Questionnaire (PACQ; Abdollahpour et al., 2017). Three measures were developed for other populations: the Scale for Positive Aspects of Caregiving Experience (SPACE; Kate et al., 2012; people with severe mental illnesses), Positive Appraisal of Care Instrument (PCI; Yamamoto-Mitani et al., 2001; carers of older Japanese people), and the Gain Through Group Involvement Scale (GIS, Kaye, 1996; older women).

Measures Developed or Validated with Family Carers (PAC, PCRS, GAIN, GAC & PACQ). The GAIN obtained the highest total with a good score (10/18). The remaining measures all obtained moderate scores. Although all measures obtained full marks for content validity, the GAC comprehensively evidenced this through in-depth qualitative and quantitative engagement with a multitude of family carers.

Internal consistency was reported via Cronbach's α in all five papers. Four of the measures ranged from acceptable (PACQ; α = 0.76) to good (PAC, PCRS, GAIN; α = 0.88-0.89). The GAC's internal consistency was α = 0.95 which indicated multicollinearity. Test-retest reliability was reported for three of the measures (GAIN,

PCRS, & PACQ). Only the GAIN however, evidenced a methodologically sufficient retest to fully satisfy the reliability criterion.

Convergent validity was presented for all measures with all reported results being significant. The PAC had small correlations with burden ($r = 0.23$) and wellbeing ($r = 0.15$). The GAIN had a large correlation with the PAC ($r = .68$) and a small correlation with burden ($r = -.15$). The GAC possessed small correlations with life satisfaction ($r = .26$) and depression ($r = .24$). The PACQ moderately correlated with self-rated health ($r = .34$) and had a small correlation with burden ($r = -.29$). The PCRS obtained moderate correlations with depression ($r = -.30$) and burden ($r = -.35$; Table 4).

The PAC scored one point on the interpretability criterion through providing means and standard deviations of four different samples of participants. As it did not provide a minimally important change (MIC) score, it could not receive full credit for this criterion. The GAIN was the only measure to report an absence of floor and ceiling effects within this construct.

Measures Developed or Validated with Other Populations (SPACE, PCI & GIS). All three measures scored poorly. The SPACE however, demonstrated good content validity through detailed consultations with its target population.

All cited internal consistency using Cronbach's α which varied from acceptable to excellent. The SPACE and PCI shared the highest score ($\alpha = 0.92$). The SPACE did not conduct its factor analysis with enough participants, limiting its internal consistency score. No measures demonstrated criterion or convergent validity. The SPACE was the only measure to demonstrate test-retest reliability but did so with less than fifty participants indicating a methodological limitation.

Table 1
Quality appraisal scores for positive aspects to caregiving measures

| Outcome measure | Content validity | Internal consistency | Criterion Validity | Construct validity | Reproducibility | | Responsiveness | Floor & ceiling effects | Interpretability | Total score |
					Agreement	Reliability				
Positive Aspects to Caregiving (Tarlow et al., 2004)	2	2	0	1	0	0	0	0	1	6
Perceived Caregiver Rewards Scale (Picot et al., 1997)	2	2	0	2	0	1	0	0	0	7
Gain in Alzheimer Care Instrument (Yap et al., 2010)	2	2	0	2	0	2	0	2	0	10
Gains Associated with Caregiving (Fabà et al., 2017)	2	0	1	2	0	0	0	0	0	5
Positive aspects of caregiving questionnaire (Abdollahpour et al., 2017)	2	2	0	2	0	1	0	1	0	8

Table 1
Quality appraisal scores for positive aspects to caregiving measures

| Outcome measure | Content validity | Internal consistency | Criterion Validity | Construct validity | Reproducibility | | Responsiveness | Floor & ceiling effects | Interpretability | Total score |
					Agreement	Reliability				
Scale for Positive Aspects of Caregiving Experience (Kate et al., 2012)	2	1	0	0	0	1	0	0	0	4
Positive Appraisal of Care Instrument (Yamamoto-Mitani et al., 2001)	2	2	0	0	0	0	0	0	0	4
Gain Through Group Involvement Scale (Kaye, 1996)	0	1	0	0	0	0	0	0	0	1

Table 2
Quality appraisal scores for resilience measures

Outcome measure	Content validity	Internal consistency	Criterion Validity	Construct validity	Reproducibility		Responsiveness	Floor & ceiling effects	Interpretability	Total score
					Agreement	Reliability				
Shortened resilience scale (Wilks, 2008)	0	0	0	2	0	0	0	0	1	3
Resilience Scale (Wagnild & Young, 1993)	2	2	0	2	0	1	0	0	1	8
Connor Davidson - Resilience Scale (Connor & Davidson, 2003)	0	2	0	1	0	1	1	0	1	6
Dispositional Resilience Scale (Bartone et al., 1989)	0	0	0	0	0	0	0	0	0	0
Brief Resilience Scale (Smith et al., 2008)	1	2	1	1	0	0	0	0	1	6
Brief Resilient Coping Scale (Sinclair & Wallston, 2004)	0	0	0	2	0	0	1	0	0	3

Table 3
Quality appraisal scores for self-efficacy measures

| Outcome measure | Content validity | Internal consistency | Criterion Validity | Construct validity | Reproducibility | | Responsiveness | Floor & ceiling effects | Interpretability | Total score |
					Agreement	Reliability				
Revised Scale for Caregiving self-efficacy (Steffen et al., 2002)	2	2	0	2	0	1	0	0	0	7
Scales for Caregiving Self-Efficacy (Zeiss et al., 1999)	1	1	0	2	0	0	0	0	0	4
RIS Eldercare SE Scale (Gottlieb & Rooney, 2003)	0	2	0	2	0	0	0	0	0	4
The Family Caregiver's Self-Efficacy for Managing Dementia Scale (Fortinsky et al., 2002)	0	2	0	2	0	0	0	0	0	4
Self-Efficacy Scale (Kuhn & Fulton, 2004)	0	0	0	0	0	0	1	0	0	1

Table 3
Quality appraisal scores for self-efficacy measures

Outcome measure	Content validity	Internal consistency	Criterion Validity	Construct validity	Reproducibility		Responsiveness	Floor & ceiling effects	Interpretability	Total score
					Agreement	Reliability				
Caregiver Efficacy Scale (Crellin et al., 2014)	0	2	1	2	0	0	0	2	1	8
Self-Efficacy Questionnaire for Chinese Family Caregivers (Zhang et al., 2012, 2013)	2	1	0	2	0	1	0	0	0	6
Short Sense of Competence Questionnaire (Vernooij-Dassen et al., 1999)	0	1	0	1	0	2	0	0	0	4
Exercise Self Efficacy Scale (Garcia & King, 1991)	0	1	0	0	0	0	0	0	0	1
Coping Self-Efficacy Scale (Chesney et al., 2006)	2	2	0	0	0	0	0	0	0	4

Table 3

Quality appraisal scores for self-efficacy measures

Outcome measure	Content validity	Internal consistency	Criterion Validity	Construct validity	Reproducibility		Responsiveness	Floor & ceiling effects	Interpretability	Total score
					Agreement	Reliability				
Overall Exercise Self-Efficacy (McAuley, 1993)	0	1	0	0	0	0	0	0	0	1

Table 4
Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
Positive Aspects to Caregiving (Tarlow et al., 2004)	Carers of people with dementia (n = 1229; mean age = 62.2) (USA)	9-items with 2 subscales; (i) self-affirmation & (ii) outlook on life. Scored on a 5-point Likert scale.	α = 0.89 Exploratory Factor Analysis (EFA): Two components with variable loadings of > 0.45 Convergent: Satisfaction of support (r = .15, p < 0.05) Burden (r = -.23, p < 0.05)	Significant responsiveness demonstrated: Czaja et al. (2018; n = 146); Cheung et al. (2015; n = 201); Czaja et al. (2013; n = 110); Beauchamp et al. (2005; n = 299).	25
Perceived Caregiver Rewards Scale (Picot et al., 1997)	Study 1: Female carers of people with dementia (n = 83; mean age = 58.9) Study 2: Carers of non-institutionalized elders with anytime of illness/ disability. (n = 256; mean age = 53.1) (USA)	25-item measure. Scored from 0-4	α = 0.88 Test-retest reliability: α = 0.75 (n = 20) Convergent: Burden (r = -.35, p < *0.0001*) Depression (r = -.30, p < 0.0001)	N/A	0
Gain in Alzheimer Care Instrument (Yap et al., 2010)	Carers of people living with dementia (n= 321; mean age = 50.1) (Singapore)	10-item measure scored on 5-point Likert scale. Higher scores reflect greater gains	α = 0.89 Test-retest reliability: α = 0.79 Convergent: PAC (r = .68, p < 0.0001) Burden (r = -.15, p = 0.02)	N/A	1

Table 4

Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
Gains Associated with Caregiving (Fabà et al., 2017)	Carers of people living with dementia (*n* = 152; mean age = 65) (Spain)	22-item measure scored from 0-3.	α = 0.95 Criterion/ Convergent: GAIN (*r* = .75, *p* < 0.001) Depression (*r* = -.24, *p* < 0.01) Life Satisfaction (*r* = .26, *p* < 0.001 Excellent content validity	N/A	0
Positive aspects of caregiving questionnaire (Abdollahpour et al., 2017)	Carers of people with dementia in Iran. (*n* = 132; mean age = 51.5) (Iran)	10-item measure. Scored from 0-4	α = 0.76 Test-retest reliability: α = 0.91 (*n*= 20) EFA: 2 factors accounted for 47% of the variance in PAC Convergent: Burden (*r* = -.29, *p < 0.001*) Self-rated health (*r* = .34, *p* < 0.01)	N/A	0
Scale for Positive Aspects of Caregiving Experience (Kate et al., 2012)	Primary carers of patients with severe mental illness (SMI; *n* = 203; mean age = 47) (India)	50-item measure with a 5-point Likert scale. 4-factors including (i) caregiving personal gains, (ii) motivation for caregiving role, (iii) caregiver satisfaction and (iv)	α = 0.92 Test-retest reliability (*n* = 37): Each subscale: (i) α = 0.98, (ii) α = 0.95, (iii) α = 0.99 & (iv) α = 0.90 Factor analysis: not enough participants needed to complete factor analysis (minimum = 7 * no. of measure questions) Convergent:	N/A	1

Table 4
Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
		self-esteem and social aspects of caring	None identified		
Positive Appraisal of Care Instrument (Yamamoto-Mitani et al., 2001)	Carers of elderly Japanese care recipients ($n = 337$; mean age = 60.4) (Japan)	21-item measure with 4 domains: (i) relationship satisfaction, (ii) role confidence, (iii) consequential gain and (iv) normative fulfilment	$\alpha = 0.92$ Test-retest reliability: $\alpha = 0.64$ ($n = 80$) Convergent: None	N/A	0
Gain Through Group Involvement Scale (Kaye, 1996)	'Older women' ($n = 225$; mean age = 76.1 years) (USA)	15-item measure scored on a 15-point scale Social, intellectual and emotional gain measured	$\alpha = 0.79$ Convergent: None	N/A	1
Shortened resilience scale (Wilks, 2008)	Alzheimer's carers ($n = 229$). Sample 1 ($n = 115$; mean age = 44.9); Sample 2 ($n = 114$, mean age = 44.6) (USA)	15-item instrument. 7-point Likert scale.	$\alpha = 0.96$ Convergent: Stress ($r = -.60$, p < 0.01) Social support ($r = .34$, $p < 0.01$)	N/A	0
Resilience Scale (Wagnild	Study 1: Development paper with women who had adapted	25-item instrument. 7-point Likert scale	$\alpha = 0.89$ (study 1); $\alpha = 0.91$ (study 2) Test-retest reliability: $\alpha = 0.67$-0.84	Significant responsiveness demonstrated: MacCourt et al., (2017; $n = 123$)	6

Table 4

Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
& Young, 1993)	successfully following a major life event (*n* = 24). Study 2: Psychometric properties explored within older adults (*n* = 810; mean age = 71.1) (USA)		Factor analysis: Yielded two factors, (i) personal competence & (ii) acceptance of self and life. Accounted for 44% of the variance Convergent: Life satisfaction (*r* = .30, *p* < 0.001) Depression (*r* = -.37, *p* < 0.001) Content Validity: Comprehensive methodology to establish instrument		
Connor Davidson - Resilience Scale (Connor & Davidson, 2003)	No specific targeted population (General population, *n* = 577; Primary care outpatients, *n* = 139; Psychiatric outpatients, n = 43; subjects in a GAD trial, n = 25; subjects in a PTSD trial, n = 22; total, n = 806) (USA)	25-item instrument. 5-point Likert scale.	α = 0.89 Test-retest reliability: α = 0.87 Convergent: Hardiness (*r* = .83, *p* < 0.0001) Stress (*r* = -.76, *p* < 0.001)	N/A	3
Dispositional Resilience Scale (Bartone et al., 1989)	Survivor assistance officers within the Army following a plane crash (*n* = 131) (USA)	45-item instrument. 3 domains: commitment, challenge and control	α = 0.85. DPS subscales: α = 0.62-0.82 Convergent: Hardiness (*r* = .93, *p* < 0.05)	N/A	1

Table 4
Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
Brief Resilience Scale (Smith et al., 2008)	American population (n = 354). Included undergraduates (n = 192), cardiac rehab patients (n= 112), women with fibromyalgia (n = 20) and healthy controls (n = 30) (USA)	6-item instrument. 5-point Likert scale	α = 0.80-0.91 Test-retest reliability: α = 0.62 (n= 61; cardiac rehab sample) and 0.69 (n = 48; undergraduate sample) Criterion/ convergent: CD-RISC (r = .59, $p < 0.01$) Optimism (r = .45 – .69, $p < 0.01$) Depression (r = -.41 - -0.66, $p < 0.01$) Predictive: BRS predicted outcomes in stress, depression, anxiety, physical symptoms and positive affect (p < 0.01)	N/A	1
Brief Resilient Coping Scale (Sinclair & Wallston, 2004)	2 samples of individuals with rheumatoid arthritis (sample 1, n = 90, mean age = 46; sample 2, n = 140, mean age = 57.8) (USA)	4-item instrument. 5-point Likert scale	α = 0.69 Test-retest correlation: r = 0.71 (baseline to 5/6 weeks) and r = 0.68 (3 months). Convergent: Each sample correlated as expected with measures of coping, pain coping and psychological wellbeing (overall scales not reported)	N/A	1

Table 4

Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
			Sensitivity to change: Significant effects across the 4 measurement periods ($F = 7.78$; *df* (1, 81) $p < 0.01$) Paired t-tests confirmed increase in mean from pre-post intervention ($t = 2.12$; *df*. 89; $p < 0.01$)		
Revised Scale for Caregiving self-efficacy (Steffen et al., 2002)	Female carers of people with dementia. ($n = 145$; mean age = 77.3 years) (USA)	The RSSE (15 items) measures three subscales; (i) obtaining respite (ii) responding to disruptive patient behaviours and (iii) controlling upsetting thoughts. Scored on a Likert scale (0-100).	α = 0.82 - 0.86 across subscales Test-retest reliability: α = 0.70 - 0.76 across 3 subscales Convergent: Depression ($r = -.38$, $p < 0.001$) Anger ($r = -0.45$, $p < 0.001$) Social support ($r = .47$, $p < 0.001$) Good content validity	Significant responsiveness demonstrated: Easom et al., (2018; $n = 123$); Czaja et al. (2018; $n = 146$); Steffen & Gant (2016; $n = 74$); Lorig et al. (2012; $n = 60$); Marziali & Garcia (2011; $n = 91$); Ducharme et al. (2011; $n = 62$); Au et al. (2010; $n = 13$); Glueckauf et al. (2007; $n = 14$); Gant et al., (2007; $n = 32$); Waelde et al., (2004; ; $n = 12$) No responsiveness demonstrated: Cristancho-Lacroix et al. (2015; $n = 25$); Williams et al. (2010; ; $n = 59$)	33

Table 4

Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
Scales for Caregiving Self-Efficacy (Zeiss et al., 1999)	Carers of frail or cognitively impaired elders (n = 217; mean age = 59.5) (USA)	SCSE consisted of two subscales: (i) problem solving (4 items) and (ii) self-care (10 items). Scored between 0-100. Higher scores reflect higher self-efficacy	α = 0.76 - 0.83 (2 subscales) Convergent: Social support (r = .30, p < 0.0001) Use of coping strategies (r = .19, p < 0.05)	Significant responsiveness demonstrated: Coon et al., (2003; n = 169)	2
RIS Eldercare SE Scale (Gottlieb & Rooney, 2003)	Carers of people with dementia (n = 146; mean age = 61) (Canada)	10-item instrument (3 subscales): (i) relational self-efficacy, (ii) instrumental self-efficacy and (iii) self-soothing self-efficacy. Scored on a 5-point Likert scale. Higher scores reflect higher self-efficacy	α = 0.72 - 0.79 Test-retest reliability: r = 0.48 – 0.69 (p < 0.0001; n = 105) Convergent: Optimism (r = .41, p < 0.001) Anger (r = .35, p < 0.001)	N/A	3
The Family Caregiver's Self-Efficacy for Managing Dementia Scale	Carers of people with dementia (n = 197; mean age = 56) (USA)	10-item instrument with 2 subscales. (i) Symptom management (ii) community support service use	α = 0.77 – 0.78 Convergent: Competence (r = .49, p < 0.01).	Significant responsiveness demonstrated: Boots et al., (2018; n = 41); Lewis et al., (2015; n = 51)	11

Table 4

Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
(Fortinsky et al., 2002)			High scores on self-efficacy associated with fewer depressive symptoms (β = -.17, $p < 0.05$) Content validity: No description of item selection given	Non-significant responsiveness: Stockwell-Smith et al., (2018; $n = 88$); Fortinsky et al. (2014; $n = 19$)	
Self-Efficacy Scale (Kuhn & Fulton, 2004)	Family carers of people with dementia ($n = 45$; mean age = 74.3) (USA)	15-item instrument. 5-point Likert scale	$\alpha = 0.90$ Convergent: None	Demonstrated significant responsiveness: Ducharme et al. (2011; $n = 62$) Demonstrated no responsiveness: Wawrziczny et al. (2019; $n = 51$)	2
Caregiver Efficacy Scale (Crellin et al., 2014)	Family carers of people with dementia ($n = 245$; mean age = 66) (UK)	Items depended on how participant responded on the Neuropsychiatric Inventory (Cummings et al., 1994)	$\alpha = 0.79$ Factor analysis: 3 factors: (i) mood and hyperactivity, (ii) psychosis and night-time disturbance and (iii) euphoria. These factors accounted for 49.85% of the variance Criterion/ convergent: RSSE subscales; SE for obtaining respite (r = -.27, $p < 0.001$), responding to disruptive behaviour	N/A	0

Table 4
Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
			($r = -.39$, $p < 0.001$) and controlling upsetting thoughts ($r = -.39$, $p < 0.001$) Depression (r = .36, $p < 0.001$) Anxiety ($r = .38$, p < 0.001)		
Self-Efficacy Questionnaire for Chinese Family Caregivers (Zhang et al., 2012, 2013)	Study 1: Family carers ($n = 10$). Study 2: Family carers of people living with dementia ($n = 196$; mean age = 63; study 2) (China)	27-item questionnaire with 5 domains: (i) gathering information (ii) obtaining support (iii) responding to behavioural disturbances (iv) managing household and care (v) managing distress	$\alpha = 0.94$ Test-retest reliability: $\alpha = 0.74$ ($n = 24$) Convergent: Neuropsychiatric symptoms negatively associated with ability to manage distress ($\beta = -.30$, p < 0.001) Social support significantly influenced the SEQ after controlling for neuropsychiatric symptoms ($\beta = -.42$, p < 0.001	Demonstrated significant responsiveness: Hou et al. (2014; $n= 70$; caregivers of people with chronic conditions); Kwok et al. (2013; $n = 18$) Demonstrated no responsiveness: Kwok et al. (2014; $n = 36$)	8
Short Sense of Competence Questionnaire (Vernooij-Dassen et al., 1999)	Carers of people with dementia ($n = 141$; Mean age = 63) (The Netherlands)	7-item instrument. 5-point Likert scale	$\alpha = 0.76$ Test-retest reliability: $\alpha = 0.68 – 0.87$ Criterion/ Convergent: Sense of Competence ($r = .88$, no reported *p* value)	Demonstrated significant responsiveness: van Knippenberg et al. (2018; $n = 76$) Demonstrated no responsiveness: Hattink et al. (2015; $n = 142$)	7

Table 4

Description of statistics for included instruments

Measure	Sample population (country of origin)	Description	Reliability & Validity	Responsiveness studies	Number of citations
Exercise Self Efficacy Scale (Garcia & King, 1991)	Not stated (USA)	14-items. (0-100% confidence scoring)	$\alpha = 0.90$ Test-retest reliability: $r = 0.67$ (n= 62) Convergent: None	Demonstrated significant responsiveness: Connell & Janevic (2009; $n = 74$)	1
Coping Self-Efficacy Scale (Chesney et al., 2006)	HIV-seropositive men ($n = 348$; mean age = 40.5) (USA)	13-item instrument (11-point scoring scale). 3 sub-scales: (i) problem focused coping, (ii) stop unpleasant emotions and thoughts and (iii) getting support	$\alpha = 0.80 - 0.91$ Test-retest reliability: $r = 0.49 - r = 0.80$ Convergent: Anxiety ($r = -.27$, $p < 0.0001$) Positive morale ($r = .23$, $p < 0.0001$)	N/A	5
Overall Exercise Self-Efficacy (McAuley, 1993)	'Middle-aged adults' ($n = 65$; mean age = 55) (USA)	An exercise specific self-efficacy scale and a self-efficacy scale measuring continued exercise participation were combined to create the OESE. No further details were provided	$\alpha = 0.90$ Convergent: None	N/A	1

Use in Carer Research. The PAC was the only measure to demonstrate responsiveness in interventions (n = 4; Table 4). All other texts citing positive aspects to caregiving measures had utilised them in cross-sectional research (PAC; n = 21; GAIN, n = 1; SPACE, n = 1; GIS, n = 1).

Positive Aspects to Caregiving Summary and Recommendation. The GAC was the only measure that attempted to obtain criterion validity through a large positive significant relationship with the GAIN. The PAC was the only positive aspects to caregiving measure to demonstrate sensitivity to change within interventions. Owing to its comprehensive development, it is recommended that the GAIN be used to measure positive aspects to caregiving. It fully satisfied the content validity, construct validity, reliability and the floor and ceiling effects criteria.

(ii) Resilience

The Resilience Scale 15 (RS15; Wilks, 2008) was the only measure found to have been developed or validated with family carers. Five instruments measuring resilience were found to have been developed or validated for populations other than family carers; the Resilience Scale (RS; Wagnild & Young, 1993; women who adapted after a serious life event), the Connor-Davidson Resilience Scale (CD-RISC; Connor & Davidson, 2003; different American populations), the Dispositional Resilience Scale (DRS; Bartone et al., 1989; survivor assistance officers in the army), the Brief Resilience Scale (BRS; Smith et al., 2008; different American populations) and the Brief Resilient Coping Scale (BRCS; Sinclair & Wallston, 2004; people with rheumatoid arthritus).

Measures Developed or Validated with Family Carers (RS15): The RS15 (Wilks, 2008) scored poorly. It did not evidence content validity sufficiently and no points were given for internal consistency as the reported Cronbach's α indicated multicollinearity. Convergent validity was achieved via a large significant correlation with perceived stress (r = -.60) and a moderate significant correlation with social support (r = .34; Table 4).

Measures Developed or Validated with Other Populations (RS, CD-RISC, DRS, BRS & BRCS). The RS was the highest scoring measure (8/18) within the resilience construct. The remaining measures varied from poor to moderate. The RS was the only instrument to involve the target population sufficiently to satisfactorily illustrate content validity. The RS, CD-RISC and BRS were the only measures to report acceptable Cronbach's α scores (α = 0.81 – 0.91). The RS and CD-RISC reported test-retest reliability but did not obtain the full score for methodological reasons.

Convergent validity was present for all measures with all reported results being significant. The RS obtained moderate correlations with life satisfaction (*r* = .30) and depression (*r* = -.37). The CD-RISC possessed large correlations with hardiness (*r* = .83) and perceived stress (*r* = -.76). The DPS had a large correlation with hardiness (*r* = .93). The BRS attained moderate correlations with depression (*r* = -.41) and optimism (*r* = .45; Table 4). The BCRS obtained a small negative correlation with depression and a small positive correlation with life satisfaction (not reported by authors).

Use in Carer Research. The RS was the only resilience measure to demonstrate sensitivity to change with 123 family carers (MacCourt et al., 2017). It found a significant increase in resilience following a grief intervention. All other texts citing

resilience measures used them within cross-sectional research (RS, n = 5; CD-RISC, n = 3; DRS, n = 1; BRS, n = 1; BRCS, n = 1; Table 4).

Resilience Summary and Recommendation. No resilience measures evidenced the absence of floor and ceiling effects or provided information on minimally important change (MIC) to aid their interpretability score. Despite this measure not being developed for family carers, the RS is the recommended instrument to measure resilience. It possessed acceptable content validity, internal consistency, construct validity in addition to partially satisfying the reliability and interpretability criteria. It was the highest scoring resilience measure and some research has demonstrated it is responsive to intervention with family carers.

(iii) Self-Efficacy

There were eight measures developed for family carers; the Revised Scale for Caregiving Self-Efficacy (RSSE; Steffen et al., 2002), the Scales for Caregiving Self-Efficacy (SCSE; Zeiss et al., 1999), the RIS Eldercare Self-Efficacy Scale (RIS; Gottlieb & Rooney, 2003), Caregivers' Self-Efficacy for Managing Dementia Scale (MDS; Fortinsky et al., 2002), the Self-Efficacy Scale (SES; Kuhn & Fulton, 2004), the Caregiver Efficacy Scale (CES; Crellin et al., 2014), the Self-Efficacy Questionnaire for Chinese Family Caregivers (SEQ; Zhang et al., 2013) and the Short Sense of Competence Questionnaire (SSCQ: Vernooij-Dassen et al., 1999). Three measures were found which were developed for other populations; the Exercise Self-Efficacy Scale (ESES; Garcia & King, 1991; population not specified), the Coping Self-Efficacy Scale (CSES; Chesney et al., 2006; men with Human Immunodeficiency Virus) and the Overall Exercise Self-Efficacy instrument (OESE; McAuley, 1993; middle aged adults).

Measures Developed or Validated with Family Carers (RSSE, SCSE, RIS, MDS, SES, CES, SEQ, SSCQ). The scores varied from poor (SES; 1/18) to moderate (CES; 8/18). The RSSE and SEQ were the only measures to evidence content validity effectively. The other instruments failed to involve the target population or experts in their respective development phase.

Every measure cited Cronbach's α, varying from acceptable to excellent. The SEQ obtained the highest score for internal consistency (α = 0.94) and the RIS obtained the lowest (α = 0.72 -0.79). The absence of factor analyses limited the internal consistency score for the SCSE and SES. The SEQ completed a factor analysis with too few respondents to adequately satisfy the internal consistency criterion. Test-retest reliability was not suitably demonstrated for a single measure. Those that attempted to measure test-retest reliability were marred by methodological shortcomings such as a small sample size.

Convergent validity was evidenced for all measures (excepting the SES and SSCQ) with all reported results being significant. The RSSE obtained moderate correlations with depression ($r = -.38$) and social support ($r = .47$). The SCSE attained small correlations with social support ($r = .30$) and logical analysis ($r = .19$). The RIS possessed moderate correlations with optimism ($r = .41$) and anger expression ($r = -.35$). The MDS had a moderate correlation with sense of competence ($r = .49$). The CES obtained moderate sized correlations with depression ($r = -.36$) and anxiety ($r = -.38$; Table 4). The SEQ found social support significantly affected all of its subscales whilst controlling for the care recipient's neuropsychiatric symptoms. Multiple regression confirmed care recipients' neuropsychiatric symptoms were negatively associated with the caregivers' ability to manage distress. The SSCQ inferred convergent validity

through a large positive significant correlation with its longer version counterpart; the Sense of Competence Questionnaire (SCQ; Vernooij-Dassen, 1993; $r = 0.88$).

Measures Developed with Other Populations (ESES, CSES, OESE). All three measures scored poorly. Both exercise instruments obtained a point for reporting on the Cronbach's α score which were acceptable. The CSES demonstrated good content validity, internal consistency and factor structure. However, the measures did not evidence any other psychometric properties contributing to their collective poor scores.

Use in Carer Research. The RSSE had nine citing texts indicating responsiveness to an intervention, and two papers indicating an effect had not been found. On the balance of probabilities, the RSSE is likely to be responsive to change. Other measures with citing texts indicating responsiveness were the SCSE (n = 1), MDS (n = 2), SES (n = 1), SSCQ (n = 1), ESES (n = 1) and the SEQ (n = 2). There were a variety of texts citing these measures that used them within cross-sectional research (RSSE, n = 24; MDS, n = 7; CSES, n = 5; SEQ, n = 5; SSCQ, n = 5; RIS, n = 3; SCSE, n = 1; OESE, n = 1; Table 4).

Self-Efficacy Summary and Recommendation. Both the exercise self-efficacy scales and the SES were the lowest scoring instruments within this construct (1/18). No measures evidenced agreement or interpretability. The CES obtained the highest score (8/18). It was the only measure to evidence a lack of floor and ceiling effects and possessed good internal consistency and construct validity. It also partially satisfied the criterion validity and interpretability criteria. Owing to how the CES is linked with the Neuropsychiatric Inventory (NPI, Cummings et al, 1993), the CES is recommended for

use when measuring carers' self-efficacy ratings in managing behavioural and psychological symptoms of dementia (BPSD). As carers were not involved in the development of the CES, it could not score a point in the content validity criterion. However, that is not to suggest there is no content validity within the measure. Its development was informed by relevant self-efficacy literature (Bandura, 2006).

If self-efficacy in managing BPSD is not the focus of interest, then the RSSE would be recommended as the next highest scoring self-efficacy measure (7/18). The RSSE comprehensively demonstrated its content validity through a robust process of item selection and factor analysis. It evidenced good construct validity and could have scored higher on its reliability criterion had it used an ICC statistic for the test-retest. Additionally, it was the most frequently cited self-efficacy measure and many studies demonstrated its responsiveness (n=10). This is important as it suggests that if an intervention does lead to a significant increase in self-efficacy, then the RSSE is likely to detect it.

Discussion

Twenty-five positive psychology outcome measures were identified as having been developed, validated or in use with family carers of people living with dementia. These measures represented the constructs of positive aspects to caregiving, resilience and self-efficacy. This review extends previous research in several ways. It identified, examined and recommended one measure for each examined construct. It is also the first review that includes data regarding each measure's responsiveness. This is an often-overlooked psychometric property that has important implications for interventional research.

All measures except the GAIN obtained a 'poor' or 'moderate' score in the quality appraisal. No measure satisfactorily evidenced criterion validity, agreement, responsiveness or interpretability. Many measures adequately evidenced content validity which has been suggested to be the most valuable criterion (Terwee et al., 2007). The highest overall scoring measure was the GAIN (Yap et al., 2010) mirroring the finding of a related review (Stansfeld et al., 2017). The GAIN measured positive aspects to caregiving. The highest scoring resilience measure was the RS (Wagnild and Young, 1993). The highest scoring self-efficacy instrument was the CES (Crellin et al., 2014).

Clinicians or researchers employing a positive psychology outcome measure may choose to utilize the recommended measures from this review. The choice of measure may also depend on the instrument required. There is ample variability with respect to how instruments measured their corresponding construct. For example, the positive aspects to caregiving construct includes instruments measuring rewards (PCRS), gains in caregiving (GAC, GAIN) and positive appraisals in caregiving (PCI).

Limitations

This review had initially identified 130 positive psychology outcome measures suitable for a quality appraisal. As this was beyond the scope of a single review, the additional criteria to exclude certain constructs and select the three most prevalent were added. An alternate method could have selected the twenty most frequently cited measures. This would not have resulted in a single recommended measure per construct, however. Through choosing the three most frequent constructs, it was hoped that the review captured the most prevalent and thus, the most important.

It is possible that instruments not originally developed for family carers included in this review had been validated for family carers elsewhere. It was beyond the scope of this review to make additional searches for these measures to determine this. This, however, could be the focus of a future systematic review.

The review's search strategy used the key search term 'intervention'. The intention was to locate measures within interventional research to provide information on responsiveness. Many of the papers however were cross-sectional in nature. The wide-ranging search terms used may have contributed to this. Given the vast number of instruments found, it was assumed that an appropriate number of measures had been accumulated and no further refinements were made. It is possible that this assumption was incorrect and subsequently, other positive measures could have been missed.

This review was influenced by Seligman's definition (1998) of positive psychology. It is conceivable that other positive psychology frameworks such as Ryff's Scales of Psychological Wellbeing (1989) could have yielded other results.

The GAIN was found to be the highest scoring measure but had only one citation within the search results. Several explanations could account for this. It is possible that as the measure was developed in Singapore, research citing it could be published in another language. Conversely, it is conceivable that researchers are not using the GAIN, or English translations of it do not exist.

This review was somewhat limited by the quality appraisal criteria. The content validity criterion for example had many measures scoring the maximum score. As this is a crucial criterion to assess, this could be scored on a scale from 0-3 as opposed to 0-2. This would enable instruments to be differentiated from each other with respect to how they satisfied this criterion. The GAIN and the GAC for instance both satisfied the content validity criterion fully, and yet, the GAC possessed more content validity than

the GAIN. The GAIN made use of a focus group of twelve Singaporean carers to inform the items on the measure. In contrast, the GAC made use of over 150 carers to obtain quantitative approval ratings for 62 potential questions. Having a wider range of scoring could distinguish between measures on criteria where scores are predominantly homogeneous.

A further limitation involved not including data measuring responsiveness for each instrument in the quality appraisal. This was omitted for several reasons. Firstly, the quality appraisal criteria infer that this responsiveness data should be collected within the development paper. Secondly, few papers citing the measures in this review demonstrated significant differences using the smallest detectable difference or minimally important change equations (Terwee et al., 2007). Thus, it was beyond the scope of this review to contact the authors of such papers to request the necessary data to calculate them. This would however be a valuable piece of work for a future review.

This paper collated and evaluated positive measures in use with family carers regardless of whether they had been developed with that population. This decision was made with the aim to determine the quality of all measures in use. Ideally however, all measures used with family carers should have been developed or validated with this population. This is to ensure that each measure possesses adequate content validity. The measures representing self-efficacy (CES) and positive aspects of caregiving (GAIN) were developed specifically for family carers. However, the highest scoring resilience measure (RS) did not.

Future Research

The screening process elicited 130 eligible instruments for appraisal. The additional criterion of analysing the three most cited constructs left 93 unanalysed measures. Such

constructs included 'coping', 'wellbeing' and 'satisfaction' and could undergo a similar quality appraisal process.

As this review highlighted a lack of focus on responsiveness across all instruments, future reviewers could complete a multi-lingual review. This would focus on finding interventions where instruments have been used in languages other than English to bring together a cohesive global account of measures.

It was noted that many measures could have increased their quality appraisal score through evidencing additional psychometric properties. For example, only two of twenty-five measures provided data detailing an absence of floor and ceiling effects. It is possible that many more instruments lacked floor or ceiling effects, but due to a lack of evidence, could not satisfy this criterion. Therefore, future authors could design instruments fulfilling criteria on established quality appraisal tools (e.g.Terwee et al., 2007) when developing measures. This can serve as a checklist to ensure all criterion have the potential to be evidenced.

Future research could also involve a validation study of the RS with a family carer sample. Although the author of the RS15 (Wilks, 2008) attempted to do this, the psychometric properties were not sufficient, so further work is needed.

The recommended measure for self-efficacy was the CES, with the RSSE (Steffen et al., 2002) falling just behind. A major difference between the two instruments being that the RSSE was the most cited measure within the review. The RSSE possessed ten citations where significant differences in self-efficacy were present in an intervention. The CES had no citations at all. If future researchers were to utilize the CES within research settings, this could determine whether it has suitable responsiveness properties. Due to the nature of the CES, this could look like an

intervention seeking to increase the self-efficacy of carers managing behaviour that challenges in those that they care for.

Implications for practice

Psychosocial research with family carers often include instruments measuring burden, stress or depression. Measuring these constructs can be an important part of a clinician's assessment. This could also have the unintended consequence of reinforcing the narrative that caring is inevitably linked with depression, feeling burdened and burned out. This emphasis may be incompatible with those carers who subscribe to a different narrative. For example, carers who experience personal growth in looking after their loved ones (Wong et al., 2009), or who still harbour hopes and dreams about the future.

When planning interventions with family carers, it is beneficial to use of a variety of instruments, reflecting a range of different constructs. This reflects the complexity of the human experience and the capacity to experience a full range of emotions. This could give researchers and carers opportunities to reflect on both positive and negative emotions in relation to any given intervention.

The adoption of less robust measures will have important clinical implications. These will vary in accordance with each unevidenced criterion. For instance, measures that did not adequately evidence convergent validity cannot be assumed to hold construct validity. Also, clinicians who find changes using a measure across time that has not had its minimally important change score evidenced sufficiently would face complexities in analysing these differences. The differences could be due to the intervention or to error found within the measure.

Conclusions

This review demonstrates that there are some positive psychology outcome measures with desirable psychometric properties in use for family carers. There were also many measures with methodological flaws. The most psychometrically sound positive aspects to caregiving measure was the GAIN (Yap et al., 2010). The highest scoring resilience measure was the RS (Wagnild and Young, 1993). The highest scoring self-efficacy measure was the CES (Crellin et al., 2014). Of these three recommended measures, only the RS demonstrated responsiveness in a single study (MacCourt et al., 2017). Thus, adopting all three measures in future interventions for further evaluation will give valuable information pertaining to their responsiveness. Through targeting an increase in such positive domains, it is hoped specific interventions will be developed to specifically strengthen them and thus, increase quality of life.

Positive psychology advocates have appealed for the development of behaviour-based domain specific measures (Seligman et al., 2005) but research using such measures can lack rigorous methodology or use instruments that are not robust (Stoner, 2019). This highlights the importance of both measure development and the need to be selective when choosing measures for clinical or research use.

References

Abdollahpour, I., Nedjat, S., Noroozian, M., Salimi, Y., & Majdzadeh, R. (2017). Positive aspects of caregiving questionnaire: A validation study in caregivers of patients with dementia. *Journal of Geriatric Psychiatry and Neurology*, *30*, 77–83. https://doi.org/10.1177/0891988716686831

Au, A., Li, S., Lee, K., Leung, P., Pan, P. C., Thompson, L., & Gallagher-Thompson, D. (2010). The coping with caregiving group program for Chinese caregivers of patients with Alzheimer's disease in Hong Kong. *Patient education and counseling*, *78*, 256–260. https://doi.org/10.1016/j.pec.2009.06.005

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan and F. Pajares (eds.), *Self-efficacy beliefs of adolescents.* Information Age Publishing.

Bartone, P. T., Ursano, R. J., Wright, K. M., & Ingraham, L. H. (1989). The impact of a military air disaster on the health of assistance workers. A prospective study. The Journal of Nervous and Mental Disease, 177, 317 - 328. https://doi.org/10.1097/00005053-198906000-00001

Beauchamp, N., Irvine, A. B., Seeley, J., & Johnson, B. (2005). Worksite-based internet multimedia program for family caregivers of persons with dementia. *The Gerontologist*, *45*, 793-801. https://doi.org/10.1093/geront/45.6.793

Boots, L. M., de Vugt, M. E., Kempen, G. I., & Verhey, F. R. (2018). Effectiveness of a blended care self-management program for caregivers of people with early-stage dementia (partner in balance): Randomized controlled trial. Journal of Medical Internet Research, 20, e10017. https://doi.org/10.2196/10017

Chesney, M. A., Neilands, T. B., Chambers, D. B., Taylor, J. M., & Folkman, S. (2006). A validity and reliability study of the coping self-efficacy scale. *British Journal of Health Psychology*, *11*, 421–437. https://doi.org/10.1348/135910705X53155

Cheung, K. S., Lau, B. H., Wong, P. W., Leung, A. Y. Lou, V. W. Q., Chan, G. M., & Schulz, R.(2015). Multicomponent intervention on enhancing dementia caregiver well-being and reducing behavioral problems among Hong Kong Chinese: a translational study based on REACH II. *International Journal of Geriatric Psychiatry, 30*, 460-469. https://doi.org/10.1002/gps.4160

Cohen, C. A., Colantonio, A., & Vernich. L. (2002). Positive aspects of caregiving: Rounding out the caregiver experience. *International Journal of Geriatric Psychiatry, 17,* 184- 188. https://doi.org/10.1002/gps.561

Connell, C. M., & Janevic, M. R. (2009). Effects of a telephone-based exercise intervention for dementia caregiving wives: A randomized controlled trial. *Journal of Applied Gerontology*, *28*, 171-194. https://doi.org/10.1177/0733464808326951

Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety*, *18*, 76–82. https://doi.org/10.1002/da.10113

Coon, D. W., Thompson, L., Steffen, A., Sorocco, K., & Gallagher-Thompson, D. (2003). Anger and depression management: psychoeducational skill training interventions for women caregivers of a relative with dementia. *The Gerontologis*t, *43*, 678-689. https://doi.org/10.1093/geront/43.5.678

Crellin, N., Charlesworth, C., & Orrell, M. (2014). Measuring family caregiver efficacy for managing behavioral and psychological symptoms in dementia: A psychometric evaluation. *International Psychogeriatrics*, *26*, 93– 103. https://doi.org/10.1017/S1041610213001646

Cristancho-Lacroix, V., Wrobel, J., Cantegreil-Kallen, I., Dub, T., Rouquette, A., & Rigaud, A. S. (2015). A web-based psychoeducational program for informal

caregivers of patients with Alzheimer's disease: A pilot randomized controlled

trial. *Journal of Medical Internet Research*, *17*, 117.

https://doi.org/10.2196/jmir.3717

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

Gombein, J. (1994). The neuropsychiatric inventory: Comprehensive assessment

of psychopathology in dementia. *Neurology*, *44*, 2308–2314.

https://doi.org/10.1212/WNL.44.12.2308

Czaja, S. J., Loewenstein, D., Schulz, R., Nair, S. N., & Perdomo, D. (2013). A

videophone psychosocial intervention for dementia caregivers. *American

Journal of Geriatric Psychiatry*, *21*, 1071-1081.

https://doi.org/10.1016/j.jagp.2013.02.019

Czaja, S. J., Chin, C. L., Perdomo, D., Loewenstein, D., Bravo, M., Moxley, J. H., &

Schulz, R. (2018). Community REACH: An implementation of an evidence-

based caregiver program. The Gerontologist, 58, 130-

137. https://doi.org/10.1093/geront/gny001

Dam, A., Boots, L., Van Boxtel, M., Verhey, F., & De Vugt, M. (2018). A mismatch

between supply and demand of social support in dementia care: A qualitative

study on the perspectives of spousal caregivers and their social network

members. *International Psychogeriatrics, 30*, 881-892.

https://doi.org/10.1017/S1041610217000898

Dow, J., Robinson, J., Robalino, S., Finch, T., McColl, E., & Robinson, L. (2018). How

best to assess quality of life in informal carers of people with dementia: A

systematic review of existing outcome measures. *PLoS ONE*, *13*, 1–18.

https://doi.org/10.1371/journal.pone.0193398

Ducharme, F. C., Lévesque, L. L., Lachance, L. M., Kergoat, M., Legault, A. J.,

Beaudet, L. M., & Zarit, S. H. (2011). "Learning to become a family caregiver"

efficacy of an intervention program for caregivers following diagnosis of

dementia in a relative. *The Gerontologist, 51*, 484-494.

https://doi.org/10.1093/geront/gnr014

Easom, L., Cotter, E., & Ramos, A. (2018). Comparison of African American and

Caucasian caregiver self-efficacy. *Journal of Gerontological Nursing*, *44*, 16-21.

https://doi.org/10.3928/00989134-20171023-01

Faba, J., Villar, F., & Giuliani, M. F. (2017). Development of a measure to evaluate

gains among spanish dementia caregivers: The gains associated with caregiving

(GAC) scale. *Archives of Gerontology and Geriatrics*, *68*, 76–83.

https://doi.org/10.1016/j.archger.2016.09.004

Fortinsky, R. H., Kercher, K., & Burant, C. J. (2002). Measurement and correlates of

family caregiver self-efficacy for managing dementia. *Aging & Mental Health*,

6, 153–160. https://doi.org/10.1080/13607860220126763

Fortinsky, R. H., Delaney, C., Harel, O., Pasquale, K., Schjavland, E., Lynch, J.,

Kleppinger, A., & Crumb, S. (2014). Results and lessons learned from a nurse

practitioner-guided dementia care intervention for primary care patients and

their family caregivers. *Research in Gerontological Nursing*, *7,* 126-137.

https://doi.org/10.3928/19404921-20140113-01

Gant, J. R., Steffen, A. M., & Lauderdale, S. A. (2007). Comparative outcomes of two

distance-based interventions for male caregivers of family members with

dementia. *American Journal of Alzheimer's Disease and Other Dementias*, *22,*

120-128. https://doi.org/10.1177/1533317506298880

Garcia, A. W., & King, A. C. (1991). Predicting long-term adherence to aerobic

exercise: A comparison of two models. *Journal of Sport and Exercise*

Psychology, *13*, 394–410. https://doi.org/10.1123/jsep.13.4.394

Gaugler, J. E., Kane, R. L., & Newcomer, R. (2007). Resilience and transitions from

dementia caregiving. *The Journals of Gerontology: Series B: Psychological*

Sciences and Social Sciences, *62*, 38-44.

https://doi.org/10.1093/geronb/62.1.P38

Glueckauf, R. L., Jeffers, S. B., Sharma, D., Massey, A. J., Davis, W. S., Wesley, L. M.,

Byrd, V.,McIntyre, V., Stine, C., & Martin, C. (2007). Telephone-based

cognitive-behavioral intervention for distressed rural dementia caregivers.

Clinical Gerontologist, 31, 21-41. https://doi.org/10.1300/J018v31n01_03

Gottlieb, B. H., & Rooney, J. A. (2003). Validation of the RIS eldercare self-efficacy

scale. *Canadian Journal on Aging*, *22*, 95–107.

https://doi.org/10.1017/S0714980800003767

Hattink, B., Meiland, F., van der Roest, H., Kevern, P., Abiuso, F., Bengtsson, J.,

Giuliano, A., Duca, A., Sanders, J., Basnett, F., Nugent, C., Kingston, P., &

Droes, R. (2015). Web-based STAR e-learning course increases empathy and

understanding in dementia caregivers: Results from a randomized controlled

trial in the Netherlands and the United Kingdom. Journal of Medical Internet

Research, 17, e241. https://doi.org/doi:10.2196/jmir.4025

Hickman, H., Clarke, C., & Wolverson, E. (in press). A qualitative study of the shared

experience of humour between people living with dementia and their partners.

Dementia. https://doi.org/10.1177/1471301218805895

Hou, R. J., Wong, S. Y., Yip, B. H., Hung, A., Lo, H., Chan, P. H., Lo, C. S., Kwok, T.,

Tang, W., Mak, W. W., Mercer, S. W., & Ma, S. H. (2014). The effects of

mindfulness-based stress reduction program on the mental health of family

caregivers: a randomized controlled trial. *Psychotherapy and Psychosomatics*,

83, 45-53. https://doi.org/10.1159/000353278

Kate, N., Grover, S., Kulhara, P., & Nehra, R. (2012). Scale for positive aspects of

caregiving experience: Development, reliability, and factor structure. *East Asian

Archives of Psychiatry*, *22*, 62–69.

Koerner, S. S., Kenyon, D. B., & Shirai, Y. (2009). Caregiving for elder relatives:

Which caregivers experience personal benefits/gains? *Archives of Gerontology

and Geriatrics*, *48*, 238–245. https://doi.org/10.1016/j.archger.2008.01.015

Kramer, B. J. (1997). Gain in the caregiving experience: Where are we? What next? *The

Gerontologist, 37*, 218-232. https://doi.org/10.1093/geront/37.2.218

Kwok, T., Wong, B., Ip, I., Chui, K., Young, D., & Ho, F. (2013). Telephone-delivered

psychoeducational intervention for Hong Kong Chinese dementia caregivers: A

single-blinded randomized controlled trial. *Clinical Interventions in Aging*, *8*,

1191-1197. https://doi.org/10.2147/CIA.S48264

Kwok, T., Au, A., Wong, B., Ip, I., Mak, V., & Ho, F. (2014). Effectiveness of online

cognitive behavioral therapy on family caregivers of people with dementia.

Clinical Interventions in Aging, 9, 631-636.

https://doi.org/10.2147/CIA.S56337

Kuhn, D., & Fulton, M. A. (2004). Efficacy of an educational program for relatives of

persons in the early stages of Alzheimer's disease. *Journal of Gerontological

Social Work, 42,* 109-130. https://doi.org/10.1300/J083v42n03_07

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer

publishing company.

Lewis, F., Karlsberg Schaffer, S., Sussex, J., O'Neill, P., & Cockcroft, L. (2014). The trajectory of dementia in the UK - making a difference. *Office of Health Economics*. Alzheimer's Research UK.

Lewis, V., Bauer, M., Winbolt, M., Chenco, C., & Hanley, F. (2015). A study of the effectiveness of MP3 players to support family carers of people living with dementia at home. *International Psychogeriatrics*, *27*, 471-479. https://doi.org/10.1017/S1041610214001999

Lloyd, J., Patterson, T., & Muers, J. (2016). The positive aspects of caregiving in dementia: A critical review of the qualitative literature. *Dementia*, *15*, 1534–1561. https://doi.org/10.1177/1471301214564792

Lorig, K., Thompson-Gallagher, D., Traylor, L., Ritter, P. L., Laurent, D. D., Plant, K., Thompson, L. W., & Hahn, T. J. (2012). Building better caregivers: A pilot online support workshop for family caregivers of cognitively impaired adults. *Journal of Applied Gerontology*, *31*, 423–437. https://doi.org/10.1177/0733464810389806

MacCourt, P., McLennan, M., Somers, S., & Krawczyk, M. (2017). Effectiveness of a grief intervention for caregivers of people with dementia. Omega: Journal of Death and Dying, 75, 230–247. https://doi.org/10.1177/0030222816652802

Marziali, E., & Garcia, L. J. (2011). Dementia caregivers' responses to 2 internet-based intervention programs. *American Journal of Alzheimer's Disease and Other Dementias*, *26*, 36-43. https://doi.org/10.1177/1533317510387586

McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *Journal of Behavioral Medicine*, *16*, 103–113. https://doi.org/10.1007/BF00844757

Moher, D., Liberate, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting item

    carers: a comparitive study of carers of demented amd nondemented older

    persons. *International Journal of Geriatric Psychiatry, 20*, 591 - 592.

Pearlin, L. I., Mullan, J. T., Semple, S. J., & Skaff, M. M. (1990). Caregiving and the

    stress process: an overview of concepts and their measures. *The Gerontologist*,

    30, 583–594. https://doi.org/10.1093/geront/30.5.583

Picot, S.J.F., Youngblut, J., & Zeller, R. (1997). Development and testing of a measure

    of perceived caregiver rewards in adults. *Journal of Nursing Measurement*, *5*,

    33-52.

Pope, N. D., Baldwin, P. K., & Lee, J. J. (2018). "I didn't expect to learn as much as I

    did": Rewards of caregiving in young adulthood. *Journal of Adult Development,*

    25, 186–197. https://doi.org/10.1007/s10804-018-9284-2

Prince, M., Knapp, M., Guerchet, M., McCrone, P., Prina, M., Comas-Herrera, A., &

    Salimkumar, D. (2014). *Dementia UK: Update.* Alzheimer's Society.

Ryff, C. D.(1989). Happiness is everything, or is it? Explorations on the meaning of

    psychological well-being. *Journal of Personality and Social Psychology, 57*,

    1069–1081. https://doi.org/10.1037/0022-3514.57.6.1069

Sanders, S. (2005). Is the glass half empty or half full? *Social Work in Health Care*, *40*,

    57–73. https://doi.org/10.1300/J010v40n03_04

Seligman, M. E. P. (1998). The president's address. *American Psychologist*, *54*, 559-

    562.

Seligman, M. E. (2002). *Authentic happiness.* Free Press.

Seligman, M. E., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology

    progress: Empirical validation of interventions. *American Psychologist*, *60*, 410–

    421. https://doi.org/10.1037/0003-066X.60.5.410

Sinclair, V. G., & Wallston, K. A. (2004). The development and psychometric evaluation of the brief resilient coping scale. *Assessment*, *11*, 94–101. https://doi.org/10.1177/1073191103258144

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: Assessing the ability to bounce back. *International Journal of Behavioral Medicine*, *15*, 194–200. https://doi.org/10.1080/10705500802222972

Smith, K. J., George, C., & Ferriera, N. (2018). Factors emerging from the "Zarit Burden Interview" and predictive variables in a UK sample of caregivers for people with dementia. *International Psychogeriatrics*, *30*, 1671–1678. https://doi.org/10.1017/S1041610218000315

Stansfeld, J., Stoner, C. R., Wenborn, J., Vernooij-Dassen, M., Moniz-Cook, E., & Orrell, M. (2017). Positive psychology outcome measures for family caregivers of people living with dementia: a systematic review. *International Psychogeriatrics*, *29*, 1281–1296. https://doi.org/10.1017/S1041610217000655

Steffen, A. M., McKibbin, C., Zeiss, A. M., Gallagher-Thompson, D., & Bandura, A. (2002). The revised scale for caregiving self-efficacy: Reliability and validity studies. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *57*, 74–86. https://doi.org/10.1093/geronb/57.1.P74

Steffen, A. M., & Gant, J. R. (2016). A telehealth behavioral coaching intervention for neurocognitive disorder family carers. International Journal of Geriatric Psychiatry, 31, 195-203. https://doi.org/10.1002/gps.4312

Stockwell-Smith, G., Moyle, W., & Kellett, U. (2018). The impact of early psychosocial intervention on self-efficacy of care recipient/carer dyads living with early-stage

dementia-A mixed-methods study. Journal of Advance Nursing, 74, 2167-2180.

https://doi.org/10.1111/jan.13710

Stoner, C. R., Orrell, M., & Spector, A. (2015). Review of positive psychology outcome

measures for chronic illness, traumatic brain injury and older adults:

Adaptability in dementia? *Dementia and Geriatric Cognitive Disorders, 40,*

340-357. https://doi.org/10.1159/000439044

Stoner, C. R., Stansfeld, J., Orrell, M., & Spector, A. (2017). The development of

positive psychology outcome measures and their uses in dementia research: A

systematic review. *Dementia, 0,* 1-22.

https://doi.org/10.1177/1471301217740288

Stoner, C. (2019). Positive psychiatry/ psychology for older adults: A new and

important movement but robust methodology is essential. *International*

Psychogeriatrics, 31,163-165. https://doi.org/10.1017/S1041610218002223

Tarlow, B. J., Wisniewski, S. R., Belle, S. H., Rubert, M., Ory, M. G., & Gallagher-

Thompson, S. R. (2004). Positive aspects of caregiving: Contributions of the

REACH project to the development of new measures for Alzheimer's

caregiving. *Research on Aging*, *26*, 429–453.

https://doi.org/10.1177/0164027504264493

Teahan, A., Lafferty, A., McAuliffe, E., Phelan, A., O'Sullivan, L., O'Shea, D.,

Nicholson, E., & Fealy, G. (2020). Psychosocial interventions for family carers

of people with dementia: A systematic review and meta-analysis. *Journal of*

Aging and Health. https://doi.org/10.1177/0898264319899793

Terwee, C. B., Bot, S. D., de Boer, M. R., van Der Windt, D. A., Knol, D. L., Dekker,

J., & de Vet, H. C. (2007). Quality criteria were proposed for measurement

properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*, 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

van Knippenberg, R. J. M., de Vugt, M. E., Ponds, R. W., Myin-Germeys, I., & Verhey, F. R. J. (2018). An experience sampling method intervention for dementia caregivers: Results of a randomized controlled trial. *American Journal of Geriatric Psychiatry, 26*, 1231-1243. https://doi.org/doi:10.1016/j.jagp.2018.06.004

Vernooij-Dassen, M. J. (1993). *Dementia and homecare: Determinants of the sense of competence of primary caregivers and the effect of professionally guided caregiver support*. Swets & Zeitlinger.

Vernooij-Dassen, M. J., Felling, A. J., Brummelkamp,E., Dauzenberg, M. G., Bos, G. A., & Grol, R. (1999). Assessment of caregiver's competence in dealing with the burden of caregiving for a dementia patient: a short sense of competence questionnaire (SSCQ) suitable for clinical practice. *Journal of the American Geriatrics Society*, *47*, 256–257.

Waelde, L. C., Thompson, L., & Gallagher-Thompson, D. (2004). A pilot study of a yoga and meditation intervention for dementia caregiver stress. *Journal of Clinical Psychology, 60,* 677-687. https://doi.org/10.1002/jclp.10259

Wagnild, G., & Young, H. (1993). Development and psychometric evaluation of the resilience scale. Journal of Nursing Measurement, 1, 165–178.

Wawrziczny, E., Larochette, C., Papo, D., Constant, E., Ducharme, F., Kergoat, M., & Pasquier, F., & Antoine, P. (2019). A customized intervention for dementia caregivers: A quasi-experimental design. *Journal of Aging and Health, 31*, 1172-1195. https://doi.org/10.1177/0898264318770056

Wilks, S. E. (2008). Psychometric evaluation of the shortened resilience scale among Alzheimer's caregivers. *American Journal of Alzheimer's Disease and Other Dementias*, *23*, 143–149. https://doi.org/10.1177/1533317507313012

Williams, V. P., Bishop-Fitzpatrick, L., Lane, J. D., Gywther, L. P., Ballard, E. L., Vendittelli, A. P., Hutchins, T. C., & Williams, R. B. (2010). Video-based coping skills to reduce health risk and improve psychological and physical well-being in Alzheimer's disease family caregivers. Psychosomatic Medicine, 72, 897-904. https://doi.org/10.1097/PSY.0b013e3181fc2d09

Wong, T.W.k., Ussher, J., & Perz, J. (2009). Strength through adversity: Bereaved cancer carers' accounts of rewards and personal growth from caring. *Palliative and Supportive Care*, *7*, 187-196. https://doi.org/10.1017/S1478951509000248

Yamamoto-Mitani, N., Sugishita, C., Ishigaki, K., Hasegawa, K., Maekawa, N., Kuniyoshi, M., & Hayashi, K. (2001). Development of instruments to measure appraisal of care among Japanese family caregivers of the elderly. *Scholarly Inquiry for Nursing Practice*, *15*, 137-141.

Yap, P., Luo, N., Ng, W. Y., Chionh, H. L., Lim, J., & Goh, J. (2010). Gain in Alzheimer care INstrument: A new scale to measure caregiving gains in dementia. *The American Journal of Geriatric Psychiatry*, *18*, 68–76. https://doi.org/10.1097/JGP.0b013e3181bd1dcd

Zeiss, A., Gallagher-Thompson, D., Lovett, S., Rose, J., & McKibbin, C. (1999). Self-efficacy as a mediator of caregiver coping: Development and testing of an assessment model. *Journal of Clinical Geropsychology*, *5*, 221–230. https://doi.org/10.1023/A:1022955817074

Zhang, S. Y., Edwards, H., Yates, P., Ruth, E., & Guo, P. (2012). Development of self-efficacy questionnaire for Chinese family caregivers. *International Journal of*

Mental Health Nursing, *21*, 358–365. https://doi.org/10.1111/j.1447-0349.2011.00791

Zhang, S. Y., Edwards, H., Yates, P., Ruth, E., & Guo, P. (2013). Preliminary reliability and validity testing of a self-efficacy questionnaire for Chinese family caregivers. *Aging & Mental Health*, *17*, 630–637. https://doi.org/10.1080/13607863.2013.771615

Part 2: Empirical Paper

The development of the PPOM-C for family carers of people with dementia:

Psychometric properties and factor structure

Abstract

Background: Positive psychology is an area of interest when considering the experiences of family carers of people living with dementia. The research is replete with studies measuring burden, stress or depression; with comparatively fewer quantifying positive emotions. Robust instruments that can measure positive states in family carers are important for reliability, validity and for comparisons amongst studies. **Methods**: This study utilised online methods to administer the Positive Psychology Outcome Measure (PPOM), a measure of hope and resilience (originally developed for people with dementia), to 267 family carers of people living with dementia. In addition, respondents completed the HADS-Depression subscale, the MSPSS and the SF-12 quality of life measure to establish the PPOM's convergent validity. **Results**: Two items were deleted on the PPOM to improve its internal consistency and factor structure. The newly formed PPOM-C is a 14-item measure validated for carers of people with dementia. It possesses excellent internal consistency, good convergent validity, adequate factor structure and excellent test-retest reliability. **Conclusions**: These robust psychometric properties suggest that it is ready for further use in research and practice. Future research could evaluate its responsiveness to change following intervention.

Introduction

Background

There are an estimated 850,000 people living with dementia in the UK (Prince et al., 2014). This figure is estimated to rise to over one million by 2025 and to two million by 2050. Many people with dementia live in community settings cared for by unpaid carers comprising of friends and family. Throughout this paper, these carers will be referred to as 'family carers'. There are approximately 700,000 such carers in the UK (Lewis et al., 2014).

Family carers have been estimated to save the UK economy 11 billion pounds per year (Prince et al., 2014). However, depression has been shown to affect one in three carers (Shoenmakers et al., 2010) and is often associated with a plethora of negative consequences. Given the exceptional contributions family carers offer, it follows that their wellbeing should be a 'national and international policy priority' (Farina et al., 2017, p.2).

Many interventions with family carers evaluate their efficacy based on the reduction of negative emotions or states as measured by outcome measures. Frequently, these include burden, depression and stress (Teahan et al., 2020). It is often implied that lessening the degree of these states contribute to increasing one's quality of life. Although quality of life instruments offer a valuable insight into the lives of carers, they do not tell the whole story.

What is missing from many interventions are robust instruments measuring positive psychological emotions or states. This is important as if intervention aims only to alleviate negative emotional states, an opportunity to design interventions intended to facilitate meaningful positive changes to people's lives is missed (Seligman, 1998).

Positive Psychology

An alternative method to designing and appraising interventions derives from the positive psychology approach (Seligman, 1998). Positive psychology researchers seek to determine which positive states or emotions are assisting the individual (Seligman et al., 2005). For example, a sense of fulfilment has been shown to lessen the sense of burden in family carers (Grant & Nolan, 1993).

A recent systematic review sought to identify and evaluate positive measures developed or validated for family carers (Stansfeld et al., 2017). The search criteria used positive psychology nomenclature including strengths, virtues and positive emotions (Seligman et al., 2005). Twelve instruments were found and represented positive constructs such as self-efficacy, spirituality, rewards, meaning and resilience.

A newly established positive measure in use for people living with dementia is the Positive Psychology Outcome Measure (PPOM; Stoner et al., 2017). It was developed following a systematic review of the positive psychology literature and interviews with people with dementia and their carers. The PPOM includes two subscales measuring hope and resilience.

Hope

Hope has been defined as 'a confident yet uncertain expectation of achieving future good, which to the hoping person is realistically possible and personally significant' (Dufault & Martocchio, 1985, p. 380). This seems relevant to the caring context, and indeed, is a key resource for family carers (Snyder et al., 2000). Qualitative interviews with carers suggest an important coping mechanism involves renewing their sense of hope on a daily basis. This process has been speculated to include acceptance, looking for the positives and seeing possibilities for the future (Duggleby et al., 2009). Despite

hope being viewed by family carers as valuable, it has been used infrequently in research with this population (Duggleby et al., 2013).

Resilience

Resilience is a multi-faceted construct with differing definitions (Windle, 2011). The Oxford Online Dictionary ("Resilience", 2020) describes resilience as the ability of a person to recover quickly after an unpleasant event. Other definitions embed resilience within social and cultural contexts in addition to resources and assets possessed by the individual (Windle & Bennett, 2011). Resilience is a relevant factor for family carers as the stressors for this population have been well documented (Pinquart & Sorenson, 2003). Despite resilience serving to mitigate stressors faced by carers, it is not routinely measured (Gaugler et al., 2007).

Importance of Psychometric Properties

There has been a lack of consensus in outcome measurement in dementia care, making comparisons amongst interventions problematic (Moniz-Cook et al., 2008). In addition, Moniz-Cook et al. noted a scarcity of robust positive measures available for family carers. This study aimed to fill this gap by administering the PPOM to a sample of family carers to conduct a psychometric analysis. It is necessary to analyse factors such as reliability, floor and ceiling effects, convergent validity and sensitivity to change (Rothrock et al., 2011). These data provide meaningful information to researchers about the validity of the measure and can be helpful when inferring change has occurred within interventions. In this context, validity refers to 'building a case that an instrument functions effectively in a particular population for a specific purpose' (Rothrock et al., 2011, p. 3).

There are a variety of different methods to determine the validity and reliability of health status measurement instruments. Many researchers have differing ideas on what constitutes acceptable psychometric properties for a measure (Stansfeld et al., 2017). This validation paper therefore made use of the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN; Mokkink et al., 2010) to guide which analyses would be conducted. The COSMIN study consulted with 57 experts to determine which analyses would be the most appropriate for each psychometric property.

The Positive Psychology Outcome Measure

The validity and reliability of the PPOM was analysed with 225 people with dementia (Stoner et al., 2018). An exploratory factor analysis and subsequent confirmatory factor analysis (CFA) yielded a two-factor structure with acceptable fit indices. Each factor contained eight items and were labelled hope and resilience. The PPOM achieved good temporal stability in a test-retest sample of forty-eight people and excellent internal consistency ($\alpha = 0.94$). Construct validity was achieved through statistically significant correlations with the Control, Autonomy, Self-realisation and Pleasure scale (Hyde et al., 2003), Geriatric Depression Scale short form (Yesavage et al., 1983) and the Quality of Life in Alzheimer's Disease Scale (QoL-AD; Logsden et al., 1999).

Aims

The purpose of this study was to complete a series of psychometric analyses assessing the validity and reliability of the PPOM in family carers. This analysis included a CFA to determine if the sixteen-item, two-factor model previously found (Stoner et al., 2018) remained consistent in a sample of family carers. It was

hypothesised that as the PPOM has previously demonstrated good factor structure, excellent internal consistency and 'good' test-retest reliability (Stoner et al., 2018), that these results would be replicated here.

Construct validity was determined by convergent validity testing with theoretically linked constructs, a binary logistic regression, and testing whether the means of subgroups on the PPOM significantly differed according to depression severity. It was hypothesised the PPOM and its subscales would possess large significant negative correlations with the HADS-D given previous findings (Stoner et al., 2018). In addition, it was hypothesised that the PPOM and its subscales would obtain large positive significant correlations with both subscales of the SF-12 based on previous research (Stoner et al., 2018). Finally, it was predicted that the PPOM would obtain significant small to medium positive correlations with the MSPSS based on research measuring the same constructs (Hatami et al., 2019). It was hypothesised that the MSPSS, SF-12 and HADS-D categories would combine to accurately predict whether a person would obtain a high or low score on the PPOM. It was also predicted that those who obtained 'normal' HADS-D scores would possess significantly higher means on the PPOM than the 'borderline' or 'clinical caseness' HADS-D groups.

No hypotheses were made with respect to the direction of any relationships between scores on the PPOM with sex, age and whether or not the carer lived with the person they provided care for.

Method

Design

A cross-sectional online study was conducted with family carers. Ethical approval was obtained from the University College London Research Ethics Committee (Ref:

15139/001; Appendix A). This study consisted of a battery of measures for respondents to complete at baseline and a retest for a subsample four weeks later. All measures were self-completed online by participants using Qualtrics. All questions were deemed required (i.e. non skippable) by Qualtrics to prevent incomplete data.

This was a joint project with another University College London trainee (Cartwright, 2020). See Appendix B for an explanation of the contributions by each trainee to this project.

Participants

Participants were recruited from the Join Dementia Research (JDR) register (www.joindementiaresearch.nihr.ac.uk) and two UK charities (Age UK and the Alzheimer's Society). The JDR is a joint venture between the National Institute for Health Research and several Alzheimer's research organisations. It promotes dementia research with affected persons and their carers. To participate, respondents must have met the criteria below.

Inclusion Criteria

- UK residents

- Self- identified as an unpaid family carer for a person living with dementia

- Over eighteen years old

- Able to communicate in English

Exclusion Criteria

- Paid carers (e.g. in care homes)

Procedure

The inclusion criteria were applied to the JDR database and a list of eligible volunteers was obtained. Volunteers were contacted via email inviting them to participate. Within

the advert was a link directing them to the battery of measures hosted on Qualtrics. This advert was also circulated to other organisations including Age UK and the Alzheimer's Society. These organisations independently disseminated the link to its members.

After respondents answered questions relating to the inclusion criteria, they were presented with information regarding the General Data Protection Regulation (GDPR, 2018) and the Data Protection Act (2018). All participants had summary sheets (Appendix C) relating to these policies appear on screen and multiple opportunities to download them. Participants were also given the contact details of the research team and encouraged to ask questions if needed prior to giving consent. Following this, participants were asked to provide their consent (Appendix D) to participate. The battery of outcome measures were then presented in a counterbalanced fashion to control for order effects. Each question was mandatory with the participant automatically alerted to complete any incomplete question. Following completion of the battery, participants were thanked for their time and were shown a selection of online resources suitable for family carers (Appendix E).

Outcome Measures

Respondent Demographics

Data were collected with respect to the respondent's age, gender, ethnicity, marital status, employment, education, kinship, and the length of time they had supported the person with dementia. Information about the person with dementia was also collected. Respondents stated what dementia diagnosis they had in addition to how long they had lived with dementia for. Respondents were asked whether they lived with the person they provided care for and if they identified as the primary carer. The primary carer was defined as the main person who provided care to the person with dementia.

The Positive Psychology Outcome Measure (PPOM)

The PPOM (Appendix F) is a measure of hope and resilience and was developed for people with dementia. It was established using qualitative and quantitative methods involving people with dementia, their carers, healthcare professionals and experts in the field (Stoner et al., 2017). It has excellent internal consistency α = 0.94 (Stoner et al., 2018). It demonstrated significant positive correlations with the Quality of Life Alzheimer's Disease scale (QoL-AD; Logsdon et al., 1999; $r = 0.627$) and significant negative correlations with the Geriatric Depression Scale (GDS; Yesavage et al., 1983; $r = -0.699$). Test-retest reliability showed a 'good' Intraclass Correlation Coefficient (ICC) for the PPOM (ICC = 0.88, 95%CI = 0.79, 0.93), and the hope subscale (ICC = 0.78, 95%CI = 0.63, 0.88). The resilience subscale displayed excellent test-retest reliability (ICC = 0.91, 95%CI = 0.83, 0.95). Items are scored on a five-point Likert scale (0-4) and are based on the last month. A sample item from the hope subscale is 'I have a positive outlook on life'. A sample item from the resilience subscale is 'I can bounce back'. A score of '0' indicates this is 'not true at all' whilst a score of 4 implies this is 'true nearly all of the time'.

Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983)

The HADS is a 14-item instrument measuring anxiety and depression. In a recent psychometric analysis article, the anxiety subscale was found to not be appropriate for family carers (Stott et al., 2017). Thus, only the depression subscale was used in this study. The HADS depression (HADS-D) subscale is a self-report measure consisting of seven items. It has good internal consistency (α = 0.85; Stott et al., 2017). Each item is scored from 0-3 with several questions negatively scored. Scores between 0 and 7 are

labelled 'normal' whilst scores between 8 and 10 are considered 'borderline'. A score of 11 or over indicates 'clinical caseness' for depression.

The Short Form Health Survey (SF-12; Ware et al., 1996)

Given the lack of a gold standard quality of life measure in use with family carers, the SF-12 was selected as it has been used extensively with this population (Farina et al., 2017). The SF-12 is the short form questionnaire of the SF-36 (Ware & Sherbourne, 1992). The SF-12 is a health-related quality of life scale measuring eight concepts in health questionnaires; physical functioning, role functioning, physical pain, general health, vitality, social functioning, emotional and mental health. It consists of two subscales; the physical component score and the mental component score. The physical (ICC = 0.86) and mental component scores (ICC = 0.77) have displayed good test-retest reliability (Ware et al., 1996).

The Multidimensional Scale of Perceived Social Support (MSPSS; Zimet et al., 1988)

The MSPSS is designed to measure perceived social support. There are 12 items which are scored on a 7-point Likert scale ranging from 1 (very strongly disagree) to 7 (very strongly agree). These scores are summed with higher scores indicating greater perceived levels of social support. The MSPSS was developed using a sample of undergraduates (Zimet et al., 1988) but has been used with family carers (Charlesworth et al., 2008). It possesses good internal consistency ($\alpha = 0.88$), and good test-retest reliability ($r = 0.85$; Zimet et al., 1988).

Statistical Power

There are no recommendations with respect to the number of respondents required to validate an outcome measure. Power calculations can give researchers an estimate of how many participants are needed to find statistically significant results for a given effect size. As this study made use of Pearson's r correlations to ascertain convergent validity, G*Power 3.1 (Faul et al., 2007) was used to calculate that a minimum 193 respondents would be required to detect a small effect size ($r = 0.2$). This sample size provides 80% power with a corresponding alpha level of 0.05. When conducting factor analyses, Terwee et al. (2007) recommend multiplying the number of items in the measure by seven (n = 112) to obtain the necessary sample size required. These sample sizes were therefore considered the minimum number of participants required for this study.

Test-retest Procedure

Emails were sent to participants who had completed the initial battery to fill out a subset of the measures four weeks later. Invitations were sent until 50 responses had been collected. A four-week timeframe was selected to reduce the chance of practice effects and new life stressors interfering with scores at the second timepoint.

Statistical Analyses

All analyses were computed using SPSS version 26. If respondents terminated their participation before completing the battery, this was assumed to be a withdrawal of consent and all data for the participant was deleted. The rationale for selected analyses was influenced by the COSMIN criteria (Mokkink et al., 2010) and a quality appraisal tool for evaluating patient reported outcome measures (Terwee et al., 2007).

The COSMIN checklist includes ten criteria to be satisfied when evaluating measurements of health status. They include sufficient measurements of internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, responsiveness and interpretability. The planned analyses met all of these criteria excluding content validity, cross-cultural validity, criterion validity and responsiveness. Content validity had been established with a related population (Stoner et al., 2018). As this measure was not being translated, cross-cultural validity was unnecessary. Criterion validity could not be evaluated due to the lack of a gold standard instrument in this area. Finally, due to the study not including an intervention component, responsiveness could not be analysed.

Descriptive Statistics

Descriptive and frequency statistics including means and standard deviations were obtained for all demographic information and scores on the outcome measures. Scores on the PPOM were analysed for any possible skew or kurtosis within the distribution. Floor and ceiling effects were evaluated by analysing the percentages of the lowest and highest scores on the PPOM and its subscales. Floor and ceiling effects were considered absent if less than 15% of participants scored the minimum or maximum score possible (Terwee et al., 2007).

Means and standard deviations for a selection of subgroups within the sample were also calculated in an effort to establish the PPOM's interpretability. Terwee et al. (2007) recommend a minimum of four subgroups to satisfy this criterion. The means and standard deviations for men and women were presented. The means and standard deviations of the PPOM were also calculated for the three different categories of the

HADS-D ('normal', 'borderline' and 'clinical caseness' for depression). Finally, the means and standard deviations for carers who both lived and did not live with the person they provided care for were calculated.

Reliability

Internal Consistency. The PPOM and its subscales were assessed for their internal consistency using Cronbach's alpha. Internal consistency would be deemed excellent ($\alpha = 0.90 < 0.95$), good ($\alpha = 0.80 - 0.89$), acceptable ($\alpha = 0.70 - 0.79$) or poor ($\alpha < 0.70$) (George & Mallery, 2003).

Temporal Stability. To assess temporal stability, an Intraclass Correlation Coefficient (ICC) was used to assess test-retest reliability over a four-week period. ICC scores would be deemed excellent (> 0.90), good ($0.75 - 0.90$), moderate ($0.50 - 0.75$) or poor (< 0.50) (Koo & Li, 2016).

Measurement Error

The standard error of measurement (SEM) was used to evaluate the variability of error within the PPOM. A person's observed score on any measure equals their true score in addition to the SEM (Leong & Huang, 2010). The SEM was calculated using the following equation: $\text{SEM} = \sigma (\sqrt{1} - \alpha)$. Where σ is the standard deviation of the PPOM and α is its Cronbach's alpha reliability statistic. 95% confidence intervals were then calculated using, $95\%\text{CI} = \text{Score} \pm (1.96 * \text{SEM})$.

Smallest Detectable Change (SDC)

The SEM can be used to determine the SDC$_{ind}$ and SDC$_{group}$. The SDC$_{ind}$ refers to the smallest within-person change occurring that can be attributed to a treatment or intervention with the *p* value set at .05 (Beckerman et al., 2001). The SDC$_{ind}$ is calculated using the following equation: $SDC = 1.96 * \sqrt{2} * \text{SEM})$. The SDC$_{group}$ refers to the smallest within-group change that could be attributed to a treatment or intervention with the *p* value set at .05. The SDC$_{group}$ is calculated through, SDC$_{ind}$ / $\sqrt{n}$.

Validity

Structural Validity. Structural validity was established using a CFA. The PPOM has previously undergone exploratory factor analysis and CFA in its development (Stoner et al., 2018). Therefore, another CFA was deemed appropriate for this dataset. CFA analysis used SPSS Amos version 25. Model of fit indices attained included the (i) chi-square and *p* value (low chi-square and $p > .05$ indicate good fit), (ii) standardised root mean square error of approximation (RMSEA; RMSEA < 0.08 indicates good fit), (iii) comparative fit index (CFI; > 0.90 indicates good fit), (iv) standardised root mean square residual (SRMR; SRMR > 0.08 indicate good fit) and (v) Average Value Explained (AVE; > 0.50 indicates acceptable amount of variance explained by factor/s) (Hooper et al., 2008; O'Rourke et al., 2013).

Convergent Validity. Pearson's *r* correlations were used to assess convergent validity of the PPOM. Convergent validity refers to the magnitude in which theoretically linked constructs are observed to be correlated through instruments of these constructs. To control for type 1 error, Bonferonni corrections were used throughout all convergent validity testing.

A binary logistic regression was conducted to augment the PPOM's convergent validity. Binary logistic regressions transform continuous variables (such as scores on the PPOM, HADS-D, MSPSS and SF-12) into categorical 'high' and 'low' scoring groups using the median of each variable. It then uses binary logistic regression to predict whether scoring low or high in one would group would successfully predict whether a person would score low or high on another.

Finally, a one-way analysis of variance (ANOVA) was conducted to determine if there were any significant differences between the means of the PPOM grouped by the HADS-D categories ('normal', 'borderline' and 'clinical caseness' for depression). Post-hoc testing using a Bonferroni correction was used to control for type 1 error.

The Relationship Between the PPOM and Respondent Characteristics

Pearson's *r* correlations and independent samples t-tests were used to explore the relationships between the PPOM and its subscales alongside age, sex and whether or not the person lived with the person they provided care for. Bonferroni corrections were utilised to control for type 1 error.

Results

Results were assessed against the COSMIN checklist (Mokkink et al., 2010) designed to evaluate the methodological quality of studies on measurement properties. See Appendix G for the completed version of this checklist.

Participants

Two hundred and sixty-seven carers (174 female, 93 male; Table 1) agreed to participate in the study. The average age of participants was 60.51 years of age (SD =

14.37). The participants were mostly white British (92.1%) and married (69.7%). Daughters of those with dementia were the most frequent responders (40.4%; Table 2). The most frequently reported length of time being a carer was three to four years (31.5%). Approximately half the carers lived with the person they cared for. The respondents were also highly educated with 57.7% having completed an undergraduate degree or postgraduate study. Four weeks after the initial battery was completed, 52 carers completed the retest. There were no significant differences between the subsample who completed time two and those who only completed at time one (Tables 1-2).

Initial Internal Consistency of the PPOM

The internal consistency of the PPOM indicated multicollinearity with $\alpha = 0.951$. The hope and resilience subscales demonstrated excellent internal consistency however, with scores of $\alpha = 0.915$ and $\alpha = 0.927$ respectively. The analysis demonstrated that deleting any item would have brought its internal consistency to an acceptable score of less than $\alpha = 0.95$ (Terwee et al., 2007). As such, further analysis in the form of a CFA was needed to identify the most suitable items to be deleted.

Table 1
Carer Demographics

	Total sample (*n* = 267)	Retest sample (*n* = 52)
Gender *n* (%)		
Female	174 (65.2%)	67.3 (67.3%)
Male	93 (34.8%)	17 (32.7%)
Age *M* (*SD*) Range	60.51 (14.37) 20 - 92	62.63 (11.08) 41 - 91
Marital status *n* (%)		
Married	186 (69.7%)	40 (76.9%)
Single	33 (12.4%)	4 (7.7%)
In a relationship	30 (11.2%)	2 (3.8%)
Divorced	9 (3.4%)	3 (5.8%)
Widowed	3 (1.1%)	1 (1.9%)
Separated	3 (1.1%)	-
Other	3 (1.1%)	2 (3.8%)
Ethnicity *n* (%)		
White (British)	246 (92.1%)	48 (92.3%)
White (other)	8 (3%)	2 (3.8%)
White (Irish)	4 (1.5%)	-
Mixed ethnic	4 (1.5%)	-
Black	2 (0.7%)	1 (1.9%)
Asian	2 (0.7%)	-
Other ethnic group	1 (0.4%)	1 (1.9%)
Highest Level of Education		
Undergraduate degree	83 (31.1%)	16 (30.8%)
Postgraduate degree	71 (26.6%)	12 (23.1%)
Higher education	42 (15.7%)	9 (17.3%)
A-Level (or equivalent)	28 (10.5%)	4 (7.7%)
GCSE's (or equivalent)	23 (8.6%)	6 (11.5%)
Other qualifications	9 (3.4%)	3 (5.8%)

Table 2
Carer Characteristics

	Total sample (*n* = 267)	Retest sample (*n* = 52)
Relation to Person with Dementia *n* (%)		
Daughter	108 (40.4%)	21 (40.4%)

Husband	58 (21.7%)	10 (19.2%)
Wife	43 (16.1%)	9 (17.3%)
Son	26 (9.7%)	6 (11.5%)
Granddaughter	10 (3.7%)	-
Other	11 (4.1%)	4 (7.7%)
Daughter-in-law	5 (1.9%)	2 (3.8%)
Son-in-law	5 (1.9%)	-
Grandson	1 (0.4%)	-
Dementia Diagnosis		
Alzheimer's disease	127 (47.6%)	28 (53.8%)
Dementia of mixed aetiology	60 (22.5%)	10 (19.2%)
Vascular dementia	31 (11.6%)	3 (5.8%)
Frontotemporal dementia	15 (5.6%)	3 (5.8%)
Other	10 (3.7%)	2 (3.8%)
Dementia with Lewy bodies	6 (2.2%)	1 (1.9%)
Posterior Cortical Atrophy	5 (1.9%)	-
Parkinson's dementia	4 (1.5%)	2 (3.8%)
Length of time as a carer n (%)		
0-12 months	15 (5.6%)	6 (11.6%)
1-2 years	40 (15%)	4 (7.7%)
3-4 years	84 (31.5%)	19 (36.5%)
5-6 years	53 (19.9%)	9 (17.3%)
7-8 years	23 (8.6%)	1 (1.9%)
9-10 years	19 (7.1%)	4 (7.7%)
10 years +	33 (12.4%)	9 (17.3%)
Does the carer live with the person they provide care for? n (%)		
Yes	133 (49.8%)	24 (46.2%)
No	132 (49.4%)	28 (53.8%)
Other	2 (0.7%)	

Confirmatory Factor Analysis (CFA)

Initial goodness of fit indices on the PPOM had room for improvement (Table 3). To improve the fit and the PPOM's Cronbach's alpha statistic, an item from each subscale was omitted to ensure they were evenly weighted. A variety of different strategies were employed to identify which deleted items benefited the fit indices the most. All of these different permutations with their associated fit indices can be found in Appendix H. There is justification for omitting items when they increase the parsimony and fit indices of the model (Frohlich, 2002; Voss et al., 2003).

The most effective method found to improve the model fit was to delete the items possessing the lowest regression weight onto their respective factor (Hope Q5 and Resilience Q2). After deleting the items, the CFA was rerun (Figure 1). Hope question five was 'I can give and receive care/ love' and resilience question two was 'I am able to deal with whatever happens'.

This resulted in a 14-item PPOM designated the PPOM-C (The Positive Psychology Outcome Measure for Carers of People with Dementia; Table 3). Factor loadings indicated all items loaded onto their respective factors successfully (Figure 1). The CFA also demonstrated that hope and resilience shared some variance. The CFI, SRMR and AVE statistics all fell within acceptable ranges (Table 3). The RMSEA and chi-square statistics however indicated poor model fit. Factor loadings varied from 0.755 to 1.408. The average value explained by each factor was $R^2 = 0.61$ for hope and $R^2 = 0.62$ for resilience. Model fit for the PPOM-C was considered adequate and all subsequent analyses were conducted on this 14-item version.

Internal Consistency of the PPOM-C

The internal consistency of the PPOM-C was excellent ($\alpha = 0.948$). The hope and resilience subscales also had excellent internal consistency with scores of $\alpha = 0.912$ and $\alpha = 0.918$ respectively (Table 3). No deleted items would have improved the internal consistency score.
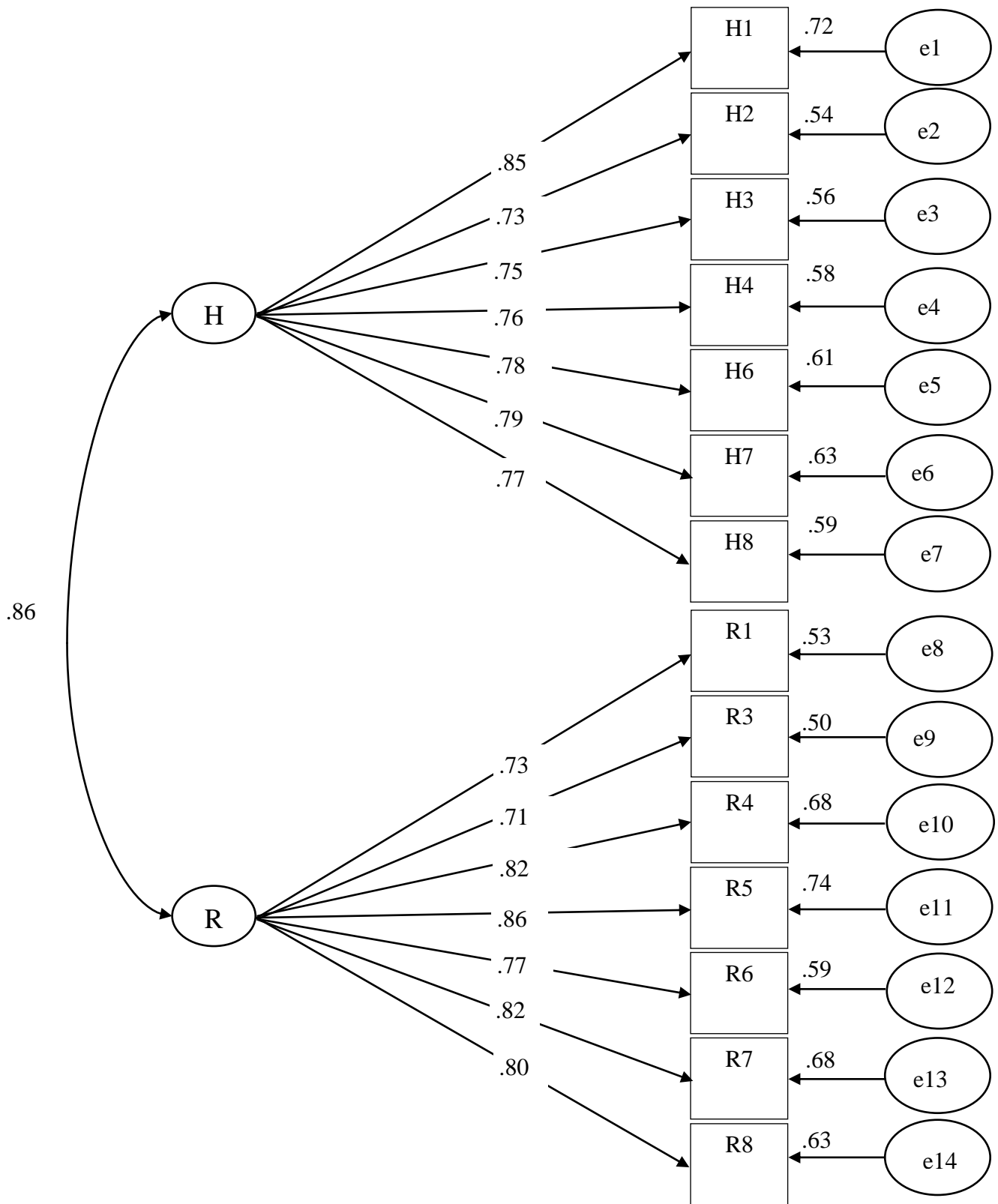
Figure 1: CFA Factor Loadings

Table 3
Variations in fit indices on two forms of the PPOM

		PPOM (16-items)	PPOM-C (14-items)
Internal Consistency (α)	Full measure	0.951	.948
	Hope subscale	.915	.912
	Resilience Subscale	.927	.918
Chi square and significance		$430.55, p < .001$	$340.95, p < .001$
CFI		.896	.904
SRMR		.06	.057
RMSEA		.109	.114
AVE	Hope subscale (R^2)	.58	.61
	Resilience subscale (R^2)	.62	.62

Descriptive Statistics

The possible score of the PPOM-C ranged from 0 to 56. The observed range varied from 2-56, with a mean of 38.65 and standard deviation of 10.98. The distribution of the PPOM (Figure 2), showed the skewness of the data to be approximately symmetric (-.455), with the kurtosis value being -0.014. No respondents (0%) scored the lowest possible score, and 13 participants (4.9%) attained the highest score. Floor and ceiling effects were therefore absent in the PPOM-C.

The possible scores of each subscale ranged from 0 to 28. The hope subscale had one respondent (0.4%) score the minimum score and twenty-six (9.7%) obtain the maximum. The resilience subscale had one respondent (0.4%) score the minimum score and sixteen (6%) obtain the maximum. Thus, floor and ceiling effects were not present within the subscales. The mean score of the hope subscale was 19.41 with a standard deviation of 5.90. The resilience subscale had a mean score of 19.24 and a standard deviation of 5.68.

Means and standard deviations of the PPOM-C were provided for a range of subgroups found within the sample (Table 4).

PPOM-C Total Score
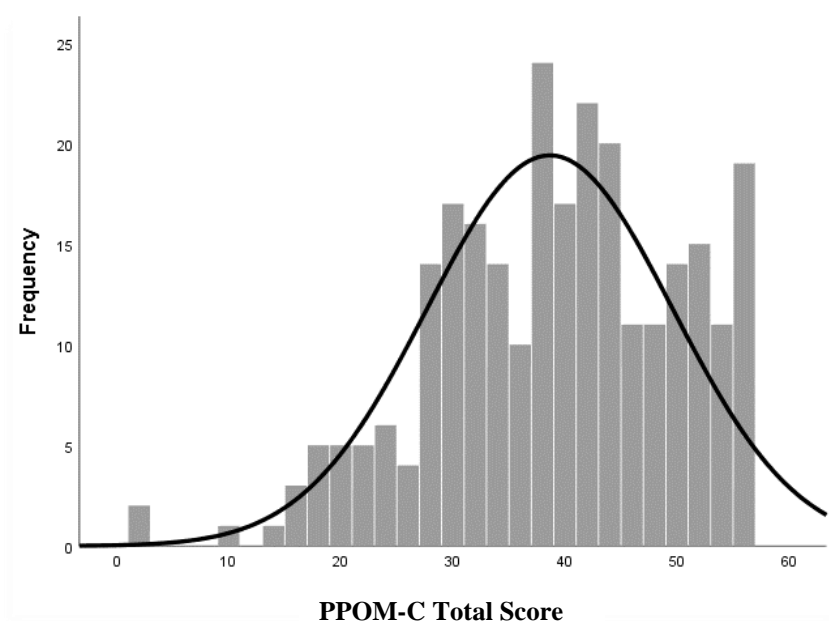
Figure 2: Distribution of the PPOM-C

Table 4

Means and standard deviations of subgroups on the PPOM-C

Subgroup	Mean (SD) of PPOM-C
Full sample (n = 267)	38.65 (10.98)
Men (n = 93)	40.57 (11.19)
Women (n = 174)	37.62 (10.75)
'Normal' scorers on HADS=D (n = 162)	43.46 (8.29)
'Borderline' scorers on HADS-D (n = 55)	33.71 (9.38)
'Clinical caseness for depression' scorers on HADS-D (n = 50)	28.48 (11.04)
Carers who lived with the person they cared for (n =_133)	39.13 (10.81)
Carers who did not live with the person they cared for (n = 131)	38.05 (11.26)

Test-retest Reliability

There were 52 respondents who completed the test-retest. The total amount of days taken to complete the retest ranged from 28-43. The majority of the respondents (75%) had completed the retest 29 days after time 1. Reliability for the PPOM-C across the four-week period was found to be 'excellent' (ICC = 0.91, 95%CI = 0.85, 0.95). Reliability was 'good' for both the hope subscale (ICC = 0.89, 95%CI = 0.82, 0.94) and the resilience subscale (ICC = 0.87, 95%CI = 0.79, 0.93).

Convergent Validity

Pearson's r correlations were employed with an adjusted p-value of .004 (.05/12), calculated using the Bonferroni correction. The PPOM-C was negatively correlated with HADS-D (r = -.66, $p < .001$), in addition to the hope and resilience subscales (r = -.67, $p < .001$; r = -.58, $p < .001$). The hope and resilience subscales were positively correlated with the SF-12 mental component score (r = .62, $p < .001$; r = .57, $p < .001$) in addition to the PPOM-C (r = .63, $p < .001$).The PPOM-C, and its hope and resilience subscales were positively correlated with the SF-12 physical component score (r = .19, p = .002; r = .17, p = .004; r = .19, p = .002). Lastly, the MSPSS was significantly correlated with the PPOM-C (r = .39, $p < .001$), the hope (r = 0.45, $p < .001$) and resilience (r = 0.29, $p < .001$) subscales.

To conduct the binary logistic regressions, the PPOM-C, HADS-D, MSPSS and the physical and mental health component scores (SF-12) were split into low and high scoring categories using their respective medians. The PPOM-C (Median = 39) had 135 'low' and 132 'high' scorers. The HADS-D (Median = 6) had 123 'low' and 144 'high' scorers. The MSPSS (Median = 58) had 137 'low' and 130 'high' scorers. The physical

component score (Median = 50.62) had 133 'low' and 134 'high' scorers. Finally, the mental component score (Median = 38.98) had 133 'low' scorers and 134 'high' scorers. The HADS-D, MSPSS, and the mental and physical component scores (SF-12) were then used as variables in a binary logistic regression model to predict the PPOM-C category. The model was significant with $\chi2(4) = 98.39$, $p < .001$. The model was correct in predicting 76.4% of cases on the PPOM-C and accounted for 41.1% of the variance. The HADS-D (Wald = 12.62, $p < .001$), MSPSS (Wald = 4.72, $p = .03$) and the mental component score (Wald = 30.44, $p < .001$) were significant variables within the model. The physical component score (SF-12) was not a significant variable (Wald = 3.62, $p = .057$).

A one-way ANOVA was calculated to determine whether there were significant differences in the PPOM-C grouped HADS-D categories. This analysis was significant, $F (2, 264) = 62.21$, $p < .001$. Post-hoc Bonferroni comparisons were carried out. The mean score for 'normal' scorers on the HADS-D (M = 43.46, SD = 8.29) was significantly higher than the 'borderline' (M = 33.71, SD = 9.38) and 'clinically depressed' (M = 28.48, SD = 11.04) groups ($p < .001$). 'Borderline' scorers were also found to be significantly higher than 'clinically depressed' cases ($p = .011$).

Measurement Error

The standard error of measurement (SEM) was calculated to be 2.53 ($SEM =$ $10.98 (\sqrt{1 - 0.948})$). The SEM can determine confidence levels of scores obtained using the PPOM-C. The 95% confidence interval for the PPOM-C is thus, 95%CI = [PPOM-C Score] ± 4.96. Applying this to the PPOM-C's mean score (38.65) yields the following 95% confidence interval, $38.65_{95\%CI} = [33.75, 43.55]$.

Smallest Detectable Change (SDC)

The $SDC_{ind} = 1.96 * \sqrt{2} * 2.53 = 6.99$. This indicates there would need to be a minimum difference of 7 points on an individual level for the change to have been considered a result of a treatment or intervention. The equation to determine the average level of change within a group would be, $SDC_{group} = 6.99 / \sqrt{n}$.

The Relationship Between the PPOM-C and Respondent Characteristics

Age and the PPOM-C

A scatter plot (Figure 3) between the PPOM-C and age indicated there may be a small relationship between these variables. A Pearson's *r* correlation coefficient confirmed this small, yet significant relationship ($r = 0.23$, $p < .001$). This suggests older carers are associated with slightly higher levels of hope and resilience.
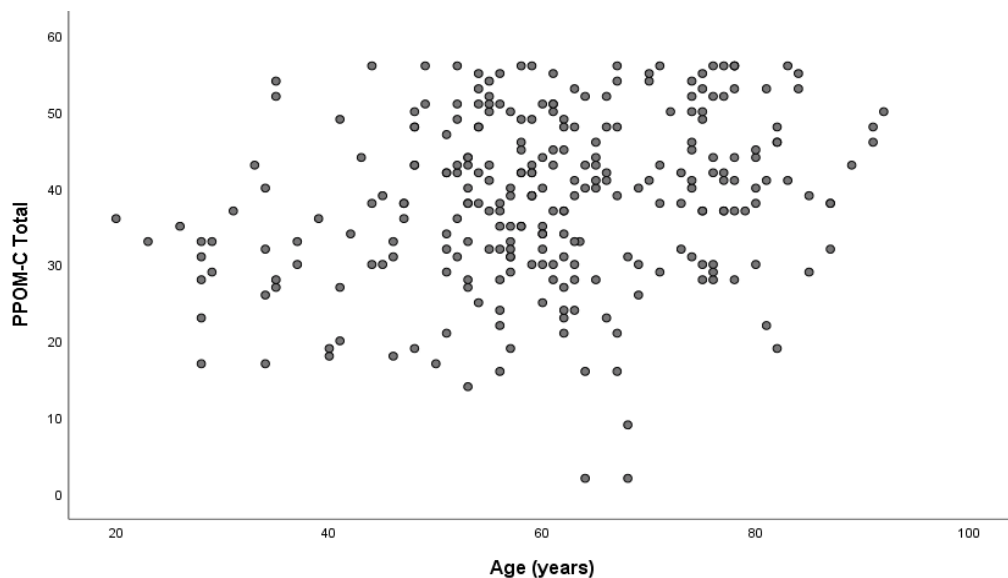

Figure 3: A scatter plot showing the relationship between age and the PPOM-C

Gender and the PPOM-C

A series of independent samples t-tests were carried out between men and women's scores on the PPOM-C and its subscales. The level of alpha was adjusted to .017 (.05/3) using a Bonferroni correction. There were no significant differences found on the PPOM-C (p = .036) or on the hope subscale (p = .127). There was however, a significant difference found in the resilience between men (M = 20.41, SD = 5.70) and women (M = 18.61, SD = 5.59), $t(265)$ = 2.48, p = .014.

The PPOM-C and the carer's living arrangements

An independent samples t-test was carried out to determine if scores on the PPOM-C were significantly different between carers who lived with the person they cared for (M = 39.13) versus those who did not (M = 38.05). There were no significant differences found (p = .43).

Discussion

The newly developed 14-item PPOM-C possesses desired psychometric properties and is ready for use in further research with family carers. Two items from the PPOM were dropped to improve both the internal consistency and the goodness of fit indices. Internal consistency of the PPOM-C and its subscales were excellent. It also possessed excellent test-retest reliability over a four-week period. This suggests that over a period of four weeks, family carers' levels of hope and resilience remained stable.

All correlation coefficients were statistically significant and in the expected direction whilst using a Bonferroni correction to control for type 1 error. The correlation

coefficient between the PPOM-C and the depression construct was similar to previous findings (Stoner et al., 2018). It was noted the PPOM-C held larger correlations with the mental component score of the SF-12 than the physical component score. This suggests hope and resilience have more in common with a person's mental health quality of life than physical health quality of life. Correlations between the PPOM-C, the HADS-D and the mental component score were all in excess of $r = 0.60$ and therefore, considered strong (Evans, 1996). The PPOM-C and MSPSS also obtained a moderate strength correlation which fit with the stated hypothesis.

The binary logistic regression also contributed to the convergent validity of the PPOM-C. The overall model containing the HADS-D, MSPSS, physical component score and mental component score categories as variables was significant in predicting the PPOM-C category a carer obtained. Whilst the HADS-D, MSPSS and mental component score were significant variables within this model, the physical component score was not. It follows that hope and resilience may be more accurately predicted by instruments designed to measure mental health constructs and social support as opposed to a physical health measure. As the use of positive psychology measures is sparse in the literature, it is not clear whether this is generally true in family carers of people with dementia or not.

The one-way ANOVA demonstrated significant differences between the mean scores on the PPOM-C based on the three categories of the HADS-D. This was an expected finding given that depression includes feelings of hopelessness (Kroenke et al., 2001) and the established negative relationship between depression and resilience (Komiti et al., 2003).

The PPOM-C demonstrated acceptable structural validity. The comparative fit index (CFI), standardised root mean square residual (SRMR) and the average value

explained (AVE) by each factor each indicated good model fit. In addition, all items loaded significantly onto their respective factors. Only the chi-square and standardised root mean square error of approximation (RMSEA) statistics suggested poor model fit. The significant chi-square test (indicating poor model fit) is a common phenomenon within CFA as it is influenced by sample size (Kline, 2011). Some structural equation theorists have proposed models should be evaluated according to the theory underlying the measures in addition to using fit indices as a guide (Hooper et al., 2008). The PPOM has previously demonstrated satisfactory evidence of its items being developed by theories underpinning hope and resilience (Stoner et al., 2017). Thus, on the balance of the evidence collated in this study, the two-factor model proposed for the PPOM-C was concluded to be adequate.

Two items were deleted from the PPOM to create the PPOM-C to improve the internal consistency and goodness of fit indices. These items were 'I can give and receive care/ love' (item from the hope subscale) and 'I am able to deal with whatever happens' (item from the resilience subscale). These items clearly possess face validity with respect to a carer's responsibilities. There are several explanations as to why these items were identified by the CFA as items to be deleted. Firstly, the PPOM was a measure originally developed for people living with dementia. It is feasible that giving this measure to a different population led the items to be interpreted differently. For instance, confidence in being 'able to deal with whatever happens' may elicit different realities of difficulty for a person with dementia compared to their carer. This in turn, may influence how these questions are answered differently by different populations. Secondly, it is possible that these two items were tapping into an undiscovered third factor that may be specifically relevant for caregivers as opposed to people living with dementia. Future researchers could explore this through administering the PPOM to

carers and conducting an exploratory factor analysis to determine how many factors fit the data.

Implications for Practice

This paper is to the author's knowledge the first time an instrument measuring two positive psychology constructs has been validated in family carers. Given that few positive measures are available for family carers (Stansfeld et al., 2017), the arrival of an instrument capable of measuring two constructs is timely. Due to a lack of robust alternate measures, adopting the PPOM-C in practice could aid in comparing outcomes across interventions (Moniz-Cook et al., 2008).

The PPOM-C represents a prompt method of assessing family carers who may be low in hope or resilience. This information could be vital for auxiliary services to mobilise around such carers and offer intervention or support. Conversely, it offers an opportunity to learn more from carers with high levels of hope or resilience. Qualitative approaches could shed further light on which psychosocial resources and assets lead to differing levels of hope and resilience. On some occasions, the PPOM-C could measure high levels of hope or resilience that may not reflect a carer's reality. For example, this could look like a carer who arguably has an unrealistic view of an optimistic future given the circumstances that they, and the person they provide care for are in. This could subsequently lead to this carer potentially feeling overwhelmed if this hopeful future were to not materialise. This highlights the importance of viewing the PPOM-C as one piece of a jigsaw in an assessment with a carer. This could fit alongside other jigsaw pieces such as a clinician's clinical interview and other assessment measures.

There were significant correlations between the HADS-D, MSPSS, the mental component score of the SF-12 and the PPOM-C subscales indicating good convergent

validity. The binary logistic regression model and one-way ANOVA added further evidence to this claim. This indicates that hope and resilience are linked to the wellbeing of family carers. This suggests interventions targeting such positive constructs may also bring about positive changes in quality of life, perceived social support and depression. The literature comprehensively supports this given that hope is a key psychosocial factor in being a family carer (Snyder et al., 2000), managing dementia (Moniz-Cook et al., 2009), acquired immune deficiency syndrome (Moon & Snyder, 2000) and cancer (Magaletta & Oliver, 1999. Resilience has also been shown to be useful for family carers (Gaugler et al., 2007), safeguarding against post-traumatic stress disorder (Waysman et al., 2001) and preventing developmental psychopathology (Masten et al., 1990).

A series of independent samples t-tests exhibited no significant differences between the sexes on the PPOM-C or hope. A small yet significant difference suggested that men possessed higher levels of resilience than women, however. Very little research has been conducted in measuring resilience between the genders. One study conducted in Sweden and Thailand with older adults found no such differences (Choowattanapakorn et al., 2010).

Future Research

This study was unable to establish responsiveness within the PPOM-C. Responsiveness is an important psychometric property that determines whether an outcome measure can detect clinically important changes following intervention (Terwee et al., 2007). Including the PPOM-C in research-based interventions would ascertain if this criterion is present. Future researchers will benefit from using the included standard error of measurement (SEM) to ascertain confidence levels of scores on the PPOM-C. In

addition, the smallest detectable change (SDC) has been calculated to aid researchers.

The SDC can be used to establish whether change following intervention can be

associated to chance or treatment conditions using the SDC_{group} equation. For a

treatment condition of 20 patients, the SDC_{group} mean would need to change by 1.55

points on the PPOM-C to illustrate a significant change.

 The PPOM is currently being used in the promoting independence in dementia

'PRIDE' research programme (Shafayat et al., 2019). This may indicate whether the

PPOM has adequate responsiveness in people living with dementia. Research has found

hope and resilience are responsive to intervention in family carers (Duggleby et al.,

2018; MacCourt et al., 2017). Thus, the PPOM-C is also expected to demonstrate this

quality.

 One therapeutic approach aligned with positive psychology is Acceptance and

Commitment Therapy (ACT; Hayes & Strosahl, 2004). ACT is a third wave cognitive

behavioural therapy (CBT) approach. It teaches the individual to defuse from

unpleasant thoughts and feelings through its cognitive defusion techniques. Having

learned this, the person is encouraged to first consider and then act on their values in

their life. It has promising results in family carers (Losada et al., 2015), outperforming

traditional CBT in a randomised controlled trial. Given the values-based principles of

ACT, including such measures as the PPOM-C in interventions could yield valuable

information about hope, resilience and traditional psychopathology measures.

 The small, yet significant correlation coefficient between age and the PPOM-C

suggests older carers may possess more hope and resilience than their younger

counterparts. Future studies could examine this relationship further. Using the PPOM-C

longitudinally in a sample of family carers would yield more information with respect to

whether hope or resilience are qualities that are acquired, learned or develop naturally in

this group of people. Pairing this research with a qualitative component would give more information on this topic. This research could also provide more data on whether hope or resilience differs between the sexes in family carers.

Methodological Problems and Limitations

The respondents within this survey aligned with what Henrich et al. (2010) classified as WEIRD (White, Educated, and lived in Industrialised, Rich, and Democratic countries). The demographics of the sample did not reflect the United Kingdom diversity with respect to race or education levels (Office for National Statistics, 2016). For example, 96.6% of this sample were white compared to 86% of the UK population. In an attempt to counteract this, invitations were sent to all black and minority ethnic carers listed on the JDR database midway through the study.

The recruited sample stayed in education for longer compared to UK census data (Office for National Statistics, 2016) with 57.7% educated to degree level or higher. This figure contrasts sharply to 27% of the UK population having completed a similar level of education. It is therefore not known how generalisable these findings are to other populations within the UK.

The sample also indicated that approximately half of the participants did not live with the person they provided care for. This is important, as it could be reasonably considered that the experiences of carers who live with the person they care for versus those that don't would be different. The independent samples t-test however demonstrated no significant differences between these two groups on the PPOM-C. Given the aforementioned points regarding how this sample was recruited, it is unknown whether this finding would be generalisable to the general family carer population in the UK.

This study did not satisfy the content validity criterion of the COSMIN criteria (Mokkink et al., 2010). Previous research (Stoner et al., 2017) had evidenced this for people living with dementia. Thus, future research could assess the PPOM-C for its content validity with family carers. This could involve assessing whether the items on it represent relevant questions for carers. In addition, the PPOM-C could be circulated to experts in the field to ascertain feedback on whether the remaining items are relevant for this population.

To fulfil the convergent validity criterion, the HADS-D was administered without its anxiety subscale counterpart. The HADS-D was chosen based on research suggesting the anxiety subscale was not appropriate for family carers (Stott et al., 2017). In only administering the depression subscale, it is possible that this represented a threat to its validity as a sole measure of depression. It may have been more beneficial to use the PHQ-9 (Kroenke et al., 2001), which is a well validated measure for a wide range of different populations and ages.

The criterion validity criterion could have also been fulfilled through administering the PPOM-C with established gold standard measures of hope and resilience. Examining the resulting correlation coefficients would have given data regarding how the PPOM-C compares to established measures. As including more measures would have increased the time burden on participants in the study, I decided against this."

Conclusions

The PPOM-C is a short self-report measure that gives insights into the degree of hope and resilience of family carers. It possesses excellent internal consistency, excellent temporal stability, good convergent validity and adequate structural validity. The

PPOM-C illustrated significant correlations with quality of life, perceived social support

and depression indicating that hope and resilience play a role in the wellbeing of family

carers. Utilising the PPOM-C in interventions will give data regarding its

responsiveness. It is hoped that adopting the PPOM-C will influence future

interventions to be more strengths focused in their approach.

References

Beckerman, H., Roebroeck, M. E., Lankhorst, G. J., Becher, J. G., Bezemer, P. D., & Verbeek, A. L. M. (2001). Smallest real difference, a link between reproducibility and responsiveness. *Quality of Life Research, 10*, 571 – 578.

Cartwright, A. V. (2020). *Validation of the multidimensional scale of perceived social support (MSPSS) for family carers of people with dementia* (Unpublished clinical psychology doctoral thesis). University College London, London.

Choowattanapakorn, T., Aléx, L., Lundman, B., Norberg, A., & Nygren, B. (2010). Resilience among women and men aged 60 years and over in Sweden and in Thailand. *Nursing & Health Sciences, 12,* 329 – 335. https://doi.org/10.1111/j.1442-2018.2010.00534.x

Dufault, K., & Martocchio, B. (1985). Hope: Its spheres and dimensions. *Nursing Clinics of North America*, *20*, 379–391.

Duggleby, W., Williams, A. M., Wright, K., & Bollinger, S. (2009). Renewing everyday hope: The hope experience of family caregivers of persons with dementia. *Issues in Mental Health Nursing, 30*, 514-521. https://doi.org/10.1080/01612840802641727

Duggleby, W., Schroeder, D., & Nekolaichuk, C. (2013). Hope and connection: the experience of family caregivers of persons with dementia living in a long-term care facility. *BMC Geriatrics, 13***,** 112. https://doi.org/10.1186/1471-2318-13-112

Duggleby, W., Ploeg, J., McAiney, C., Peacock, S., Fisher, K., Ghosh, S., Markle-Reid, M., Swindle, J., Williams, A., Triscott, J. A., Forbes, D., & Jovel, K. (2018). Web-based intervention for family carers of persons with dementia and multiple

chronic conditions (my tools 4 care): Pragmatic randomized controlled trial.

Journal of Medical Internet Research, *20*. https://doi.org/10.2196/10484

Evans, J. D. (1996). *Straightforward statistics for behavioural sciences*. Brooks/Cole

Publishing.

Farina, N., Page, T. E., Daley, S., Brown, A., Bowling, A., Basset, T., Livingston, G.,

Knapp, M., Murray. J., & Bannerjee, S. (2017). Factors associated with the

quality of life of family carers with people with dementia: A systematic review.

Alzheimer's & Dementia. 13, 572-581.

https://doi.org/10.1016/j.jalz.2016.12.010

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible

statistical power analysis program for the social, behavioral, and biomedical

sciences. *Behavior Research Methods*, *39*, 175-191.

Frohlich, M. T. (2002). E-integration in the supply chain: barriers and performance.

Decision Sciences, *33*, 537-556.

Gaugler, J. E., Kane, R, L., & Newcomer, R. (2007). Resilience and transitions from

dementia caregiving. *Journal of Gerontology, 62,* 38-44.

https://doi.org/10.1093/geronb/62.1.p38

George, D., & Mallery, P. (2003). SPSS for windows step by step: A simple guide and

reference. 11.0 update (4th ed.). Allyn & Bacon.

Grant, G., & Nolan, M. (1993). Informal carers: Sources and concomitants of

satisfaction. *Health & Social Care in the Community*, *1*, 147–159.

https://doi.org/10.1111/j.1365-2524.1993.tb00210.x

Hatami, A., Ghalati, Z. K., Badrani, M. R., Jahangirimehr, A., & Hemmatipour, A.

(2019). The relationship between resilience and perceived social support with

hope in hemodialysis patients: A cross-sectional study. *Journal of Research in Medical and Dental Science, 7*, 14-20.

Hayes S.C., & Strosahl K.D. (2004). *A practical guide to acceptance and commitment therapy* (eds.). Springer.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). *"The weirdest people in the world?, "working paper series of the German council for social and economic data*. German Council for Social and Economic Data (RatSWD).

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal on Business Research Methods*, *6*, 2.

Hyde, M., Wiggins, R. D., Higgs, P., & Blane, D. B. (2003). A measure of quality of life in early old age: The theory, development and properties of a needs satisfaction model (CASP-19). *Aging & Mental Health*, 7, 186–194. https://doi.org/10.1080/1360786031000101157

Kline, R. B. (2011). *Principles and practice of structural equation modelling* (3rd ed.). The Guilford Press.

Komiti, A., Judd, F., Grech, P., Mijch, A., Hoy, J., Williams, B., Street, A., & Lloyd, J. H. (2003). Depression in people living with HIV/AIDS attending primary care and outpatient clinics. *Australian and New Zealand Journal of Psychiatry*, *37*, 70-77. https://doi.org/10.1046/j.1440-1614.2003.01118.x

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606-613.

Leong, F. T. L., & Huang, J. L. (2010). Standard error of measurement. In Salkind. N. J. (Ed.), *Encyclopaedia of research design.* Sage Reference.

Lewis, F., Karlsberg-Schaffer, S., Sussex, J., O'Neill, P., & Cockcroft, L. (2014). *Trajectory of dementia in the UK – making a difference, report produced the office of health economics for Alzheimer's research UK.* Dementia UK: 2014 edition. Alzheimer's Society, Carer information and Support Programme.

Logsdon, R. G., Gibbons, L. E., McCurry, S. M., & Teri, L. (1999). Quality of life in Alzheimer's disease: Patient and caregiver reports. *Journal of Mental Health and Aging*, *5*, 3–6. https://doi.org/10.1097/00006842-200205000-00016

Losada, A., Márquez-González, M., Romero-Moreno, R., Mausbach, B.T., Lopez, J., Fernandez-Fernandez, V., & Nogales-Gonzalez, C. (2015). Cognitive-behavioral therapy (CBT) versus acceptance and commitment therapy (ACT) for dementia family caregivers with significant depressive symptoms: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *83*, 760-772. https://doi.org/10.1037/ccp0000028

MacCourt, P., McLennan, M., Somers, S., & Krawczyk, M. (2017). Effectiveness of a grief intervention for caregivers of people with dementia. *Omega: Journal of Death and Dying*, *75*, 230–247. https://doi.org/10.1177/0030222816652802

Magaletta, P., & Oliver, J. (1999). The hope construct, will, and ways: Their relations with self-efficacy, optimism and general wellbeing. *Journal of Clinical Psychology, 55*, 539–551.

Masten, A. S., Best, K. M., & Garmezy, N. (1990). Resilience and development: Contributions from the study of children who overcome adversity. *Development and Psychopathology*, *3*, 425–444. https://doi.org/10.1017/S0954579400005812

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., Bouter, L., & De Vet. H. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology, 10*, 22. https://doi.org/10.1186/1471-2288-10-22

Moon, C., & Snyder, C. (2000). Hope and the journey with AIDS. In C. Snyder (Ed.), *Handbook of hope: Theory, measures and applications* (pp.341–353). Academic Press.

Moniz-Cook, E., Vernooij-Dassen, M., Woods, R., Verhey, F., Chattat, R., Vught, M. D., Mountain, G., O'Connell, M., Harrison, J., Vasse, E., Droes, R. M., Orrell, M., & INTERDEM Group. (2008). A European consensus on outcome measures for psychosocial intervention research in dementia care. *Aging & Mental Health, 12*, 14-29.

Moniz-Cook, E., Gibson, G., Harrison, J., & Wilkinson, H. (2009). Timely psychosocial interventions in a memory clinic In E. Moniz-Cook & J. Manthorpe (Eds.), *Early psychosocial interventions in dementia: Evidence-based practice.* Jessica Kingsley.

O'Rourke, N., Psych, R., & Hatcher, L. (2013). *A step-by-step approach to using SAS for factor analysis and structural equation modelling.* SAS Institute.

Office for National Statistics, National Records of Scotland, Northern Ireland Statistics and Research Agency. (2016). 2011 census aggregate data. UK Data Service. http://dx.doi.org/10.5257/census/aggregate-2011-1

Pinquart, M., & Sorensen, S. (2003). Differences between caregivers and non-
caregivers in psychological health and physical health: A meta-analysis.
Psychology and Aging, *18*, 250-267. https://doi.org/10.10.1037/0882-
7974.18.2.250

Prince., M, Knapp, M., Guerchet, M., McCrone, P., Prina, M., Comas-Herrera, A.,
Wittenberg, R., Adelaja, B., Hu, B., King, D., Rehill, A., & Salimkumar, D.
(2014). *Dementia UK: update second edition report.* King's College London
and the London School of Economics for the Alzheimer's Society.

Resilience. (2020). In *Oxford online dictionary*.
https://www.oxfordlearnersdictionaries.com/definition/english/resilience?q=resil
ience

Rothrock, N. E., Kaiser, K. A., & Cella, D. (2011). Developing a valid patient-reported
outcome measure. *Clinical pharmacology and therapeutics*, *90*, 737–742.
https://doi.org/10.1038/clpt.2011.195

Seligman, M. E. P. (1998). The president's address. *American Psychologist*, *54*, 559-
562.

Seligman, M. E., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology
progress: empirical validation of interventions. *American Psychologist*, *60*, 410–
421. https://doi.org/10.1037/0003-066X.60.5.410

Shafayat, A., Csipke, E., Bradshaw, L., Charlesworth, G., Day, F., Leung, P., Moniz-
Cook, E., Montgomery, A. A., Morris, S., Mountain, G., Ogollah, R., Sprange,
K., Yates, L., & Orrell, M. (2019). Promoting independence in dementia
(PRIDE): protocol for a feasibility randomised controlled trial. *Trials, 20*, 709.
https://doi.org/10.1186/s13063-019-3838-x

Shoenmakers, B., Buntinx, F., & Delepeleire. J. (2010). Factors determining the impact
of care-giving on caregivers of elderly patients with dementia. A systematic
literature review. *Maturitas.* 66, 191, 200.

Snyder, C., Feldman, D., Taylor, J., Schroeder, L., & Adams, V. (2000). The roles of
hopeful thinking in preventing problems and enhancing strengths. *Applied and
Preventative Psychology, 15*, 262–295.

Stansfeld, J., Stoner, C. R., Wenborn, J., Vernooij-Dassen, M., Moniz-Cook, E., &
Orrell, M. (2017). Positive psychology outcome measures for family caregivers
of people living with dementia: a systematic review. *International
Psychogeriatrics*, *29*, 1281–1296. https://doi.org/10.1017/S1041610217000655

Stoner, C, R., Orrell, M., Long, M., Csipke, E., & Spector, A. (2017). The development
and preliminary psychometric properties of two positive psychology outcome
measures for people with dementia: the PPOM and the EID-Q. *BMC Geriatrics,
17*, 72. https://doi.org/10.1186/s12877-017-0468-6

Stoner, C, R., Orrell, M., & Spector, A. (2018). The positive psychology outcome
measure (PPOM) for people with dementia: Psychometric properties and factor
structure. *Archives of Gerontology and Geriatrics, 76*, 182–187.
https://doi.org/10.1016/j.archger.2018.03.001

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. *Geriatric Psychiatry,
32,* 805-813. https://doi.org/10.1002/gps.4530

Tarlow, B. J., Wisniewski, S. R., Belle, S. H., Rubert, M., Ory, M. G., & Gallagher-
Thompson Stephen R. (2004). Positive aspects of caregiving: Contributions of
the REACH project to the development of new measures for Alzheimer's

caregiving. *Research on Aging*, *26*, 429–453.

https://doi.org/10.1177/0164027504264493

Teahan, A., Lafferty, A., McAuliffe, E., Phelan, A., O'Sullivan, L., O'Shea, D.,

Nicholson, E., & Fealy, G. (2020). Psychosocial interventions for family carers

of people with dementia: A systematic review and meta-analysis. *Journal of*

Aging and Health. https://doi.org/10.1177/0898264319899793

Terwee, C. B., Bot, S. D., de Boer, M. R., van Der Windt, D. A., Knol, D. L., Dekker,

J., & de Vet, H. C. (2007). Quality criteria were proposed for measurement

properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*,

34–42.

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and

utilitarian dimensions of consumer attitude. *Journal of Marketing Research*, *40*,

310-320.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey

(SF-36). I. Conceptual framework and item selection. *Medical Care*, 30, 473–

483.

Ware, J. E, Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey:

construction of scales and preliminary tests of reliability and validity. *Medical*

Care, 34, 220-233. https://doi.org/10.1097/00005650-199603000-00003

Waysman, M., Schwartzwald, J., & Soloman, Z. (2001). Hardiness: An examination of

its relationship with positive and negative long term changes following trauma.

Journal of Traumatic Stress, 14, 531-548.

https://doi.org/10.1023/A:1011112723704

Windle, G. (2011). What is resilience? A review and concept analysis. *Reviews in*

Clinical Gerontology, 21, 152-169. https://doi.org/10.1017/S0959259810000420

Windle, G., & Bennett, K.M. (2011). Resilience and caring relationships. In M. Ungar (Ed.), The social ecology of resilience (pp. 219_232). Springer.

Yesavage, J. A., Brink, T. L., Rose, T. I., & Adey, M. (1983). Development and validation of a geriatric depression rating scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37–49. https://doi.org/10.1016/0022-3956(82)90033-4

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*, 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment, 52*, 30-41. https://doi.org/10.1207/s15327752jpa5201_2

Zimet, G. D., Powell, S. S., Farley, G. K., Werkman, S., & Berkoff, K. A. (1990). Psychometric characteristics of the multidimensional scale of perceived social support. *Journal of Personality Assessment, 55*, 610-617. https://doi.org/10.1080/00223891.1990.9674095

Part 3: Critical Appraisal

Introduction

The following is a critical appraisal of both the literature review and empirical paper undertaken as a part of my doctorate in clinical psychology. The appraisal begins through outlining what it was about this project which appealed to me. After, I will detail the specific challenges faced whilst completing both papers. I will also consider the broader implications of quantitative measurement within the positive psychology context, before reflecting on the process.

Interest in the Project

Before obtaining a place on the doctorate, I had worked in the Improving Access to Psychological Therapies (IAPT; Layard et al., 2008) initiative for eight years. IAPT is the largest national initiative in history to offer a range of therapeutic interventions within primary care. There is a heavy emphasis on data collection using the IAPT minimum data set. This minimum data set is comprised of the Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001), the Generalised Anxiety Disorder-7 questionnaire (GAD-7; Spitzer et al., 2006) and the Work and Social Adjustment Scale (WSAS; Mundt et al., 2002). I was interested in how IAPT both defined and measured 'caseness' using the minimum data set in addition to what made a patient meet 'recovery'. Caseness refers to whether a person's difficulties measured by the minimum data set warrants an assessment by an IAPT service. Typically, this refers to a PHQ-9 score of above 9 and a GAD-7 score in excess of 7. Being in recovery refers to when someone who initially met caseness then has a PHQ-9 score of less than 9 and a GAD-7 score less than 7.

Something I had noticed whilst working in this context as a cognitive behavioural therapist was the principal focus on the negative aspects of wellbeing. In

this case, measures of depression, generalised anxiety and a measure of how much the person's difficulties were affecting work, home or education. I lamented the fact that instruments measuring constructs such as quality of life, wellbeing, improvement, coping, hope or resilience were not also administered. If the addition of these instruments were adopted at a national level, it struck me that much could be learned about the relationship between those in primary care therapy settings, the minimum data set and positive measures of wellbeing.

Of interest to me being a therapist trained in Acceptance and Commitment Therapy (ACT; Hayes & Strosahl, 2004), was anecdotal evidence that ACT was more effective in increasing resilience over the long term when compared to CBT. This anecdotal evidence is often offered by ACT trainers. Their rationale frames ACT as a transdiagnostic approach that can apply to virtually any distress-causing difficulty. Learning to defuse from distressing thoughts and feelings whilst living a values-led life are the key components of ACT. If recipients of ACT learn these principles, it follows that benefits would be experienced over the short and longer term. One study (Losada et al., 2015) has found this to be the case with family carers when comparing ACT to CBT, but more research would be welcome. Administering a measure of resilience at assessment, the midpoint, the end of sessions and at follow up would give more information on this hypothesis.

I was also interested in the wellbeing of carers of people with dementia given the remarkable efforts they undertake to care for their loved ones. As this project involved the evaluation of measurements that could ultimately lead to the enhancement of family carers' wellbeing, it certainly held appeal for me.

Systematic Review Challenges

The Research Question

The systematic review aims were broad and pertained to the evaluation of positive psychology outcome measures in use with family carers of people living with dementia. This question was influenced by a related review (Stansfeld et al., 2017) who had evaluated positive measures specifically developed or validated for family carers. Stansfield et al. (2017) recommended an additional review be conducted to evaluate all positive measures in use with carers. This included measures that had not been specifically developed or validated with family carers. This was suggested because researchers often use measures not specifically developed for their target population because those measures are perceived as robust, useful or convenient.

The aim of this review was therefore to evaluate how many positive measures were in use with family carers and evaluate their psychometric properties. The search culminated in finding 130 suitable measures. This was an unforeseen finding as I was not expecting so many positive measures to be in use. This presented a dilemma with respect to which instruments to include.

Selecting Instruments to Evaluate

As there was not enough time or space in the paper to subject 130 measures for quality appraisal, decisions had to be made with respect to which to exclude. These decisions subsequently shaped and refined the nature of the review. There were many ways in which the measures could have been included or excluded. For instance, the top twenty instruments with the most citations within my search could have been evaluated. The rationale for this approach could have been that they are the most used and therefore the most significant. An alternate method could have been to simply choose the constructs I

wanted to evaluate. In consultation with a supervisor, I opted to include measures linked to the three most frequently cited constructs. The justification for this was that comparisons of instruments found within these constructs could be compared against each other to determine which was the most robust.

Closer inspection of the measures revealed that not all constructs were exemplars of positive psychology as defined by Seligman (Seligman, 1998; Seligman et al., 2005). Examples of this included 'coping' and 'wellbeing'. Though 'coping' and 'wellbeing' can be useful constructs to measure, they did not seem as relevant to positive psychology as resilience or positive aspects to caregiving. Excluding constructs such as these was a difficult and subjective decision to make. This exclusion process leaves this study open to criticism given the innumerable ways in which the measures could have been chosen for evaluation. However, it is my hope that I or other colleagues in the future will come back to the excluded constructs, which could be evaluated through further systematic reviews.

Evaluating the Instruments

Evaluation of the 25 instruments made use of a quality appraisal framework for health-related outcome measures (Terwee et al., 2007). I noted whilst using this tool how at times it utilised rigid rules in conjunction with subjective judgements. With respect to the rigidity, the guidelines for scoring the Cronbach's alpha internal consistency score best exemplifies this. A score between $\alpha = 0.70$ and $\alpha = 0.949$ would indicate acceptable internal consistency. On one occasion, a measure (Faba et al., 2017) cited a Cronbach's alpha of $\alpha = 0.95$ and therefore did not achieve any score for internal consistency. Had the internal consistency score been 0.01 lower, this measure would have scored full marks for the reliability criterion.

Despite the rigid criteria, using this framework has produced different results measuring several overlapping instruments with the Stansfeld et al. (2017) review. This indicates that some criteria are more subjective with respect to whether the appraiser is convinced that the development paper has satisfied construct validity (where 75% of stated hypotheses need to be correct), content validity (where a clear description of the measure aim, target population, concepts being measured, item selection and consultation with the target population alongside experts were adequately described) along with criterion validity (where convincing arguments are given that a gold standard is in fact the gold standard).

Reflections on the Appraisal Process

Throughout the evaluation process, I was surprised at the differing ways in which authors attempted to evidence the reliability and validity of their instruments. Not one development paper cited a framework such as Terwee et al. (2007) that influenced the design or analysis of their measure. In addition, most texts concluded that the instrument's psychometric properties were adequate, when in fact, many properties were missing. This is evidenced by the fact that not one measure fully demonstrated criterion validity, agreement, responsiveness or interpretability criteria throughout this review.

This could reflect that for many development papers, quality appraisal tools such as Terwee et al.'s (2007) and the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN; Mokkink et al., 2010) were not yet developed. Such guidelines serve to collate and disseminate existing knowledge on psychometric measurement theory in a user-friendly manner. This is important as often the source texts for such topics can be difficult to interpret. These guidelines can also be

contradicted by other texts advising different statistics or thresholds to meet reliability or validity criteria. It is therefore possible that researchers had access to a myriad of information on measure development and potentially used different ideas on how to develop and validate them.

Even within such quality appraisal framework guidelines however, there are incongruences with respect to the language used to describe psychometric properties. For example, factor analyses are scored within 'reliability' in one framework (Terwee et al., 2007) and considered 'structural validity' in the other (Mokkink et al., 2010). A further example of this is highlighted by the development of the COSMIN criteria. Many experts in psychometrics were invited to participate in a Delphi consensus study to form the COSMIN guidelines (Mokkink et al., 2010). What I found interesting and enlightening is that even amongst experts in the field, there were still numerous disagreements with respect to how analyses should be conducted.

It is my view that unless a framework such as the COSMIN criteria are followed and integrated within the design and planning stages of measure development, there will continue to be variability in the quality of instruments. This is particularly important for positive psychology measures given their recent emergence into the field. If such measures can be developed or validated to a high standard, it would make incorporating them and analysing their data easier for future researchers.

Whilst completing this review I learned about the importance of narrowing down my research question. I had underestimated the number of positive measures in use for family carers. If I had the chance to recomplete this, I would set out from the beginning to locate measures only related to specific constructs. This approach would have negated the need to filter out certain constructs post-searching.

Empirical Paper Critical Appraisal

Hope & Resilience in the Caregiving Context

My first thoughts when reflecting on the prospect of validating the Positive Psychology Outcome Measure (PPOM; Stoner et al., 2017) in a sample of family carers were mixed. Regretfully, I had not considered the degree to which hope and resilience could be present in caregivers given the deteriorating nature of dementia in the person they supported. I wondered whether this is an attitude shared by others. This is evident through negative measures thoroughly outweighing positive instruments in the literature with this population.

I was also aware of not wanting to invalidate carers' difficult experiences through disseminating a measure asking questions that perhaps carers are rarely asked (e.g. I believe each day has potential, PPOM question 7). Conversely, I was aware of the buffering effect of constructs such as hope and resilience against psychopathology. I also believe there are opportunities to learn from those carers who remain hopeful and resilient in the face of stressors.

The PPOM-C and its Epistemological Position

The definitions of hope and resilience are multi-faceted. As such, researchers often use a variety of definitions for each construct. Wolverson et al. (2009) in a literature search on hope noted 18 differing definitions with 13 associated instruments. Such definitions included;

- Hoping for a better future (Obayuwana & Carter, 1982).
- Hope being activated when triggered by stressful situations such as an illness (Fryback, 1991).

- Hope relating to domains of life that are valued by the individual (Averill et al., 1990).

- Hope linked to specific, measurable and achievable goals (Dufault & Martocchio, 1985).

- Hope linked to desiring a life worth living (Farran et al., 1995).

- Hope propelling people towards their valued actions (Snyder et al., 1991).

- Hope being linked to relationships with others and occurring in systems around and between people (Miller, 1986).

Similarly, Windle (2011) set out to define the resilience construct following a literature search and summarised resilience as follows:

'Resilience is the process of effectively negotiating, adapting to, or managing significant sources of stress or trauma. Assets and resources within the individual, their life and environment facilitate this capacity for adaptation and 'bouncing back' in the face of adversity. Across the life course, the experience of resilience will vary'. (Windle, 2011, p. 163).

I found the above multi-faceted definitions thought-provoking, given that corresponding measures do not always map onto them precisely. A limitation of the PPOM-C and perhaps hope and resilience measures in general is the assumption of where hope and resilience are located. Do the items on the measure indicate the construct resides within an individual, the system, their environment, or a combination? Frequently, such instruments locate the construct within an individual. Perhaps this is because it is harder to quantitatively measure hope and resilience within the system or

environment. Nevertheless, a valid criticism of the PPOM-C would be its lack of questionnaire items addressing contextual factors around the individual.

Study Sample

It was apparent that the sample were not representative of the UK population. Efforts were made to ensure a variety of ethnic backgrounds, age and gender were fairly represented in the study. This was able to happen due to the check on the data that was conducted once 50 respondents had completed the study. The participant characteristics revealed the main responders were older white British women. To obtain a more representative sample, efforts were made to recruit younger carers, male carers and carers from black and minority ethnic backgrounds. Additional younger carers and men ultimately completed the survey. However, the final percentages of the sample revealed that the proportion of white British participants was higher than the last census (Office for National Statistics, 2016). This is a common phenomenon in research (Henrich et al., 2010). It suggests the convenience sampling method using the Join Dementia Research organisation had inherently biased the study towards white socio-economically privileged people. This hinders the generalisability of the findings of this measure, and is a limitation of the PPOM-C.

The Inception of the PPOM-C

The internal consistency of the PPOM indicated multicollinearity. This presented a methodological dilemma to negotiate. One option included acknowledging that there was multicollinearity but falling back on the fact that it had previously fallen within an acceptable range (Stoner et al., 2018). Another perspective is that in this target

population, perhaps fewer items would reliably measure the construct. I decided on the latter and omitted one item per subscale to ensure they were evenly weighted.

Wieland et al. (2017) described several procedures through which to omit items following a review into the literature. They justify deleting items if it improves the reliability, validity or parsimony of the measure. Within reliability, items are recommended to be deleted if the individual item reliability, the average variance extracted, the item-total correlation or the Cronbach's alpha are too low. Within validity, items should be deleted if correlations amongst items are too low within the same construct or if convergent validity testing fails to obtain high enough correlations with theoretically related constructs. In addition, items should be deleted if they have been shown not to represent the linked construct in a CFA. The parsimony criteria suggest omitting an item if the inter-item correlations are too high or if the number of items per construct is too high. Finally, the parsimony criteria suggest deleting an item if it results in increasing the goodness of fit indices (Frohlich, 2002; Voss et al., 2003).

As I am not experienced in factor analysis, deciding which route to take was fraught with difficulties. There were positives and negatives across the range of different approaches. Eventually, I opted for Frohlich (2002) and Voss et al.'s (2003) approach in using a variety of criteria to delete one item for each subscale. Doing this enabled me to test out a comprehensive range of strategies to ascertain which increased the goodness of fit indices the most (Appendix H).

Eventually, this led me to delete hope, question 5 and resilience, question 2 from the PPOM. These items were chosen as they possessed the lowest regression weights onto their respective constructs. Deleting these items had the greatest benefit on goodness of fit indices. Whenever items are omitted from a measure, there is a risk that its validity has been compromised somewhat. Given the other satisfactory analyses the

PPOM-C achieved, my hope is that its validity remains intact. Other researchers may have adopted a different method in undertaking this. One such approach could have utilised item response theory (IRT).

IRT could have been an effective way in which to identify redundant items within the measure. However, large sample sizes in excess of 500 are recommended (Zanon et al., 2016). IRT models typically compare whether observed scores conform to what the IRT model predicts. Fit indices can then be improved by taking several different approaches to deleting redundant items if necessary. Firstly, fit could be assessed through evaluating whether standardised residuals approximately fit a normal distribution pattern. Secondly, items with standardised residuals falling between -2 and 2 indicate poor fit and would warrant removal. The application of IRT practices such as Rasch models are typically filled with complexity and in some cases produce similar results as CFA classical test theory (Prieto et al., 2003).

CFA Goodness of Fit Indices

The goodness of fit indices of the PPOM-C were adequate. Only the chi-square and standardised root mean square error of approximation (RMSEA) indicated poor model fit. Part of the difficulty for a newcomer to factor analysis are the sheer number of fit indices available. Debate within the field points to occasions where fit indices claim a well-fitting model when in fact, sections of the model may be poor (Reisinger & Mavondo, 2006; Tomarken & Waller, 2003). This is further complicated by contradictory guidance as to what constitutes an acceptable level of fit. For example, the RMSEA has previously been recommended to vary between 0.5 and .10 (MacCallum et al., 1996), but this guidance now fluctuates between 0.06 – 0.08 (Steiger, 2007).

Some researchers advocate covarying error terms within the CFA to improve fit. This route can be taken only if there are good justifications for doing so. Whilst analysing my own CFA, I found that I could dramatically improve goodness of fit through covarying error terms with high modification indices. I opted against this as it felt like a shortcut to better outcomes with little justification. Others have cautioned against using this strategy (Gerbing & Anderson, 1984) for similar reasons.

Reflections on Quantitative Measurement

I recognise that quantitative measurement instruments are a useful research and clinical tool. They can inform both researcher and client of any clinically significant changes in addition to assisting in diagnosis. More generally, measures are useful in informing researchers about the effectiveness of interventions in research settings. Conversely, I'm aware that such a measure can reduce a person's state or trait to what is essentially a number.

This number cannot replace the rich conversations with people about what their strengths and resources are in addition to their difficulties and problems. Pragmatically, this qualitative information could be very useful for the individual therapist. This therapist could additionally use measures to instigate conversations around meaningful changes for the client. For example, if a significant change in the level of hope or resilience was noted, a conversation could be had about what changes led to this difference. In research intervention settings, obtaining such qualitative information for a large sample size may be impractical. Thus, a robust measure could capture useful information pertaining to pre and post changes for a larger number of people.

Conclusions

The use of outcome measures for research and clinical purposes is useful for many different reasons. It can be used with a practice-based evidence ethos to ensure improvements and setbacks can be talked about collaboratively.

Measuring negative states or emotions has become the norm for many different research and therapeutic settings within the UK. In some cases, this has helped advance the evidence base and the reputation of certain therapeutic orientations such as CBT.

This thesis has focused on positive psychology outcome measures in the family caregiving context. Predominantly, the measures in use at present with family carers focus on the measurement of negative states or emotions. I hope this thesis can facilitate further research around positive measures used when working with this population. Measures such as the PPOM-C can highlight the hopefulness or resilience of carers. Adopting measures such as this could be paired with designing new interventions trying to help make meaningful positive changes to people's lives (Seligman, 1998).

The process of validating and shortening a measure has taught me that despite frameworks guiding researchers, there are still a myriad of subjective decisions to be made. This will encourage me to be critical when examining the results of such measures in the future. It is possible to have a reliable measure that lacks validity and vice versa. All of these psychometric properties interlink to reliably measure the construct they set out to measure.

Future researchers using the PPOM-C should consider its limitations in its use. Whilst, it measures a person's hope and resilience, it arguably does this through assuming they reside in the individual as opposed to their network and environment. Researchers may wish to supplement the PPOM-C with a measure of social support to better understand their social capital and resources.

References

Averill, J. R., Catlin, G., & Chon, K.K. (1990). *Rules of hope.* Springer.

Dufault, K., & Martocchio, B. (1985). Hope: Its spheres and dimensions. *Nursing Clinics of North America*, *20*, 379–391.

Faba, J., Villar, F., & Giuliani, M. F. (2017). Development of a measure to evaluate gains among spanish dementia caregivers: The gains associated with caregiving (GAC) scale. *Archives of Gerontology and Geriatrics*, *68*, 76–83. https://doi.org/10.1016/j.archger.2016.09.004

Farran, C., Herth, K., & Popovich, J. (1995). *Hope and hopelessness: Critical clinical constructs.* Sage Publications Ltd.

Frohlich, M. T. (2002). E-integration in the supply chain: barriers and performance. *Decision Sciences*, *33*, 537-556. https://doi.org/10.1111/j.1540-5915.2002.tb01655.x

Fryback, P. (1991). *Perceptions of health by persons with a terminal disease: Implications for nursing* (Unpublished doctoral dissertation). University of Texas, Austin.

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, *11*, 572-80. https://doi.org/10.1086/208993

Hayes S. C., & Strosahl K. D. (2004). *A practical guide to acceptance and commitment therapy* (eds.). Springer.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). *"The weirdest people in the world?," working paper series of the German council for social and economic data.* German Council for Social and Economic Data (RatSWD).

Kroenke, K., Spitzer, R. L., & Williams, J, B. (2001). The PHQ-9: validity of a brief

depression severity measure. *Journal of General Internal Medicine, 16*, 606-

613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Layard, R., Bell, S., Clark, D. M., Knapp, M., Meacher, M., & Priebe, S. (2006). *The*

depression report: A new deal for depression and anxiety disorders. London

School of Economics.

Losada, A., Márquez-González, M., Romero-Moreno, R., Mausbach, B.T., Lopez, J.,

Fernandez-Fernandez, V., & Nogales-Gonzalez, C. (2015). Cognitive-behavioral

therapy (CBT) versus acceptance and commitment therapy (ACT) for dementia

family caregivers with significant depressive symptoms: Results of a

randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *83*,

760-772. https://doi.org/10.1037/ccp0000028

MacCallum, R.C., Browne, M.W., & Sugawara, H., M. (1996). Power analysis and

determination of sample size for covariance structure modeling. *Psychological*

Methods, 1, 130-49. https://doi.org/10.1037/1082-989X.1.2.130

Miller, J. (1986). Inspiring hope. *American Journal of Nursing, 85*, 22–25.

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L.,

Bouter, L., & De Vet. H. (2010). The COSMIN checklist for evaluating the

methodological quality of studies on measurement properties: A clarification of

its content. *BMC Medical Research Methodology, 10*, 22.

https://doi.org/10.1186/1471-2288-10-22

Mundt, J. C., Marks, I. M., Shear, K., & Greist, J. H. (2002). The work and social

adjustment scale: A simple measure of impairment in functioning. *The British*

Journal of Psychiatry, 180, 461-464. https://doi.org/10.1192/bjp.180.5.461

Obayuwana, A., & Carter, A. (1982). The anatomy of hope. *Journal of the National Medical Association, 74*, 229–234.

Office for National Statistics, National Records of Scotland, Northern Ireland Statistics and Research Agency. (2016). 2011 census aggregate data. UK Data Service. http://dx.doi.org/10.5257/census/aggregate-2011-1

Prieto, L., Alonso, J., & Lamarca, R. (2003). Classical test theory versus rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes, 1*. https://doi.org/10.1186/1477-7525-1-27

Reisinger, Y., & Mavondo, F. (2006). Structural equation modeling: Critical issues and new developments. *Journal of Travel and Tourism Marketing*, *21*, 41-71. https://doi.org/10.1300/J073v21n04_05

Seligman, M. E. P. (1998). The president's address. *American Psychologist*, *54*, 559-562.

Seligman, M. E., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist*, *60*, 410–421. https://doi.org/10.1037/0003-066X.60.5.410

Snyder, C., Harris, C., Anderson, J., Holleran, S., Irving, L., Sigmon, S. T., Yoshinobu, J., Gibb, J., Langelle, C., & Harney. P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60*, 570–585. https://doi.org/10.1037//0022-3514.60.4.570

Spitzer, R, L., Kroenke, K., Williams, J. B., and Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*, 1092-7. https://doi.org/10.1001/archinte.166.10.1092

Stansfeld, J., Stoner, C. R., Wenborn, J., Vernooij-Dassen, M., Moniz-Cook, E., & Orrell, M. (2017). Positive psychology outcome measures for family caregivers of people living with dementia: a systematic review. *International Psychogeriatrics*, *29*, 1281–1296. https://doi.org/10.1017/S1041610217000655

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modelling. *Personality and Individual Differences*, *42*, 893-98. https://doi.org/10.1016/j.paid.2006.09.017

Stoner, C, R., Orrell, M., & Spector, A. (2018). The positive psychology outcome measure (PPOM) for people with dementia: Psychometric properties and factor structure. *Archives of Gerontology and Geriatrics, 76*, 182–187. https://doi.org/10.1016/j.archger.2018.03.001

Terwee, C. B., Bot, S. D., de Boer, M. R., van Der Windt, D. A., Knol, D. L., Dekker, J., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*, 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Tomarken, A. J., & Waller, N. G. (2003). Potential problems with "well fitting" models. *Journal of Abnormal Psychology*, *112*, 578-98. https://doi.org/10.1037/0021-843X.112.4.578

Voss, K.E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and utilitarian dimensions of consumer attitude. *Journal of Marketing Research*, *40*, 310-320. https://doi.org/10.1509/jmkr.40.3.310.19238

Wieland, A., Durach, C. F., Kembro, J., & Treiblmaier, H. (2017). Statistical and judgmental criteria for scale purification. *Supply Chain Management*, *22*, 324-328. https://doi.org/10.1108/SCM-07-2016-0230

Windle, G. (2011). What is resilience? A review and concept analysis. *Reviews in Clinical Gerontology, 21*, 152-169. https://doi.org/10.1017/S0959259810000420

Wolverson, E., Clarke, C., & Moniz-Cook, E. (2009, January). Hope – is it a relevant concept for dementia research? *Psychology Specialists: Working with Older People: British Psychological Society*, 11-16.

Zanon, C., Hutz, C. S., Yoo, H., & Hambleton, R. K. (2016). An application of item response theory to psychological test development. *Psicologia: Reflexão e Crítica*, *29,* 18.

Appendix A

UCL Ethics Approval Letter

8th March 2019

Professor Aimee Spector
Department of Clinical, Educational and Health Psychology
UCL

Dear Professor Spector,

Notification of Ethics Approval with Provisos
Project ID/Title: 15139/001: Online validation of the Multidimensional Scale of Perceived Social Support (MSPSS) and Positive Psychology Outcome Measure (PPOM) for family carers of people with dementia

Further to the review of your re-submitted application at the February meeting of the UCL REC, I am pleased to confirm in my capacity as Joint Chair of the UCL Research Ethics Committee (REC) that your study has been ethically approved by the UCL REC until 1st March 2020.

Ethical approval is subject to the following conditions:

Notification of Amendments to the Research
You must seek Chair's approval for proposed amendments (to include extensions to the duration of the project) to the research for which this approval has been given. Each research project is reviewed separately and if there are significant changes to the research protocol you should seek confirmation of continued ethical approval by completing an 'Amendment Approval Request Form'
http://ethics.grad.ucl.ac.uk/responsibilities.php

Adverse Event Reporting – Serious and Non-Serious
It is your responsibility to report to the Committee any unanticipated problems or adverse events involving risks to participants or others. The Ethics Committee should be notified of all serious adverse events via the Ethics Committee Administrator (ethics@ucl.ac.uk) immediately the incident occurs. Where the adverse incident is unexpected and serious, the Joint Chairs will decide whether the study should be terminated pending the opinion of an independent expert. For non-serious adverse events the Joint Chairs of the Ethics Committee should again be notified via the Ethics Committee Administrator within ten days of the incident occurring and provide a full written report that should include any amendments to the participant information sheet and study protocol. The Joint Chairs will confirm that the incident is non-serious and report to the Committee at the next meeting. The final view of the Committee will be communicated to you.

Final Report

At the end of the data collection element of your research we ask that you submit a very brief report (1-2 paragraphs will suffice) which includes in particular issues relating to the ethical implications of the research i.e. issues obtaining consent, participants withdrawing from the research, confidentiality, protection of participants from physical and mental harm etc.

In addition, please:

- ensure that you follow all relevant guidance as laid out in UCL's Code of Conduct for Research: http://www.ucl.ac.uk/srs/governance-and-committees/resgov/code-of-conduct-research
- note that you are required to adhere to all research data/records management and storage procedures agreed as part of your application. This will be expected even after completion of the study.

With best wishes for the research.

Yours sincerely
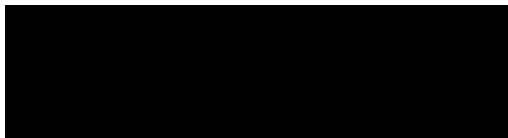
Professor Michael Heinrich
Joint Chair, UCL Research Ethics Committee

Cc: Anna Cartwright

Appendix B

Joint project statement of contributions to the project

The empirical paper of this thesis was completed jointly with another trainee clinical psychologist from University College London.

Systematic Review

The systematic reviews were independently conducted. However, in an effort to increase the quality of the analysis, my research partner and I collaborated in analysing each other's results. Therefore, my research partner assisted this review via independently completing psychometric evaluations of the 25 measures selected for appraisal. We met to discuss and agree upon differences in scoring using the quality appraisal criteria.

Empirical Paper

The main way in which my research partner and I collaborated on this paper was in the design and recruitment of the study. We searched the literature to locate instruments that would be appropriate for the convergent validity testing aspect of the study. Our studies were then designed to validate a different measure in a sample of family carers of people with dementia. My research partner aimed to validate a measure of perceived social support (the MSPSS) whilst I aimed to validate the PPOM. We then worked together in recruiting for the study. As mentioned in the methods section, we utilised the Join Dementia Research organisation to contact volunteers who might be interested in participating. At the end of recruitment, we worked together to obtain a joint data set. Subsequently, all analyses and write up were completed separately.

Appendix C

Participant information sheet

**RESEARCH DEPARTMENT OF CLINICAL,
EDUCATIONAL AND HEALTH PSYCHOLOGY
UNIVERSITY COLLEGE LONDON
1-19 TORRINGTON PLACE
LONDON
WC1E 6BT**

Participant Information Sheet for family carers of people with dementia
UCL Research Ethics Committee Approval ID Number: 15139/001

Title of Study: Online validation of the Multidimensional Scale of Perceived Social Support (MSPSS) and Positive Psychology Outcome Measure (PPOM) for family carers of people with dementia
Department: Research Department of Clinical, Educational and Health Psychology.
Name and Contact Details of the Researcher(s): Anna Cartwright (a.cartwright.17@ucl.ac.uk); Richard Pione (r.pione.17@ucl.ac.uk)
Name and Contact Details of the Principal Researcher: Professor Aimee Spector (a.spector@ucl.ac.uk)

1. **Invitation to participate**
   You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information and discuss it with others if you wish. Feel free to contact us if anything is not clear or if you would like more information. Thank you for reading this information sheet.

2. **What is the project's purpose?**
   This study is aiming to determine the reliability and validity of two questionnaires that are used in research with carers of people with dementia and to increase understanding of the roles that social support, hope and resilience play in caring for a person living with dementia. We aim to better understand how these factors may be related to mood and quality of life.

3. **Why have I been chosen?**
   You have been invited to take part in this research, as you have had contact with a service or website that supports dementia-related research. **To complete these questionnaires, you must identify as an adult (age 18+) unpaid family carer of a person with dementia, have access to the internet to complete questionnaires online and be able to understand written English.**

4. **Do I have to take part?**
   It is up to you to decide whether or not to take part. If you do decide to take part, you will be asked to agree to a consent form on the next page. You are free to withdraw at any time without giving a reason. If you decide to withdraw, you can contact the researchers to request for your data to be deleted.

5. **What will happen to me if I take part?**

   You will be asked to provide some information about yourself, such as your age, gender, ethnicity and marital status and to complete four questionnaires online. The questionnaires include questions relating to hope and resilience, social support, mood and quality of life. This should take approximately 25-35 minutes to complete. We will also ask you to provide an email address, so that we can send an email invitation to some participants approximately four weeks later, asking you to complete two of the questionnaires again. This second set should take approximately 10 minutes to complete.

6. **What are the possible disadvantages and risks of taking part?**

   We do not consider that taking part will involve any disadvantages or specific risks or that it would cause you any harm.

7. **What are the possible benefits of taking part?**

   If you do decide to take part in this project, you will be making a valuable contribution in helping us improve understanding of how personal strengths and social support are related to mood and quality of life in carers of people living with dementia.

8. **What if something goes wrong?**

   If you are unhappy or dissatisfied with any aspect of your participation, we would ask you firstly to speak to one of the researchers, so that we can try to address your concerns and find a solution. Alternatively, you can speak to the project supervisor; Professor Aimee Spector (see contact details below). Should you feel that a complaint has not been handled to your satisfaction, then you can contact the Chair of the UCL Research Ethics Committee at: ethics@ucl.ac.uk

9. **Will my taking part in this project be kept confidential?**

   All the information that we collect about you will be kept strictly confidential. Any personal details (i.e. your email address) are kept securely on an encrypted device and stored separately from the information collected during the study. All other data will be pseudoanonymised: a unique code will be linked to your data, from which you will not be identifiable. You will not be identified in any reports or publications.

   We will also ask your permission for us to store, in anonymised form, the information that we collect from you for long-term use, in data archives provided by University College London. This would mean that information could be shared with accredited researchers in future so that they might understand more and learn from the information we collect. No one would be able to identify you from this archived information.

10. **Limits to confidentiality**

    Confidentiality will be respected subject to legal constraints and professional guidelines.

11. **What will happen to the results of the research project?**

    The results of this study will be presented within Doctorate theses and published in scientific journals. Participants can be kept informed of the progress of the project and results, should they express a desire to receive further information.

12. Data Protection Privacy Notice
Notice:

The controller for this project will be University College London (UCL). The UCL Data Protection Officer provides oversight of UCL activities involving the processing of personal data, and can be contacted at data-protection@ucl.ac.uk

This 'local' privacy notice sets out the information that applies to this particular study. Further information on how UCL uses participant information can be found in our 'general' privacy notice:

If you would like to find out more about this privacy notice, please click here.

The information that is required to be provided to participants under data protection legislation (GDPR and DPA 2018) is provided across both the 'local' and 'general' privacy notices.

The categories of personal data used will be as follows: email address

Your personal data will be processed for the purposes outlined in this notice. It will be collected and stored using the Qualtrics questionnaire programme. Qualtrics are compliant with all GDPR regulations. You can find out more about how Qualtrics store data at:
https://www.qualtrics.com/uk/platform/gdpr/

The legal basis that would be used to process your personal data will be performance of a task in the public interest.

The lawful basis used to process *special category personal data* will be for scientific and historical research or statistical purposes.

Your personal data will be processed so long as it is required for the research project. If we are able to anonymise or pseudonymise the personal data you provide we will undertake this and will endeavour to minimise the processing of personal data wherever possible.

If you are concerned about how your personal data is being processed, please contact UCL in the first instance at data-protection@ucl.ac.uk. If you remain unsatisfied, you may wish to contact the Information Commissioner's Office (ICO). Contact details, and details of data subject rights, are available on the ICO website at: https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/individuals-rights/

13. Who is organising and funding the research?
The study is being organised by Anna Cartwright and Richard Pione; trainee clinical psychologists at University College London (UCL). The research is being funded by the Department for Clinical, Educational and Health Psychology, University College London.

14. Contact for further information
For more information about this research, please contact:

FAO Anna Cartwright / Richard Pione, Trainee Clinical Psychologist
Research Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place

WC1E 7HB
Tel: 020 7679 4466
Email: a.cartwright.17@ucl.ac.uk; r.pione.17@ucl.ac.uk

If you have any concerns or complaints about anything to do with this study, please contact:

Prof. Aimee Spector
Research Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place
London
WC1E 7HB
Tel: 020 7679 1844
Email: a.spector@ucl.ac.uk

Thank you for taking the time to consider taking part in this research study.

Appendix D

Consent form

**RESEARCH DEPARTMENT OF CLINICAL,
EDUCATIONAL AND HEALTH PSYCHOLOGY
UNIVERSITY COLLEGE LONDON
1-19 TORRINGTON PLACE
LONDON
WC1E 6BT**

CONSENT FORM FOR FAMILY CARERS FOR PEOPLE WITH DEMENTIA IN RESEARCH STUDIES

Please complete this form after you have read the Information Sheet.

Title of Study: Online validation of the Multidimensional Scale of Perceived Social Support (MSPSS) and Positive Psychology Outcome Measure (PPOM) for family carers of people with dementia
Department: Department of Clinical, Educational and Health Psychology
Name and Contact Details of the Researcher(s): Anna Cartwright (a.cartwright.17@ucl.ac.uk) and Rich Pione (r.pione.17@ucl.ac.uk).
Name and Contact Details of the Principal Researcher: Aimee Spector (a.spector@ucl.ac.uk)
Name and Contact Details of the UCL Data Protection Officer: Lee Shailer (l.shailer@ucl.ac.uk)
This study has been approved by the UCL Research Ethics Committee: Project ID number: 15139/001

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I confirm that I understand that by ticking each box below I am consenting to this element of the study. I understand that it will be assumed that unticked boxes means that I DO NOT consent to that part of the study. I understand that by not giving consent for any one element that I may be deemed ineligible for the study.

		Tick Box
1.	I confirm that I have read and understood the Information Sheet for the above study. I have had an opportunity to consider the information and what will be expected of me. I have also had the opportunity to ask questions which have been answered to my satisfaction and would like to take part in the online questionnaires.	
2.	I consent to participate in the study. I understand that my personal information (i.e. my email address) will be used for the purposes explained to me. I understand that according to data protection legislation, 'public task' will be the lawful basis for processing.	
3.	I understand that my participation is completely voluntary and that I am free to withdraw at any time, without giving reason. I understand that withdrawing will not affect my healthcare or legal rights. I understand that if I withdraw from the study, I can contact the researchers to request that they delete any data already provided.	
4.	**Use of the information for this project only** I understand that all personal information will remain confidential and that my email address will be stored securely and separately from the data collected during the study. All other data will be pseudoanonymised and stored securely. All efforts will be made to ensure I cannot be identified from the data. It will not be possible to identify me in any publications.	
5.	I understand that my information may be subject to review by responsible individuals	

		from the University for monitoring and audit purposes.	
	6.	I understand that no promise of guarantee of benefits have been made to encourage me to participants.	
	7.	I understand that the data will not be made available to any commercial organisations but is solely the responsibility of the researchers undertaking this study.	
	8.	I understand that I will not benefit financially from this study or from any possible outcome it may result in in the future.	
	9.	I agree that my anonymised research data may be used by others for future research. [No one will be able to identify you when this data is shared.]	
	10.	I understand that the information I have submitted will be published as a report and that I can contact the researchers to request further information and the progress and findings of the project.	
	11.	I hereby confirm that I understand the inclusion criteria as detailed in the Information Sheet and confirm that I fall under these inclusion criteria.	
	12.	I am aware of who I should contact if I wish to log a complaint.	
	13.	**Use of information for this project and beyond** I would be happy for the data I provide to be archived, in anonymous form, at University College London (UCL) and understand that other authenticated researchers will have access to my anonymised data.	
	14.	I voluntarily agree to participate in this study	

Appendix E

The end of survey message to carers

Thank you for taking the time to complete our questionairre! Your responses and time spent completing these are very much appreciated.

You can download the participant information sheet here and the consent form here if you would like to do so.

We have included some links below that we thought could be useful for carers of people living with dementia.

Support for carers of pople with dementia

NHS guidance on caring for people living with dementia

Search for your local NHS therapy, counselling service

Looking for support, advice or informatiom about dementia? Call the National Dementia Helpline

Find out more about Carers UK

Contact other carers on the Carers UK forum

Appendix F

The Positive Psychology Outcome Measure (PPOM)

The Positive Psychology Outcome Measure (PPOM)

Instructions

We would like to know how you have been feeling over the **past month**. Please answer the below questions by **circling one number** (0, 1, 2, 3 or 4) that most closely reflects how you have felt for each question. Please answer all the questions. If you are unsure, circle the number that is your best guess.

PPOM

	Not true at all	Rarely true	Sometimes true	Often true	True nearly all the time
1. I have a positive outlook on life	0	1	2	3	4
2. I can see positive things in difficult situations	0	1	2	3	4
3. I can recall happy/ joyful times	0	1	2	3	4
4. I have inner strength	0	1	2	3	4
5. I can give and receive care/ love	0	1	2	3	4
6. I have a sense of direction in life	0	1	2	3	4
7. I believe that each day has potential	0	1	2	3	4
8. My life has value and worth	0	1	2	3	4
9. I am able to adapt to things	0	1	2	3	4
10. I am able to deal with whatever happens	0	1	2	3	4
11. I am able to see the humorous side	0	1	2	3	4
12. I can cope with stress well	0	1	2	3	4
13. I can bounce back	0	1	2	3	4
14. I can stay focused	0	1	2	3	4

15. I am an emotionally strong person	0	1	2	3	4
16. I can handle unpleasant feelings	0	1	2	3	4

Appendix G

The COSMIN Checklist

For all checklists below, green indicates 'yes', red indicates 'no' and black indicates not applicable'

Box A. Internal Consistency	
Criterion	Yes/ No/ N/A
Does the scale consist of effect indicators, i.e. is it based on a reflective model?	Y
Design Requirements	
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the internal consistency analysis adequate?	Y
Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Y
Was the sample size in the unidimensionality analysis adequate?	Y
Was an internal consistency statistic calculated for each (unidimensional) sub (scale) separately?	Y
Were there any important flaws in the design or methods of the study?	N
Statistical methods	
For classical test theory (CTT), was Cronbach's alpha calculated?	Y
For dichotomous scores: was Cronbach's alpha or KR-20 calculated?	(black)
For IRT: Was a goodness of fit statistic at a global level calculated?	(black)

Box B. Reliability	
Criterion	Yes/ No/ N/A
Design Requirements	
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the analysis adequate?	Y
Were at least two measurements available?	Y
Were the administrations independent?	Y
Was the time interval stated?	Y
Were patients stable in the interim period on the construct to be measured?	Y
Was the time interval appropriate?	Y
Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions.	Y
Were there any important flaws in the design or methods of the study?	N

Statistical methods	
For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	Y
For dichotomous/ nominal/ ordinal scores: was kappa calculated?	
For ordinal scores: was a weighted kappa calculated?	
For ordinal scores: was the weighing scheme described? E.g. linear, quadratic.	

Box C. Measurement error: absolute measures	
Criterion	Yes/ No/ N/A
Design Requirements	
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the analysis adequate?	Y
Were at least two measurements available?	Y
Were the administrations independent?	Y
Was the time interval stated?	Y
Were patients stable in the interim period on the construct to be measured?	Y
Was the time interval appropriate	Y
Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions.	Y
Were there any important flaws in the design or methods of the study?	N
Statistical methods	
For (CTT): Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	Y

Box D. Content Validity	
Criterion	Yes/ No/ N/A
General requirements	
Was there an assessment of whether all items refer to the relevant aspects of the construct to be measured?	N
Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting).	N
Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/ or predictive).	N
Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	N

Were there any important flaws in the design or methods of the study?	Y

Box E. Structural validity

Criterion	Yes/ No/ N/A
Does the scale consist of effect indicators, i.e. is it based on a reflective model?	Y
Design Requirements	Y
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the analysis adequate?	Y
Were there any important flaws in the design or methods of the study	N
Statistical methods	
For (CTT): Was exploratory or confirmatory factor analysis performed?	Y
For IRT: were IRT tests for determining the (uni) dimensionality of the items performed?	

Box F. Hypothesis testing

Criterion	Yes/ No/ N/A
Design Requirements	
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the analysis adequate?	Y
Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)?	Y
Was the expected direction of correlations or mean differences included in the hypotheses?	Y
Was the expected or relative magnitude of correlations or mean differences included in the hypotheses?	Y
For convergent validity: was an adequate description provided of the comparator instruments?	Y
For convergent validity: were the measurement properties of the comparator instruments adequately described?	Y
Were there any important flaws in the design or methods of the study	N
Statistical methods	
Were design and statistical methods adequate for the hypotheses to be tested?	Y

Box G. Cross-cultural validity

Not included due to measure not being validated cross-culturally.

Box H. Criterion validity

Criterion	Yes/ No/ N/A
Design Requirements	
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the analysis adequate?	Y
Can the criterion used or employed be considered as a reasonable gold standard?	N
Were there any important flaws in the design or methods of the study	N
Statistical methods	
For continuous scores: were correlations, or the area under the receiving operating curve calculated?	N
For dichotomous scores: were sensitivity and specificity determined?	■

N.B. This study did not make use of a gold standard measure within this study to ensure a brief time for respondents to complete the battery.

Box I. Responsiveness

Not included due to measure not being used within in intervention in the current study.

Box J. Interpretability

Criterion	Yes/ No/ N/A
Was the percentage of missing items given?	Y
Was there a description of how missing items were handled?	Y
Was the sample size included in the analysis adequate?	Y
Was the distribution of the (total) scores in the study sample described?	Y
Was the percentage of the respondents who had the lowest possible (total) score described?	Y
Was the percentage of the respondents who had the highest possible (total) score described?	Y
Were scores and change scores (i.e. means and SD) presented for relevant (sub) groups? E.g. for normative groups, subgroups of patients, or the general population?	Y
Was the minimal important change (MIC) or the minimal important difference (MID) determined?	Y
Were there any important flaws in the design or methods of the study	N

Appendix H

Permutations of different assorted model fits

Strategy employed	Items deleted	Chi Square & significance	CFI	SRMR	RMSEA	AVE (R^2, hope, resilience)
Lowest regression weight against their respective factors deleted (PPOM-C)	Hope Q5, resilience Q2	340.95, $p < .001$	.904	.057	.114	.61, .62
Original 16-item PPOM	None	430.55, $p < .001$	.896	.060	.109	.58, .62
Highest regression weight against their respective factor deleted	Hope Q6 & resilience Q4	368.89, $p < .001$	.889	.056	.120	.61, .57
Lowest squared multiple correlations items deleted	Hope Q5 & resilience Q3	350.67, $p < .001$	.902	.056	.117	.61, .64
Highest squared multiple correlations items deleted	Hope Q1 & resilience Q5	333.55, $p < .001$	.897	.060	.113	.56, .61

Lowest standardized regression weight items deleted	Hope Q5 & resilience Q3	350.67, *p* < .001	.902	.056	.117	.61, .64
Highest standardized regression weight items deleted	Hope Q1 & resilience Q5	333.55, *p* < .001	.897	.060	.113	.56, .61
Lowest implied covariance items with factor deleted	Hope Q5 & resilience Q1	338.10, *p* < .001	.903	.058	.114	.61, .63
Highest implied covariance items with factor deleted	Hope Q6 & resilience Q4	368.89, *p* < .001	.889	.056	.120	.57, .61
Item correlating the lowest with respective factor deleted	Hope Q5 & resilience Q3	350.67, *p* < .001	.902	.056	.117	.61, .64
Item correlating the highest with respective factor deleted	Hope Q1 & resilience Q5	333.55, *p* < .001	.897	.060	.113	.56, .61