International Psychogeriatrics (2016), 28:9, 1471–1480 © International Psychogeriatric Association 2016
doi:10.1017/S1041610216000338

A biopsychosocial vignette for case conceptualization in dementia (VIG-Dem): development and pilot study

Aimee Spector,[1] Molly Hebditch,[2] Charlotte R. Stoner[1] and Luke Gibbor[3]

[1]*Department of Clinical, Educational and Health Psychology, University College London, 1–19 Torrington Place, London WC1E 7HB, UK*
[2]*Faculty of Health and Medical Sciences, Duke of Kent Building, University of Surrey, Guildford, Surrey, GU2 7XH, UK*
[3]*Clinical Trials Facility, Cognitive Impairment and Dementia Service, Lakeside Mental Health Unit, West Middlesex Hospital Site, Twickenham Road, Isleworth, TW7 6AF, UK*

ABSTRACT

Background: The ability to identify biological, social, and psychological issues for people with dementia is an important skill for healthcare professionals. Therefore, valid and reliable measures are needed to assess this ability. This study involves the development of a vignette style measure to capture the extent to which health professionals use "Biopsychosocial" thinking in dementia care (VIG-Dem), based on the framework of the model developed by Spector and Orrell (2010).

Methods: The development process consisted of Phase 1: Developing and refining the vignettes; Phase 2: Field testing ($N = 9$), and Phase 3: A pilot study to assess reliability and validity ($N = 131$).

Results: The VIG-Dem, consisting of two vignettes with open-ended questions and a standardized scoring scheme, was developed. Evidence for the good inter-rater reliability, convergent validity, and test–retest reliability were established.

Conclusions: The VIG-Dem has good psychometric properties and may provide a useful tool in dementia care research and practice.

Key words: vignette, biopsychosocial, dementia, reliability, validity, measure, outcome

Introduction

Conceptual models of dementia

Purely medical models of dementia (Lyman, 1989) explain symptoms and behavior in dementia as a result of organic change. These have been increasingly superseded by approaches that encompass psychosocial aspects (Kitwood, 1990; Downs *et al.*, 2008). Such models promote a more positive approach to dementia care by not simply ruling behavior or mood as an inevitable consequence of dementia but as something that may be amenable to change.

More recently, Spector and Orrell (2010) created a "Biopsychosocial model of dementia," an amalgamation of prominent theories and intended to be a useful clinical tool to help formulate interventions and identify potential issues in practice. The model recognizes that the level of actual functioning is a combination of the severity of impairment plus the interaction of a number of limiting factors, which restrict a person's potential level of functioning. The difference between "possible" and "actual" functioning can be thought of as "excess disability" that may be amenable to change with the consideration of the limiting factors that are "tractable." These tractable factors include psychosocial factors (such as mental stimulation, mood, and social psychology) and biological factors (such as physical health and sensory impairment).

Psychosocial outcome measures for care staff

An obstacle in the successful evaluation of training is inadequate outcome measures. Measures exploring constructs around person centered approaches or biopsychosocial factors are minimal (Spector *et al.*, 2012). Vignettes are descriptions of people or circumstances that can be real or fictional, and are used in research as a tool to investigate participant responses in relation to a construct or idea. They may be used to explore beliefs and attitudes or evaluate practice (Lapatin *et al.*, 2012) and have

Correspondence should be addressed to: Dr Aimee Spector, Department of Clinical, Educational and Health Psychology, University College London, 1–19 Torrington Place, London WC1E 7HB, UK. Phone: +0207-679-1844. Email: a.spector@ucl.ac.uk. Received 27 Sep 2015; revision requested 9 Dec 2015; revised version received 2 Feb 2016; accepted 3 Feb 2016. First published online 7 April 2016.

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

been shown to be one of the most effective ways of assessing clinician competence (Peabody *et al.*, 2000). This is due to a number of strengths: they are reasonably cheap and quick to administer, they provide a realistic case study from which to work that has no issues with ethics (Hughes and Huby, 2002), and they allow natural answers rather than forcing responses or testing passive knowledge.

Development of vignette-based measures

Whilst there is no established framework for developing vignettes, previous studies have outlined recommendations. Lapatin *et al.* (2012) broke the process down into a number of "iterative steps" including (1) identification and prioritization of content and format, (2) draft of vignettes by suitable team, and (3) reliability test and calibration process. Barter and Renold (1999) published an overview of recommendations: vignettes must be plausible, related to target audience, and contain sufficient depth, as well as considering format and relative rareness of examples with aims in mind. There are a number of approaches to scoring vignettes including qualitative analysis or use of a standardized scoring scheme, with the development of the scoring scheme largely depending on the aim of the measure. Standardized scoring structures can include asking for interpretation of the vignettes in the form of multiple choice answers (Kitchener and Jorm, 2002) or levels, for example one study awarded marks by rating formulations on six levels representing increasing depth and relevance (McClain *et al.*, 2004).

Current study

Whilst a number of measures for dementia care staff exist, none specifically tap into the skill of thinking about biological, psychological, and social factors, nor do they demonstrate one's application of knowledge to real world scenarios. This study aimed to create a biopsychosocial vignette for case conceptualization in dementia care and explore its inter-rater reliability, test–retest reliability and construct validity. For the purposes of this study, "thinking in a biopsychosocial way" refers to the ability to distinguish issues in dementia care beyond symptoms or biological factors, to create a wider picture of issues faced, and therefore potential actions.

Regarding construct validity, it was hypothesized that scores after training in the biopsychosocial model would be higher (over practice effects explored by test–retest with no intervention). Regarding convergent validity, it was hypothesized that those with more background experience, greater knowledge of dementia and more person centered and hopeful attitudes to dementia would score higher.

Methods

The process of development involved a number of sequential phases, each addressing different aspects of reliability and validity. Ethics were approved by University College London (UCL) ethics.

Phase 1: Vignette development

Review of existing measures

A review of previous vignette measures within dementia and vignettes in clinically related fields was conducted. This led to indications that the measure; (1) should be relatively brief in order to increase ease of application; (2) should include sufficient content to be able to draw multiple points with a variation in difficulty; (3) should present fairly typical issues, in order for healthcare professionals to relate them to current clients and problems; (4) should include examples from two different settings (care home and community setting) and contain different features, e.g. age and gender, which would increase generalizability.

Development of preliminary vignettes

Two fictional vignettes were created: "Mary" and "John" (see Figure 1). The biopsychosocial model (Spector and Orrell, 2010) was used to create a brief outline of possible problems, which were then built upon to create a coherent "story." Resources used at this stage included researchers own professional experience, and expert texts including the Camberwell Assessment of Needs for the Elderly (CANE) manual (Orrell and Hancock, 2004). For both cases, issues were presented in a way that would require varying levels of interpretation to identify them. Preliminary vignettes were presented to two psychologists and one dementia services manager. These were refined following their feedback.

Phase 2: Field testing

Aim

To check general stylistic issues and language suitability, and generate maximum content validity.

Procedure

Using convenience sampling, 12 people were approached and invited to complete the vignette measure online using Qualtrics (2001), an online

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

Case Study one: 'Mary'

Mary is 83 years old and has moderate Alzheimer's. Her husband died six years ago and she moved to a care home last year, having struggled to live independently. Mary left school at an early age, never worked and led a quiet life as a dedicated wife and mother to her two children. They both now live abroad. Mary rarely interacts with other residents, most of whom spend time in a large circle in the main lounge. There is a programme of entertainment in the home, yet Mary always refuses to join in if asked. When she first moved to the home, she joined a bingo session and burst into tears. She spends much of her time in the quiet lounge or walking in the hallway saying that she is 'trying to go home'. When in the lounge, Mary will often face the wall rather than the television. Staff have tried to engage her in photo albums or magazines, but she will only flick through them quickly. Recently, she has been found awake in the middle of the night. Staff report that she has been unusually resistant during personal care, grimacing and occasionally shouting at staff.

Question 1) What problems or *possible* issues can you identify for Mary?

Please list as many as you can think of. Keep each point brief and number them 1, 2,...etc.

Question 2) What potential actions would you suggest to address each of the problems or issues you have identified?

Please list as many as you can think of. Keep each point brief and number them 1, 2, ...etc

Case study two: 'John'

John, aged 62, suffered a stroke two years ago and was later diagnosed with Vascular Dementia. He lives at home with his wife Elaine. John was previously a mechanic who ran his own garage. He is no longer able to drive or work and has passed his business to his only son. John has always been a physically active man who enjoyed gardening, DIY and cars. John has increasingly spent more time watching TV and comfort eating. Recently, John went to a car rally with some friends, where he experienced 'angry outbursts'. He has not been out since. Elaine has tried to take the strain off John; hiring a gardener as he is unsteady on his feet and getting their son to help out with admin and practical jobs. Elaine reports that John often ignores the world around him, for example, he rarely answers the phone and frequently 'doesn't listen to her'. She feels that he does not appreciate her help and is always angry, leading to frequent arguments.

Question 1 and 2 as above.

Figure 1. A Biopsychosocial vignette for case conceptualization in dementia; an extract.

survey platform. Participants were additionally asked for basic demographic data and to provide feedback (in the form of open ended questions with prompts, for example, "how realistic did you find these case studies?").

Outcome

Nine people completed the exercise; one community psychiatric nurse (CPN), one dementia adviser, one postdoctoral researcher, one registered nurse, one senior clinical lecturer, two trainee clinical psychologists, and two research assistants. Feedback from responses was positive; with most people indicating that they felt the exercise was "realistic." Those who expressed that they had found the exercise difficult related this to their lack of experience within a particular setting or client group.

The content and instructions of the vignette were refined according to feedback. First, the answers generated indicated that all of the factors were identifiable because they were each referred to at least once; however they were not consistently identified by each respondent. This suggested that there might be variation in the ease of reporting or interpreting factors within the vignettes, resulting in adequate variation in responses. In terms of content, it was suggested that the vignettes could be made more succinct to increase readability. This was done with minor alterations to the issues included. Small changes to the instructions to add clarity to questions were included, for example "Please list as many as you can think of" was added.

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

Phase 3: Pilot

Procedure

A convenience sample was generated using established connections with health professionals known to researchers. This included current colleagues, those presently attending training and previous attendees of a training course who had given permission to be contacted about future research. Inclusion criteria were that people (1) currently worked with people with dementia and (2) could speak English. Potential participants were approached via email or in person to complete the exercise.

Measures

The VIG-Dem consisted of two small fictional case studies named "John" and "Mary" of approximately 200 words each. They were both preceded by a set of instructions and each of the vignettes were presented in turn and accompanied by two questions. The aim of the questions was to allow participants to demonstrate their knowledge by interpreting the vignettes in terms of "issues" they may contain and suggesting appropriate "interventions." The final scoring scheme structure contained a list of potential items that could be awarded either one or two marks depending on the level of interpretation and could be summed to give an overall score with a maximum score of 56. These items are arranged into four sections "part one: issues" and "part two: interventions" for both "Mary" and "John" (see Figures 2 and 3).

Participants were given the option of completing it as a paper-based measure or online via Qualtrics (2001). The content was identical in each format. The word "Biopsychosocial" was deliberately omitted from all materials so that respondents would not be cued into eliciting a particular style of response.

The Approaches to Dementia Questionnaire (ADQ) (Lintern *et al.*, 2000) is a 19 item rating scale that was developed to indicate positive, person centered attitudes of care staff. It contains two sub scales, "Hope" and "Person-centeredness." Statements are rated on a 5-point likert rating scale ranging from one (strongly agree) to five (strongly disagree), with higher scores indicating higher hopeful and person centered approaches. Previous validation of this scale has shown good reliability (Cronbach's $\alpha > 0.80$) and validity. The ADQ was included to provide information on convergent validity.

The DK-20 (Shanahan *et al.*, 2013) was added after an interim analysis, to further explore convergent validity. Hence, DK-20 scores are only available for a sub-section of participants ($N = 43$). It consists of 20, five-option multiple-choice questions on a mixture of dementia knowledge and care attitudes. The measure has two subscales, "Dementia Core Knowledge" and "Dementia Care Knowledge." Moderate reliability has been found for this measure (Cronbach's $\alpha > 0.60$).

General demographics (age and gender) and background variables were recorded. This included number of years working with people with dementia, name of current working role, current and previous types of work settings, and highest level of education.

Training effects and test–retest reliability

A sub-sample ($N = 30$) included those who were attending training specifically on the biopsychosocial model. They were asked to complete the measure at two time points; before and after training, with an interval of approximately five weeks. This was to explore the measure's sensitivity to change following training. In order to assess test–retest reliability a further sample ($N = 10$) agreed to complete the measure at a second time point approximately five weeks after initially completing the exercise. These participants were included to assess consistency of the measure over time.

Scoring scheme development

When developing the scoring scheme, there were specific aims: achieving good inter-rater reliability and ensuring that it is user friendly, that is, not too long or difficult to mark. The process involved several steps: (i) a brief heading was created for each of the initial points that were designed to be included in (and therefore could be pulled out of) the vignettes; (ii) each pilot answer was explored by putting each statement under the most appropriate heading and a new heading created if it did not fit (i.e. expressed a new appropriate issue or intervention). This was conducted with all answers from the field test and continued with responses from the main pilot ($N = 5$), until it was felt that all possible *relevant* answers had been represented; (iii) the list of possible items were explored in terms of lower level observations and interpretations and then regrouped into meaningful categories (using the biopsychosocial factors) to form a provisional structured marking frame; (iv) two alternative marking frame structures were explored: hierarchical (i.e. levels awarded on depth and relevance) and accumulative (i.e. marks awarded for specific content). An accumulative structure was chosen, as it was the easiest to implement on the types of response generated by this exercise;

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

MARY **ISSUES IDENTIFIED**	
Instructions 1) Tick each box for each item that is covered in their answer. 2) Do not tick more than one box for a single statement. If one statement seems related to two boxes, tick only one box that you think is the most relevant. 3) You cannot tick a box more than once 4) Square boxes (☐) warrant one mark and circular boxes (◯) warrant two marks. 5) The total possible marks available for each section are listed in the right hand column. Markers should add these to create a 'total' score.	**Marks**
Mental Stimulation ☐ Identifies a lack of activities/ mental stimulation/ boredom ◯ Elaborates on perhaps why e.g. activities not suitable due to personal history; 'quiet life'/ limited education/ interests	**/3**
Social Psychology ☐ Identifies lack of family contact ◯ Elaborates on significance; loneliness/ isolation/ lack of support.	**/3**
Reaction to life events ☐ Notes loss of husband ☐ Notes move to care home/ unfamiliar environment.	**/2**
Mood ☐ Identifies sleeping difficulties ◯ Identifies possible low mood/ depression/ anxiety.	**/3**
Personal Psychology ☐ Notes disorientation/ distress at 'looking for home' ◯ Mentions Mary may feel a lack of role/ loss of identity/ loss of independence.	**/3**
Physical Health ☐ Identifies discomfort at personal care ◯ Suggests need to explore cause/ gives possible explanation/ asks questions Identifies possible pain.	**/3**
INTERVENTIONS IDENTIFIED	
Instructions in addition to above 1) Participants should have linked their 'intervention' to an 'issue' using numbering. 2) The same 'intervention' listed more than once under a different number should not be considered a new statement. 3) If a new 'issue' (not previously stated) is listed in the intervention section, you may go back and award marks within the 'issues' section.	**Marks**
Psychosocial Interventions ***Engagement in activities*** ☐ Suggests engagement in activities/ a general comment of initiating/ encouraging more activity ☐ Specific suggestion of an appropriate activity or ways to encourage/ facilitate inclusion (including a cognitive intervention) ◯ These suggestions are person-centred i.e. drawing on previous history/ including Mary in the process.	**/4**
Social contact ☐ Suggests facilitating family contact or ways to include them in her life ☐ Suggests a different approach to bringing meaningful social interactions with staff/ residents/ friends (e.g. one-to-one, smaller groups).	**/2**
Exploration of behaviour ◯ Suggests possible strategies to explore behaviour (e.g. staff training/ asking Mary what is wrong).	**/2**
Practical support ☐ Proposes practical suggestions / strategies to cope with moving into new home (e.g. signs/ making it familiar/ more choice).	**/1**
Emotional support ☐ Suggests that Mary has someone to talk to about her grief or changes in life.	**/1**
Biological Interventions ***Medical assessments*** ☐ Suggests an assessment from a healthcare professional for possible mood/ behaviour issues ☐ Suggests possible medication/ review of medication/ treatment (e.g. talking therapy/ psychosocial intervention).	**/2**
Resistance to care ☐ Suggests medical assessment for location and reason of pain	**/1**
TOTAL SCORE FOR MARY:	**/32**

Figure 2. Mary scoring scheme.

(v) this scoring scheme was used to mark three pilot responses by the researcher and a second marker, in order to check the usability of the scheme as well as indicate potential inter-rater reliability; (vi) discrepancies between marks, as well as ambiguity in its interpretation were discussed and clarified, resulting in a second revision. The finalized scoring scheme was used by the researcher to mark all participant responses. Twenty-five responses were then re-marked by a second rater

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

JOHN ISSUES IDENTIFIED	
Instructions 1) Tick each box for each item that is covered in their answer. 2) Do not tick more than one box for a single statement. If one statement seems related to two boxes, tick only one box that you think is the most relevant. 3) You cannot tick a box more than once 4) Square boxes (☐) warrant one mark and circular boxes (◯) warrant two marks. 5) The total possible marks available for each section are listed in the right hand column. Markers should add these to create a 'total' score.	**Marks**
Mental Stimulation ☐ Identifies a lack of activities/ mental stimulation/ boredom.	**/1**
Social Psychology ☐ Notes deterioration in relationship with wife ◯ Provides further explanation e.g. why this may be happening/ Elaine unable to manage/ understand. ☐ Identifies a lack of other social contact.	**/4**
Reaction to life events ☐ Mentions significance of non-specific life changes.	**/1**
Mood ☐ Identifies possible low mood/ depression ☐ Identifies anger/ frustration.	**/2**
Personal Psychology ☐ Notes lack of motivation/ withdrawn from activities e.g. not leaving the house ◯ Mentions that John may feel a lack of role/ loss of identity/ loss of independence.	**/3**
Physical Health ☐ Identifies that John has mobility problems	**/1**
INTERVENTIONS IDENTIFIED	
Instructions in addition to above 1) Participants should have linked their 'intervention' to an 'issue' using numbering. 2) The same 'intervention' listed more than once under a different number should not be considered a new statement. 3) If a new 'issue' (not previously stated) is listed in the intervention section, you may go back and award marks within the 'issues' section.	**Marks**
Psychosocial Interventions ***Engagement in activities*** ☐ Suggests engagement in activities/ a general comment of initiating/ encouraging more activity ☐ Specific suggestion of an appropriate activity or ways to encourage/ facilitate inclusion (including a cognitive intervention) ◯ These suggestions are person-centred i.e. drawing on previous history/ includes John in the process.	**/4**
Social contact ☐ Suggests increasing social contact e.g. connections with old friends/ ways to initiate new social interaction.	**/1**
Interventions for wife ☐ Suggests education/ information for wife ☐ Suggests emotional/ practical support for wife	**/2**
Emotional support for John ☐ Suggests support to help with adjustment for the changes that John is experiencing ◯ Suggests helping John to build self-esteem/ support to regain independence/ regain a role	**/3**
Biological Interventions ***Medical assessments*** ☐ Suggests an assessment from a healthcare professional for possible mood/ behaviour issues ☐ Suggests possible medication/ review of medication/ treatment (e.g. talking therapy/ psychosocial intervention).	**/2**
TOTAL SCORE FOR JOHN:	**/24**

Figure 3. John scoring scheme.

to explore inter-rater reliability. All researchers were blind to participant characteristics at the time of marking; (vii) the initial scoring scheme for the vignette measure had a possible range from 0–113, with participant total Scores ranging from 6–62 with a mean of 25.75 (SD 10.36). Coverage of this scoring scheme was 49% for these participants. Therefore, the scoring scheme was further developed by analyzing possible redundant items i.e. the items that were rarely awarded marks on the scoring scheme reflecting the most difficult points to identify in the vignettes. It was decided that by using an 80% cut-off, the resulting scoring scheme would have greater utility in that it would be quicker to mark, allow greater coverage and produce a normal distribution.

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

Table 1. Participant demographics ($N = 131$)

DEMOGRAPHICS AND BACKGROUND STATISTICS	N(%) OR MEAN (STANDARD DEVIATION, RANGE)
Gender ($N = 129$):	
Male	16 (13%)
Females	113(87%)
Age ($N = 130$)	
18–24	9(7%)
25–34	30(23%)
35–44	26(20%)
45–54	44(34%)
55+	21(16%)
Years experience ($N = 125$)	9.86 (SD = 8.60)
Current role ($N = 130$):	
General nurse	4(3%)
Mental health nurse	27(21%)
Student nurse	4(3%)
OT	13(10%)
Clinical psychologist	2(1%)
Trainee clinical psychologist	10(8%)
Physiotherapist	3(2%)
Speech and language therapist	6(5%)
Team managers/coordinators	26(20%)
Social worker	3(2%)
Support worker	15(11%)
Healthcare assistant	10(8%)
Dementia worker (unqualified)	7(6%)
Number of settings ($N = 100$)	1.93 (SD = 1.21)
Highest level of education ($N = 120$)	
Doctorate	4(3%)
Master's degree	15(13%)
Undergraduate degree	54(45%)
Diploma/professional certificate	37(31%)
None of the above	10(8%)

Results

One-hundred-thirty-one completed responses were collected; the demographics of this sample can be seen in Table 1. Due to convenience sampling, used response rates were not possible to calculate.

Data were inputted, coded, and analyzed on SPSS version 21.0 and an alpha level of 0.05 was used for all statistical tests.

Scoring properties

The scoring scheme for the vignette measure had a possible range from 0–56, with participant total ccores ranging from 4–42 with a mean of 25.8 (SD 10.36). A Shapiro–Wilk test indicated that the distribution of total scores was normal (W (131) = 0.990, $p = 0.426$.

Content validity

This was ensured by a rigorous development process. Spector and Orrell (2010) conducted a thorough literature review to create a biopsychosocial model summarizing factors in dementia care and therefore by using this as a framework it was ensured that a wide range of biological, social, and psychological factors were considered. The use of expert opinions and feedback from healthcare professionals ensured its applicability to real world practice.

Face validity

Targeted questions were asked during field-testing to assess people's perceptions of the exercise, while general comments were sort throughout. Field-testing indicated that people felt the vignettes were "realistic" and had positive attitudes to the exercise. Further comments throughout the study continued to show this trend; one participant commented "it was useful doing the exercise and illustrated that I need to brush up my knowledge in some areas." Another commented "I enjoyed the exercise as it's the type of work I do."

Effect of format

The 69 participants who completed the exercise online had a mean score of 23.2 (SD 7.08) and the 62 who completed hard copies had a mean score of 17.2 (SD 7.31). The difference between total scores (Sqrt) for these groups was significant (t (129) = 4.77, $p < 0.001$) indicating that those completing the exercise online scored higher on average.

Internal consistency

The mean score for "Mary" was 11.5 (SD 4.76) and "John" was 8.8 (SD 3.74). Spearman's correlation showed that scores for the two vignettes had a strong, significant correlation (ρ (129) = 0.68, $p <$ 0.001). Cronbach's α coefficient was 0.73 for the 44 items on the scoring scheme indicating good internal consistency.

Inter-rater reliability

The intra-class correlations indicated substantial inter-rater reliability looking at absolute agreement for total score for both data sets marked by two raters ($N = 25$, ICC(2,k) = 0.94, $p < 0.001$).

Convergent validity: existing measures

Total scores (sqrt) had a positive significant correlation with the ADQ total score (ρ (127) = 0.37, $p < 0.001$), Hope sub-scale (ρ (127) = 0.38, $p < 0.001$), and Person-centeredness sub-scale (ρ (128) = 0.27, $p = 0.002$). For the sub-sample

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

of DK-20 data, total scores (sqrt) had a positive significant correlation with the DK-20 total score (ρ (41) = 0.49, p= 0.001), Care Knowledge sub-scale (ρ (41) = 0.45, p = 0.002), and the Core Knowledge sub-scale (ρ (41) = 0.46, p = 0.002).

Background characteristics

The different roles of the healthcare professionals were recoded into 13 categories. Analysis of variance showed that the differences between Roles were significant for total scores (sqrt) (F(12,117) = 3.55, p = 0.001). *Post hoc* comparisons using the Tukey HSD test indicated that occupational therapists (M= 26, SD = 8.27) scored significantly higher than support workers (M = 15.6, SD = 7.34) and healthcare assistants (M = 12.3, SD = 6.73). Trainee psychologists (M = 24.9, SD = 6.84) scored higher than healthcare assistants; dementia workers (M = 25.9, SD = 8.45) scored better then support workers and healthcare assistants. Analysis of variance revealed a significant difference between levels of education for total scores (sqrt) (F(4,115) = 3.33, p = 0.01). *Post hoc* Tukey HSD test did not reveal significant relationships between these groups. Spearman's correlations found that total score was not significantly correlated to the number of years experience working in dementia care (ρ (123) = 0.10, p = 0.26), nor number of settings worked in (ρ (98) = 0.18, p = 0.07).

Construct validity: sensitivity to change.

There was a significant difference between total scores (sqrt) before and after the biopsychosocial learning intervention (t (29) = 3.57, p < 0.001) with post-test scores on average 5.8 points higher (SD = 9.29).

Test–retest reliability

The mean total score for time point one was 26.3 and 25.7 for time point two. The intra-class correlations indicated excellent test–retest reliability for total Score between time points one and two (N = 10, ICC(2,k) = 0.96, p < 0.001).

Discussion

Summary of findings

This study has described the development of a vignette exercise and scoring scheme and assessed its reliability and validity as an outcome measure. The scores achieved in this study had an adequate range of 38 covering 67% of the possible scheme. Scores were skewed to the lower end of the scale; this could potentially be an artifact of this sample not possessing the knowledge to reach the higher scores, but is more likely to be a consequence of the scoring scheme design. Previous studies have shown that when scoring schemes are based on a comparison to a model answer coverage may be low (Dudley *et al.*, 2010). For example, Abbas *et al.* (2012) found participant scores covered only 40% of the total possible range when scoring formulations against model answers even after training. Good internal consistency for the scoring scheme was established and the scores for each of the two vignettes were significantly correlated. This suggests that the quality of answers generated by the different vignettes do not vary greatly. There was good inter-rater reliability between two raters.

A number of indicators for convergent validity were found; two existing outcome measures, the ADQ and DK-20 were shown to have a strong positive relationship with vignette scores.

Background variables were found to have significant relationships with participant scores in the expected direction; total scores were on average higher for those with more senior roles. This would be expected as an increased knowledge base of working with people with dementia, which they would be able to apply during this exercise. This adds support that the measure is testing a knowledge-based construct.

This study also identified that it is sensitive to change in knowledge as subjects scored higher on the measure after completing a training course. Positive test–retest findings suggest this is due to learning rather than practice effects.

Experience in years and number of settings worked in did not correlate with total score. Possible explanations are (1) that those with more experience may have less motivation to "test" their knowledge or (2) that some healthcare professionals who received training years ago may have been exposed to medical models and be less familiar with biopsychosocial approaches. To support this idea, Zimmerman *et al.* (2005) showed that person centered and hopeful attitudes were associated with working fewer (0–2 years) rather than more years in the field.

The format of response may have influenced scores, with those completing the questionnaire online having significantly higher scores on average. The response rate for those replying online from a database was low and hence it is possible that those replying online had a stronger motivation to complete the task. Online completers also did so in a non-controlled environment where there could have been a number of extraneous variables; time to complete task, using reference texts, or discussion amongst colleagues.

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

Strengths and limitations

Several limitations are of note. First, a rigorous assessment of test–retest reliability was beyond the scope of this study resulting in a limited sample. While excellent reliability over time was indicated, further data collection would strengthen this conclusion. Another potential limitation was time restraints: participants commented that there was not enough time, and in some cases, it was evident that they had taken longer than requested (both on online and hard copies). Therefore, time is a potential confounding variable that was not controlled in this study. It may be that the time given to participants was simply not adequate for them to complete the exercise to a standard that represents their ability, and that increasing the time available to participants may lead to higher scores, and therefore increased coverage of the scoring scheme.

There may have been biases within the sample. Demographics of non-completers were not recorded and therefore it is not known whether there were any systematic differences between those who did and did not respond to the invitation to take part. Although the sample was predominantly female, this is a reasonable representation of the gender mix within healthcare professionals.

Implications for future directions

Whilst results suggest that the measure may be used in its current format, potential revisions are also indicated. While the scoring scheme generates a standardized score, the generated answers have value in themselves and could be explored in different ways. For instance, qualitative questions could also be included as a way of indicating certain attitudes or approaches, for example "does this response take Mary or John's wishes in to account?" Further work on the scoring scheme could extend its use.

It was originally felt that using two vignettes with different backgrounds would increase generalizability. However, as the scores for "Mary" and "John" were significantly correlated, it may be possible that the measure could be shortened by using only one vignette at one time. This could (1) allow participants more time to focus on their answer, and (2) enable each vignette to be used at a different time point, thus reducing practice effects. To establish feasibility of this approach, a number of investigations would be needed including identification of why scores for "Mary" were consistently higher. This may be the result of it being presented first, or due to easier content or scoring.

Implications for clinical practice

As this measure employs a standardized scoring scheme, it enables the assessment between and within individuals over time. The ability to approach dementia care from a biopsychosocial framework is an increasingly valued skill and is important in order to reach an acceptable standard of quality care. This measure could be used to indicate healthcare professionals' level of understanding, for example, when recruiting staff to different roles. It could help establish adequate staff approaches to dementia care and highlight training needs. It has the potential application as an outcome measure for assessing training or other interventions for care workers aimed at increasing a holistic approach. It may also offer researchers a tool to explore other constructs related to biopsychosocial approaches in healthcare staff. As no current measure exists to explore this particular aspect of healthcare knowledge, it could have a number of important functions.

Conclusion

This study has developed what is known to be the first vignette-based measure with a standardized scoring scheme, to assess "Biopsychosocial thinking" in dementia care. Preliminary findings for reliability and validity of the measure show promise of establishing a useful and credible tool. This measure may play a useful part in assessing competence and the impact of interventions.

Conflict of interest

None.

Description of authors' roles

Aimee Spector conceived of the idea, oversaw the whole program of research, supervised both research projects and helped to write/commented on drafts of the paper. Molly Hebditch designed the vignette and collected and analyzed data for a large proportion of the sample, for her MSc research project. She was involved in data entry, analysis, and writing drafts of the paper. Luke Gibbor collected and analyzed further data for his BSc research project. He was involved in data entry, analysis, and writing drafts of the paper. Charlotte Stoner helped devise the new scoring system and rescored and reanalyzed the full dataset, using this scoring system. She has also commented on the final draft of the paper.

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press

Acknowledgments

The authors would like to thank the London Alzheimer's Society, Richmond upon Thames Memory Services for their support in undertaking this research. Also, to residents at Ash Court Nursing Home and Clara Nehab House and the managers of these homes: Steven Anderson and Grace Ramalho, respectively.

References

Abbas, M., Walton, R., Johnston, A. and Chikoore, M. (2012). Evaluation of teaching an integrated case formulation approach on the quality of case formulations: randomised controlled trial. *Psychiatric Bulletin*, 36, 140–145.

Barter, C. and Renold, E. (1999). The use of vignettes in qualitative research. *Social Research Update*, 25, 1–6.

Downs, M., Clare, L. and Anderson, E. (2008). Dementia as a biopsychosocial condition: implications for practice and research. In R. Woods and L. Clare (eds.), *Handbook of the Clinical Psychology of Ageing*, 2nd edn. (pp. 145–160). Chichester: Wiley.

Dudley, R., Park, I., James, I. A. and Dodgson, G. (2010). Rate of agreement between clinicians on the content of a cognitive formulation of delusional beliefs: the effect of qualifications and experience. *Behavioural and Cognitive Psychotherapy*, 38, 185–200.

Hughes, R. and Huby, M. (2002). The application of vignettes in social and nursing research. *Journal of Advanced Nursing*, 37, 382–86.

Kitchener, B. A. and Jorm, A. F. (2002). Mental health first aid training for the public: evaluation of effects on knowledge, attitudes and helping behaviour. *BioMed Central Psychiatry*, 2, 10.

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

Lapatin, S. *et al.* (2012). Lessons from the use of vignettes in the study of mental health service disparities. *Health Services Research*, 47, 1345–1362.

Lintern, T., Woods, B. and Phair, L. (2000). Before and after training: a case study of intervention. *Journal of Dementia Care*, 8, 15–17.

Lyman, K. A. (1989). Bringing the social back in: a critique of the biomedicalization of dementia. *Gerontologist*, 29, 597–605.

McClain, T., O'Sullivan, P. S. and Clardy, J. A. (2004). Biopsychosocial formulation: recognizing educational shortcomings. *Academic Psychiatry*; 28, 88–94.

Orrell, M. and Hancock, G. A. (2004). *The Camberwell Assessment of Need for the Elderly (CANE)*. London: Gaskell.

Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R. and Lee, M. (2000). Comparison of vignettes, standardized patients, and chart abstraction. *Journal of the American Medical Association*, 283, 1715–1722.

Qualtrics. (2001). *Qualtrics (Version 47059). [Software]*. Provo, UT: Qualtrics. Available at: http://www.qualtrics.com; last accessed 22 March 2016.

Shanahan, N., Orrell, M., Schepers, A. K. and Spector, A. (2013). The development and evaluation of the DK-20: a knowledge of dementia measure. *International Psychogeriatrics*, 25, 1–9.

Spector, A. and Orrell, M. (2010). Using a biopsychosocial model of dementia as a tool to guide clinical practice. *International Psychogeriatrics*, 22, 957–965.

Spector, A., Orrell, M., Schepers, A. and Shanahan, N. (2012). A systematic review of "knowledge of dementia" outcome measures. *Ageing Research Reviews*, 11, 67–77.

Zimmerman, S. *et al.* (2005). Attitudes, stress, and satisfaction of staff who care for residents with dementia. *Gerontologist*, 45, 96–105.

https://doi.org/10.1017/S1041610216000338 Published online by Cambridge University Press