Title: Dementia Knowledge Assessment Scale (DKAS): Confirmatory factor analysis and comparative subscale scores among an international cohort.

Authors:

Michael Annear (michael.annear@gmail.com)
Christine Toye (C.Toyex@curtin.edu.au)
Kate-Ellen Elliott (KateEllen.Elliottx@utas.edu.au)
Frances McInerney (Fran.McInerneyx@utas.edu.au)
Claire Eccleston (Claire.Ecclestonx@utas.edu.au)
Andrew Robinson (Andrew.Robinsonx@utas.edu.au)

Version: 1 Date: 26 Apr 2017

Author’s response to reviews:

27 April 2017

Re: Response letter to BMC Geriatrics Editorial staff (manuscript BGTC-D-16-00488)

To whom it may concern

Thank you for the opportunity to make revisions to our manuscript: "Dementia Knowledge Assessment Scale (DKAS): Confirmatory factor analysis and comparative subscale scores among an international cohort" (BGTC-D-16-00488). Changes made (additions, deletions, and moved text) are highlighted within the manuscript using the track changes function of Windows. Below, we have also provided a point-by-point response to the reviewer comments. Please don’t hesitate to contact us if you have any further questions or comments.

Yours sincerely

Dr Michael Annear (on behalf of all the authors)
Reviewer 1 comments

Reviewer comment 1: The paper attempts to establish the psychometric validity of a dementia knowledge questionnaire. Overall it is well structured and well written, with an adequate justification as to why the study may be of interest. There are, however, some comments on confirmatory factor analysis which are at least debateable but these largely stem from the previous use of principal component analysis. In the view of many PCA is not a good starting point for analysis and also there are good reasons to believe that its results will not always be supported by CFA. I think the authors should acknowledge this, even though in this case the CFA largely supports the PCA.

Author response 1: As reviewer one acknowledges the use of PCA as a hypothesis-generating mechanism to inform CFA is a subject of debate. An argument for the use of PCA as a preliminary, hypothesis generating approach is that PCA is an emergent approach to factor/component generation with a focus on item reduction. PCA is not theoretically driven as Exploratory Factor Analysis (EFA) is (Tabachnick and Fiddell, 2014). We did not begin the development of the DKAS with a-priori theoretical assumptions about which factors/components would emerge or which items would load on factors within the scale. Our intention was to develop a reliable and valid measure that addresses a diversity of evidence-informed knowledge parameters about dementia. The content of the measure (item statements) were not known in advance, but developed out of a literature review and international Delphi study published in BMC Geriatrics (Annear et al., 2015). Data reduction was important as the initial analysis began with a set of 57 items, so PCA was preferred for its data reduction capacity. Moreover, results of PCA are frequently identical or very similar to those obtained via EFA (Field, 2013), so there is little reason to believe that an EFA would have resulted in a different hypothesised factor structure to be tested via subsequent CFA (the subject of the present paper). Indeed, the present CFA results showed good support for the findings of the prior PCA. It should also be noted that the DKAS development process that was employed in this paper has precedent in the international gerontological and methodological literature where others have also employed PCA as a justifiable and preliminary step prior to CFA in scale development procedures. See the manuscripts listed below, which followed this analytical process. The exploratory PCA referred to in the manuscript has also been peer-reviewed and published in the Journal of the American Geriatrics Society (Annear et al., 2015) and forms the hypothesised model for testing via a CFA. In order to address the reviewer’s comment, we have added text in the methods section to address the debate in the literature and provide additional rationale for the utilisation of PCA as a foundational step (see page 9, lines 163-171). Exemplar manuscripts concerning research studies that have employed PCA as a preliminary step in scale development with subsequent CFA procedures are outlined below.


Reviewer comment 2: I would also argue that instead of considering a four correlated factors solution it makes sense to also consider a bi-factor solution in which there is a general factor on which all items load and then four other factors. This should further be supported by the calculation of omega and omega h for the factors rather than the use of Cronbach's alpha. See McDonald and Revelle among others for reasons for using omega rather than Cronbach.

Author response 2: A number of possible factor solutions were explored in a previously published principal components analysis (PCA) as part of preliminary exploratory analyses. We accept that there is an overarching construct that our data fit, which has previously been conceptualised as ‘dementia knowledge’, but our primary interest in the present analysis is in understanding fine-grained differences in knowledge that can be drawn out at the domain or factor level. Hence, the elucidation of subscales through the CFA was the primary concern. Our data suggest that the DKAS holds together well as a scale and supports four constituent subscales. Additional factorisation would, we contend, be an unnecessary step that was not indicated by our earlier analysis and hypothesis. We have elected not to pursue a bi-factor model for several reasons that have been addressed in the literature. In particular, we have concerns that a bi-factor model may be overly restrictive and fail to accurately reflect the structure of item response data in the population (Reise, 2012). Furthermore, recent literature concerning the use of bi-factor models suggests that the practice is fraught with potential difficulty: “the potential parameter distorting effects of forcing small cross-loadings to zero, and accommodating items with substantial cross-loadings on group factors, remains an issue in the use of bi-factorisation in confirmatory models (see also Finch, 2011)” (Reise, 2012). Additionally, there are clear distinctions between the four factors, which are readily interpretable. It is questionable what value a common factor would add when items are highly varied across factors, with statements addressing the underlying pathology of dementia and care requirements, for example. We, therefore, consider the four factors identified to be more than residuals to a common factor and
the exposition of a common factor to be a largely redundant technical exercise in the context of this analysis.

In order to comprehensively address the reviewer’s query, we investigated the potential of a common factor within AMOS for SPSS. We tested a model wherein a common factor was hypothesised to be correlated with each of the scale items and inter-correlated with the four previously identified factors. Upon analysis, however, we found this solution to be unsatisfactory. The solution resulted in a reduction of GFI from .974 to .926 and an increase in RMSEA .040 to .067 (where values above .06 are considered unacceptable). Thus, it appears that estimation of a common factor reduces model fit to an unacceptable degree and cannot be accommodated in the present analysis.

While the use of Cronbach’s alpha is conventional and widely supported in the international literature concerning reports of scale reliability, we have also calculated McDonald’s Omega for the total score, which reinforces the acceptable reliability (ωh = .87) of the overall scale and aligns with the alpha score. It should be noted, however, that it is not possible to calculate Omega statistics using AMOS for SPSS – our primary package for SEM and CFA. We have had to use R to generate this statistic. Moreover, we have not calculated omega statistics for subscales as this would require an entirely new CFA analysis within R, which we view as an unnecessary step when we believe that we have satisfactorily performed a CFA using the widely used AMOS package for SPSS. The reporting of CFA results certainly conforms to international norms and other published reports.

Reviewer comment 3: I think it also makes sense to make it easier to see what the items are and their pattern of correlation. Perhaps it is only in my version but there were very sporadic details of both the scoring of the items and their content. There was more detail on items that did not fit than those that did.

Author response 3: We have added an appendix to the end of the manuscript, which shows the content of each item, its relationship to the domain confirmed via the CFA procedure, and the standardised factor coefficients/loadings (See Appendix 1; pages 32-33).

Reviewer comment 4: Lastly I think it would be useful to have detail on the breakdown of gender and dementia care, so that we can see how representative the sample is, and this might be true for other classifications.

Author response 4: We have included the gender of participants in Table 1. In addition to the gender of the overall sample, we have also added statistics for the main respondent groups (see Table 1, page 12). We have further addressed the issue of the gendered nature of dementia and
care provision in the limitations section of the paper (see page 21, lines 386-395). We acknowledge that female representation is considerably higher than male representation in this study, although this is consistent with reported composition of professional and familial carers for older adults with dementia. In this way, we believe that our sample provides representation of a heavily gendered profession and care type.

Reviewer 2 comments

Reviewer comment 1: This was an enjoyable and clear paper. I have several observations and suggestions (below). I think an excellent addition would be to look at the strength of a common factor in a bi-factor model approach.

Author response 1: As per our response to the first reviewer, the purpose of a CFA is to confirm or refute a hypothesised model based on prior analysis, not to explore other potential factorisations. Moreover, we do not believe that the addition of a common factor will add to the utility and interpretability of the scales in any practical sense. Moreover, after undertaking an analysis of a common factor at the reviewer’s request, the results indicated poor model fit. See response to reviewer one above (comment 2, response 2).

Reviewer comment 2: Removing the depression item to meet the 0.80 correlation threshold felt like a weak-spot to me so could be justified further.

Author response 2: All decisions made in the survey development process were based on established statistical criteria. Decisions to retain or remove items during scale development were not, and should not, be based on subjective consideration. We have added further justification for the removal of the depression item within the text and have acknowledged the tension in scale development between conceptual and statistical integrity (see pages 17-18, line 307-317). To clarify, the decision to remove the depression-related item (people with dementia are unlikely to experience depression) was undertaken to improve the model summary statistics (GFI and RMSEA) as well as improving the correlation between factors based on established criteria. The decision was not taken lightly and was made in consultation with a team of dementia experts. It should also be noted that there is another item within the measure that also addresses depression (symptoms of depression can be mistaken for symptoms of dementia), although we acknowledge that it is slightly different in composition.

Reviewer comment 3: Abstract: what are the two alpha values - it's not clear.
Author response 3: We have added additional information in the abstract to clarify what each of the Alpha values refer to (see track changes in revision abstract).

Reviewer comment 4: Abstract: alpha of 0.65 is low. 0.7 is usual minimum value of alpha seen.

Author response 4: The alpha of .65 for the final subscale is indeed lower the often-quoted criterion value of 0.7 for internal reliability. However, we consider that this value is approaching the 0.7 criterion and suggest that it may achieve 0.7 when further testing is undertaken with large, random samples. DeVellis (2016) has noted previously that α values between 0.65 and 0.69 can be considered as minimally acceptable during scale development, and Kline (1999) suggests that values below 0.7 are often expected when developing scales in social and psychological research due to the diversity of constructs. Further testing is also required to assess subscale reliability for individual cohorts as part of future analyses. We have addressed this in the limitations section of the manuscript (see page 20, lines 383-386).


Reviewer comment 5: p6: Consider adding a box with a brief overview of the DKAS, intended domains and items in each.

Author response 5: We have added a summary of DKAS items, domains, and factor loadings as an appendix (see Appendix 1).

Reviewer comment 6: p6: I may have missed this, but were these all true/false questions? Multiple choice?

Author response 6: Dementia knowledge items in the DKAS are presented with a Likert-type scale that allows respondents to rate how truthful they believe each statement to be. Possible responses include: false, probably false, probably true, true, and I don’t know. Each statement was developed from a detailed literature review and Delphi study with international dementia experts (previously published in BMC Geriatrics; Annear et al., 2015), so each statement can be considered to be factually correct or incorrect based on current best evidence. To address the reviewer’s comment, we have added clarification to the methods section about the design of the scale (see page 8, lines 146-150). We have also added more information about items in Appendix 1.

Reviewer comment 7: p6: I thought the critique of existing measures should come before the description of the DKAS development (para beginning line 106). The logical flow is that need for the new measure comes first.

Author response 7: We have changed the order of the introductory material such that the critique of existing measures is now presented before the description of the DKAS.

Reviewer comment 8: p7: Some repetition under "Measure".

Author response 8: we have reviewed the Measure section of the Methods and removed potentially duplicative text (see page 8, lines 143-146).

Reviewer comment 9: p8: Consider splitting this paragraph into two parts. Is the approach you mention suitable for ordinal data (in stata this is generalised SEM, not sure about AMOS).

Author response 9: We have split the section as the reviewer suggested to improve readability (see pages 8-10). AMOS is an auxiliary package for SPSS, which is used for structural equation modelling (SEM) and CFA. We have also added text to denote that CFA was performed using a package suitable for SEM (see page 9, lines 163-164).

Reviewer comment 10: p9: Is there any information at all about non-respondents?

Author response 10: It is not possible to derive detailed information about non-respondents as the completion of demographic details were a voluntary part of completion of the DKAS measure. Some demographic details were collected for the total sample; however, due to the design of the data collection and the anonymity of respondents, it was not possible to compare response and non-response groups directly.

Reviewer comment 11: p11/12 & p15: Is there anything else you can add about the importance of reducing correlation below 0.80? The removal of item 20 and the loss of construct validity seems a high price to pay for meeting this 0.8 threshold. What are the consequences?
Author response 11: All decisions made in the survey development process were based on established statistical criteria. Decisions to retain or remove items during scale development were not, and should not, be based on a subjective consideration. We have added further justification for the removal of the depression item within the text and have acknowledged the tension in scale development between conceptual and statistical integrity (see pages 17-18, lines 302-317). To clarify, the decision to remove the depression-related item (people with dementia are unlikely to experience depression) was undertaken to improve the model summary statistics (GFI and RMSEA) as well as improving the correlation between factors based on established criteria. The decision was not taken lightly and was made in consultation with dementia experts. It should also be noted that there is another item within the measure that also addresses depression (symptoms of depression can be mistaken for symptoms of dementia), although we acknowledge that it is slightly different in composition from the removed item.

Reviewer comment 12: p11/12 & p15: Linked to this, I wonder if a bi-factor model would have been more appropriate for these data to see how much variance is accounted for by the common factor? The four-factor model in PCA with ordinal data may have recovered too many factors (see http://dx.doi.org/10.1371/journal.pone.0118900).

Author response 12: See response to first reviewer’s comment (comment 2, response 2).

Reviewer comment 13: p12-14: I wondered if all the p-values were needed, given the n of several thousands... focus more on size of differences rather than whether they are statistically significant.

Author response 13: Standard reporting conventions were followed here, so we believe that the reporting of p values in conjunction with Chi-square metrics are justified. We acknowledge, however, that larger sample do naturally increase the expected significance of the results. Magnitudes of differences between cohort scores are evident in Figure 2.

Reviewer comment 14: p12: I got lost at a step. Why was the correlation between factors around 0.80 in some of the results, and 0.40 in later sections? Can you clarify?

Author response 14: A criterion value of 0.80 was used to ascertain the acceptable level of estimated factor correlation between the four confirmed factors within the model. Higher values suggest too much redundancy, although correlation is expected as the measure should be expected to hang together as a scale (values between 0.3 and 0.8 are generally accepted). In order to improve comprehensibility we have removed reference to the Spearman correlation, which is based on a different set of measured parameters (independent factor relationships) outside of the
CFA context and which, we concur, can create some difficulties in interpreting the results for readers. See changes in text and deletion (page 14; lines 257-259).

Reviewer comment 15: p16: I wasn't clear why validity and reliability is assured with all the different groups of respondents mentioned. To make that statement, we would need to see some tests of coefficient stability between sub-groups.

Author response 15: To address the reviewer’s comment, we have amended the language in the discussion section to ensure that comments concerning the final validity and reliability of the DKAS relate to the aggregated sample of health professionals and consumers, rather than particular sub-cohorts (see page 18, lines 323-324).

Reviewer comment 16: p17-19: That's a long paragraph.

Author response 16: We have split this paragraph into several smaller statements to improve readability.