Author's response to reviews

Title: The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ)

Authors:

Richard H Osborne (richard.osborne@deakin.edu.au)
Roy W Batterham (roy.batterham@deakin.edu.au)
Gerald R Elsworth (gerald.elsworth@deakin.edu.au)
Melanie Hawkins (melanie.hawkins@deakin.edu.au)
Rachelle Buchbinder (rachelle.buchbinder@monash.edu)

Version: 3 Date: 1 May 2013

Author's response to reviews:

We thank the reviewers for their thoughtful comments. We have considered each in turn and have dealt with each as outlined below.

Reviewer 1

1. In the abstract (conclusion) it is suggested that the measurement instrument assesses health literacy needs.

This is correct. The conclusion that the HLQ “assess the needs and challenges of a wide range of people and organisations”

Introduction:

In section 3.2.1 they use the terms health literacy capacity. By describing the goals and applications of the instrument it may become clear whether they want to measure the health literacy ‘status’, and use it for an evaluative purposes, or whether the instrument is primarily meant to assess individuals health literacy needs or the capacities. These are all slightly different constructs. The paragraph in the discussion on the applications may be informative and helpful if it is positioned in the introduction. This might also clarify the requirements for the items as discussed in the last section of the introduction.

The starting point for the development of the HLQ was the WHO definition: “the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health.”

The HLQ therefore measures an individual’s ability to access, understand and use information. The status of individuals is assessed by the questionnaire. When a person is found to have a low score, such a score can be described as reflecting a person’s ‘health literacy needs’.

These (status or needs) are not variations in constructs. The interpretation of
questionnaires ‘scores’ can simply reflect ‘status’, and if the score is low, a potential ‘need’ can be revealed.

As the reviewer suggests, we have improved the introduction (end of last paragraph) to include clearer information on the purpose of the research:

“We sought to develop a tool that was capable of detecting a wide range of health literacy needs of people in the community, and that could be used for a variety of purposes from describing the health literacy of the population in health surveys through to measuring outcomes of public health and clinical interventions designed to improve health literacy.”

We prefer to keep discussion about the potential applications of the HLQ for the discussion section where we think it will be better understood. We would prefer to keep the introduction for providing information on the context of health literacy and the needs for a new and improved measure.

Methods

These are very well described.

3. It is informative to mention the seeding statements in the Methods section. This enhances the understanding why the constructs do or don’t fit within the definition.

We thank the reviewer for this suggestion. The seeding statement is now in the Methods:

“…This first brainstorming step involved participants responding to the following seeding statement; “Thinking broadly about your experiences in trying to look after your health, what abilities does a person need to have in order to get, understand, and use health information to make informed decisions about their health?”…

4. Did the sentinel themes results from the analyses or were they introduced by the researchers themselves?

The sentinel themes arose directly from the data. On reflection, the term ‘sentinel theme’ is potentially confusing as they were extant sub-constructs (or themes). We have made this point clearer in the text by removing the idea of ‘sentinel theme’.

5. The critical RMSEA value 0.06 is mentioned here, while in the results section the value 0.05 is mentioned.

The RMSEA critical value for ‘good fit’ was set at 0.06 in this study following the most recent recommendation for critical values for the close fit indices (West SG, Taylor AB, Wu W: Model fit and model selection in structural equation modeling. In: Handbook of Structural Equation Modeling. edn. Edited by Hoyle RH. New York, London: The Guilford Press; 2012: 209-231.). The mention of 0.05 in the results section was an error and has been corrected.
Results

It is nice to have first a description of the content of the items and constructs, followed by the results of the quantitative analysis, and then again followed by an interpretation involving the content of the items/constructs.

Thank you.

6. The Tables are not in accordance with the text. It seems that they are not the latest tables or texts that are included. E.g. I could not find part A and B see section 4.2). And probably the constructs are numbered differently in the end. Moreover, I count 56 items instead of 0.56. The range of factor loadings are all different.

We apologize for the confusion. Part A and Part B referred to the first 5 scales (with agree/disagree response options) and the following 4 scales (with cannot do / very easy) response options. The A and B denotations had inadvertently been removed from Table 3, and in any case were not useful, and so have been removed from the manuscript.

The number of surviving items is 44 and this is shown in the abstract, Table 1 and Results.

The tables have been updated

7. The constructs Cognitive barriers and Understanding and Critical appraisal had much overlap and were combined. However, the first construct belonged to the theme “being resourced” and the second one to the theme “About one-self” Is it possible to combine both and what is the resulting theme?

This is an important insight. However, given that the final analysis strongly supports the separation of all of the nine constructs, in both a large construction sample and a large replication sample we strongly suggest that further data exploration is unwarranted at this stage. The overarching themes will be an important part of our (and/or others) future work where Path Analysis or other forms of structural equation modeling can be used with these themes as a priori groupings to explore mechanisms by which health literacy might impact on health and equity outcomes.

Discussion

8. I miss a remark on the very low response (13.7%) of the 3000 people. Do the authors have any data or any idea on whether this was a very selective sample?

We apologise for this omission. We have added the following to the Discussion:

“While the replication dataset had a low response rate (13.7%) the administration of the questionnaire was passive (one letter posted upon return from a visit to an emergency department) with no follow up or reminders. The setting was a regional public hospital (45% of Australians do not have private health insurance and so use public hospitals), with a large proportion of immigrants and refugees.
While 13.7% is a low response rate, given the hospital’s catchment, and the purpose of this phase of the questionnaire development process, the data provide a reasonable challenge to the psychometric structure of the questionnaire.

9. Does this instrument apply to Western countries, and also to developing countries. Note that one of the persons who reviewed the items had experience in “indigenous health and development”. The number of health illiterate persons will be lower. Will the instrument be able to distinguish between low and very low health literacy?

We expect the questionnaire will be a precise evaluative tool and will distinguish between low and very low health literacy. We put great effort into creating items and sets of items to achieve this. Although Australia is regarded as a Western country, the HLQ has already been assessed by experts in several Western regions (UK, USA, South America, Nordic countries, and mainland EU), Eastern regions (Taiwan) and indigenous cultures (Australia) to be an appropriate tool. Given that the tool was designed to be administered orally (as well as paper based), we expect that it will be widely used.

We have inserted the following statement in the discussion:

“…Given this inclusive starting point and our wide consultation, we expect that the HLQ will be a suitable tool in many Western and Eastern cultures, however it will be necessary to undertake rigorous studies to confirm its applicability in each setting [29].”

10. In the third section the authors state that the scales cover a broad range of issues pertinent to an individual’s life and reflect both intrinsic and extrinsic dimensions of health a) and b). Do these correspond in one way or another to the three overarching themes that were mentioned previously? Isn’t it confusing to introduce another classification in the discussion?

This is a challenge that can arise when undertaking questionnaire development research using grounded approaches. We believe that it is important to explain our thinking and processes throughout the development of the tool, to reveal, as far as possible, all methods and processes so they can be further tested and refined in the future. Hence we suggest leaving in the frameworks that emerged during the development. When we carefully considered the final constructs, we see that further valuable interpretations of the constructs emerged. In the Discussion we raise the idea that the interpretation of the HLQ scales may reflect intrinsic and extrinsic factors i.e., a) intrinsic factors reflecting issues attributes of the respondent, and b) attributes of the organisations the respondent engages with. These are intended to be ideas to assist users with interpretation of the scales, not another way to classify the constructs.

We have added some minor clarifications in the text to improve clarity

11. I suggest to remove the section of the application of the HLQ from the
discussion to the introduction.

Thank you to the review for this suggestion. We agree that it is important to orientate the reader towards how the questionnaire might be used in the introduction. To improve the introduction according to this advice, we include an outline of how the HLQ may be used at the end of the Introduction.

12. Does the omission of some constructs require an adaptation of the definition of health literacy? affect the definition of

We have taken a grounded approach and an operationalisation of the WHO definition has emerged. We feel that specific testing and application of the HLQ in a wide range of settings is required before definitions of health literacy could be revisited. One construct (Being health focused) was not included in final set of HLQ constructs because the psychometric properties of this scale were not sufficiently strong. We point out in the Discussion that further work is required:

“Finally, the Being health focused scale (later renamed Engagement in decisions) did not survive our validity-driven approach to scale development and is a gap in the HLQ that requires further work.”

Minor essential Revisions:
- Abstract, last sentence: HLQ is likely
  Corrected

Introduction.
You refer to initial concept mapping data without a reference. Is it reference 26 or 27?
Corrected

Method
- Some references are not included Error! Reference source not found.
Corrected
- “concepts 1 to 3” is not clear what you refer to.
Corrected
- do the authors refer to the results of the focus groups when they use the term ‘consultation data’?

Yes, consultation data are data from concept mapping groups. This is clarified in the Abstract.

Results
Section 4.3: some figures, like items difficulty score = 0.36 I tried to find in Table 4. Later I found out that the results from these section (results of the analysis on
the calibration sample) are not presented in one of the Tables. This might be mentioned here.

This has been clarified with the following statement:

“Only a summary of the extensive calibration dataset analysis is presented here.”

RMSEA values > 0.5 and < 0.8 are mentioned here. I think that 0.05 and 0.08 is meant here. Note that in the Results section, they present a critical value of 0.06. The remark that potential cross-loadings were fixed precisely to 0.0 I haven’t found in the methods section.

We thank the reviewer for noticing this error in the statement of the critical values for the RMSEA. This has been corrected. Please also see our earlier comment on our use of 0.06 rather than 0.05 as the critical value for ‘good fit’. We have also needed to make some minor corrections to the summary of the model fit in the text at this point. The statement relating to the fixing of cross-loadings ‘precisely to 0.0’ has been moved to the method section as suggested.

The text says that item difficulties of the other scales ranged form 10 to 30%. How does this correspond the difficulty values presented in Table 4: difficulty mostly higher than 0.75.

We thank the reviewer for pointing out the potential confusion in our tabling of ‘difficulty’ estimates and confidence intervals. The tabled values in the original text indexed the ‘easiness’ of the items (higher values were associated with a greater probability of an ‘agree/strongly agree’ or ‘quite easy/very easy’ response) although the column was headed ‘difficulty’. We have amended Table 4 so that the tabled values (point estimates and confidence intervals) represent the level of ‘difficulty’ (higher values associated with a ‘disagree/strongly disagree’ or ‘quite difficult to cannot do’ response). The table values now correspond directly to those mentioned in the text of the results section.

Reviewer 2

Methods

Overall, the structure and presentation of the method is confusing and it is quite difficult to follow. For example, on page 6, the authors have mentioned that “…with structured processes governing the movement from consultation data to measurement tool. This initially involved an elaboration phase to …… The second phase….”. What about 3.3 and 3.5? Should these be part of development process? The section 3.2 should be discussed separately from the process. It is necessary to have a summary of the process before you start to discuss them individually.

We agree that the flow was not quite right and the references to the section numbers may have made the manuscript hard to follow. A clear flow chart of the processes has been improved. The introduction to the methods section was
broken up with our reference to previous work and Ethics. This has been moved to the bottom of this section. This should make it easier to follow.

Under 3.3, what purpose of the review and what you expect to get from the consultation and review? Do you have an outline for the review panels? Have you discussed the results from the review? They are not clear to me. It is needed some work to clarify these.

The purpose has now been made clear. The review by clinical staff and managers was informal. This has been made clearer in the text:

‘The purpose of these reviews was for experts from a wide range of backgrounds to provide feedback on the relevance and appropriateness of the items and concepts to their settings.’

Under 3.4, you have mentioned that Mplus Version was used for your data analysis. You then mentioned SPSS and SAS as well, which you didn’t use them for your analysis. I don’t think that it is necessary to include in your text. In method, you just need to state what you have been used and done for your study.

We have removed the references to other statistical programs. We have retained the rationale for why specific statistical routines are used as many readers may be unfamiliar with these.

Results
The results are not presented clearly and are mixed with methods and discussion. For example, on page 16, how to calculate a difficulty level of items should be discussed in the method. On page 21, the paragraph “Given the tendency for multi-scale CFA models with all potential cross-loadings fixed precisely to 0.0 to provide inflated estimates of inter-factor correlations, these correlations were regarded as acceptable”, can be discussed in methods section.

The methods describing calculation of difficulty and inter-factor correlations have been relocated into the Methods.

You talked about Part A and Part B that are labelled in table 1 on page 16, but I couldn’t find it.

This has been corrected, see above.

Table 5 and 6 (no title at all, I assume it is Table 6 after Table 5) are not mentioned in the text. What are you trying to present in table 6? No result in relation to the table (6?) is discussed in the article.

Table 5 is referred to in the Discussion
Table 6 is actually Figure 1. We apologise for not uploading this correctly with Title. This Figure is a critical part of describing the methods and is now more integrated at the start of the Methods section.
Overall, the presentation of the tables except table 2, 5 and 6 are presented well. Table 2 can be presented more clearly by using individual column with subtitles, e.g. scale/item, high and low. Table 6 needs a title. The structure of Tables should be consistent.

We have improved the structure of tables 2, 5 and 6. Table 2 is now much improved using the advice of Reviewer 2. Table 6 is intended to be Figure 1 and this is now properly labeled.

Discussion and conclusions:
Both background and discussion, authors provided some evidence from previous studies and clearly acknowledged the work. However, some part of discussion should be introduced in background. For example, on page 25, the last paragraph, the authors have discussed “We used the definition of health literacy proposed by the WHO….. through the concept mapping exercise”. This should be defined in the background.

The WHO definition is already defined in the introduction (first sentence), this is made clearer by including the acronym (WHO).

The study limitations are not clearly stated.

We have improved the limitations section. See above.

Overall comments:
- The structure of the writing is not acceptable.
- Some references are not correctly cited, e.g page 6, 7....
- It is quite unusual to number headings and subheadings, which may not acquire the BMC formatting requirements.

We hope these have been improved to the satisfaction of the reviewers.

We have also corrected the name of the Critical appraisal scale – it should have been labeled Active appraisal. This decision was made in the final item selection phase (replication) but we omitted to include this decision in the manuscript.