Patient Education and Counseling 91 (2013) 228–235

Contents lists available at SciVerse ScienceDirect

Patient Education and Counseling

journal homepage: www.elsevier.com/locate/pateducou

Health Literacy

The Health Literacy Management Scale (HeLMS): A measure of an individual's capacity to seek, understand and use health information within the healthcare setting

Joanne E. Jordan [a,d], Rachelle Buchbinder [b,c], Andrew M. Briggs [a,d], Gerald R. Elsworth [e], Lucy Busija [a], Roy Batterham [e], Richard H. Osborne [a,e,*]

[a] Department of Medicine (Royal Melbourne Hospital), The University of Melbourne, Melbourne, Australia
[b] Monash Department of Clinical Epidemiology, Cabrini Hospital, Victoria, Australia
[c] Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia
[d] School of Physiotherapy and Curtin Health Innovation Research Institute, Curtin University of Technology, Perth, Australia
[e] Public Health Innovation, Deakin University, Melbourne, Australia

ARTICLE INFO

Article history:
Received 15 June 2012
Received in revised form 4 January 2013
Accepted 12 January 2013

Keywords:
Health literacy
Measurement
Questionnaire design
Patient communication
Patient reported outcomes

ABSTRACT

Objective: Health literacy refers to an individual's ability to seek, understand, and use health information. This paper describes the development and psychometric testing of the Health Literacy Management Scale (HeLMS).
Methods: Content areas were identified from a conceptual framework derived from interviews and concept mapping. Items were generated from statements from concept mapping participants. Construction ($N = 333$) and replication ($N = 350$) samples were participants in chronic disease self-management programs and emergency department attendees. Factor analysis was used to refine constructs and define psychometric properties.
Results: Consultations generated 8 scales each with 4–5 items: Understanding health information, Accessing GP healthcare services, Communication with health professionals, Being proactive and Using health information, Patient attitudes towards their health, Social support, and Socioeconomic considerations. Confirmatory factor analyses indicated good fit of the data with the model (RMSEA = 0.07, SRMR = 0.05, CFI = 0.97) and all domains had high internal consistency (Cronbach alpha > 0.82).
Conclusion: The HeLMS has acceptable psychometric properties and assesses a range of health literacy constructs important to patients when seeking, understanding and using health information within the healthcare system.
Practice implications: The HeLMS presents a new approach to assessing health literacy in healthcare settings.

1. Introduction

Health literacy refers to an individual's ability to seek, understand, and use health information to make appropriate decisions regarding their health [1]. Because of its relevance and importance to patient-centred healthcare and health reform, health literacy is receiving increasing attention from governments, researchers, clinicians and patients [2–5]. Providing individuals with information, and actively involving them in decisions about their health, are key components of patient-centred healthcare [6,7]. Active involvement is, however, unlikely to be achieved if patients have suboptimal health literacy.

The majority of available health literacy measures [8–14] focus on assessing reading, comprehension and numeracy skills, and do not capture many underlying concepts of health literacy [15]. There is a growing consensus that health literacy encompasses a broader range of attributes other than just literacy skills such as abilities to interact within broader social and environmental contexts [16–20]. Results from available health literacy measures have shown that people with suboptimal health literacy have:

* Corresponding author at: Public Health Innovation, Population Health Strategic Research Centre, Deakin University, Melbourne, Australia. Tel.: +61 3 9244 6292; fax: +61 3 9244 6624.
E-mail address: Richard.osborne@deakin.edu.au (R.H. Osborne).

0738-3991/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.pec.2013.01.013

difficulty understanding health information [21,22]; poorer knowledge of their condition [23–25]; and lower utilization of preventive health services [26–28]. These tools measure a limited subset of the health literacy abilities that clinicians and health service planners need to understand to improve outcomes.

A clear conceptual framework is essential to the development of a questionnaire that has validity and clinical utility. Without a conceptual grounding, the content of a measure may not accurately reflect the phenomenon under study, and it may be unclear if the most relevant elements have been identified [29]. Conceptualizations of health literacy have been developed from the perspective of researchers, health professionals and literacy experts, with minimal consultation with patients. Modern approaches to questionnaire development consider that the inclusion of patients' views is a crucial foundation to ensure content validity [30]. We now build on our previous work covering the conceptualizing of health literacy from the patient perspective [31], and detail the development of the Health Literacy Management Scale (HeLMS).

2. Methods

The development of the HeLMS is outlined in Fig. 1. This included two main components, the development of the conceptual framework (Section 2.1) and the development and testing of the measure (Sections 2.2–2.7). This study was approved by The University of Melbourne, Cabrini Health, and Melbourne Health Human Research Ethics Committees.

2.1. Development of a conceptual framework

A conceptual framework was developed using data from in-depth consultations with patients who were likely to represent 'critical cases' in terms of their health literacy needs [32]. Two approaches were taken; in depth interviews and concept mapping. The synthesis of the interviews has been described elsewhere [31] and involved 48 people: those with chronic conditions, from the general community or had recently attended the emergency department at the Royal Melbourne Hospital (a public tertiary hospital in Victoria with a low socioeconomic and high culturally and linguistically diverse catchment). The interviews sought to reveal elements of health literacy through patients' narratives around their 'healthcare journey'; from identification of a health problem to resolution in seeking, understanding and using health information [31]. The interview transcripts were analyzed using the grounded theory analysis techniques described by Strauss [33]. These included 'in vivo' codes where possible (code labels derived from statements in the data rather than from theory) and, where this was not possible, the use of neologisms or brief descriptive phrases for coding in order to avoid imposing theoretical pre-conceptions on the data.

In addition to in depth interviews, the conceptual model development was guided by concept mapping, a structured group process that integrates qualitative and quantitative approaches to derive a graphical representation of major ideas about the phenomenon of interest [34]. The authors have used this technique previously in the development of questionnaires [35–38]. It involved five steps:

i. **Participant recruitment**

Participants were recruited to two workshops using critical case sampling [32] to select groups at extremes of the disease and socioeconomic continuums, namely (a) individuals ($N = 8$) with a chronic condition from suburbs with low socioeconomic status as defined by the proportion of residents with low income, low education, unemployed, and dwellings without

Fig. 1. Steps used in the development of the HeLM.

motor vehicles [39]; and (b) individuals ($N = 7$) with no known condition from a high socioeconomic area in Melbourne. Participants were recruited from a University of Melbourne database of individuals with chronic conditions who had participated in education programs [40], and from a population-based health literacy survey [41]. If the results of the concept mapping were substantially different from the interview results then additional groups would have been conducted but this was not the case.

ii. **Brainstorming statements**

The following seeding statement was used to generate statements describing health literacy: "Thinking about your experiences in trying to look after your health (or the health of your family), what abilities does a person need to have in order to get and to use all of the information they need?" Prompts were also used: Think about "a person's own ability" and "connections that may help them to get the information, understand and act on it".

iii. **Statement sorting**

Statements were printed on individual cards and participants were asked to sort them into separate piles in a manner that made sense to them [34].

iv. **Analysis of the sort data**

Analysis was conducted using concept mapping software (Trochim WMK (1989) The Concept System, Ithaca, New York, Cornell University, Department of Human Services Studies). A similarity matrix was created and then analyzed by two-dimensional multidimensional scaling. The similarity matrix highlighted the similarities among statements as indicated by the number of participants who sorted any pair of statements into the same pile. This procedure generated points on a map where each point represented a brainstormed statement, while the distance between the points represented the likelihood that statements were sorted into the same pile by different people [34,42]. Hierarchical cluster analysis was then used to create a map that combined similarly-sorted statements into clusters, based on their distance from each other. The clusters represent major concepts related to the topic of interest that were identified by the participants [34,36,43].

v. **Participant interpretation of the concept map**

The workshop participants were presented with a printout of all the statements grouped per cluster. They were asked to review the statements and groupings by themselves and consider whether or not each statement fitted in the clusters they were assigned and whether they might be better placed elsewhere, and also consider whether some clusters should be joined or split. Participants were also asked to provide a cluster label, using their own words, covering the general theme of the statements. A final name for each cluster was agreed by the participants.

2.2. Specification of hypothesized dimensions for measurement

The potential areas of measurement of health literacy were generated from previously reported synthesis of in depth interviews [31] and the above concept mapping. To assist with the consolidation of these data, descriptive meta-matrices were applied [44]. These are visual displays that allow for the systematic presentation and analysis of information collected from multiple and different settings [44,45]. One author (JEJ) undertook the primary analysis, which was then reviewed by two other authors (RB, RHO).

The selection of content areas was guided by the following criteria: (a) potential domains were applicable across a range of diseases, (b) potentially modifiable, and (c) important to individuals when seeking, understanding and using health information within the healthcare setting.

2.3. Item and response scale generation

Items for each content area were generated using statements from the concept mapping workshops and interviews. Items were required to contain one concept and suitable for verbal or written administration. A five-point Likert-style response format was developed in response to the content of the items: 1 = unable to do, 2 = very difficult, 3 = with some difficulty, 4 = little difficulty, 5 = without difficulty. Item content and domains were evaluated for clarity, conciseness, grammar, and face and content validity by health service researchers and clinicians.

Draft items were mailed to 542 individuals (construction sample) known to have a chronic condition and listed in the database described in Section 2.1. Individuals who had participated in the interviews and concept mapping were excluded.

2.4. Comparing the grounded factor structure with other common structuring methods

We did some brief, supplementary exploratory factor analysis (EFA), using MPlus version 5.0 [46] to compare the dimensionality suggested by the factor analysis with that derived from our previous qualitative processes [47]. This involved:

(i) estimating the number of factors (underlying constructs) potentially present in the data;
(ii) determining the optimal number of factors to retain for the measurement model.

The number of factors in the data set was estimated using two methods: eigenvalue > 1 [48] and Scree test [49]. The first of these is based on the inspection of factor eigenvalues, while the second considers the residual errors associated with the whole multifactor solution [50].

Factor rotation was applied to further clarify the structure of the hypothesized measurement model. Factors were extracted using means and variance adjusted weighted least squares extraction (WLSMV) with oblique (GEOMIN) rotation [46]. Model fit was assessed using a combination of absolute and incremental fit indices: root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and comparative fit index (CFI). Cut-off criteria for good model fit were RMSEA < 0.08, SRMR < 0.09, and CFI > 0.95, based on suggestions in the literature [51,52].

2.5. Confirmatory factor analysis

Confirmatory factor analysis (CFA) was used with both a construction and a replication sample (see Section 2.4) to determine whether the hypothesized latent variables representing domains loaded on the specified items. A sequence of one-factor models, all possible 'pairwise' models, and a final multi-factor model were fitted to the data [53]. MPlus was used with WLSMV estimator for categorical data [46,54].

Using the model fit parameters above, the following criteria were used to identify problematic items:

- Consistent intra-factor correlated errors across factors (>3 factors)
- Cross loading (>0.3) on a factor other than the hypothesized factor
- Cross loading (>0.2) on two or more other factors
- Inter-factor correlated errors (>0.1)

2.6. Cognitive interviews

This involved asking individuals to recount how they derived their answers to ascertain whether the items were understood as intended [55]. Twelve participants took part from the Royal Melbourne Hospital Emergency Department. Each item was covered at least three times.

2.7. Model replication and examination of item characteristics

The refined items were then posted to emergency department attendees ($N = 500$), or people with a chronic condition ($n = 250$) listed in the research database (excluding individuals who had previously participated). CFA was conducted to confirm the structure of the measurement model that emerged from the construction sample.

Differential item functioning analysis was undertaken to determine that items performed in similar ways regardless of group, country of birth (Australia vs elsewhere), age or sex. Analysis was undertaken using MPlus within the CFA framework with group, age and gender as covariates. Significant direct effects of a covariate on items were interpreted as the presence of differential item functioning.

The model was initially specified with all direct effects of covariates on items fixed at 0.0 [47,56]. Modification indices were then inspected to identify significant effects of the covariate on items, as indicated by a modification index > 2.5. Following inspection of modification indices, the direct effects of the covariates on items were incorporated into the model, starting with the parameter with the highest value. The model was then re-estimated to examine the possibility of differential item functioning among the remaining items [56]. Only persons with complete data were included in the analysis.

2.8. Scale score calculation and test–retest reliability

Scale scores were calculated as the mean score of each domain. Mean differences between groups were explored with analysis of variance (ANOVA) and with age as a covariate (ANCOVA) using SPSS 17.0.

Test–retest reliability was explored in a convenience sample of 100 individuals attending a private a metropolitan physiotherapy practice in Perth, Western Australia. The HeLMS was administered twice over a two-week period. It was anticipated that their health literacy would not change during this time. Test–retest reliability was estimated though a one-way random effects intra-class correlation coefficient (ICC) using SPSS 17.0.

3. Results

3.1. Conceptual framework

The previously analyzed 48 interviews revealed seven key abilities [31]: knowing when to seek health information; knowing where to seek health information; verbal communication skills; assertiveness; literacy skills; capacity to process and retain information; and application skills.

Concept mapping workshops were undertaken with 8 people with a chronic condition (Workshop 1) and 7 without a chronic condition (Workshop 2), see Table 1. Participants in Workshop 1 produced 45 statements across 9 concepts and Workshop 2, 36 statements arose across 9 concepts. The labels for each cluster are shown in Table 2.

Data from the interviews and concept mapping were initially analyzed by one author (JEJ) using descriptive meta-matrices which were then reviewed by two other authors (RB, RHO). This process revealed key overarching components of health literacy which defined a two-level conceptual framework: (i) six core individual abilities required to seek, understand and use information in the healthcare setting, and (ii) eleven extrinsic and intrinsic contextual factors that underpin these abilities. See Fig. 2 for the full list of components of the abilities and contextual factors.

3.2. Hypothesized domains of health literacy

The application of the pre-specified criteria by which we sought to operationalize a measure of health literacy (Section 2.2) identified ten potential content areas for measurement development. These included the six individual abilities and only four of the eleven enabling factors: Social support, Socioeconomic, Attitudes towards health, and Lay knowledge. The remaining seven enabling factors, such as Linguistic background, Healthcare setting, and Educational background were excluded as they were regarded as principally non-modifiable.

3.3. Item generation

Using the item writing guideline (Section 2.3) items were constructed across the 10 content areas. They were based on what patients reported in workshops or interviews. For example the statement *Regular visits to doctors when required* was converted to a question format such as *Are you able to see a doctor when you need to*. The process of item generation resulted in further scrutinisation of the content areas to derive the final constructs. Substantial overlap in content was revealed in two pairs of content areas and these were combined: Lay knowledge was merged with Attitudes towards health and, Capacity to retain and process information was merged with Application skills to manage health. For each domain between 6 and 13 draft items were generated (total 66) and applied in the field. Table 3 outlines the scope of the measure.

3.4. Comparing the grounded factor structure with other common structuring methods

For the initial testing of the 66 items, 333 (61%) of the 542 surveys were returned; 75% female, mean age 67 years (range 25–93), and 4% were from culturally and linguistically diverse backgrounds. Among the 66 items there were between 4 and 10 factors. The eigenvalue method suggested 9–10 factors while the Scree plot suggested 4–6.

Table 1
Participant demographics from concept mapping workshops.

Descriptor	Workshop 1	Workshop 2	Overall
Participants (*n*)	8	7	15
Age (years)			
40–49	2	0	2
50–59	0	1	1
60–69	3	3	6
70–79	3	2	5
80+	0	1	1
Sex (% female)	63	100	80
Education: *n* (%)			
Completed some primary school	1 (13)	1 (14)	2 (13)
Completed some high school	2 (24)	4 (57)	6 (40)
Completed high school	3 (38)	0 (0)	3 (20)
Completed university	2 (25)	2 (29)	4 (27)
Socioeconomic index for area (decile of socioeconomic disadvantage[a])			
Decile – 1–2	6	0	6
Decile – 3–4	2	0	2
Decile – 5–6	0	0	0
Decile – 7–8	0	0	0
Decile – 9–10	0	7	7

[a] Deciles taken from Australian Bureau of Statistics Socio-economic Indexes for Areas (SEIFA) report 2006.

Table 2
Outputs from concept mapping workshops showing themes across groups.

Cluster descriptor	Number of statements
Workshop 1	
Cost and transport	4
Access to health professionals	9
Communication skills	3
Emotional skills and support	2
Emotional and psychological issues necessary to help yourself (require support)	9
Positive attitude	2
Support and support groups	4
Education, access to relevant information	8
Relevant up to date information (information needed at a particular time)	4
Workshop 2	
Finding sources of information/skills to enhance information exchange	8
Support networks	4
Patient awareness while in the healthcare setting	3
Social skills	2
Self-help	2
Communication and questioning	6
Trust and confidence in doctor	4
Self-assertion and self-worth	5
Listening skills	2

Using GEOMIN factor rotation, an 8 factor model had the best fit: RMSEA = 0.07, SRMR = 0.03, CFI = 0.97 compared to 4-, 5-, 6-, 7-, 9- and 10-factor models. Substantial similarities were identified between constructs in the proposed EFA model and the specified domains and respective items derived from the qualitative data; i.e., items largely loaded on factors that they were expected to be associated with. The EFA did not change the domains and no items were re-allocated solely on the basis of the EFA. The EFA did however suggest some potential cross-loadings and other issues that were investigated further in the subsequent CFA.

3.5. Confirmatory factor analysis

CFA was then applied to the same data set used for EFA. Single factor models for the 8 hypothesized constructs indicated moderate to good fit, although there was some item redundancy on several factors, i.e., high intra-factor correlated errors and high residual error (>0.4). Pairwise models identified items that cross-loaded, had inter-factor correlated errors or high residuals. Overall, 29 items were excluded based on pre-specified criteria (see Section 2.5). For the remaining 37 items, single-factor models showed good fit for 7 domains and moderate fit for 1 (Communication with health professionals). Using a restrictive confirmatory approach, a full 8-factor model was evaluated. No cross loadings or correlated errors were allowed and a good model fit resulted: RMSEA = 0.08; SRMR = 0.06 and CFI = 0.96.

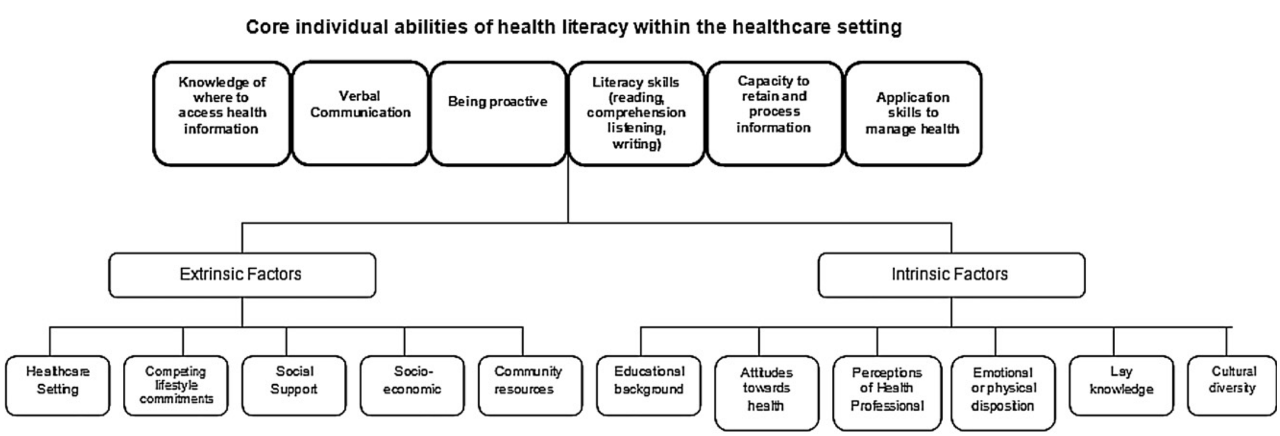
3.6. Cognitive interviews

Twelve individuals participated: mean age 68 years (range 34–87) and 5 were males. Three individuals disclosed that they were illiterate. The responses strongly reflected the intention of items and no items were revised. No participant interviewed indicated any difficulty in selecting an appropriate response option.

3.7. Replication survey and analysis

Overall, 350 (47%) of 750 surveys were returned. The mean age was 56 (23–94 years), 249 (71%) were female, and 27 (8%) of individuals were from culturally and linguistically diverse backgrounds.

Six of the 8 factors had excellent fit in single factor CFA models. Understanding health information and Being proactive had poor fit. Individual item fit revealed a poorly-fitting item (high intra-factor correlated errors >0.3) in each factor. Excluding these items resulted in improved fits for both. Cronbach alpha was >0.82 for all factors.



Core individual abilities	Description
Knowledge of where to access health information	This relates to knowing how to access the healthcare system as well as the ability to source health information from community resources such as the internet.
Verbal communication	Verbal communication skills are essential for the patient to effectively express their needs, concerns to obtain the assistance they need.
Being proactive	Relates to an individual's ability to take 'proactive' steps to source and understand health information in order to better address a health issue.
Literacy skills	This encompasses reading, comprehension, listening and writing skills.
Capacity to retain and process information	This refers to an individual's ability to comprehend and extract meaning from health information in the context of patient-health professional interaction.
Application skills to manage health	This is concerned with an individual's ability to apply information and incorporate it into their own lifestyle.

Fig. 2. Conceptual framework of health literacy from the patient perspective.

Table 3
Scope and content of the Health Literacy Management Scale (HeLMS).

Purpose
To measure generic and potentially modifiable abilities of an individual and their broader environment to determine their ability to seek, understand and use health information within the healthcare setting.

Target population
Adult population

Domains and exemplar items (number of items/Cronbach alpha)
1: Patient attitudes towards their health (4/0.87)
  Make time for things that are good for your health
  Pay attention to your health needs
2: Understanding health information (4/0.82)
  Read written information e.g. leaflets given to you by a doctor
  Read health information brochures found in hospitals or at a doctor's clinic
3: Social support (4/0.87)
  Ask someone to go with you to a medical appointment
  Ask family or friends for help to understand health information
4: Socioeconomic considerations (3/0.83)
  Pay to see a doctor
  Afford transport to medical appointments
5: Accessing GP healthcare services (4/0.87)
  Know where a doctor can be contacted
  Know how to get a doctor's appointment
6: Communication with health professionals (3/0.86)
  Ask a doctor questions to help you understand health information
  Get the information you need when seeing a doctor
7: Being proactive (3/0.89)
  Change to a different doctor to get better care
  Get a second opinion about your health from a health professional
8: Using health information (4/0.86)
  Use information from a doctor to make decisions about your health
  Follow instructions that a doctor gives you

Pairwise analyses confirmed the hypothesized 8-factor structure of the measurement model. Poorly-fitting items were considered for deletion while considering the content and breadth of each construct. Another 6 poorly fitting items were omitted: (1) on the basis of high cross-loading (>0.5), (2) due to high intra-factor correlated errors (>0.3), and (3) due to inter-factor correlated errors.

The final model with 8 factors and 29 items had good fit: RMSEA = 0.07, SRMR = 0.05 and CFI = 0.97 with no correlated errors or cross-loadings.

3.7.1. Health literacy scores

Table 4 presents demographics and domain scores for the groups (emergency department attendees and individuals with a chronic condition) in the replication survey. Respondents in the chronic condition group were older ($p < 0.001$). Individuals from the emergency department reported higher scores in 5 domains (Understanding health information, Social support, Socioeconomic considerations, Communication with health professionals, and Using health information). With adjustment for age, significantly higher scores were also observed in the domains of Patient attitudes towards health and Accessing GP healthcare services for respondents from the emergency department. Differential item functioning analysis showed no difference in how items performed across age, gender, country of birth and group. The highest value of modification index for differential item functioning was 2.1 for item 10 across country of birth but was below the critical cut-off value of 2.5.

3.8. Test–retest reliability

Seventy-nine (79%) people completed the test–retest. Their mean age was 37 (25–80 years), 68% were female, and 14% were from culturally and linguistically diverse backgrounds. Good to excellent test–retest reliability was demonstrated with ICC ranging from 0.73–0.96 across the 8 domains. Five of the domains had an ICC > 0.90. Understanding health information had the lowest although acceptable [57] reliability (0.73).

4. Discussion and conclusion

4.1. Discussion

The Health Literacy Management Scale assesses individuals' abilities, and their broader social and environmental contexts, to determine overall capacity to seek, understand and use health information within the healthcare setting. The HeLMS is based upon a clear conceptual framework of health literacy, which was derived from in-depth consultations with diverse patient groups. Through careful attention to item content, and factor analyses, the HeLMS was designed to have strong content, face and construct validity.

The framework derived in this study represents health literacy as an interaction between individual abilities and factors at a personal, healthcare system and broader community level. This is consistent with previous models in the literature [16,17]. Explicit criteria were set for the inclusion of content areas. To ensure cultural appropriateness, items were developed using language and terminology contained in patient statements describing health literacy.

While the concept mapping and interview data suggested potential measurement scales, the process of identifying themes in the interview data and of moving from the qualitative data to

Table 4
Mean difference and 95% confidence interval (95% CI) between groups for each domain of the HeLMS, expressed unadjusted and adjusted for age (years).[a]

Descriptor	Chronic condition	RMH	Unadjusted		Adjusted for age	
			Mean difference	95% CI for difference	Mean difference	95% CI for difference
Participants (N)	112	238				
Age (mean years)	63	53				
Female (%)	61	71				
Domain scores: mean (SD)						
(1) Patient attitudes towards their health	3.96 (0.91)	4.08 (0.76)	0.13	−0.07–0.32	0.20	0.00–0.38
(2) Understanding health information	4.52 (0.83)	4.78 (0.42)	0.26	0.09–0.43	0.22	0.08–0.36
(3) Social support	4.13 (1.12)	4.47 (0.71)	0.35	0.12–0.58	0.35	0.15–0.55
(4) Socioeconomic considerations	3.99 (1.00)	4.39 (0.81)	0.40	0.19–0.61	0.36	0.15–0.56
(5) Accessing GP healthcare services	4.78 (0.62)	4.86 (0.32)	0.09	−0.04–0.21	0.12	0.01–0.22
(6) Communication with health professionals	4.41 (0.87)	4.63 (0.56)	0.23	0.05–0.40	0.29	0.13–0.45
(7) Being proactive	4.07 (1.09)	4.27 (0.89)	0.20	−0.02–0.41	0.22	−0.11–0.44
(8) Using health information	4.45 (0.73)	4.76 (0.49)	0.51	0.02–0.31	0.21	0.07–0.34

[a] Each domain score ranges between 1 and 4 and related directly to the following response options: 1 = unable to do, 2 = very difficult, 3 = with some difficulty, 4 = little difficulty and 5 = without difficulty.

hypothesized scales and draft items involved researcher judgment at many points. We also made a number of additional decisions such as judgments about potential modifiability of a scale and about the practical usefulness of distinctions between scales. These discussions involved the whole research team and conclusions were reviewed by critical associates but are nonetheless potentially fallible. They are also an inevitable part of all instrument development. It is partly for this reason that we applied multiple processes to test the conclusions including cognitive interviews, two rounds of confirmatory model testing and checking the final content back against the initial qualitative data. Continuing validation of these decisions will occur through the ongoing collection of evidence related to the construct validity and utility of the tool.

EFA and CFA were undertaken to identify and confirm the presence of underlying constructs. The HeLMS was tested with a range of patients drawn from both healthcare (primary and acute) and general community settings. The final measurement model demonstrates good statistical fit and all eight domains have very good to excellent psychometric properties. While good to excellent precision was demonstrated across the HeLMS domains, the sample selected for this test was a convenience sample of people attending a private physiotherapy clinic who were likely to have adequate health literacy. Future work should explore test–retest in a more representative sample.

The scoring of the HeLMS is designed to provide a clear indication of the level of difficulty across components of health literacy. An average score is provided for each of the eight domains aligning with the original 5-point Likert-style response format (1 = unable to do to and 5 = without difficulty), for ease of interpretation for users. A lower score reflects a higher degree of difficulty experienced. Replication results identified significant statistical differences in domain scores between groups however the reason for this is unclear and would need to be explored using prospective studies. It also remains to be determined whether these differences are clinically meaningful. Further studies should focus on refining the scoring system and establishing clinically verified benchmarks.

While these initial data suggests that the HeLMS has strong properties, further demonstration of its construct validity and utility in different countries, cultures and across a range of clinical populations is required. While we attempted to ensure a wide representation of individuals in the conceptual development and validation phases, men were somewhat under represented. The differential item functions analysis indicated no significant bias between scores provided by men and women, however it cannot be excluded that constructs particularly relevant to men are missing. The validation of a multi-dimensional measure of health literacy is challenging given there is no single objective measure that it can be compared to. Future testing should include comparing the HeLMS with other indicators such as the health beliefs of general practice patients' questionnaire [59] and existing health literacy measures [9]. Future work must also focus on determining whether the scales and response options are responsive to change when used to evaluate specific interventions and whether these changes are linked to better health outcomes.

4.2. Conclusions

This study provides initial evidence that the HeLMS provides a broad, multidimensional understanding of an individual's health literacy. Its constructs go beyond the assessment of reading, comprehension and numeracy skills to include knowledge of where to seek health information; ability to be proactive in seeking or understanding health information; verbal communication skills; and capacity to retain, process, and apply information. The components represent abilities and contextual factors that patients have identified as critical to effectively seek, understand and use health information within the healthcare setting. A key difference between the HeLMS and previous measures is that it assesses generic and potentially modifiable factors that influence these abilities. The concept of health literacy in this study implies that an individual's health literacy is not fixed, and is dependent on a combination of circumstances, several of which may be outside the control of the individual. Therefore in addressing health literacy, the focus should not lie solely with the patient.

A limitation of the HeLMS is that not all components will be relevant in other healthcare systems. For instance the domain relating to accessing GP healthcare services may not be as applicable in countries where individuals register with one GP within a catchment area [60]. Therefore the cross-cultural applicability of each domain will need to be determined in international settings.

4.3. Practice implications

Application the HeLMS may provide new insights into the prevalence of suboptimal health literacy across populations. A range of potentially modifiable factors and intervention points may be identified. For instance, researchers and policymakers could use the HeLMS to inform interventions or health policy development.

Acknowledgements

The authors thank Dr Sandra Nolte and those who participated in this research. RB was supported in part by a National Health and Medical Research Council (NHMRC) Practitioner Fellowship, AB was supported in part by an NHMRC Postdoctoral Fellowship, and RHO was supported in part by a NHMRC Population Health Fellowship.

All patient/personal identifiers have been removed or disguised so people described are not identifiable and cannot be identified through the details of the story.

References

[1] USA Department of Health and Human Services. Healthy people 2010: understanding and improving health. Washington, DC: US Government Printing Office; 2000.
[2] Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008;67:2072–8.
[3] UK Department of Health. In: DoH UK, editor. Choosing health: making healthier choices easier. London, England: UK Department of Health; 2004.
[4] Commonwealth of Australia. A healthier future for all Australians – final report of the national health and hospitals reform commission. ACT: Commonwealth of Australia; 2009.
[5] Kickbusch IS. Health literacy: addressing the health and education divide. Health Promot Int 2001;16:289–97.
[6] Coulter A. After Bristol: putting patients at the centre. Brit Med J 2002;324:648–51.
[7] Dahlberg K, Todres L, Galvin K. Lifeworld-led healthcare is more than patient-led care: an existential view of well-being. Med Health Care Philos 2008.
[8] Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med 1993;25:391–5.
[9] Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med 1995;10:537–41.
[10] Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3:514–22.
[11] Chew L, Griffin JM, Partin MR, Noorbalochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacyin a large VA outpatient population. J Gen Intern Med 2008;23:561–6.
[12] Morris NS, MacLean CD, Chew LD, Littenberg B. The single item literacy screener: evaluation of a brief instrument to identify limited reading ability. BMC Fam Pract 2006;7:21.
[13] Ishikawa H, Takeuchi T, Yano E. Measuring functional, communicative, and critical health literacy among diabetic patients. Diabetes Care 2008;31:874–9.

[14] Hanchate AD, Ash AS, Gazmararian JA, Wolf MS, Paasche-Orlow MK. The demographic assessment for health literacy (DAHL): a new tool for estimating associations between health literacy and outcomes in national surveys. J Gen Intern Med 2008;23:1561–6.
[15] Jordan JE, Osborne RH, Buchbinder R. Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. J Clin Epidemiol 2010.
[16] Nielsen-Bohlman L, Panzer AM, Kindig DA. Health literacy: a prescription to end the confusion; 2004.
[17] Lee SY, Arozullah AM, Cho YI. Health literacy, social support, and health: a research agenda. Soc Sci Med 2004;58:1309–21.
[18] Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. Am J Health Behav 2007;31:S19–26.
[19] Buchbinder R, Batterham R, Ciciriello S, Newman S, Horgan B, Ueffing E, et al. Health literacy: what is it and why is it important to measure? J Rheumatol 2011;38:1791–7.
[20] Sorensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80.
[21] Schillinger D, Bindman A, Wang F, Stewart A, Piette J. Functional health literacy and the quality of physician–patient communication among diabetes patients. Patient Educ Couns 2004;52:315–23.
[22] Mancuso CA, Rincon M. Asthma patients' assessments of health care and medical decision making: the role of health literacy. J Asthma 2006;43:41–4.
[23] Williams MV, Baker DW, Parker RM, Nurss JR. Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. Arch Intern Med 1998;158:166–72.
[24] Gazmararian JA, Williams MV, Peel J, Baker DW. Health literacy and knowledge of chronic disease. Patient Educ Couns 2003;51:267–75.
[25] Powell CK, Hill EG, Clancy DE. The relationship between health literacy and diabetes knowledge and readiness to take health actions. Diabetes Educ 2007;33.
[26] Scott TL, Gazmararian JA, Williams MV, Baker DW. Health literacy and preventive health care use among Medicare enrollees in a managed care organization. Med Care 2002;40:395–404.
[27] Guerra C, Krumholz M, Shea JA. Literacy and knowledge, attitudes and behaviour about mammography in latinas. J Health Care Poor Underserved 2005;16:152–66.
[28] von Wagner C, Semmler C, Good A, Wardle J. Health literacy and self-efficacy for participating in colorectal cancer screening: the role of information processing. Patient Educ Couns 2009;75:352–7.
[29] Clark L, Watson D. Construction validity: basic issues in objective scale development. Psychol Assess 1995;7:309–19.
[30] Lasch KE, Marquis P, Vigneux M, Abetz L, Arnould B, Bayliss M, et al. PRO development: rigorous qualitative research as the crucial foundation. Qual Life Res 2010;19:1087–96.
[31] Jordan JE, Buchbinder R, Osborne RH. Conceptualising health literacy from the patient perspective. Patient Educ Couns 2010;79:36–42.
[32] Patton M. Qualitative research and evaluation methods, 3rd ed., California: Sage Publications; 2002.
[33] Strauss A, Corbin J. Basics of qualitative research – techniques and procedures for developing grounded theory; 1998.
[34] Trochim W. An introduction to concept mapping for planning and evaluation. Eval Program Plann 1989;12:1–16.
[35] Batterham R, Southern D, Appleby N, Elsworth G, Fabris S, Dunt D, et al. Construction of a GP integration model. Social Sci Med 2002;54:1225–41.
[36] Osborne RH, Elsworth GR, Whitfield K. The health education impact questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007;66:192–201.
[37] Osborne RH, Norquist JM, Elsworth GR, Busija L, Mehta V, Herring T, et al. Development and validation of the influenza intensity and impact questionnaire (FluiiQ). Value Health 2011;14:687–99.
[38] Buchbinder R, Batterham R, Elsworth G, Dionne CE, Irvin E, Osborne RH. A validity-driven approach to the understanding of the personal and societal burden of low back pain: development of a conceptual and measurement model. Arthritis Res Ther 2011;13:R152.
[39] Australian Bureau of Statistics. Socioeconomic indexes for areas – state suburb code. Canberra: Commonwealth of Australia; 2006.
[40] Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. The extent and breadth of benefits from participating in chronic disease self-management courses: a national patient-reported outcomes survey. Patient Educ Couns 2007;65: 351–60.
[41] Barber MN, Staples M, Osborne RH, Clerehan R, Elder C, Buchbinder R. Up to a quarter of the Australian population may have suboptimal health literacy depending upon the measurement tool: results from a population-based survey. Health Promot Int 2009;24:252–61.
[42] Trochim W, Cook JA, Setze RJ. Using concept mapping to develop a conceptual framework of staff's views of a supported employment program for persons with severe mental illness. J Consult Clin Psychol 1994;62:766–75.
[43] Southern D, Young D, Dunt D, Appleby N, Batterham R. Integration of primary health care services: perceptions of Australian general practitioners, non-genteral practitioner health service providers and consumers at the general practice – primary care interface. Eval Program Plann 2002;25: 47–59.
[44] Miles M, Huberman M. Qualitative data analysis: an expanded sourcebook; 1994.
[45] Huberman M, Miles M. Drawing valid meaning from qualitative data: some techniques of data reduction and display. Qual Quant 1983;17:281–339.
[46] Muthen L, Muthen B. MPlus user's guide; 1998–2007.
[47] Brown TA. Confirmatory factor analysis for applied research. New York: Guilford Press; 2006.
[48] Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 1960;20:141–51.
[49] Cattell RB. The Scree test for the number of factors. Multivar Behav Res 1966;1:245–76.
[50] Rust J, Golombok S. Modern psychometrics: the science of psychological assessment; 1999.
[51] Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equation Model 1999;6: 1–55.
[52] Marsh H, Balla J, McDonald R. Goodness-of-Fit indexes in confirmatory factor analysis: the effect of sample size. Psychol Bull 1988;103:391–410.
[53] Joreskog K. Testing structural equation models. In: Bollen KL, Long JC, editors. Testing structural equation models. Newbury Park, CA: Sage; 1993.
[54] Muthen B. Goodness of Fit with categorical and other non-normal variables. In: Bollen KL, Long JS, editors. Testing structural equation models. Newbury Park, CA: Sage; 1993.
[55] Beatty P, Willis G. Research synthesis: the practice of cognitive interviewing. Public Opin Q 2007;71:287–311.
[56] Gallo JJ, Anthony J, Muthén B. Age differences in the symptoms of depression: a latent trait analysis. J Gerontol B Psychol Sci Soc Sci 1994;49:251–64.
[57] Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? J Clin Epidemiol 1992;45:1341–5.
[59] Cockburn JPF, Sanson-Fisher R. Construction and validation of a questionnaire to measure the health beliefs of general practice patients. Fam Pract 1987;4:108–16.
[60] Cook S. News: Health secretary announces end to GP catchment areas. Brit J Med 2009;339:b3860.