

Original Paper

# A Multidimensional Tool Based on the eHealth Literacy Framework: Development and Initial Validity Testing of the eHealth Literacy Questionnaire (eHLQ)

Lars Kayser<sup>1</sup>, MD, PhD; Astrid Karnoe<sup>1,2</sup>, MSc (Health Informatics); Dorthe Furstrand<sup>1,3</sup>, MD, MSc (Health Informatics); Roy Batterham<sup>4,5</sup>, BAppSc, MEd; Karl Bang Christensen<sup>1</sup>, PhD; Gerald Elsworth<sup>4</sup>, PhD; Richard H Osborne<sup>1,4</sup>, PhD

<sup>1</sup>Department of Public Health, University of Copenhagen, Copenhagen, Denmark

<sup>2</sup>The Danish Multiple Sclerosis Society, Valby, Denmark

<sup>3</sup>Danish Cancer Research Center, The Danish Cancer Society, Copenhagen, Denmark

<sup>4</sup>Health Systems Improvement Unit, Centre for Population Health Research, Faculty of Health, Deakin University, Geelong, Australia

<sup>5</sup>Faculty of Public Health, Thammasat University, Bangkok, Thailand

**Corresponding Author:**

Lars Kayser, MD, PhD  
Department of Public Health  
University of Copenhagen  
Øster Farimagsgade 5  
Copenhagen, DK-1014  
Denmark  
Phone: 45 28757291  
Email: [lk@sund.ku.dk](mailto:lk@sund.ku.dk)

## Abstract

**Background:** For people to be able to access, understand, and benefit from the increasing digitalization of health services, it is critical that services are provided in a way that meets the user's needs, resources, and competence.

**Objective:** The objective of the study was to develop a questionnaire that captures the 7-dimensional eHealth Literacy Framework (eHLF).

**Methods:** Draft items were created in parallel in English and Danish. The items were generated from 450 statements collected during the conceptual development of eHLF. In all, 57 items (7 to 9 items per scale) were generated and adjusted after cognitive testing. Items were tested in 475 people recruited from settings in which the scale was intended to be used (community and health care settings) and including people with a range of chronic conditions. Measurement properties were assessed using approaches from item response theory (IRT) and classical test theory (CTT) such as confirmatory factor analysis (CFA) and reliability using composite scale reliability (CSR); potential bias due to age and sex was evaluated using differential item functioning (DIF).

**Results:** CFA confirmed the presence of the 7 a priori dimensions of eHLF. Following item analysis, a 35-item 7-scale questionnaire was constructed, covering (1) using technology to process health information (5 items, CSR=.84), (2) understanding of health concepts and language (5 items, CSR=.75), (3) ability to actively engage with digital services (5 items, CSR=.86), (4) feel safe and in control (5 items, CSR=.87), (5) motivated to engage with digital services (5 items, CSR=.84), (6) access to digital services that work (6 items, CSR=.77), and (7) digital services that suit individual needs (4 items, CSR=.85). A 7-factor CFA model, using small-variance priors for cross-loadings and residual correlations, had a satisfactory fit (posterior predictive  $P$  value: .27, 95% CI for the difference between the observed and replicated chi-square values: -63.7 to 133.8). The CFA showed that all items loaded strongly on their respective factors. The IRT analysis showed that no items were found to have disordered thresholds. For most scales, discriminant validity was acceptable; however, 2 pairs of dimensions were highly correlated; dimensions 1 and 5 ( $r=.95$ ), and dimensions 6 and 7 ( $r=.96$ ). All dimensions were retained because of strong content differentiation and potential causal relationships between these dimensions. There is no evidence of DIF.

**Conclusions:** The eHealth Literacy Questionnaire (eHLQ) is a multidimensional tool based on a well-defined a priori eHLF framework with robust properties. It has satisfactory evidence of construct validity and reliable measurement across a broad range

of concepts (using both CTT and IRT traditions) in various groups. It is designed to be used to understand and evaluate people's interaction with digital health services.

(*J Med Internet Res* 2018;20(2):e36) doi: [10.2196/jmir.8371](https://doi.org/10.2196/jmir.8371)

## KEYWORDS

eHealth; health literacy; computer literacy; questionnaire design

## Introduction

Modern health promotion and health care with increasing digitalization of information and services have become increasingly challenging for both community members and providers [1,2]. Community members can be delivered a panoply of messages from many media and may have access to information from a rapidly growing World Wide Web of information and service providers [3].

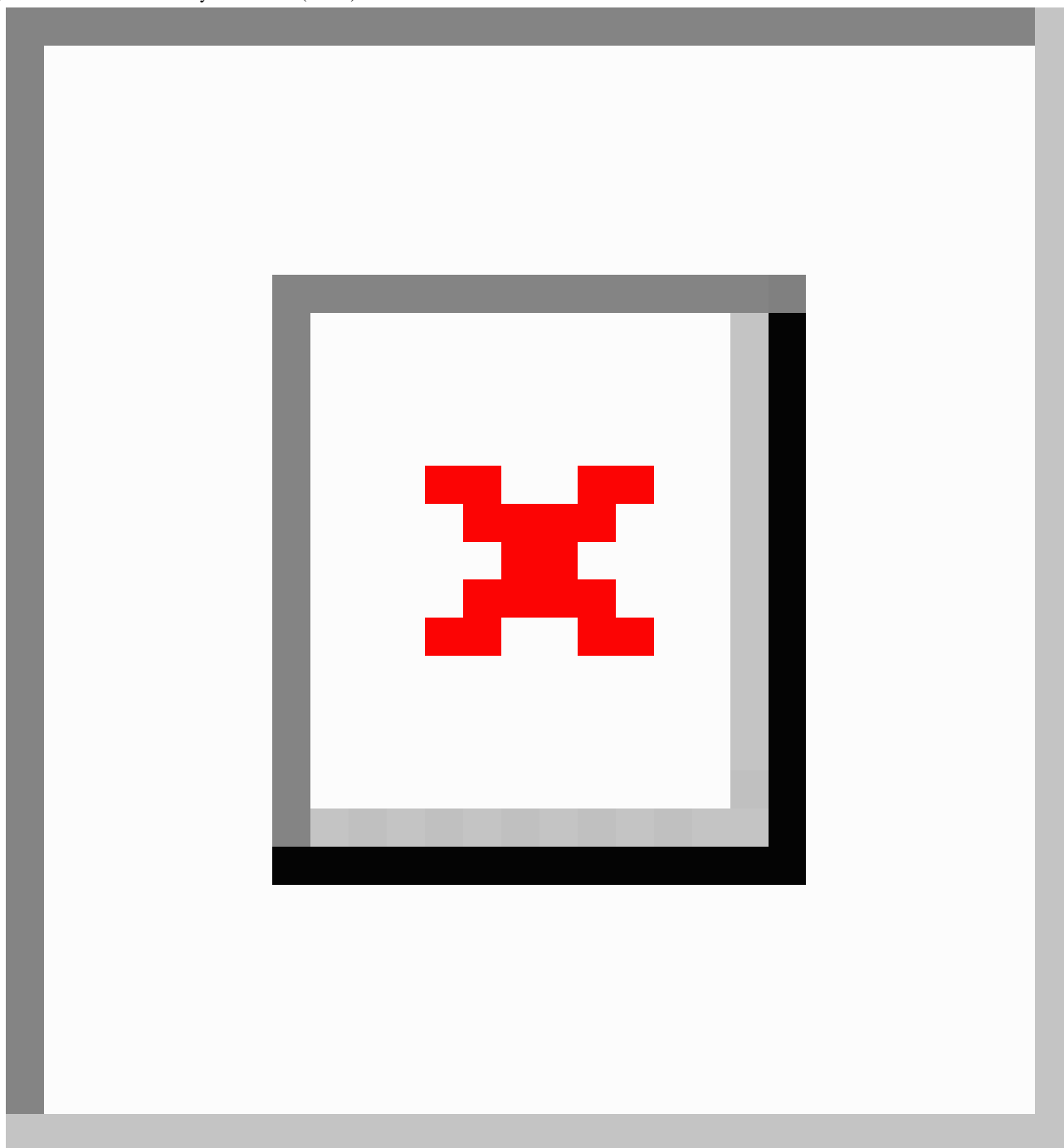
For people to be able to effectively and equitably access health services, it is critical that such services are provided in a way that generates appropriate actions and that the recipient benefits in the intended way. If people have a range of health literacy limitations, that is, limitations across “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” [4], they are at risk of having reduced access to care, poor self-management skills, increased hospitalization, and decreased life span [5].

With the increasing digitalization of health care through electronic services, including health portals and health records, which are accessed by people from their homes, a new level of complexity has been added to the ways health care systems and the community have to interact.

The increased complexity demands a range of digital competencies among users, and this then calls for new ways to describe and evaluate users' digital capabilities and experiences in this rapidly changing health context.

Consequently, the concept of eHealth literacy (or digital health literacy) has emerged. Norman and Skinner (2006) described it as “the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem” [6]. Norman and Skinner's concept has since been updated by others to include elements related to users' cognitive skills [7,8] and dimensions such as communication, cultural elements, and social elements [9]. Despite increasing interest in this eHealth literacy concept, limited evidence exists regarding whether addressing people's eHealth literacy improves health outcomes [10]. This may be due to either a lack of appropriate instruments to measure eHealth literacy or, until now, low adoption of technology when providing health care [10,11].

In 2015, we identified a new concept for eHealth literacy: a model based on systematic and inductive methods that sought to identify the full range of elements relevant to individuals attempting to understand and use eHealth technologies and digital services [12].

**Figure 1.** The eHealth literacy framework (eHLF).

This model, the eHealth Literacy Framework (eHLF), consists of 7 dimensions that describe the attributes of the users (information and knowledge about their health); the intersection between users and the technologies (their feeling of being safe and in control and their motivation); and users experience of systems (they work and are accessible, and suits users' needs). The eHLF provides a comprehensive map of the individual technology user's health literacy that covers his or her knowledge and skills, the system, and how the individual interacts with the system (see [Figure 1](#)).

The eHLF was specifically designed to inform the development of a conceptually and psychometrically sound questionnaire measure of eHealth literacy. The aim of this study was to create

and test items and scales that capture the 7 eHLF dimensions using the validity-driven approach [13].

## Methods

### Initial Development

Development of the eHLQ followed the validity-driven approach [13] that has been used to develop several widely used and highly robust questionnaires [14-16]. The objective of the eHLQ development was to create an instrument that captures the 7 hypothesized dimensions of eHLF. Items were based on 450 statements obtained from Norgaard et al's (2015) conceptual development of eHLF [12]. These statements had been collected during 8 concept mapping sessions, which included participation

of patients (recruited from patient organizations and primary care clinics), computer scientists, academics, and health professionals [12]. Items for the questionnaire were simultaneously written in Danish and English as recommended by Eremenco (2005) [17] to avoid words or phrases that may be difficult to translate into other languages. To provide a rich range of candidate items for the construction of each scale, up to 9 items were drafted for each of the 7 eHLF dimensions. The statements were written to relate to a 4-point set of response options: strongly disagree, disagree, agree, and strongly agree, assigned a value of 1 to 4, respectively. Items were also written to represent a range of difficulty that were broadly guided by Bloom's taxonomy (ie, related to knowledge and remembering, understanding, applying, analyzing, and evaluating) [18] so that the full range of the construct could be covered. The items were extensively discussed in Australian, Danish, and other contexts by the multidisciplinary team whose members have extensive experience in writing questionnaire items and applying them across cultures.

Danish draft items were tested in 7 cognitive interviews to check whether respondents understood the items as intended [19]. The cognitive testing involved initial administration of items using paper and pen format with careful observation of each respondent. The interviewer then reviewed items with the respondent and asked specific questions about items the respondent had hesitated on or appeared to have found difficult in answering. Respondents were asked "What were you thinking about when you were answering that question?" This process elicited the cognitive process behind the answers. A prompt was used if needed: "Why did you select that response option?" The items were adjusted based on the inputs with a particular focus on how the informants understood terms or concepts in relation to providers of health care and technologies.

### Recruitment of Participants

Respondents were recruited from a wide range of sociodemographic settings to broadly represent the targets for the application of the questionnaire in the future. Individuals were included if they were above age 18 years and able to read or understand Danish. Potential respondents were randomly approached by trained interviewers in a variety of locations in the broader community, such as in libraries, private sector workplaces, a hospital, nursing homes, health centers, and an outpatient clinic. To ensure inclusion of people who may have low literacy, potential respondents were given the option of completing the questionnaire themselves or to have it read aloud in an interview. If respondents did not have time to finish the questionnaire, they were encouraged to complete it at home and were provided with a reply-paid envelope. They also had the option of completing a Web-based questionnaire.

Demographic data including age, sex, educational background, self-reported health condition, and presence of chronic conditions were also collected to evaluate whether the resulting scales were invariant to these exogenous factors and thus provided unbiased estimates of mean differences across these groups.

The administration of the questionnaires also included the administration of a validation version of the eHealth literacy

assessment toolkit, which is reported elsewhere (personal communication, Karnoe 2017). Respondents did not receive any payment for filling out the questionnaire.

### Item Analysis and Selection

The first step in the analysis for the items was to examine item characteristics. It was intended that each scale would have the smallest number of items necessary to capture the meaning of the construct in the most efficient manner while ensuring adequate coverage of the construct. Each item and each set of items forming a scale, as well as associations between and across items and scales, were tested using psychometric procedures afforded by the conventions of both classical test theory [20] and item response theory (IRT) [21].

### Statistical Analysis

Descriptive statistics were generated for each item to determine missing values, floor and ceiling effects, interitem correlations, correlation with scale score, scale reliability (composite reliability, Cronbach alpha, person separation index), estimated item location, and *P* value for item fit tests. Given that hypothesized constructs were specified a priori within the eHLF [12], confirmatory factor analysis (CFA) was used.

Results from each of these analyses were used to assist with decisions about optimizing the number of items in each scale. Of central importance to the item deletion retention strategy was ensuring that the retained items properly represented the intended a priori construct from the eHLF. Items that performed poorly on the above criteria were earmarked for deletion. We also sought to generate scales that had a minimum reliability of .8 but where items had no excessively high interscale correlations, violations of local independence, or high correlated residuals. Where the content of items had substantial overlap, and tended to perform well on a range of indicators, the item selection strategy then included criteria to improve the diversity of item locations, that is, selection of a range of items within a scale that range across all difficulty levels of the construct that the scale measures.

CFA was conducted with Mplus versions 7.4 and 8 (Muthén & Muthén, Los Angeles, CA, USA) using Bayesian Structural Equation Modeling (BSEM) [22,23]. BSEM is a specific application of Bayesian statistical analysis to factor analysis and structural equation modeling. In a conventional CFA usage, for example, maximum likelihood analysis, model parameters that represent possible cross-loadings and correlations between item residuals are typically set to be exactly zero. When these restrictions are placed on CFA models used to confirm, a multifactor structure poor model fit often results, and a large number of stepwise modifications (estimation of parameters originally constrained to zero) can be required to achieve a reasonably fitting result. This approach may not, however, necessarily lead to a single solution depending on the sequence of choices of modifications made along the way, and there is frequently a strong upward bias in the estimation of interfactor correlations, leading to erroneous conclusions about scale discriminant validity. By using small-variance priors, BSEM allows models to be fitted that have the flexibility to estimate small variations from these strictly zero constraints. For the

final BSEM analysis, after a little systematic experimentation ([Multimedia Appendix 1](#)), the variance of the Bayesian priors for the cross-loadings was set at 0.01 such that there was a 95% probability that the cross-loadings would be within the range  $\pm 0.20$ . Similarly, the variance for the residual correlations was set to give a 95% probability that the correlations were within the range of  $\pm 0.2$  [22].

The approach to model fit in BSEM differs from that in conventional CFA. The conventional CFA fit indices (eg, chi-square, comparative fit index, and root mean square error of approximation) are not used. Rather, fit in BSEM is assessed using a procedure called “posterior predictive checking” that results in a “posterior predictive *P* value” (PPP value). In a very well-fitting model, the PPP value is expected to be close to .5, whereas a value of  $< 0.05$  (or  $< 0.10$  or  $< 0.01$  if less or greater stringency is applied) is typically regarded as indicating unsatisfactory fit. A PPP value of  $> 0.05$  was chosen as the threshold for satisfactory fit in this study. Additionally, a fit statistic derived from the Bayesian model-fitting process is calculated with 95% CIs. A pattern of symmetrical upper and lower CIs centered around zero indicates excellent fit, whereas a lower 95% CI that is positive suggests that the fit is not satisfactory [22–24].

The IRT model-based evaluation of item fit used a unidimensional IRT model—the generalized partial credit model (GPCM) [25]. For each scale, we compared the observed item scores with the scores expected under the model using an implementation in SAS statistical software version 9.4 (SAS Institute, Cary, NC, USA) [26]. This provides a graphical test of item fit where the item’s mean score is plotted against the group’s scale mean and evaluated against values.

For each item, the estimated item discrimination parameter (the IRT-equivalent of a factor loading) and the estimated item location (computed as the average of the item thresholds) were reported.

Differential item functioning (DIF) is a statistical characteristic of an item indicating the extent to which the item can be said to measure the same construct across subpopulations [27]. We tested DIF using logistic regression techniques [28–30] addressing the elevated risk of type I error due to the larger number of tests performed and by the Benjamini and Hochberg procedure [31] controlling the false discovery rate results.

### Ethical Considerations

According to Danish law, when survey-based studies are undertaken in accordance with the Helsinki Declaration, specific approval by an ethics committee and written informed consent are not required. Potential respondents were provided with information about the survey and its purpose, including that participation was voluntary. The completion of the survey by participants was then considered to be implied consent.

## Results

### Item Construction and Refinement

Between 7 and 9 draft items per scale were generated (58 items in total). Cognitive testing was undertaken with 7 individuals

(4 women), aged between 16 and 74 years from different cultural backgrounds and with varying educational levels. The interviews found that almost all items were understood as intended; 1 item was removed due to unclear text, whereas only some small refinements were made to other items to improve clarity and simplicity. The refinements mainly related to getting the clearest possible Danish words related to the core concept of technology, digital tools, or electronic tools. The term “digital” was preferred across demographic groups. Moreover, the best Danish term to express “health” and “health conditions” was tested, and the term “health” was found to work the best. The final number of items for going to the field was 57.

### Item Number Reduction and Scale Construction

The refined items were randomized and administered to 475 individuals—100 outpatients who all filled in a paper version. Out of the 375 people in the community a total of 328 filled in the paper version and 47 filled in the digital version.

The hypothesized 7-factor BSEM model for the initial 57 items showed a satisfactory fit to the data (PPP value .79, 95% CI for the difference between the observed and replicated chi-square values  $-239.75$  to  $99.85$ ). A total of 10 items had low ( $< 0.5$ ) standardized factor loadings; however, 3 items showed evidence of factorial complexity having cross-loadings  $> 0.25$ . There were also 15 pairs of items with residual correlations  $> 0.30$ . Of these item pairs, 12 correlated residuals were between items that were associated with different target factors, suggesting that, from an IRT perspective, the assumption of local independence of the hypothesized scales was largely satisfied.

After inspection of the results of the initial BSEM analysis and the parallel IRT analyses, items that performed poorly or were clearly redundant to others (ie, had highly similar content and measurement properties) across analyses were iteratively removed, resulting in 35 items in 7 scales comprising 4, 5, or 6 items in each (see [Multimedia Appendix 2](#)).

A final BSEM model ([Multimedia Appendix 1](#)) was then fitted to the data for these items. Model fit was satisfactory (PPP value .27, 95% CI for the difference between the observed and replicated chi-square values  $-63.7$  to  $133.8$ ). With 4 exceptions, all factor loadings were  $> 0.50$  (see [Multimedia Appendix 3](#)). All loadings  $< 0.5$  were  $> 0.4$ .

There were no statistically significant cross-loadings, and there were 8 residual correlations  $\geq 0.3$ . All but one of these larger residual correlations was negative with 4 associated with scale (1) using technology to process health information. The only positive residual correlation  $\geq 0.3$  was between 2 contiguous items, former 44 and 45. They were, therefore, separated and ended up in the final version as item numbers 26 and 35. There were no within-scale positive residual correlations  $\geq 0.3$ .

The fit of the GPCM to the data was excellent in all of the dimensions ([Multimedia Appendix 4](#)).

Estimates of composite scale reliability are also shown in [Multimedia Appendix 3](#) (95% CIs in parentheses, and, for comparison, Cronbach alpha) for the final 7 scales. All 7 scales had acceptable composite reliability ( $> 0.7$ ), and 5 scales have a



reliability above .8; however, scales 2 and 6 had somewhat lower reliability than intended (.75 and .77, respectively).

Interfactor correlations ranged from .31 (factors 3 and 4) to .97 (factors 6 and 7) with the next highest being .96 (factors 1 and 5), suggesting satisfactory discrimination between the majority of the scales with the exception of the following: (6) access to digital services that work; (7) digital services that suit individual needs; (1) using technology to process health information; and (5) motivated to engage with digital services (Table 1).

The DIF analysis showed no evidence of influence of age or sex on the item scores. The item locations, item discriminations, and factor loadings are shown in Multimedia Appendix 3. As expected, the order of factor loadings and the order of the item discriminations are very similar. The spread in the item locations indicate that, for each scale, the items represent a range of difficulty, indicating that the full range of the construct is covered. The analyses generated scales that were well targeted

to the respondent sample (Table 1) and showed good distributional properties (Table 2).

An analysis of the construct representation, that is, the completeness of the match between the intended construct (first column in Multimedia Appendix 2) and the content of the resulting scales (second column, Multimedia Appendix 3) indicated strong concordance between the intended and generated constructs for 5 of the 7 constructs. For construct 1, ability to process information, the specific elements of fundamental reading, writing, and cognitive ability were missing; however, they were represented through higher-order functions such as application and use of such skills. This scale was renamed as (1) using technology to process health information. For construct 2, engagement in own health, the surviving items focused more on whether the respondents perceived that they had adequate knowledge and understanding of their health in general terms. The intended construct of eHLF had a broader focus on awareness and engagement in and utilization of information and knowledge about health.

**Table 1.** Interfactor correlation between the eHealth Literacy Questionnaire scales.

Factor name	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Factor 2. Understanding of health concepts and language	0.69					
Factor 3. Ability to actively engage with digital services	0.90	0.61				
Factor 4. Feel safe and in control	0.36	0.49	0.31			
Factor 5. Motivated to engage with digital services	0.95	0.61	0.83	0.47		
Factor 6. Access to digital services that work	0.77	0.57	0.73	0.69	0.83	
Factor 7. Digital services that suit individual needs	0.78	0.45	0.74	0.58	0.84	0.96

**Table 2.** eHealth Literacy Questionnaire scales and descriptive statistics.

No	Scale	n (%) (N=475)	Mean (SD)	Median (IQR <sup>a</sup> )
1	Using technology to process health information	462 (97.3)	2.55 (0.66)	2.60 (2.20-3.00)
2	Understanding of health concepts and language	466 (98.1)	2.97 (0.55)	3.00 (2.60-3.40)
3	Ability to actively engage with digital services	465 (97.9)	2.81 (0.69)	2.80 (2.40-3.20)
4	Feel safe and in control	466 (98.1)	2.61 (0.66)	2.60 (2.20-3.00)
5	Motivated to engage with digital services	466 (98.1)	2.55 (0.65)	2.60 (2.00-3.00)
6	Access to digital services that work	466 (98.1)	2.52 (0.55)	2.50 (2.17-2.83)
7	Digital services that suit individual needs	457 (96.2)	2.42 (0.62)	2.33 (2.00-3.00)

<sup>a</sup>IQR: interquartile range.

The construct included, potentially, a very wide range of elements around physiological functions, risk factors, and elements of the health care system. This disparate range of elements would require an inventory to capture in complete breadth of the construct; however, the eHLQ items generated and surviving the item reduction phase, captured a somewhat higher order assessment of the respondent's understanding and engagement in health information, which is more suitable for a psychometric scale, rather than an inventory. The scale was renamed as (2) understanding of health concepts and language.

## Discussion

### Principal Findings

We sought to develop and test a new measure of eHealth literacy using both classical and modern psychometric approaches to questionnaire development. Using a robust conceptual model, developed through extensive local and international community consultation [12] and through a validity-driven approach [12,13], 7 new psychometrically robust scales to comprehensively measure the broad concept of eHealth literacy were developed. Construction and validity testing in a broad range of target

groups generated clear evidence of construct validity, discriminant validity, and scale reliability. This initial validity testing indicates that the eHLQ is likely to be valuable for the characterization and understanding of digital health literacy.

This research introduced some highly rigorous and innovative elements to questionnaire development and validation. First, the data to generate the eHLF model were obtained using concept mapping in 2 cultures (Danish and English) and through an international e-consultation. Concept mapping has been found to be a highly robust technique for developing conceptual models and for questionnaire development [32]. Additionally, when the items were written, they were simultaneously constructed in Danish and English. With this step, we sought to uncover and remove idiomatic expressions in either language so that when it is translated to further languages, only concrete concepts are presented for translation with fidelity and to support international equivalence of the constructs. We also employed extensive and rigorous application of the two, often opposing, traditions of psychometric analysis: CTT and IRT. The items surviving the scrutiny of these procedures, as well as from rigorous qualitative methods, have satisfactory psychometric properties in all the scales.

The eHLQ is now ready for application, and for further testing, in a wide range of settings and purposes; these include the following:

- Evaluation of interventions. The eHLQ will provide insight into users' perceptions and experiences when using digital health solutions. As with previous tools developed using the same methods as the eHLQ [14,15], we expect that, with further testing, it will be a robust patient-reported outcome measure and sensitive to change.
- Implementation and adoption of digital health services. We also expect the scales to provide insights into why digital health services implementations work or fail (ie, understanding intermediate or process outcomes). Given that the scales cover individual user attributes, attributes of the system, and the interaction between the two, the eHLQ is expected to uncover mechanisms that determine adoption outcomes.
- Community and population surveys. The eHLQ is expected to be useful for local, regional, and population surveys. This information will inform policy makers, program managers, and service providers about the profile of needs and strengths of individuals across the population and demographic subgroups.

The eHLQ also offers an opportunity to stratify users for inclusion in design processes [33] to explore the reactions of these groups to new digital solutions or, more generally, to document the strengths and weaknesses among user subgroups. Information about subgroups with respect to their eHLQ profile (of strengths and weaknesses) and other characteristics may be used to design interventions or inform designers through creation of archetypes or personas as described in the Ophelia (OPTimising HEalth LIteracy and Access) process developed for service redesign and development of person-centered services [34,35].

A further important element of scale construction was the consideration of within-construct concept representation and item difficulty. The items we drafted sought to cover the full range of issues within a construct and to achieve a spread of item difficulties from items being easy to answer to items being hard to score highly. The statistical analyses demonstrate that these demanding targets were broadly achieved. Some subconcepts identified in the concept mapping, and detailed in the eHLF, did not survive the item development and testing process, and therefore, some scales are not as broad as initially intended. The requirement for being faithful to the codesign outcomes (the eHLF), broad item difficulties, and meeting the requirements of the 2 psychometric traditions has generated a tool that robustly captures the concept of eHealth literacy, but with some minor subcomponents underrepresented. If researchers and developers wish to capture the omitted subelements (basic functional health literacy or broader issues around knowledge and engagement), other tools should be used to supplement the eHLQ, such as the eHealth literacy assessment toolkit (personal communication, Karnoe 2017), or the digital health literacy instrument (DHLLI) by van der Vaart [36].

We found that there are 2 particularly high interfactor correlations between factors 1 and 5 ( $r=.95$ ) and factors 6 and 7 ( $r=.96$ ). The eHLF contains 7 dimensions, all of which we sought to include in the final eHLQ model. Content analysis, the careful deliberations while developing items, and the original views of patient and professional groups indicated that each of the original constructs were different. The high correlations may be a result of them being located along the same causal path. When considering the content of these scales, we expect that different interventions will result in different patterns of change in the pairs of related scales. This should be carefully monitored over time and across different settings.

The psychometric and construct representation demands we placed on the eHLQ construction process were further compounded by the need for the eHLQ to be a relatively short questionnaire. Importantly, all of the scales have acceptable reliability, despite having only 4 to 6 items. Given that the eHLQ is intended for application among people with low literacy and they may be ill, it is critical that the smallest possible number of items be included. Every scale had satisfactory loadings on its intended items, with negligible cross-loadings on other items. Although we had hoped all scales would have a reliability  $\geq .8$ , this was not quite achieved for 2 scales ( $\geq .75$  for both). Importantly, this level of reliability is acceptable for research and evaluation purposes.

Future concurrent validity and other validity tests, including predictive validity tests, will be valuable. For the most part, we have found that the concept mapping, and subsequent qualitative studies, ensures that the validity of the data the questionnaire generates is robust. The eHEALS [37] appears to be the best current measure of some elements of eHealth literacy, and it will be valuable for future researchers to compare the tools.

Moreover, future research should include further quantitative and qualitative work to develop interpretation and use arguments [38] to generate comprehensive validity testing of the eHLQ

across a wide range of contexts to support current and future users understand the data generated by the eHLQ.

The eHLQ provides a wider range of dimensions of eHealth literacy than previous tools. It covers not only an individual's competences, as in the Lily model [6] and the van der Vaart DHLI model [36], but also an individual's experiences and interactions with technologies and services. It provides a much richer understanding of the whole of the system through providing perspectives of individuals, the system, and the interaction between the two.

Recent study by van der Vaart [36] specifically expanded the digital health field by creating 7 areas that mainly focus on individual competencies. These are operational skills, navigation skills, information searching, evaluating reliability, determining relevance, adding self-generated content, and protecting privacy.

The introduction of the eHLQ's scales covering user interaction and the user's experience of engaging with the system is an important innovation for the rapidly growing digital health field. The eHLQ has the potential to provide insight into the maturity of a country's digital services. With mature digital health services, we expect that individuals find the system more responsive to their needs, and thus can engage more fully in support of achieving health and equity. This is akin to the new health concept of health literacy responsiveness [39]. We expect that in a digitally mature society, the scales covering the system and the interaction will reflect stronger: (6) access to digital services that work and (7) digital services that suit individual needs. In settings with developing systems, where there are few digital services or limited or piecemeal coverage and access, the scores on these scales are expected to reflect major challenges for individuals.

## Acknowledgments

The authors would like to thank Mr Ole Norgaard for permission granting the use of statements sampled in the process of creating the eHLF for the development of eHLQ. The authors wish to thank the sociology students who administered questionnaires and Gentofte Hospital for giving access to administer questionnaires in their outpatient clinic. Ms Emily Duminski is thanked for help in preparing the manuscript for submission. The study has received funding from the Danish Health Foundation (grant 15-B-0237). DF is a PhD fellow from the Danish Cancer Society and is currently supported by the Danish foundation TrygFonden. AK is a PhD fellow at the Danish Multiple Sclerosis Society and is also supported by the Innovation Fund Denmark. RHO is funded in part through a National Health and Medical Research Council (NHMRC) of Australia Senior Research Fellowship #APP1059122. The eHLQ is available through license by either Deakin University or by University of Copenhagen. At the discretion of Deakin and Copenhagen Universities, licenses are provided for free for noncommercial use ([Multimedia Appendix 5](#)).

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Fit statistics and summary model parameters for six Bayesian structural equation models with varying prior variances for cross-loadings and correlated residuals—model 3 chosen as a satisfactory solution.

[\[PDF File \(Adobe PDF File\), 245KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

The 7 constructs of the eHealth Literacy Framework and derived scales of the eHealth Literacy Questionnaire.

It is important to note that eHealth literacy, like the concept of health literacy, is both a reflection of the individual's knowledge and the skills he or she may employ in the cultural, social, and institutional context in which they are engaging in; therefore, it is critical to assess these domains simultaneously [9,12,40]. As we see wide-scale improvements in digital health services, we would expect to see the user interaction scales improve—that is, (4) feel safe and in control and (5) motivated to engage with digital services.

The eHLQ has already generated substantial interest in the field. It is currently being translated into Chinese, Norwegian, and Czech, and a German-speaking country initiative is under way. The English version is also undergoing validity testing. The conceptual model and the range of intended applications fit well with a wide range of current policy initiatives. These include the World Health Organisation (WHO) People Centred Health Services Framework [1] and the WHO Shanghai Declaration, which places health literacy as one of the three key elements to support the achievement of the sustainable development goals by 2030 [41]. To reach these ambitious goals, eHealth literacy responsive systems need to be implemented on a wide scale to leave no one behind. The eHLQ is currently in the field and is available with the authors.

## Conclusions

The eHLQ is a psychometrically robust multidimensional instrument with 7 scales that comprehensively cover all 7 dimensions of the eHLF. The eHLQ and the eHLF's conceptual underpinnings are likely to be a useful set of tools to support researchers, developers, designers, and governments to develop, implement, and evaluate effective digital health interventions.



[\[PDF File \(Adobe PDF File\), 178KB-Multimedia Appendix 2\]](#)

### Multimedia Appendix 3

The estimated item locations, item discriminations, and factor loadings.

[\[PDF File \(Adobe PDF File\), 213KB-Multimedia Appendix 3\]](#)

### Multimedia Appendix 4

Illustration of the fit of items applying item response theory and the generalized partial credit model. The observed items scores are compared with the scores expected under the model.

[\[PDF File \(Adobe PDF File\), 679KB-Multimedia Appendix 4\]](#)

### Multimedia Appendix 5

eHLQ Licence Agreement.

[\[PDF File \(Adobe PDF File\), 141KB-Multimedia Appendix 5\]](#)

## References

1. World Health Organization. 2016. WHO Framework on integrated people-centred health services URL: <http://www.who.int/servicedeliverysafety/areas/people-centred-care/en/> [accessed 2016-08-17] [WebCite Cache ID 6v6v9KNAf]
2. European Commission. 2016 Dec. Blueprint for a digital transformation of health and care in an ageing society URL: <https://ec.europa.eu/digital-single-market/en/blueprint-digital-transformation-health-and-care-ageing-society> [accessed 2018-01-08] [WebCite Cache ID 6wJfCdnD9]
3. European Commission. Brussels: Flash Eurobarometer 404 - TNS Political & Social; 2014. European citizens' digital health literacy URL: [http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl\\_404\\_sum\\_en.pdf](http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl_404_sum_en.pdf) [accessed 2018-01-08] [WebCite Cache ID 6wJg2K7Vu]
4. Nutbeam D. Health promotion glossary. *Health Promot Int* 1998 Jan 01;13(4):349-364. [doi: [10.1093/heapro/13.4.349](https://doi.org/10.1093/heapro/13.4.349)]
5. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med* 2011 Jul 19;155(2):97-107. [doi: [10.7326/0003-4819-155-2-201107190-00005](https://doi.org/10.7326/0003-4819-155-2-201107190-00005)] [Medline: [21768583](https://pubmed.ncbi.nlm.nih.gov/21768583/)]
6. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006 Jun 16;8(2):e9 [FREE Full text] [doi: [10.2196/jmir.8.2.e9](https://doi.org/10.2196/jmir.8.2.e9)] [Medline: [16867972](https://pubmed.ncbi.nlm.nih.gov/16867972/)]
7. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. *J Med Internet Res* 2011 Nov 17;13(4):e94 [FREE Full text] [doi: [10.2196/jmir.1750](https://doi.org/10.2196/jmir.1750)] [Medline: [22094891](https://pubmed.ncbi.nlm.nih.gov/22094891/)]
8. Chan CV, Mirkovic J, Furniss S, Kaufman DR. eHealth literacy demands and cognitive processes underlying barriers in consumer health information seeking. *Knowl Manag E-Learn Int J KMEL* 2015 Dec 26;7(4):550-575 [FREE Full text]
9. Gilstad H. Toward a comprehensive model of eHealth literacy. 2014 May 19 Presented at: Proceedings of the 2nd European Workshop on Practical Aspects of Health Information; May 19-20, 2014; Trondheim, Norway URL: <http://ceur-ws.org/Vol-1251/paper7.pdf>
10. Karnoe A, Kayser L. How is eHealth literacy measured and what do the measurements tell us? A systematic review. *Knowl Manag E-Learn Int J KMEL* 2015 Dec 26;7(4):576-600 [FREE Full text]
11. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. *J Med Internet Res* 2011 Nov 9;13(4):e86 [FREE Full text] [doi: [10.2196/jmir.1840](https://doi.org/10.2196/jmir.1840)] [Medline: [22071338](https://pubmed.ncbi.nlm.nih.gov/22071338/)]
12. Norgaard O, Furstrand D, Klokke L, Karnoe A, Batterham RW, Kayser L, et al. The e-health literacy framework: a conceptual framework for characterizing e-health users and their interaction with e-health systems. *Knowl Manag E-Learn Int J KMEL* 2015 Dec 26;7(4):522-540 [FREE Full text]
13. 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(5):R152 [FREE Full text] [doi: [10.1186/ar3468](https://doi.org/10.1186/ar3468)] [Medline: [21933393](https://pubmed.ncbi.nlm.nih.gov/21933393/)]
14. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013 Jul 16;13:658 [FREE Full text] [doi: [10.1186/1471-2458-13-658](https://doi.org/10.1186/1471-2458-13-658)] [Medline: [23855504](https://pubmed.ncbi.nlm.nih.gov/23855504/)]
15. 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 May;66(2):192-201. [doi: [10.1016/j.pec.2006.12.002](https://doi.org/10.1016/j.pec.2006.12.002)] [Medline: [17320338](https://pubmed.ncbi.nlm.nih.gov/17320338/)]

16. van Randerad-van der Zee CH, Beurskens AJ, Swinkels RA, Pool JJ, Batterham RW, Osborne RH, et al. The burden of neck pain: its meaning for persons with neck pain and healthcare providers, explored by concept mapping. *Qual Life Res* 2016 May;25(5):1219-1225 [FREE Full text] [doi: [10.1007/s11136-015-1149-6](https://doi.org/10.1007/s11136-015-1149-6)] [Medline: [26466835](https://pubmed.ncbi.nlm.nih.gov/26466835/)]
17. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof* 2005 Jun;28(2):212-232. [doi: [10.1177/0163278705275342](https://doi.org/10.1177/0163278705275342)] [Medline: [15851774](https://pubmed.ncbi.nlm.nih.gov/15851774/)]
18. Bloom BS, Engelhart M, Furst E, Hill W, Krathwohl D. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. New York, NY: Longmans Green; 1956.
19. García AA. Cognitive interviews to test and refine questionnaires. *Public Health Nurs* 2011 Sep 1;28(5):444-450. [doi: [10.1111/j.1525-1446.2010.00938.x](https://doi.org/10.1111/j.1525-1446.2010.00938.x)] [Medline: [22092429](https://pubmed.ncbi.nlm.nih.gov/22092429/)]
20. Bollen KA. *Structural Equations with Latent Variables*. New York: Wiley; 1989.
21. van der Linden WJ, Hambleton RK. *Handbook of Modern Item Response Theory*. New York, NY: Springer-Verlag; 1997.
22. Muthén B, Asparouhov T. Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychol Methods* 2012 Sep;17(3):313-335. [doi: [10.1037/a0026802](https://doi.org/10.1037/a0026802)] [Medline: [22962886](https://pubmed.ncbi.nlm.nih.gov/22962886/)]
23. Asparouhov T, Muthén B, Morin AJ. Bayesian structural equation modeling with cross-loadings and residual covariances. *J Manag* 2015 Jun 30;41(6):1561-1577. [doi: [10.1177/0149206315591075](https://doi.org/10.1177/0149206315591075)]
24. Elsworth GR, Beauchamp A, Osborne RH. Measuring health literacy in community agencies: a Bayesian study of the factor structure and measurement invariance of the health literacy questionnaire (HLQ). *BMC Health Serv Res* 2016 Sep 22;16(1):508 [FREE Full text] [doi: [10.1186/s12913-016-1754-2](https://doi.org/10.1186/s12913-016-1754-2)] [Medline: [27659559](https://pubmed.ncbi.nlm.nih.gov/27659559/)]
25. Muraki E. A generalized partial credit model: application of an EM algorithm. *Appl Psychol Meas* 1992 Jul 27;16(2):159-176. [doi: [10.1177/014662169201600206](https://doi.org/10.1177/014662169201600206)]
26. Olsbjerg M, Christensen KB. Department of Biostatistics, University of Copenhagen. 2014. LIRT: SAS macros for longitudinal ITR models URL: [https://ifsv.sund.ku.dk/biostat/annualreport/images/fd/Research\\_Report\\_14-03.pdf](https://ifsv.sund.ku.dk/biostat/annualreport/images/fd/Research_Report_14-03.pdf) [accessed 2018-01-08] [WebCite Cache ID [6wJnHEt5L](https://www.webcitation.org/6wJnHEt5L)]
27. Holland PW, Wainer H, editors. *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993.
28. Zumbo BD. *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.
29. Crane PK, Gibbons LE, Jolley L, van Belle G. Differential item functioning analysis with ordinal logistic regression techniques. *DIFdetect* and *difwithpar*. *Med Care* 2006 Nov;44(11 Suppl 3):S115-S123. [doi: [10.1097/01.mlr.0000245183.28384.ed](https://doi.org/10.1097/01.mlr.0000245183.28384.ed)] [Medline: [17060818](https://pubmed.ncbi.nlm.nih.gov/17060818/)]
30. Swaminathan H, Rogers HJ. Detecting differential item functioning using logistic regression procedures. *J Educ Meas* 1990;27(4):361-370.
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 1995;57(1):289-300 [FREE Full text]
32. Rosas SR, Ridings JW. The use of concept mapping in measurement development and evaluation: application and future directions. *Eval Program Plann* 2017 Feb;60:265-276. [doi: [10.1016/j.evalprogplan.2016.08.016](https://doi.org/10.1016/j.evalprogplan.2016.08.016)] [Medline: [27601290](https://pubmed.ncbi.nlm.nih.gov/27601290/)]
33. Kayser L, Kushnirik A, Osborne RH, Norgaard O, Turner P. Enhancing the effectiveness of consumer-focused health information technology systems through eHealth literacy: a framework for understanding users' needs. *JMIR Hum Factors* 2015 May 20;2(1):e9. [doi: [10.2196/humanfactors.3696](https://doi.org/10.2196/humanfactors.3696)] [Medline: [27025228](https://pubmed.ncbi.nlm.nih.gov/27025228/)]
34. Beauchamp A, Batterham RW, Dodson S, Astbury B, Elsworth GR, McPhee C, et al. Systematic development and implementation of interventions to Optimise Health Literacy and Access (Ophelia). *BMC Public Health* 2017 Dec 03;17(1):230 [FREE Full text] [doi: [10.1186/s12889-017-4147-5](https://doi.org/10.1186/s12889-017-4147-5)] [Medline: [28253883](https://pubmed.ncbi.nlm.nih.gov/28253883/)]
35. Batterham RW, Buchbinder R, Beauchamp A, Dodson S, Elsworth GR, Osborne RH. The OPTimising HEalth LIterAcy (Ophelia) process: study protocol for using health literacy profiling and community engagement to create and implement health reform. *BMC Public Health* 2014 Jul 07;14:694 [FREE Full text] [doi: [10.1186/1471-2458-14-694](https://doi.org/10.1186/1471-2458-14-694)] [Medline: [25002024](https://pubmed.ncbi.nlm.nih.gov/25002024/)]
36. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of health 1.0 and health 2.0 Skills. *J Med Internet Res* 2017 Jan 24;19(1):e27 [FREE Full text] [doi: [10.2196/jmir.6709](https://doi.org/10.2196/jmir.6709)] [Medline: [28119275](https://pubmed.ncbi.nlm.nih.gov/28119275/)]
37. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. *J Med Internet Res* 2006 Nov;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
38. Kane MT. An argument-based approach to validity. *Psychol Bull* 1992;112(3):527-535. [doi: [10.1037/0033-2909.112.3.527](https://doi.org/10.1037/0033-2909.112.3.527)]
39. Dodson S, Good S, Osborne RH. Dro.deakin.edu.au. 2015. Health literacy toolkit for low and middle-income countries: a series of information sheets to empower communities and strengthen health systems URL: <http://dro.deakin.edu.au/view/DU:30074618> [accessed 2018-01-10] [WebCite Cache ID [6wMz3h0XB](https://www.webcitation.org/6wMz3h0XB)]
40. Parker R. Measuring health literacy: What? So what? Now what? In: Institute of Medicine (US) Roundtable on Health Literacy, editor. *Measures of Health Literacy: Workshop Summary*. Washington, DC: The National Academies Press; 2009:91-98.

41. World Health Organization. 2016. Shanghai Declaration on promoting health in the 2030 Agenda for Sustainable Development URL: <http://www.who.int/healthpromotion/conferences/9gchp/shanghai-declaration/en/> [accessed 2017-10-02] [WebCite Cache ID 6v6vxn48G]

## Abbreviations

**BSEM:** Bayesian Structural Equation Modeling  
**CFA:** confirmatory factor analysis  
**CSR:** composite scale reliability  
**CTT:** classical test theory  
**DHLI:** digital health literacy instrument  
**DIF:** differential item functioning  
**eHLF:** eHealth Literacy Framework  
**eHLQ:** eHealth Literacy Questionnaire  
**GPCM:** generalized partial credit model  
**IRT:** item response theory  
**Ophelia:** Optimising HEalth LIteracy and Access  
**PPP:** posterior predictive *P* value

*Edited by CL Parra-Calderón; submitted 12.07.17; peer-reviewed by K Hyun, K Thies; comments to author 11.09.17; revised version received 02.10.17; accepted 18.11.17; published 12.02.18*

*Please cite as:*

*Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, Osborne RH*

*A Multidimensional Tool Based on the eHealth Literacy Framework: Development and Initial Validity Testing of the eHealth Literacy Questionnaire (eHLQ)*

*J Med Internet Res 2018;20(2):e36*

*URL: <https://www.jmir.org/2018/2/e36/>*

*doi: [10.2196/jmir.8371](https://doi.org/10.2196/jmir.8371)*

*PMID: [29434011](https://pubmed.ncbi.nlm.nih.gov/29434011/)*

©Lars Kayser, Astrid Karnoe, Dorthe Furstrand, Roy Batterham, Karl Bang Christensen, Gerald Elsworth, Richard H Osborne. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 12.02.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.