Describing progression of health literacy skills: Establishing the HLS-Q12 cut-off scores

Øystein Guttersrud (✉ oystein.guttersrud@naturfagsenteret.no)

OsloMet - storbyuniversitetet <https://orcid.org/0000-0001-8771-1021>

Christopher Le

Norwegian Directorate of health

Kjell Sverre Pettersen

Oslo Metropolitan University

Sølvi Helseth

Oslo Metropolitan University

Hanne Sjøberg Finbråten

Inland Norway University of Applied Sciences; Oslo Metropolitan University

Research article

Keywords: Benchmarking, Cut-off scores, Health literacy, HLS-EU-Q47, HLS-Q12, Proficiency scale, Progression, Rasch Measurement Theory, Standard setting

Posted Date: August 23rd, 2019

DOI: <https://doi.org/10.21203/rs.2.13456/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background The self-reported European Health Literacy Survey Questionnaire (HLS-EU-Q47) is a widely used health literacy measure. Based on confirmatory factor analyses and Rasch-modelling, the short form HLS-Q12 was developed to meet the Rasch unidimensional measurement model expectations. After its publication, there has been a worldwide call to identify HLS-Q12 cut-off scores and establish clearly delineated standards in terms of the skills assessed. This study therefore aims to identify the HLS-Q12 scores associated with statistically distinct levels of proficiency and construct a proficiency scale that expresses progression – what individuals typically know and can do at increasingly more sophisticated levels of health literacy.

Methods We applied the unidimensional Rasch measurement model for polytomous items to responses from 900 randomly sampled individuals and 388 individuals with type 2 diabetes. Using Rasch-based item calibration, we constructed a proficiency scale by locating the ordered item thresholds along the scale; by applying Wright's method for the maximum number of strata, we determined the cut-off scores for significantly different levels. By directly referring to item content that people who reached the cut-off scores view as 'easy', we could describe these gradually more advanced levels of health literacy.

Results We identified statistically distinct levels of health literacy at the empirically identified cut-off scores 27, 33 and 39 and confirmed them by analysing the responses from individuals with diabetes. Individuals who reach these cumulative benchmarks of marginal, intermediate and advanced literacy can typically access, appraise and apply information relevant to stay healthy, improve health and critically judge health claims and compare treatments, respectively.

Conclusions By revealing progression in health literacy, we extended the current models of health literacy and inform policy-makers aiming to bridge literacy gaps. Determining how to best adapt information to patients' literacy level and thereby assist them in managing their own health is a significant clinical outcome. A substantial methodological consequence is the inevitability of Rasch modelling in measurement. We also found that Wright's method identified rating scale cut-off scores consistently across independent samples. To unveil sources of potential biases, threats to validity and imprecision of benchmarks, replicating our study in other contexts is required.

Background

Research recognises limited health literacy (HL) as having major consequences for health outcomes [1–3]. In order to improve quality in the health services, reduce inequalities in health and promote factors that enhance health, the Norwegian Ministry of Health and Care Services recently published a national strategy for HL [4]. To support the strategy and provide knowledge-based policy and practice on HL, the Ministry commissioned the Norwegian Directorate of Health to map HL in the population and selected subgroups. Consequently, the Directorate appointed representatives to join The Action Network on Measuring Population and Organizational Health Literacy (M-POHL) within the European Health

Information Initiative (EHII) of WHO-Europe. The M-POHL network promotes standards and guidelines for European co-operation in the field of HL; functions as a communication platform for sharing and disseminating information and expertise among its members as well as stakeholders and policy-makers and facilitates comparative HL measurement.

The M-POHL network's internationally coordinated HL surveys are angling to become Europe's premier yardstick for evaluating the efficiency of health policies. By identifying the characteristics of health literate populations, M-POHL allows governments to identify effective knowledge-based policies and compare good practices that may address health disparities and reduce inequalities in HL, regionally, nationally and internationally.

M-POHL is firmly committed to the continuing development of their comprehensive and widely used self-reported European Health Literacy Survey questionnaire (HLS-EU-Q47). The questionnaire was applied in a population study in eight European countries in 2011 [1] and later in Asian countries [5] and further European countries, such as Norway in 2014–2015 [6, 7]. An accumulated body of corroborating evidence shows that the HLS-EU-Q47 data is not unidimensional [5–11], which gave rise to a controversy that led to the publication of the HLS-Q12 short form [7]. Subsequent to its publication, there has been an international request to establish the HLS-Q12 cut-off scores for different HL levels.

To assure justified estimates of the prevalence of marginal, intermediate and advanced HL as measured by HLS-Q12, we need to identify the HLS-Q12 cut-off scores associated with statistically distinct levels of proficiency. We will do so by applying Wright's sample-independent method for the maximum number of statistically distinct strata [12] to our data collected in 2014. To assure quality and ensure utility of results, we will evaluate the empirically derived cut-off scores by repeating the analysis in the sample drawn from a population with chronic disease (type 2 diabetes) in 2015.

To describe what people who reach the different cut-off scores typically know and can do, we need to construct a proficiency scale by directly referring to item content, that is, we need to construct what is called a Wright map based on *ordered* item thresholds or steps [13]. This proficiency scale will reveal progression in HL—what it means to move from one level to the next in terms of HLS-Q12 item content.

The objective of this paper is therefore to report how we identified the HLS-Q12 cut-off scores and in what way we developed the HLS-Q12 proficiency scale of increasingly more sophisticated and cumulative levels of HL. The results of these two distinct but related processes will answer the following two research questions:

- What are the HLS-Q12 cut-off scores for statistically distinct levels of health literacy?
- What do individuals who reach the HLS-Q12 cut-off scores typically know and can do?

Methods

Sampling

In November 2014, responses from 900 randomly sampled individuals aged 16 and above were collected in Norway using computer-assisted telephonic interviews (CATI). The international survey agency Ipsos collected the data according to the standardized interview protocol for the HLS-EU-Q47 [8]. In March 2015, responses from 388 individuals with type 2 diabetes aged 18 and above were collected in Norway using a paper-and-pencil version of the 47-item HLS-EU-Q47. The sample was drawn from the member list of the Norwegian Diabetes Association, who offered a geographically stratified sample that included the members' places of residence. In this paper, we base our analyses on the 12-item short form HLS-Q12 [6, 7]. Finbråten et al. [6, 7] thoroughly describe the sampling procedures and sample characteristics.

Definitions and conceptual models of health literacy

Definitions of HL address the lifelong learning perspective—the development of *individual* skills to access, appraise and apply health information essential to make appropriate health decisions [14–18]. Broader definitions consider HL as an *interaction* between the proficiency of individuals and the demands of health services and systems [14, 19, 20], such as knowing how to navigate health systems and between health services. The interaction aspect of HL implicitly assumes culturally sensitive health communication between patients, health educators and health providers.

Conceptual models of population or 'public' HL [21] reflect an analysis of aspects of the HL construct—a process that reveals the complexity of the construct. Conceptual models of HL also consider factors and covariates that predict an individual's HL and view HL itself as a critical determinant of health outcomes. However, these latter aspects are not essential to the measurement of the latent trait of HL.

Basic skills or 'fundamental literacy' [19], such as oral skills [14, 22], reading and numeracy [14, 23–25], writing [14] and digital skills [26], are viewed as key aspects of HL. Other key aspects are media literacy [27], the ability to apply health information to make knowledge-based decisions [23] and the ability to interact with [28] and navigate within the health system [22, 29]. In their systematic review, Sørensen et al. [18] pointed out that HL has a wide range of aspects and that there is a large variation between the conceptual models of HL. They conclude that the Nutbeam's model [30], which refers to the three categories—functional, interactive and critical HL, summarizes many of the key aspects of HL. Their claim is supported by the fact that other conceptual models take on a similar approach, like Kickbusch and Maag [31] and Manganello [27]. Based on the literature review, Sørensen et al. [18] suggested the 'integrated conceptual model of health literacy', which may be narrowed down to three health domains (HD), *health care* (HC), *disease prevention* (DP) and *health promotion* (HP), and four cognitive domains (CD), *access* (A), *understand* (B), *appraise* (C) and *apply* (D) *health information*. However, few definitions and conceptual models of HL identify stages of HL and organize a continuous progression of HL. This is a major challenge to policy-makers who aim to respond to a current major societal challenge—namely, how to bridge health literacy gaps and promote equity by improving health education and simplifying health services. Our study therefore aims to respond to this challenge.

The unidimensional Rasch measurement model for polytomously scored items

Model requirements and properties: The concept of measurement implies a linear continuum of some kind, such as height and mass. Psychological performances result from measures that have neither an absolute zero point nor equal units of measurement. In fact, we are hampered by the fact there exists no unit by which to measure. Rasch measurement models [32], which satisfy axioms of conjoint measurement [33–37], construct measures with equal-interval properties, an arbitrary zero point and ‘logit’ as unit from ordered data, such as rating-scales [38]. Conceptually, Rasch modelling involves selecting a measurement model, applying it to the data and examining the data-model fit. Rasch models *require* local independence [39], monotonicity [40] and homogeneity or parallel item characteristic curves [40, 41], and they have the *properties* of sufficiency [42], additivity [36, 43], separability [44] and specific objectivity or invariance [45, 46].

Unidimensional Rasch models: The unidimensional Rasch model [32] for dichotomous items $\log(p/(1-p)) = \beta_n - \delta_i$ models the probability p of a correct response as the distance between the proficiency β_n of individuals and the item difficulty or item threshold δ_i . The partial credit model (PCM; [47]) is a straightforward generalisation of the unidimensional Rasch model applied to polytomous data, such as pairs of increasing adjacent rating scale response categories.

Item thresholds and proficiency: Applying the PCM to the HLS-Q12 four-point rating scale, with the categories very difficult (1), difficult (2), easy (3) and very easy (4), each item i is assigned three j ‘uncentralised item thresholds’, δ_{ij} , and each individual n is assigned a proficiency estimate β_n —the most likely point estimate based on their HLS-Q12 score. A standard error (SE) quantifies the precision of this estimate, where low SE means that the random error is small and that the person measure is more reliable. SE is a reciprocal of the Fisher information, where high Fisher information indicates that the maximum is sharp, i.e., there are few nearby values with a similar log-likelihood.

Overall data-model fit: Comparing observed proportions to model expected probabilities, the RUMM2030 statistical software package [48] assesses the overall data-model fit by the summary fit chi-square (χ^2) statistic. For large sample sizes, the root mean square error of approximation ($RMSEA = \sqrt{((\chi^2/df) - 1/N - 1)}$, expected value = 0.0 for perfect data-model fit) might be calculated as a supplementary statistic to determine data-model fit [49]. Applying the PCM, the additional assumption of ordered thresholds is vital.

Variance, reliability and strata

RUMM reports a proficiency estimate with associated SE for each observed HLS-Q12 score, ranging from 12 to 48 points, the standard deviation (sd) of the distribution of these person proficiency estimates, their average measurement error variance or ‘root mean square error’ ($RMSE^2 = 1/N \sum i = 1/N SE_i^2$), their estimated ‘true’ population variance ($SD^2 = sd^2 - RMSE^2$) and reliability (Cronbach's $\alpha = SD^2 / sd^2$) when

there is complete data. Applying Wright’s sample-independent method for the maximum number of statistically distinct strata [12], we estimate the associated Wright strata reliability = $\text{Strata}^2 / (1 + \text{Strata}^2)$.

The HLS-Q12 scale

The three health domains and the four cognitive domains of the ‘integrated conceptual model of health literacy’ [18] forms a ‘3x4 matrix’, which constitutes a 12-cell matrix. This matrix or scaffold supported the HLS-EU-Q47 item development (Table 1).

Applying confirmatory factor analyses and Rasch modelling, Finbråten et al. [6, 7] developed the 12-item HLS-Q12 scale from the 47-item HLS-EU-Q47 by deleting items that displayed differential item functioning (DIF; [50]) for available person factors (gender, age and education level), items that displayed unordered response categories or reversed thresholds [51] and redundant items. Redundant items were identified as the qualitatively least relevant items from pairs of dependent items (response dependency; [52]). If more than one item was still left in a ‘cell’ in the 3x4 matrix, the least relevant item(s) were deleted based on qualitative evaluation of item content. The purpose was to obtain a conceptually balanced scale [53], consisting of 12 items in line with the ‘3x4 matrix’, which met the Rasch model expectations. Finally, the PCA/ *t*-test protocol [54, 55] was applied to test whether the HLS-Q12 scale was sufficiently unidimensional. Applying this method, a principal component analysis (PCA) of Rasch residuals was used to identify two subsets of the items; each person was assigned a proficiency estimate on each subset and the two estimates were compared using dependent *t*-tests to account for the dependence between the two measures. The HLS-Q12 scale was interpreted as unidimensional, as approximately only 5% of these *t*-tests were significant. The process of establishing the HLS-Q12 and the psychometric properties of the HLS-Q12 items are thoroughly described in Finbråten et al. [6, 7].

<PLEASE INSERT TABLE 1 HERE>

Standard setting—the process of establishing benchmark cut-off scores

Essentially, standard setting refers to the process of establishing cut-off scores on a test. Several methods exist for the same, such as the Nedelsky, Angoff and the Bookmark method [56]. These methods are time and resource consuming owing to requiring the selection and training of experts to make judgements about individual items, suffering from possible judgmental biases, failing to yield results that agree with one another (poor validity) and failing to yield the same results on repeated applications (poor reliability) [56]. Further, they do not directly apply to rating-scale items typical for the HLS-Q12 questionnaire.

Wright's sample-independent method for the maximum number of statistically distinct strata [12] identifies significantly distinct cut-off scores based on Rasch modelling. The Wright method therefore applies to rating-scale item formats with integer scoring models and is efficient in terms of costs and time commitments. However, the cut-off scores are 'found' by estimation; they are not 'set' by experts.

Since the purpose of the HLS-Q12 proficiency scale is to express proficiency *not* relative to other respondents (norm-referenced) but in terms of HL skills with reference to HLS-Q12 item content (criterion-referenced) at three or more proficiency levels (low, medium and high), the Wright method is applicable. Three or more proficiency levels will serve diverse purposes, such as formative patient education, mapping of patients' HL in the health services, controlling information for accountability and population health studies. In addition, the HLS-Q12 *continuous* person proficiency estimates can become the data for subsequent statistical analysis, such as regression analysis, where HL is the dependent variable.

Once we have calibrated the HLS-Q12 items, the person proficiency estimate and standard error corresponding to every possible HLS-Q12 sum score is computed. Having the standard errors, we can easily compute the number of statistically distinct levels of performance by using Wright's method [12]. By starting at the lowest proficiency estimate and advancing by twice the pooled standard error, we can identify the lowest score of the next proficiency level by locating the lowest sum score that is outside the confidence interval upper bound. We repeat the procedure until there is no room for another level. These score sums define 'benchmark cut-off scores' for statistically distinct proficiency levels.

Developing a proficiency scale—item mapping for criterion-referenced literacy levels

Using dichotomously scored items, like multiple choice, we would map the item on the continuum at the place where the item information is maximized, i.e., at the item location. Individuals with proficiency estimates matching the item location have a response probability value of .50. Thus, an item is mapped at the continuum where we would expect that respondents have the skills underlying the item.

For polytomously scored items like constructed response items, Huynh [57] argues that it may be more informative to emphasis the location of each separate response, i.e., the item thresholds. Transferring this idea to the HLS-Q12 rating scale items, the second item threshold, which locates the standing on the latent trait where the typical response is 'easy', is the location where we would expect respondents to typically have the skills underlying the item. Using item mapping, we can then anchor the second uncentralised HLS-Q12 item thresholds along the measurement scale and describe the Wright cut-off scores in terms of HLS-Q12 item content.

Results

Identifying the HLS-Q12 cut-off scores for statistically distinct levels of proficiency

The HLS-Q12 displayed good overall data-model fit ($RMSEA \approx 0.0$) and sufficient reliability (Wright strata reliability = 0.94) for both samples (see overall statistics in Table 2).

The lowest observed HLS-Q12 sum score in the population sample was 20 points (equivalent to the proficiency estimate $\beta = -2.39$ logits, $SE = 0.50$). Using pooled standard error ($-2.39 + 2 \cdot \sqrt{0.50^2 + 0.49^2} = -0.99$), we found that the sum score 27 points ($\beta = -0.76$, $SE = 0.49$) lies *outside* the confidence interval upper bound for 20 points ($-0.99 < 0.76$). Therefore, 27 points is a minimum measure for 'proficiency level 1' or benchmark 1. Repeating the procedure, we found that the HLS-Q12 sum scores 27, 33, 39 and 46 represent cut-off scores for statistically distinct levels of proficiency in HL.

We revised and confirmed the empirically identified cut-off scores 27, 33 and 39 among individuals with diabetes, where the lowest observed HLS-Q12 sum score was 21 points ($\beta = -2.32$, $SE = 0.53$). As the sum scores 44 and 45 points were not observed in the diabetes sample, we could not strictly confirm the upper cut-off score at 46 points.

<PLEASE INSERT TABLE 2 HERE>

Describing the HLS-Q12 cut-off scores in terms of HLS-Q12 item content

The endorsability of the response category 'easy'—the difficulty of the second item threshold within items q23, q32, q43 and q14—were located *below* benchmark 1 ($\delta_{i3} = -0.76$ equivalent to sum score 27 points). Skills auditing implies that individuals who reach benchmark 1 *typically* view it as 'easy' to 'understand why they need health screenings' (q23), 'find information on healthy activities such as exercise, healthy food and nutrition' (q32), 'judge which everyday behaviour is related to health' (q43) and 'follow the instructions on medication' (q14). The following describes typical perceived skills at level 1 which we, based on interpretation of content for items q23, q32, q43 and q14, labelled marginal HL:

Individuals with HLS-Q12 score 27 or above can typically access, appraise and apply health information relevant to stay healthy.

Likewise, the difficulty of the second score step within items q2, q7, q44, q18, q30 and q38 were located *below* benchmark 2 ($\delta_{i3} = 0.74$ equivalent to sum score 33 points) but above benchmark 1. Skills auditing implies that individuals who reach benchmark 2 *typically* can 'find information on treatments of illnesses that concern them' (q2), 'understand what to do in a medical emergency' (q7), 'make decisions to improve their health' (q44), 'find information on how to manage mental health problems like stress or depression' (q18), 'decide how they can protect themselves from illness based on advice from family and friends' (q30) and 'understand information on food packaging' (q38). The following generalized interpretation of item content describes typical skills at level 2, which we labelled intermediate HL:

Individuals with HLS-Q12 score 33 or above can typically access, appraise and apply health information and advice relevant to enhance physical and mental health.

Individuals who reach benchmark 3 *typically* master the knowledge and skills associated with items q10 and q28—the HLS-Q12 items that are hardest to endorse. It follows that the second score step within each of these items is located *below* benchmark 3 ($\beta_n = 2.19$ equivalent to sum score 39 points). This means that individuals who reach level 3 typically can ‘judge if the information on health risks in the media is reliable’ (q28) and ‘judge the advantages and disadvantages of different treatment options’ (q10). The ability to choose between treatments might indicate that individuals at level 3 can identify suitable and optimal care, i.e., navigate in the health care system and between health care services [29]. The following describes typical perceived skills at level 3, which we labelled advanced HL:

Individuals with HLS-Q12 scores 39 or above can typically access, appraise and apply health information and advice relevant to make informed healthcare choices by critically judging health claims and fairly comparing treatments.

It is vital to recognize that the HLS-Q12 proficiency scale a) is constructed such that individuals even at the *bottom* of a proficiency level typically master the knowledge and skills associated with that level, b) describes *progression* in HL—what it means to move from one proficiency level to the next and c) is *cumulative* as individuals scoring at the advanced level (sum score ≥ 39) also know and can do what is associated with performances at the intermediate level ($33 \leq \text{sum score} \leq 38$) and marginal level ($27 \leq \text{sum score} \leq 32$). Individuals who perform below the marginal level ($12 \leq \text{sum score} \leq 26$) may, for example, understand why they need health screenings (q23) like screenings for diabetic retinopathy if they have type 2 diabetes, but they do not possess a repertory of knowledge and skills needed to reduce the risk of injury, stay healthy and prevent and manage illness. We may refer to individuals performing at or below the marginal level as displaying limited HL ($12 \leq \text{sum score} \leq 32$). The HLS-Q12 performance levels ‘below marginal’, marginal, intermediate and advanced may resemble the HLS-EU-Q47 levels considered as insufficient, problematic, sufficient and excellent HL, respectively.

Estimating the proportion who reach each level of proficiency

A key finding is that almost 1 in 2 (47%) in the Norwegian population sample performed at the marginal level. As only 2 in 100 (2%) performed below the marginal level, the prevalence of limited HL in the Norwegian population is close to 50%. While almost 4 in 10 (39%) reached the intermediate level, approximately 1 in 10 (12%) performed at the advanced level. Another key finding is that 8% of individuals with diabetes performed below level 1. Among individuals with diabetes, the proportions of individuals performing at the marginal, intermediate and advanced levels were 45%, 34% and 13%, respectively.

Discussion

Applying Wright's method for the maximum number of strata, we located the benchmark cut-off scores for statistically distinct levels of health literacy (HL) at the HLS-Q12 cut-off scores 27, 33 and 39, and we found that Wright's method yielded consistent results on repeated applications in independent samples (population sample and sample of persons with diabetes). Using item mapping based on Rasch calibration of rating scale items, we developed a proficiency scale that describes what people who reach each benchmark typically know and can do in terms of the HLS-Q12 item content. We found that individuals who reach these cumulative levels of HL typically can access, appraise and apply information relevant to stay healthy (marginal HL), improve health (intermediate HL); and critically judge health claims and compare treatments (advanced HL), respectively. The resulting proficiency scale expresses progression in HL—what it qualitatively may mean to move from low to high HL.

The newly published Norwegian strategy for HL [4] responds to the growing body of international literature documenting the magnitude of limited HL [14]. We found that approximately 50% of the Norwegian population and 53% of individuals with type 2 diabetes may have limited HL. These results add to the ideas of high prevalence of limited HL among European residents [1]. A key finding is that 8% of individuals with diabetes scored less than 27 points and performed below the marginal level. It follows that these individuals most probably have difficulties with acting on health information relevant to preventing and managing diabetes-related conditions.

However, a thought-provoking result is some medical doctors' malpractice of presuming that patients with low HL will get help from their family and friends at home (item q30) if they do not understand or cannot read the written health information provided [58]. This is a kind of non-sequitur, as applying advice from family and friends are associated with HL at the intermediate level. However, the chance to benefit from your right to choose a personally preferred method of treatment may have passed when you finally receive help from family and friends to understand the written information provided, as the doctor might already have decided to undertake the most appropriate treatment perceived. Patients within the Norwegian healthcare services have a statutory right to choose a treatment institution and choose among prudent methods of treatment available [58]. Nevertheless, to navigate within the healthcare systems and services benefiting from your statutory rights, you need skills to critically judge claims about treatment quality and methods of treatment and fairly compare those—skills associated with navigation literacy and advanced HL. This means that very few people are proficient at benefiting from their right.

As the short form HLS-Q12 was developed by reducing the 47-item questionnaire HLS-EU-Q47 to a 12-item scale based on Rasch modelling and confirmatory factor analysis [6, 7], we see few sources of potential bias and threats to internal validity. By applying a recognised statistical method for identifying cut-off scores [12] and confirming these values by analysing responses from two independent samples, we believe the HLS-Q12 cut-off scores are accurate and precise.

Robust measurement properties combined with brevity make HLS-Q12 preferable to HLS-EU-Q47 at present for monitoring HL in individuals and in populations. However, a limitation of our study is that we only based our analyses on Norwegian samples. When assessing HL across countries or groups from

different cultural and linguistic backgrounds, groups at the same level of HL may not display the same probability of affirming the items. Then we would deem these items to display differential item functioning (DIF), and cultural and linguistic backgrounds would confound the measurement of health literacy.

The degree to which DIF threatens the external validity and generalizability of our findings remain unresolved. However, the upcoming M-POHL international study will justify whether HLS-EU-Q47 and its short forms, like HLS-Q12, yield comparative results. The M-POHL study will also identify causal factors associated with limited HL. Another issue that remains unresolved is the extent to which the HLS-Q12 proficiency levels can predict future ‘success’ in terms of staying healthy, improving health and making informed healthcare choices through the life course.

Conclusions

Using Rasch-based item calibration and person measurement, we mapped the items and estimated benchmark cut-off scores. The resulting HLS-Q12 proficiency scale expresses *typical* knowledge and skills at three statistically distinct proficiency levels. Based on qualitative interpretation of item content, we labelled the levels *marginal*, *intermediate* and *advanced*, respectively. It is evident from the scale’s cumulative nature that it reveals *progression* in health literacy (HL).

Our description of increasingly more sophisticated levels in HL has important theoretical outcomes as it extends the ‘integrated conceptual model of health literacy’ developed by Sørensen et al. [18]. Recognizing that there is a significant gap between ‘avoiding injury and staying healthy’ (marginal HL) on one hand and ‘knowing how to improve mental and physical health’ (intermediate HL) on the other may inform policy-makers aiming at addressing health disparities. Our study also demonstrates the supremacy of Rasch modelling when it comes to identifying benchmark cut-off scores and describing criterion-referenced proficiency levels.

Designed for constructing measurement, the efficient HLS-Q12 scale enables clinicians to gather reliable information about an individual’s HL. The goal may be to immediately tailor health information, adapt communication and inform interventions at the *personal* level. The formative aspects of being able to identify people’s HL, and knowing which skills they need to develop to improve HL, have substantial clinical value for health personnel involved in patient education. The HLS-Q12 can potentially facilitate researchers’ endeavour in evaluating risk factors and providing recommendations for resource allocation at the *local* level or providing standards with which population HL can be compared at the *global* level. For such large-scale studies, the HLS-EU-Q47 ‘indicators’ may specify how we can best locate resources to gain effect on various aspects of HL nationally and globally. Researchers may select the unidimensional short form HLS-Q12 or the multidimensional HLS-EU-Q47, depending on whether they require to construct measurement or not.

Abbreviations

A: access health information

B: understand health information

C: appraise health information

CD: cognitive domain

D: apply health information

DIF: differential item functioning

DP: disease prevention

HC: health care

HD: health domain

HL: health literacy

HP: health promotion (HP)

M-POHL: The Action Network on Measuring Population and Organizational Health Literacy

PCA: Principal Component Analysis of Rasch residuals

PCM: Partial Credit Model

RMSE: Root Mean Square Error (the average measurement error variance of person proficiency estimates)

RMSEA: Root Mean Square Error of Approximation (a chi-square based supplementary data-model fit index for accommodating large sample sizes)

RUMM: Rasch Unidimensional Measurement Model (refers to RUMM2030 software)

sd: the sample or “observed” standard deviation of person proficiency estimates

SD: the population or “true” standard deviation of person proficiency estimates

SE: Standard Error

Declarations

Ethics approval and consent to participate

This study was approved by the Norwegian Social Science Data Service (NSD), ref. 38,917. According to NSD this study did not require ethics approval from an ethics committee. Participation was voluntary, and

respondents were anonymous and aged 16 years or above. Using telephone interviews, verbal informed consent was obtained from the participants. NSD approved the procedure.

Consent for publication

Not applicable

Availability of data and materials

The data sets are available from HSF (hanne.finbraten@inn.no) upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The Norwegian Nurses' Organisation, Inland Norway University of Applied Sciences and the Public Health Nutrition research group at Oslo Metropolitan University funded the data collection. The funders had no role in the data collection other than funding the telephone interviews carried out by the survey agency Ipsos.

Authors' contributions

ØG did the analyses and drafted the manuscript. All authors (ØG, CL, KSP, SH and HSF) contributed in the process of revising and approving the final manuscript. HSF and KSP initiated the data collection. All authors have read and approved the manuscript, and agreed to submit the manuscript to BMC Health services research.

Acknowledgements

The authors thank Kristine Sørensen and the HLS-EU Consortium for their permission to report on the short form HLS-Q12, which is based on the HLS-EU-Q47.

References

1.Sørensen K, Pelikan JM, Rothlin F, Ganahl K, Slonska Z, Doyle G, Fullam J, Kondilis B, Agrafiotis D, Ueters E et al. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur

J Public Health. 2015;25(6):1053–1058.

2.Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med.* 2011;155(2):97.

3.Bostock S, Steptoe A. Association between low functional health literacy and mortality in older adults: longitudinal cohort study. *BMJ.* 2012;344:e1602.

4.Ministry of Health and Care Services. Strategi for å øke helsekompetansen i befolkningen [Strategy to increase health literacy in the population]. Ministry of Health and Care Services. Oslo;2019.

5.Duong TV, Aringazina A, Baisunova G, Nurjanah, Pham TV, Pham KM, Truong TQ, Nguyen KT, Oo WM, Mohamad E et al. Measuring health literacy in Asia: Validation of the HLS-EU-Q47 survey tool in six Asian countries. *J Epidemiol.* 2017;27(2):80–86.

6.Finbråten HS, Pettersen KS, Wilde-Larsson B, Nordström G, Trollvik A, Guttersrud Ø. Validating the European Health Literacy Survey Questionnaire in people with type 2 diabetes. Latent trait analyses applying multidimensional Rasch modelling and confirmatory factor analysis. *J Adv Nurs.* 2017;73(11):2730–2744.

7.Finbråten HS, Wilde-Larsson B, Nordström G, Pettersen KS, Trollvik A, Guttersrud Ø. Establishing the HLS-Q12 short version of the European Health Literacy Survey Questionnaire: latent trait analyses applying Rasch modelling and confirmatory factor analysis. *BMC Health Serv Res.* 2018;18(1):506.

8.Sørensen K, Van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, Kondilis B, Stoffels V, Osborne RH, Brand H. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC Public Health.* 2013;13(1):948.

9.Duong TV, Lin I-F, Sørensen K, Pelikan JM, Van Den Broucke S, Lin Y-C, Chang PW. Health Literacy in Taiwan A Population-Based Study. *Asia Pac J Public Health.* 2015;27(8):871–880.

10.Nakayama K, Osaka W, Togari T, Ishikawa H, Yonekura Y, Sekido A, Matsumoto M. Comprehensive health literacy in Japan is lower than in Europe: a validated Japanese-language assessment of health literacy. *BMC Public Health.* 2015;15(1):505.

11.Huang Y-J, Chen C-T, Lin G-H, Wu T-Y, Chen S-S, Lin L-F, Hou W-H, Hsieh C-L. Evaluating the European Health Literacy Survey Questionnaire in Patients with Stroke: A Latent Trait Analysis Using Rasch Modeling. *Patient.* 2018;11(1):83–93.

12.Wright BD. Separation, reliability and skewed distributions: statistically different levels of performance. *Rasch Measurement Transactions.* 2001;14(4):786.

13. Van Wyke J, Andrich D. A typology of polytomously scored mathematics items disclosed by the Rasch model: implications for constructing a continuum of achievement. In: Report No 2 ARC Linkage Grant LP0454080. Maintaining Invariant Scales in State, National and International Level Assessments. Perth, Western Australia. Murdoch University; 2006.
14. Nielsen-Bohlman L, Panzer A, Kindig D. Health Literacy: A prescription to end confusion. Washington, DC: Institute of Medicine of the National Academies. National Academies Press; 2004.
15. Ad hoc committee on health literacy for the Council on Scientific Affairs. Health Literacy Report of the Council on Scientific Affairs. JAMA. 1999;281(6):552–557.
16. Nutbeam D. Health promotion glossary. Health Promot Int. 1998;13(4):349–364.
17. Mancuso JM. Health literacy: a concept/dimensional analysis. Nurs Health Sci. 2008;10(3):248–255.
18. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, Brand H: Health literacy and public health: A systematic review and integration of definitions and models. BMC Public Health. 2012;12(1):1–13.
19. Zarcadoolas C, Pleasant A, Greer DS. Elaborating a definition of health literacy: a commentary. J Health Commun. 2003;8(S1):119–120.
20. Kwan B, Frankish J, Rootman I, Zumbo B, Kelly K, Begoray D, Kazanjian A, Mullet J, Hayes M. The development and validation of measures of “Health Literacy” in different populations. UBC Institute of Health Promotion Research and University of Victoria Community Health Promotion Research; 2006.
21. Freedman DA, Bess KD, Tucker HA, Boyd DL, Tuchman AM, Wallston KA. Public health literacy defined. Am J Prev Med. 2009;36(5):446–451.
22. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. Am J Health Behav. 2007;31(1):S19-S26.
23. Speros C. Health literacy: concept analysis. J Adv Nurs. 2005;50(6):633–640.
24. Yost KJ, Webster K, Baker DW, Choi SW, Bode RK, Hahn EA. Bilingual health literacy assessment using the Talking Touchscreen/la Pantalla Parlanchina: Development and pilot testing. Patient Educ Couns. 2009;75(3):295–301.
25. Stocks NP, Hill CL, Gravier S, Kickbusch L, Beilby JJ, Wilson DH, Adams RJ. Health literacy-a new concept for general practice? Aust Fam Physician. 2009;38(3):144.
26. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res. 2006;8(2):e9.

- 27.Manganello JA. Health literacy and adolescents: a framework and agenda for future research. *Health Educ Res.* 2007;23(5):840–847.
- 28.Ishikawa H, Yano E. Patient health literacy and participation in the health-care process. *Health Expectations* 2008;11(2):113–122.
- 29.Gui X, Chen Y, Pine KH. Navigating the Healthcare Service Black Box: Individual Competence and Fragmented System. *Proceedings of the ACM on Human-Computer Interaction.* 2018;2(CSCW):61.
- 30.Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int.* 2000;15(3):259–267.
- 31.Kickbusch I, Maag D. Health literacy. In: Quah S, editor. *International Encyclopedia of Public Health.* 2008;204–211.
- 32.Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests.* Expanded edn. Chicago: University of Chicago Press; 1960.
- 33.Luce RD, Tukey JW. Simultaneous conjoint measurement: A new type of fundamental measurement. *J Math Psychol.* 1964;1(1):1–27.
- 34.Brogden H. The Rasch model, the law of comparative judgment and additive conjoint measurement. *Psychometrika.* 1977;42(4):631–634.
- 35.Karabatsos G. The Rasch model, additive conjoint measurement, and new models of probabilistic measurement theory. *J Appl Meas.* 2001;2(4):389–423.
- 36.Perline R, Wright BD, Wainer H. The Rasch model as additive conjoint measurement. *Appl Psychol Meas.* 1979;3(2):237–255.
- 37.Scheiblechner H. Additive conjoint isotonic probabilistic models (ADISOP). *Psychometrika.* 1999;64(3):295–316.
- 38.Linacre JM, Wright BD. Constructing Linear Measures from Counts of Qualitative Observations. In: *Fourth International Conference on Bibliometrics, Informetric and Scientometrics.* Berlin, Germany; 1993.
- 39.Marais I, Andrich D. Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *J Appl Meas.* 2008;9(3):200–215.
- 40.Kreiner S. The Rasch model for dichotomous items. In: Christensen KB, Kreiner S, Mesbah M, editors. *Rasch Models in Health.* Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2013. p. 5–25.
- 41.de Gruijter DNM, van der Kamp LJT. *Statistical test theory for the behavioral sciences.* New York, USA: Chapman and Hall/CRC; 2007.

- 42.Andersen EB. Sufficient statistics and latent trait models. *Psychometrika*. 1977;42(1):69–81.
- 43.Andrich D. Distinctions between assumptions and requirements in measurement in the social sciences. In: Keats JA, Taft R, Heath RA, Lovibond SH, editors. *Mathematical and theoretical systems*. North Holland: Elsevier Science Publishers BV; 1989. p. 7–16.
- 44.Rasch G. On Specific Objectivity: An attempt at Formalizing the Request for Generality and Validity of Scientific Statements. *Danish Yearbook of Philosophy*. 1977;(14):58–94.
- 45.Stenner AJ. Specific objectivity - local and general. *Rasch Measurement Transactions*. 1994;8(3):374.
- 46.Andrich D. *Rasch models for measurement*. Newsbury Park, CA: Sage; 1988.
- 47.Masters GN. A Rasch model for partial credit scoring. *Psychometrika*. 1982;47(2):149–174.
- 48.Andrich D, Sheridan B, Luo G. Rasch unidimensional measurment model. RUMM2030. RUMM Laboratory Pty Ltd.; 2003
- 49.Tennant A, Pallant JF. The root mean square error of approximation (RMSEA) as a supplementary statistic to determine fit to the Rasch model with large sample sizes. *Rasch Measurement Transactions*. 2012;25(4):1348–1349.
- 50.Andrich D, Hagquist C. Real and Artificial Differential Item Functioning in Polytomous Items. *Educ Psychol Meas*. 2014;75(2):185–207.
- 51.Hagquist C, Andrich D. Is the sense of coherence-instrument applicable on adolescents? A latent trait analysis using Rasch-modelling. *Pers Individ Dif*. 2004;36(4):955–968.
- 52.Andrich D, Humphry SM, Marais I. Quantifying local, response dependence between two polytomous items using the Rasch model. *Appl Psychol Meas*. 2012;36(4):309–324.
- 53.Guttersrud Ø, Petterson KS. Young adolescents' engagement in dietary behaviour–the impact of gender, socio-economic status, self-efficacy and scientific literacy. *Methodological aspects of constructing measures in nutrition literacy research using the Rasch model*. *Public Health Nutr*. 2015;18(14):2565–2574.
- 54.Smith Jr EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas*. 2002;3(2):205–231.
- 55.Hagell P. Testing rating scale unidimensionality using the principal component analysis (PCA)/t-test protocol with the Rasch model: the primacy of theory over statistics. *Open J Stat*. 2014;4:456–465.
- 56.Cizek GJ, Bunch MB. *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: SAGE Publications Ltd; 2007.

57.Huynh H. On score locations of binary and partial credit items and their applications to item mapping and criterion-referenced interpretation. J Educ Behav Stat. 1998;23(1):35–56.

58.Le C. Pasientmedvirkning i praksis? En studie om implementering av systematisk vurdering av pasienters Health Literacy i helse- og omsorgstjenestene. [Patient Participation in Practice? A Study of Implementing Routine Patients' Health Literacy Assessment in Health and Care Services]. Master Thesis. Trondheim, Norway: Norwegian University of Science and Technology; 2019.

Tables

Due to technical limitations, tables are only available as a download in the supplemental files section

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.docx](#)