Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cross-national validation and measurement invariance of the HLS-EU-Q6 in Latin America: Evidence from adults in Ecuador, Peru and Venezuela

Gabriel Ortiz Francisco
Universidad Técnica Particular de Loja
Víctor López-Guerra
Universidad Técnica Particular de Loja
Daniel F Condor-Camara
Universidad Peruana Cayetano Heredia
Angélica María Rojas Fuenmayor
Universidad Centroccidental Lisandro Alvarado
Judith Francisco-Pérez
judithfrancisco@gmail.com

Pontificia Universidad Católica del Ecuador

Research Article

Keywords: Health Literacy, Psychometrics, Surveys and Questionnaires, Cross-Cultural Comparison

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10998246/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Cross-national validation and measurement invariance of the HLS-EU-Q6 in Latin America: Evidence from adults in Ecuador, Peru and Venezuela

Gabriel Ortiz Francisco[1], Víctor López-Guerra[1], Daniel Cóndor-Cámara[2] Angelica María Rojas Fuenmayor[3], Judith Francisco-Pérez[4]

[1]Department of Psychology, Universidad Técnica Particular de Loja, Loja, Ecuador

[2]Faculty of Nursing, Universidad Peruana Cayetano Heredia, Lima, Perú

[3]Dean's Office of Humanities and Arts. Universidad Centroccidental Lisandro Alvarado. Barquisimeto, Venezuela

[4]Faculty of Health and Well-being. Research Center for Health in Latin America (CISeAL), Pontificia Universidad Católica del Ecuador, Quito, Ecuador.

Correspondence author: judithfrancisco@gmail.com

Abstract

Background: Health literacy is a key determinant of population health, yet evidence on the cross-national validity of brief health literacy instruments in Latin America remains limited. This study evaluated the psychometric properties and measurement invariance of the Health Literacy Survey European Questionnaire–Short Form (HLS-EU-Q6) among adults from Ecuador, Peru, and Venezuela.

Methods: A cross-sectional psychometric validation study was conducted in 2,034 adults from Venezuela (n = 957), Ecuador (n = 617), and Peru (n = 460). Exploratory and confirmatory factor analyses were performed using split-sample validation. Reliability was assessed using Cronbach's alpha, McDonald's omega, composite reliability, and average variance extracted. Measurement invariance was examined across countries and sociodemographic groups and known-groups validity was evaluated using Welch analyses of variance.

Results: The HLS-EU-Q6 showed a stable unidimensional structure with satisfactory reliability (α = .830; ω = .832; composite reliability = .866) and adequate convergent validity (average variance extracted = .519). Configural and metric invariance were supported across countries, whereas scalar invariance was not, indicating that observed score differences cannot be interpreted as equivalent differences in underlying health literacy. In contrast, full invariance was supported across educational level, scalar invariance across age groups, and partial scalar invariance across sex. Known-groups validity was supported by higher health literacy among participants with higher educational attainment and older age.

Conclusions: The HLS-EU-Q6 is a reliable and valid instrument for assessing global health literacy within Ecuador, Peru, and Venezuela. However, the absence of scalar invariance limits direct comparisons of observed scores across countries. These findings support the use of country-specific normative values and highlight the importance of evaluating measurement invariance before health literacy instruments are applied in multinational surveillance, programme evaluation, or cross-country public health research.

Keywords: Health Literacy; Psychometrics; Surveys and Questionnaires; Cross-Cultural Comparison.

1. Background

Health literacy (HLS) is an essential determinant of public health, closely linked to the promotion of healthy habits and better overall health outcomes (1–3). Limited levels of HLS have been associated with increased use of curative services, higher healthcare costs, and widening health inequalities (4–7).

Evidence suggests that those most affected by low HLS are older adults, who have more difficulty managing health situations, which can lead to chronic illnesses, functional decline, and mental health problems (8–10). These limitations are also more frequent among people with low levels of education, low socioeconomic status, or who are financially deprived (9).

In recent years, HLS has evolved theoretically and methodologically towards a broader understanding, which conceives of it not only as an individual skill, but as a relational competence, linked to interaction with the environment (11). This perspective recognizes that the obstacles to adequate health literacy go beyond access to reliable information, as they are conditioned by sociocultural, educational, and economic factors that limit people's ability to understand and effectively use health information (12,13).

Valid measurement of health literacy is essential to understand how people interact with information systems and make health-related decisions. Further studies across diverse ages, educational, and social groups are needed to design more equitable and effective public health strategies (14).

The measurement of HLS should account for the social, cultural, and linguistic characteristics of each context. The European Health Literacy Survey Questionnaire (HLS-EU-Q47) is the most widely used original instrument for evaluating and performing this measurement (15); however, its length limits its applicability in certain scenarios, especially when time or resources are limited. Short versions of the instrument reduce response burden while maintaining reliability and facilitating use in population and primary care settings.

To address this limitation, abbreviated versions of the instrument have been developed to reduce respondent burden without compromising reliability or practical utility. Among these, the HLS-EU-Q6 is a brief self-administered instrument that has shown acceptable reliability and feasibility for use in population-based, clinical, and community settings.

Previous validations in Swedish and Arabic support the use of the HLS-EU-Q6 as a brief measure of HLS, with acceptable internal consistency ($\alpha = 0.77–0.79$) and high test-retest reliability ($\kappa = 0.812–0.89$) (16,17). However, other studies have shown significant variations in their psychometric properties. In Sweden, principal component analysis identified a two-factor structure, whereas in France, confirmatory factor analysis showed a poor fit (17,18). Furthermore, the French version demonstrated poor discriminative capacity among subjects with medium-to-high levels of HLS, as well as differential item functioning based on sex, age, and educational level, posing challenges for measurement invariance (18). Criterion validity has also been debated, as in the case of the Arabic version, where low agreement with the HLS-EU-Q16 scale was observed (16).

Other validations of the HLS-EU-Q6, such as the Portuguese, Brazilian, Turkish, and Bengali versions, have generally shown satisfactory psychometric properties, including acceptable internal consistency, a unidimensional structure, and adequate factor fit (19–22). In particular, the Turkish version reported $\alpha = 0.82$, $\omega = 0.84$, CR = 0.89 and AVE = 0.59, while the Bengali version showed good internal consistency ($\alpha = 0.857$), although its authors recommended using the HLS-EU-Q16 due to its superior psychometric performance (21,22). Overall, these findings suggest

that the HLS-EU-Q6 can be useful as a brief measure of HLS, while also highlighting that its performance may vary across cultural and linguistic contexts.

In addition to the HLS-EU-Q6, other abbreviated versions have been developed for optimisation and applicability purposes. In Indonesia, a 10-item model (HLS-EU-SQ10-IDN) was proposed using data mining techniques, which showed greater accuracy than other traditional abbreviated scales (23). Additionally, the HLS-Q12 was developed in Norway, it meets Rasch model assumptions and offers a brief, psychometrically robust alternative for the clinical screening of HLS (24). More recently, the Mexican version, HLS-EU-Q12M, was validated in patients with hypertension, demonstrating adequate internal consistency ($\alpha = 0.83$) and high test-retest reliability (ICC = 0.94), which supports its utility in this population (25).

To our knowledge, the HLS-EU-Q6 has not been formally validated in Latin American populations, and evidence regarding its measurement equivalence across countries in the region is lacking. This represents an important gap, because cultural, socioeconomic, educational and healthcare-system differences may influence both response to HLS items and the comparability of scores across populations (26). Therefore, this study aimed to evaluate the factor structure, reliability, convergent validity, known-groups validity, and measurement invariance of the HLS-EU-Q6 among adults from Venezuela, Ecuador and Peru. Measurement invariance was examined across countries and key sociodemographic groups to determine whether the instrument supports valid within and cross-group comparisons.

2. Method

Participants

The study included 2,034 adults from three Latin American countries: Venezuela (n = 957), Ecuador (n = 617), and Peru (n = 460). Eligible participants were aged 18 years or older, resided in one of the participating countries, were able to read Spanish, and provided informed consent before participation.

A non-probabilistic convenience sampling strategy was used to recruit participants through complementary online and community-based approaches implemented in each country. Recruitment sought to include adults with diverse demographic, educational, and health-related characteristics.

Overall, women represented 61.7% of the sample, and the mean age was 36.50 years (SD = 14.88). The largest age group comprised individuals aged 18–29 years (42.8%), and half of the participants had completed university or postgraduate education (50.0%). The sample also included participants with and without chronic diseases and with and without family members affected by chronic diseases. Detailed sociodemographic characteristics are presented in Table 1.

Because the study employed a non-probabilistic sampling strategy, the sample should not be considered representative of the adult populations of Venezuela, Ecuador, or Peru.

Table 1

Sociodemographic Characteristics of the Sample

Variable	Venezuela (n = 957)	Ecuador (n = 617)	Peru (n = 460)	Total (N = 2034)
Sex				
Female	647 (67.6)	399 (64.7)	209 (45.4)	1255 (61.7)
Male	310 (32.4)	218 (35.3)	251 (54.6)	779 (38.3)
Age group				
18–29	269 (28.1)	348 (56.4)	253 (55.0)	870 (42.8)
30–44	258 (27.0)	187 (30.3)	112 (24.3)	557 (27.4)
45–59	276 (28.8)	66 (10.7)	73 (15.9)	415 (20.4)
60 or older	154 (16.1)	16 (2.6)	22 (4.8)	192 (9.4)
Education				
Elementary or below	343 (35.8)	212 (34.4)	130 (28.3)	685 (33.7)
Technical	97 (10.1)	130 (21.1)	105 (22.8)	332 (16.3)
University/Postgraduate	517 (54.0)	275 (44.6)	225 (48.9)	1017 (50.0)
Chronic disease				
Yes	284 (29.7)	81 (13.1)	54 (11.7)	419 (20.6)
No	673 (70.3)	536 (86.9)	406 (88.3)	1615 (79.4)
Family member with chronic disease				
Yes	718 (75.0)	373 (60.5)	49 (10.7)	1140 (56.0)
No	239 (25.0)	244 (39.5)	411 (89.3)	894 (44.0)
Age*	41.87 (15.72)	31.03 (11.61)	32.68 (13.06)	36.50 (14.88)

*Note. Values are n (%). Education categories are harmonised across countries: postgraduate degrees are collapsed into the University/Postgraduate category to allow cross-country comparability. *Mean (SD).*

Study design and procedure

A multicountry cross-sectional psychometric validation study was conducted to evaluate the psychometric properties of the HLS-EU-Q6 among adults from Venezuela, Ecuador, and Peru. The three countries were selected because they share a Spanish-speaking context while differing in healthcare systems, socioeconomic conditions, and educational contexts, allowing assessment of cross-national measurement invariance. Data were collected between October 2023 and April 2024 using a self-administered online questionnaire. Regardless of the recruitment channel, all participants completed the questionnaire electronically.

Before accessing the questionnaire, participants received information about the study objectives and procedures, the voluntary nature of participation, the confidentiality of their responses, and their right to discontinue participation without consequences. Electronic informed consent was obtained before questionnaire completion. Responses were collected anonymously, and no personally identifiable information was recorded.

Ethical approval was obtained independently in each participating country before data collection commenced: the Research Ethics Committee of Universidad Peruana Cayetano Heredia, Peru (CIEI-389-36-23); the Human Research Ethics Committee of Pontificia Universidad Católica del Ecuador (CEISH-860-2023); and the Bioethics Committee of Universidad Centroccidental Lisandro Alvarado, Venezuela (CBDCS-4-2023). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Instruments

Sociodemographic questionnaire

A sociodemographic questionnaire was administered to collect information on sex, age group, educational level, chronic disease status, and the presence of chronic disease among family members.

Health Literacy Survey European Questionnaire (HLS-EU-Q6)

Health literacy was assessed using the Health Literacy Survey European Questionnaire–Short Form (HLS-EU-Q6), a six-item short form derived from the European Health Literacy Survey Questionnaire (HLS-EU-Q), developed by the HLS-EU Consortium (27). The HLS-EU-Q6 is a brief self-administered instrument designed to assess self-perceived health literacy, defined as individuals' ability to access, understand, appraise, and apply health-related information to make informed health decisions across the domains of healthcare, disease prevention, and health promotion. In this study, responses were recorded on a five-point Likert-type scale ranging from 1 (very difficult) to 5 (very easy)**.** A global HLS-EU-Q6 score was calculated as the mean of the six item responses, with higher scores indicating greater self-perceived health literacy.

Cultural adaptation

The Spanish version of the HLS-EU-Q6 underwent a standardised linguistic and cultural adaptation process in each participating country to promote semantic, conceptual, and cultural equivalence while preserving the conceptual meaning of the original items.

In each country, the questionnaire was independently reviewed by local experts in psychometrics, health literacy, and public health. The review focused on semantic and conceptual equivalence, clarity, comprehensibility, and contextual appropriateness of each item. When necessary, minor wording modifications were introduced to improve readability and cultural relevance while aiming to preserve the construct assessed by each item.

Following expert review, pilot testing was conducted with 30 adults in each participating country who shared characteristics with the target population. Participants completed the questionnaire and provided feedback on item clarity, comprehensibility, and cultural appropriateness. Pilot testing indicated that the items were generally understood and considered culturally appropriate across the three participating countries, and no further substantive modifications were required before the main data collection

Data analysis

All statistical analyses were performed using R version 4.2.3 (28). Exploratory factor analyses (EFA) were conducted using the psych package (29), whereas confirmatory factor analyses (CFA) and measurement invariance analyses were performed using the lavaan (30) and semTools (31) packages. The total sample was randomly divided into two equal subsamples stratified by country using a fixed random seed (50/50 split; seed = 123). The calibration subsample was used for EFA, whereas the validation subsample was reserved for CFA. Internal consistency, measurement invariance, and known-groups validity analyses were conducted using the full sample.

Prior to factor extraction, item distributions were examined using skewness and kurtosis statistics. Given the ordinal response format, the EFA was based on a polychoric correlation matrix. Maximum likelihood estimation with oblimin rotation was applied. The number of factors to retain was evaluated using parallel analysis (32), the minimum average partial (MAP) test (33), and the Bayesian Information Criterion (BIC) derived from the Very Simple Structure (VSS) procedure (34). Factor loadings $\geq .40$ were considered substantively meaningful.

CFA was conducted using the weighted least squares mean- and variance-adjusted (WLSMV) estimator, which is appropriate for ordinal indicators. Six competing measurement models representing alternative theoretical and empirically derived structures of the HLS-EU-Q6 were evaluated. Model fit was assessed using the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA) with its 90% confidence interval, and standardised root mean square residual (SRMR). CFI and TLI values ≥ .95, RMSEA values ≤ .06, and SRMR values ≤ .08 were used as reference guidelines (35).

However, these reference values were treated as interpretive guidelines rather than rigid decision rules. Because the HLS-EU-Q6 comprises only six indicators, the competing models had relatively few degrees of freedom, a condition under which RMSEA may produce inflated estimates and indicate poor approximate fit even when other indices suggest comparatively limited residual discrepancy. Therefore, model selection was not based on any single fit index. Instead, CFI, TLI, RMSEA, SRMR, standardised factor loadings, model admissibility, theoretical interpretability, and parsimony were considered jointly. was also given to whether alternative models produced improper solutions, weakly identified factors, item cross-loadings, or only marginal improvements in fit at the expense of interpretability. For the retained model, convergent validity was assessed using the average variance extracted (AVE), with values ≥ .50 considered indicative of adequate convergent validity (36).

Internal consistency was evaluated using Cronbach's alpha (α), McDonald's omega (ω), and composite reliability (CR), with values ≥ .70 considered acceptable (37). Cronbach's alpha and McDonald's omega were estimated for the overall sample and separately for each participating country, whereas composite reliability was calculated for the retained measurement model.

Measurement invariance was evaluated sequentially across countries at the configural, metric, scalar, and strict levels. Because sparse and empty response categories prevented stable estimation of the multicountry invariance models under WLSMV, invariance analyses were conducted using robust maximum likelihood estimation (MLR), treating item responses as approximately continuous. Changes in model fit were interpreted according to the recommendations of Cheung and Rensvold (38) and Chen (39) with absolute changes of $\Delta CFI \leq .010$ and $\Delta RMSEA \leq .015$ considered supportive of invariance. When full scalar invariance was not supported, item intercepts were examined sequentially to identify potential sources of non-invariance, and statistical significance was adjusted using the Benjamini–Hochberg procedure. The same sequential invariance-testing strategy was subsequently applied across sex, educational level, and age group.

Known-groups validity was examined by comparing HLS-EU-Q6 scores across sex, age group, educational level, chronic disease status, and the presence of chronic disease among family members. Because group variances and sample sizes were unequal, Welch's ANOVA was used, followed by Games–Howell post hoc comparisons when the omnibus test was statistically significant. Effect sizes were estimated using epsilon-squared (ε^2). Because scalar invariance across countries was not supported, latent mean comparisons were not performed. Instead, between-country differences were examined as observed-score comparisons, using eta-squared (η^2) for the omnibus comparison and Cohen's d for pairwise comparisons. These comparisons were interpreted cautiously because they may reflect both substantive differences in health literacy and cross-national differences in response thresholds.

3. Results

Preliminary item analysis

Item means ranged from 3.34 (SD = 1.08; item 3) to 3.76 (SD = 0.93; item 6) in the pooled sample (N = 2,034). Skewness ranged from −0.11 to −0.57, and no floor effects were observed. Across countries, item 3 consistently showed the lowest mean score, whereas item 6 showed the highest mean score in the pooled sample and Venezuela, and tied with item 5 for the highest mean score

in Peru (both M = 3.67); in Ecuador, item 5 had the highest mean. Table 2 presents the descriptive statistics for each item across the three participating countries.

Table 2

Descriptive Statistics for HLS-EU-Q6 Items by Country

Item	Domain	Overall M (SD)	Venezuela M (SD)	Ecuador M (SD)	Peru M (SD)	Skewness	Kurtosis
Item 1	Healthcare	3.52 (0.99)	3.67 (1.13)	3.30 (0.75)	3.50 (0.87)	-0.22	-0.49
Item 2	Healthcare	3.66 (0.97)	3.89 (1.09)	3.38 (0.72)	3.58 (0.88)	-0.33	-0.37
Item 3	Dis. prev.	3.34 (1.08)	3.42 (1.27)	3.14 (0.79)	3.43 (0.95)	-0.11	-0.71
Item 4	Dis. prev.	3.50 (1.03)	3.68 (1.18)	3.17 (0.75)	3.57 (0.91)	-0.18	-0.63
Item 5	Health prom.	3.73 (1.05)	3.88 (1.25)	3.55 (0.75)	3.67 (0.90)	-0.57	-0.27
Item 6	Health prom.	3.76 (0.93)	4.00 (1.01)	3.46 (0.72)	3.67 (0.85)	-0.46	-0.11

Note. Skewness and kurtosis computed on the pooled sample (N = 2,034). Item 3 = most difficult across all countries; Item 6 = least difficult in the pooled sample (Item 5 is least difficult in Ecuador).

Exploratory factor analysis

The data were suitable for factor analysis (KMO = .844; Bartlett's $\chi^2(15) = 2015.96$, $p < .001$). Parallel analysis and the VSS-BIC criterion supported a two-factor solution, whereas the MAP test and VSS complexity indices favoured a one-factor solution. The two-factor solution, however, produced an inadmissible communality for Item 4 ($h^2 = 1.97$), a factor largely dominated by a single item, and a substantial interfactor correlation ($\Phi = .65$), suggesting insufficient discriminant validity between the factors. Consequently, both one- and two-factor structures were evaluated using CFA. Table 3 presents the standardised factor loadings for the unidimensional exploratory solution.

Table 3

Exploratory Factor Analysis: Standarised Loadings for the Unidimensional Solution

Item	Domain	λ	h^2	u^2
Item 2	Healthcare	0,727	0,528	0,472
Item 1	Healthcare	0,680	0,462	0,538
Item 4	Disease prev.	0,704	0,496	0,504
Item 6	Health prom.	0,701	0,492	0,508
Item 3	Disease prev.	0,646	0,417	0,583
Item 5	Health prom.	0,581	0,338	0,662

Note. ML estimation, oblimin rotation, N = 1,017 (calibration half). λ = standardised loading; h^2 = communality; u^2 = uniqueness. Items sorted by loading magnitude.

Confirmatory factor analysis

Six competing measurement models representing alternative factorial structures of the HLS-EU-Q6 were evaluated (Table 4). M1 specified a unidimensional model in which all six items loaded onto a single health literacy factor. M2 and M3 represented two correlated factors derived from the exploratory factor analysis. In M2, the first factor comprised the healthcare and disease

prevention items (Items 1–4) and the second factor comprised the health promotion items (Items 5–6). M3 was identical to M2 except that Item 4 was excluded from the first factor to examine whether its removal improved model fit. M5 extended the M2 structure by allowing Item 4 to cross-load on the health promotion factor. M4 represented the theoretical three-factor structure proposed by the HLS-EU framework, with separate healthcare (Items 1–2), disease prevention (Items 3–4), and health promotion (Items 5–6) factors. Finally, M6 specified a bifactor model including one general health literacy factor loading on all six items together with three orthogonal domain-specific factors corresponding to the theoretical dimensions.

Two models were inadmissible. The theoretical three-factor model (M4) produced a Heywood case, with an interfactor correlation greater than one, whereas the bifactor model (M6) failed to converge.

The CFA findings were consistent with the exploratory analyses. Although parallel analysis initially suggested a two-factor solution, the exploratory results indicated that this structure lacked adequate psychometric support because it produced an inadmissible communality for Item 4, a dominant first factor, and a substantial interfactor correlation ($\Phi = .65$), indicating insufficient discriminant validity between the factors. This apparent discrepancy highlights the importance of integrating exploratory, confirmatory, and substantive evidence when selecting the most appropriate measurement model. Consequently, the apparent multidimensionality identified during the exploratory phase was not supported as representing distinct latent constructs.

Among the admissible solutions, the unidimensional model (M1) was retained as the most parsimonious and theoretically interpretable representation of theHLS-EU-Q6. Although some alternative models showed marginally better values for individual fit indices, these improvements were not accompanied by superior psychometric performance or greater theoretical interpretability. Instead, they either produced improper estimates, weakly identified factors, or introduced additional model complexity without providing meaningful advantages over the unidimensional solution. Furthermore, the retained model is consistent with the intended purpose of the HLS-EU-Q6 as a brief screening instrument designed to assess global health literacy rather than separate domain-specific scores.

Although the RMSEA exceeded conventional reference values, the retained model demonstrated acceptable CFI (.943), SRMR (.051), satisfactory standardised factor loadings, and adequate convergent validity (AVE = .519). Taken together, the exploratory and confirmatory findings support the interpretation of the HLS-EU-Q6 as a unidimensional measure of global health literacy in adults from Venezuela, Ecuador, and Peru. Model fit indices for all competing models are presented in Table 4.

Table 4

Confirmatory Factor Analysis: Model Fit Indices and Internal Consistency

Panel A. CFA model comparison

Model	χ^2 (df)	CFI	TLI	RMSEA	90% CI	SRMR	ΔCFI	ΔRMSEA
M1 — Unidimensional[a]	324.1 (9)	0.943	0.905	0.185	[.168, .202]	0.051	—	—
M2 — Two-factor (item 4 in F1)	292.1 (8)	0.949	0.903	0.186	[.168, .204]	0.046	.006	.001
M3 — Two-factor (without item 4)	147.2 (4)	0.961	0.903	0.187	[.161, .213]	0.040	—	—
M4 — Three-factor theoretical[b]	68.5 (6)	0.989	0.972	0.101	[.080, .123]	0.024	.046	−.084
M5 — Two-factor (cross-load item 4)	216.4 (7)	0.962	0.919	0.171	[.151, .190]	0.040	.013	−.015
M6 — Bifactor[c]	— (−3)	—	—	—	—	—	—	—

Panel B. Internal consistency — M1 (selected model)

Country	n	M	SD	Welch F (df)	p / η^2	Pairwise d (GH)
Venezuela	957	3.76	0.80	F(2, 1158.4) = 75.19	p < .001 η^2 = .061	VE – EC: d = 0.59 VE – PE: d = 0.24 EC – PE: d = 0.37
Ecuador	617	3.33	0.58			
Peru	460	3.57	0.71			

Note. M1: single general factor (all 6 items); M2: two correlated factors — F1 = healthcare + disease prevention (items 1, 2, 3, 4), F2 = health promotion (items 5, 6); M3: same as M2, item 4 excluded from F1; M4: three correlated factors — healthcare (items 1, 2), disease prevention (items 3, 4), health promotion (items 5, 6); M5: same as M2 with item 4 cross-loading on F2; M6: bifactor — general factor on all items + three orthogonal specific factors. WLSMV estimator, ordered = TRUE, holdout half (n = 1,017). [a] *Selected model.* [b] *Inadmissible: Heywood case (r_{xx}' = 1.017 between disease prevention and health promotion).* [c] *Did not converge (df = −3). AVE = .519; CR = .866 for M1. A = Cronbach's alpha; ω = McDonald's omega (full sample).*

Internal consistency

The retained unidimensional model demonstrated satisfactory reliability. In the pooled sample, Cronbach's alpha (α = .830) and McDonald's omega (ω = .832) showed excellent agreement, indicating stable internal consistency. Composite reliability (CR = .866) also exceeded the recommended threshold, providing additional evidence of score reliability. Country-specific analyses revealed acceptable reliability across Venezuela, Ecuador, and Peru, with slightly lower coefficients observed in Venezuela but all estimates remaining within recommended ranges. Collectively, these findings support the reliability of the total HLS-EU-Q6 score for assessing health literacy in the participating countries (see Table 4).

Measurement invariance across countries

Sequential measurement invariance testing was conducted across Venezuela, Ecuador, and Peru using the MLR estimator (Table 5). Configural invariance was supported (CFI = .930, RMSEA = .101), indicating that the unidimensional factor structure was consistent across countries. Metric invariance was also established (ΔCFI = −.010, ΔRMSEA = −.009), demonstrating equivalence of factor loadings across the three national samples. However, scalar invariance was not supported, as constraining item intercepts to equality resulted in a deterioration of model fit beyond the recommended threshold (ΔCFI = −.026, ΔRMSEA = .002).

To identify the source of non-invariance, item intercepts were sequentially freed relative to the scalar model. After Benjamini–Hochberg correction, five of the six items showed significant intercept non-invariance, whereas only Item 1 demonstrated an invariant intercept (ΔCFI = .001, p = .306). Because fewer than two invariant intercepts were identified, the minimum requirement for partial scalar invariance was not met (40). These findings indicate that participants from the three countries responded to most items using systematically different response thresholds, despite sharing the same underlying factor structure and equivalent factor loadings. Consequently, latent mean comparisons across countries were not supported, and subsequent between-country analyses were conducted using observed composite scores. Country-specific normative values are therefore recommended for applied interpretation of HLS-EU-Q6 scores.

Table 5

Measurement Invariance Across Countries (MLR Estimator)

Panel A. Sequential invariance models

Model	χ^2	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	Pass
1. Configural	214.64	27	0.930	0.883	0.101	0.043	—	—	—
2. Metric	249.74	37	0.920	0.903	0.092	0.049	−.010	−.009	✓
3. Scalar	330.42	47	0.894	0.898	0.094	0.058	−.026	.002	✗
4. Strict	1295.95	59	0.537	0.647	0.176	0.144	−.357	.082	✗

Panel B. Intercept non-invariance: one-at-a-time freeing

Item	Domain	ΔCFI	$\Delta\chi^2$	df	p	p (BH-adj)	Status
Item 3	Disease prev.	0.008	24.00	2	<.001	<.001	Non-invariant
Item 4	Disease prev.	0.007	20.87	2	<.001	<.001	Non-invariant
Item 6	Health prom.	0.006	19.64	2	<.001	.0002	Non-invariant
Item 5	Health prom.	0.004	13.16	2	.001	.002	Non-invariant
Item 2	Healthcare	0.003	10.12	2	.006	.008	Non-invariant
Item 1	Healthcare	0.001	2.37	2	.306	.306	Invariant

Note. MLR (Yuan-Bentler) estimator. Grouping variable: country (Venezuela, Ecuador, Peru). Invariance criteria: |ΔCFI| ≤ .010 and ΔRMSEA ≤ .015 (38,39). Panel B: each item's intercept is freed relative to the scalar model; p-values adjusted with Benjamini–Hochberg correction.

Measurement invariance by sociodemographic variables

To assess the suitability of the HLS-EU-Q6 for within-country subgroup comparisons, measurement invariance was evaluated across three sociodemographic variables — sex (male vs. female), education level (three harmonised categories), and age group (18–29, 30–44, 45–59, 60+) — using the same MLR (Yuan-Bentler) estimator and sequential procedure applied for the cross-national analysis (see Table 6, Panel A). Education level demonstrated the strongest invariance profile: full strict invariance was supported across all four constraint levels (all |ΔCFI| ≤ .009), indicating that factor loadings, item intercepts, and residual variances are equivalent across education groups. Age group supported metric and scalar invariance — the latter exactly at the conventional threshold (ΔCFI = −.010). Strict invariance, however, was not retained (ΔCFI = −.033): one-at-a-time residual freeing identified four items with non-invariant residual variances after Benjamini–Hochberg correction — Item 5 (ΔCFI = +.015, $p^{BH} < .001$), Item 6

(ΔCFI = +.008, p^{BH} = .001), Item 1 (ΔCFI = +.005, p^{BH} = .007), and Item 4 (ΔCFI = +.005, p^{BH} = .007); Items 2 and 3 showed invariant residuals (see Table 6, Panel C). These items thus measure health literacy with different precision across age groups even though factor structure and item intercepts are equivalent. Sex showed metric invariance (ΔCFI = −.006) but not scalar (ΔCFI = −.012). One-at-a-time intercept freeing identified item 5 (health promotion domain) as the sole source of non-invariance after Benjamini–Hochberg correction (ΔCFI = +.006, p^{BH} < .001; Table 6, Panel B). The remaining five items showed invariant intercepts across male and female respondents, constituting partial scalar invariance. Notably, item 5 also appeared among the items with non-invariant residual variances by age group (Panel C), marking it as the most measurement-sensitive item across multiple demographic dimensions. Direct observed-score comparisons between male and female respondents should therefore be interpreted with caution.

Table 6

Measurement Invariance Across Sociodemographic Groups (MLR Estimator)

Panel A. Sequential invariance models

Grouping variable	Model	χ^2(scaled)	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	Pas
Sex	Configural	173.70	18	.929	.882	.092	.041	—	—	—
	Metric	193.37	23	.923	.899	.085	.044	−.006	−.007	✓
	Scalar	223.15	28	.911	.905	.083	.047	−.012	−.002	✗
	Strict	229.72	34	.911	.922	.075	.049	.000	−.008	n/a
Age group	Configural	186.71	36	.932	.887	.091	.041	—	—	—
	Metric	223.41	51	.922	.909	.082	.051	−.010	−.009	✓
	Scalar	260.70	66	.912	.920	.076	.054	−.010	−.006	✓
	Strict	352.05	84	.879	.914	.079	.063	−.033	+.003	✗
Education	Configural	191.14	27	.925	.875	.095	.043	—	—	—
	Metric	219.06	37	.917	.899	.085	.048	−.008	−.010	✓
	Scalar	248.45	47	.908	.912	.080	.050	−.009	−.005	✓
	Strict	266.87	59	.905	.928	.072	.052	−.003	−.008	✓

Panel B. Intercept non-invariance: sex — one-at-a-time freeing

Item	ΔCFI	χ^2(diff)	df	p	p_BH	Status
Item 5 (health promotion)	+.006	16.35	1	.0001	< .001	Non-invariant
Item 1 (healthcare)	+.002	2.92	1	.088	.144	Invariant
Item 3 (disease prevention)	+.002	2.77	1	.096	.144	Invariant
Item 6 (health promotion)	+.002	4.22	1	.040	.120	Invariant
Item 2 (healthcare)	+.001	0.53	1	.465	.558	Invariant
Item 4 (disease prevention)	+.001	0.16	1	.690	.690	Invariant

Panel C. Residual non-invariance: age group — one-at-a-time freeing

Item	ΔCFI	χ^2(diff)	df	p	p_BH	Status
Item 5 (health promotion)	+.015	33.84	3	< .001	< .001	Non-invariant
Item 6 (health promotion)	+.008	18.70	3	< .001	.001	Non-invariant
Item 1 (healthcare)	+.005	12.99	3	.005	.007	Non-invariant
Item 4 (disease prevention)	+.005	13.46	3	.004	.007	Non-invariant
Item 2 (healthcare)	+.002	7.28	3	.064	.076	Invariant
Item 3 (disease prevention)	+.001	4.63	3	.201	.201	Invariant

Note. MLR (Yuan-Bentler) estimator, full pooled sample (N = 2,034). Grouping variables: sex (male vs. female), education level (low, medium, high), age group (18–29, 30–44, 45–59, 60+). ΔCFI and ΔRMSEA reflect change relative to the preceding model. Criterion: $|\Delta CFI| \leq .010$ (38); $\Delta RMSEA \leq +.015$ (39). Panel B: intercepts freed one at a time relative to the scalar model; Panel C: residual variances freed one at a time relative to the strict model. All p-values corrected with the Benjamini–Hochberg procedure. The Strict model for sex is shown for completeness only (marked "n/a" in the Pass column); because scalar invariance was not achieved for sex, this result should not be interpreted as confirming strict invariance.

Known-groups validity and cross-country comparisons

Known-groups validity. Health literacy differed significantly according to educational level (Welch $F(2, 916.8) = 13.12$, $p < .001$, $\varepsilon^2 = .012$). Participants with university or postgraduate education scored significantly higher than those with technical education (mean difference = 0.14, 95% CI [0.03, 0.24], $p = .006$) and elementary education or below (mean difference = 0.18, 95% CI [0.09, 0.27], $p < .001$), whereas no significant differences were observed between the technical and elementary education groups. Health literacy also varied across age groups (Welch $F(3, 673.1) = 4.55$, $p = .004$, $\varepsilon^2 = .007$), with participants aged 45–59 years and ≥60 years reporting higher scores than those aged 18–29 years, and the oldest group also scoring higher than adults aged 30–44 years. In contrast, health literacy did not differ according to chronic disease status (Welch $F(1, 600.2) = 2.34$, $p = .126$). Although statistically significant, the observed effect sizes were small across all known-groups comparisons ($\varepsilon^2 = .007$–$.012$), providing evidence of known-groups validity consistent with previous validation studies conducted in Latin American populations (see Table 7).

Cross-country comparisons. Because scalar invariance across countries was not supported, between-country comparisons were conducted using observed composite scores rather than latent means. Health literacy differed significantly across the three participating countries (Welch $F(2, 1158.4) = 75.19$, $p < .001$, $\eta^2 = .061$). Venezuela showed the highest mean score (M = 3.76, SD = 0.80), followed by Peru (M = 3.57, SD = 0.71) and Ecuador (M = 3.33, SD = 0.58). Games–Howell post hoc tests indicated that all pairwise country comparisons were statistically significant ($p < .001$), with effect sizes ranging from small to moderate (Cohen's $d = 0.24$–0.59). These findings indicate that observed health literacy scores differed consistently across the three national samples (Table 7).

Table 7

Known-Groups Validity and Cross-Country Comparisons

Panel A. Known-groups validity (Welch ANOVA)

Variable / Group	n	M	SD	Welch F (df)	p	ε²
Education level				F(2, 916.8) = 13.12	< .001	.012
University/Postgraduate	1,017	3.67	0.76			
Technical	332	3.53	0.69			
Elementary or below	685	3.49	0.73			
Age group				F(3, 673.1) = 4.55	.004	.007
18–29	870	3.55	0.68			
30–44	557	3.54	0.76			
45–59	415	3.67	0.80			
60 or older	192	3.72	0.82			
Chronic disease				F(1, 600.2) = 2.34	.126	< .001
No	1,615	3.60	0.72			
Yes	419	3.53	0.81			

Games–Howell post-hoc (education): University/Postgraduate > Elementary (difference = 0.18, 95% CI [0.09, 0.27], p < .001) and University/Postgraduate > Technical (difference = 0.14, 95% CI [0.03, 0.24], p = .006). Age: 45–59 and 60+ scored higher than 18–29 (p = .044 and p = .038); 60+ scored higher than 30–44 (p = .048). Chronic disease: no significant post-hoc (omnibus non-significant).

Panel B. Cross-country comparisons (observed composite scores)

Country	n	M	SD	Welch F (df)	p / η²	Pairwise d (GH)
Venezuela	957	3.76	0.80	F(2, 1158.4) = 75.19	p < .001 η^2 = .061	VE – EC: d = 0.59 VE – PE: d = 0.24 EC – PE: d = 0.37
Ecuador	617	3.33	0.58			
Peru	460	3.57	0.71			

Note. Welch ANOVA with Games–Howell post-hoc tests. ε² = epsilon-squared (effect size for known-groups comparisons); η² = eta-squared (country comparison). All pairwise country differences significant at p < .001. GH = Games–Howell. VE = Venezuela; EC = Ecuador; PE = Peru.

4. Discussion

This study examined the psychometric properties of the HLS-EU-Q6 in a combined Latin American sample of 2,034 adults from Venezuela, Ecuador, and Peru and is the first to evaluate its cross-national measurement invariance in this region. Overall, the findings support HLS-EU-Q6 as a reliable and structurally valid instrument for assessing global health literacy within each country. However, the absence of scalar invariance across countries indicates that observed differences in scores should not be interpreted as equivalent differences in the underlying construct, underscoring the need for country-specific normative values when applying the instrument across Latin America.

Factor structure and reliability

Evidence from both exploratory and confirmatory factor analyses consistently supported a unidimensional representation of the HLS-EU-Q6. Although the original HLS-EU framework conceptualizes health literacy as comprising the domains of healthcare, disease prevention, and health promotion, the six-item version did not reproduce this multidimensional structure. Instead, the theoretical three-factor model produced an inadmissible Heywood solution, whereas the bifactor model failed to converge. These findings are consistent with the methodological limitations inherent to short-form instruments, in which each theoretical domain is represented by only two indicators. Consequently, the HLS-EU-Q6 should be interpreted and scored as a global indicator of health literacy rather than as a multidimensional measure.

The retained unidimensional model also demonstrated satisfactory reliability. Internal consistency was adequate in the pooled sample ($\alpha = .830$; $\omega = .832$) and remained acceptable across the three participating countries, although coefficients were slightly lower in Venezuela ($\alpha = .782$). This pattern is consistent with the greater dispersion of scores observed in the Venezuelan sample, which may attenuate inter-item correlations relative to more homogeneous response distributions. Furthermore, composite reliability (CR = .866) and average variance extracted (AVE = .519) exceeded the recommended thresholds proposed by Fornell and Larcker (36), providing additional evidence of the internal consistency and convergent validity of the scale.

Comparison with previous validation studies

The present findings converge with previous psychometric evaluations of the HLS-EU-Q6 conducted in Brazil, Portugal, and Turkey. Mialhe et al.(20) validated the Brazilian Portuguese version in 783 adults recruited from four Brazilian cities and reported almost identical KMO values (.84 vs. .844 in the present study), indicating comparable sampling adequacy despite differences in recruitment strategies. Standardised factor loadings were also remarkably similar (Brazil: .53–.64; present study: .58–.73), with both investigations supporting a well-fitting unidimensional structure. Reliability estimates likewise converged, with Mialhe et al. reporting $\alpha = .87$ and $\omega = .87$ compared with $\alpha = .83$ and $\omega = .83$ in the present study.

Similarly, Pedro et al. (19) reported satisfactory internal consistency for the European Portuguese version ($\alpha = .78$) in a sample of 977 adults, although confirmatory factor analysis could not be performed in that study. More recently, Yeşildal (22) validated the Turkish version in 475 adults and reported $\alpha = .82$, $\omega = .84$, CR = .89, and AVE = .59, all of which are highly comparable to the estimates observed in the present investigation (CR = .866; AVE = .519). Collectively, these findings provide convergent evidence that the HLS-EU-Q6 maintains stable psychometric performance across culturally and linguistically diverse populations.

Importantly, the present study extends this body of evidence beyond previous single-country validations by providing the first systematic assessment of measurement invariance across countries and sociodemographic groups. This contribution strengthens the evidence supporting the structural robustness of the HLS-EU-Q6 while simultaneously identifying important limitations regarding the comparability of scores across national contexts.

Measurement invariance and comparability across countries

The most important contribution of this study concerns the cross-national measurement invariance of the HLS-EU-Q6. Sequential invariance testing supported configural and metric invariance but failed to support scalar invariance. This pattern indicates that the instrument measures the same underlying construct across Venezuela, Ecuador, and Peru and that the relationships between individual items and the latent health literacy factor are equivalent. However, participants from different countries do not use the response scale in the same way when they possess the same underlying level of health literacy.

Consequently, direct comparisons of latent means across countries are not supported, and cross-national rankings based on observed mean scores should be interpreted with caution. Rather than reflecting deficiencies in the instrument itself, scalar non-invariance suggests that national context systematically influences response thresholds, a phenomenon commonly reported in cross-cultural psychometric research (41,42). From a practical perspective, these findings indicate that applications of the HLS-EU-Q6 for screening, population surveillance, and programme evaluation should rely on country-specific normative values rather than a common regional reference. Using a single regional cut-off could lead to inaccurate estimates of the prevalence of limited health literacy and, consequently, influence the identification of priority populations and the planning of health promotion and prevention strategies.

Although Venezuela obtained the highest observed health literacy scores, followed by Peru and Ecuador, these differences cannot be interpreted as equivalent differences in latent health literacy because response thresholds differed substantially across countries. Therefore, the lower observed scores in Ecuador should not be interpreted as evidence of poorer underlying health literacy relative to the other participating countries. Instead, they should be understood within the context of measurement non-equivalence and differences in response patterns across countries. Moreover, the distribution of observed scores provides additional context for understanding the absence of scalar invariance. Ecuador showed a consistent concentration of scores below the scale midpoint across all sociodemographic subgroups, a pattern that was not explained by sample composition ($R^2 = .023$). In contrast, Venezuela displayed the greatest variability in scores across the full response range, whereas Peru exhibited an age pattern that differed from the pooled sample, with younger adults reporting higher health literacy than older adults within the national subsample ($R^2 = .047$).

Taken together, these findings suggest that the observed cross-country differences reflect not only measurement non-equivalence but also meaningful contextual variation in how health literacy is expressed and reported. Differences in educational opportunities, access to health information, digital literacy, and broader social and healthcare environments may contribute to these distinct response patterns. Accordingly, the lack of scalar invariance should not be viewed solely as a psychometric limitation but also as evidence that perceptions of health literacy are shaped by national context. From a public health perspective, these findings highlight the importance of interpreting health literacy within the social and healthcare context of each country. Developing country-specific normative values may improve the accuracy of population surveillance, facilitate the identification of groups at greater risk of limited health literacy, and support the design and evaluation of health promotion and disease prevention programmes tailored to national needs. Consequently, the use of a single regional reference for interpreting HLS-EU-Q6 scores across Latin America is not recommended.

Measurement invariance across sociodemographic groups

In contrast to the cross-national findings, measurement invariance across sociodemographic groups was generally supported. Educational level achieved full strict invariance, indicating equivalent factor loadings, intercepts, and residual variances across groups. Consequently, both observed and latent comparisons across educational attainment can be interpreted with

confidence, providing strong psychometric support for using the HLS-EU-Q6 to examine educational inequalities in health literacy.

Age-group comparisons yielded metric and scalar invariance but not strict invariance. Four items (Items 5, 6, 1, and 4) exhibited non-invariant residual variances, indicating that measurement precision differed across age groups despite equivalent latent factor structure and item intercepts. Importantly, this pattern does not compromise latent mean comparisons, which remain appropriate under scalar invariance (42). Rather, it suggests that some items are measured with different levels of precision across generations.

One plausible explanation relates to generational differences in health information practices. Older adults continue to rely primarily on healthcare professionals and traditional media, whereas younger adults increasingly access health information through digital platforms, mobile applications, and social media (43,44). Such differences may increase response variability for items requiring active information seeking and evaluation, particularly those assessing health promotion and critical appraisal. Although these findings do not threaten the construct validity of the instrument, they indicate that the reliability of specific items may vary according to age.

Sex invariance presented a more nuanced pattern. Metric invariance was supported, whereas full scalar invariance was not. Differential item functioning analyses identified Item 5 ("Finding activities that are good for your mental wellbeing, such as meditation, exercise, walks etc.") as the only item exhibiting intercept non-invariance. Men reported slightly higher observed scores than women on this item ($M = 3.82$ vs. 3.68; $t(1754.9) = 2.97$, $p = .003$, $d = 0.133$), although the effect size was small.

This finding is consistent with previous evidence indicating that sex-related differential item functioning can occur in health literacy instruments. Rouquette et al. (18) similarly identified sex-related DIF in the HLS-EU-Q16, although involving a different item. The present findings therefore suggest that differential functioning is not specific to a single indicator but may reflect broader differences between male and female respondents in the perception and reporting of health-related competencies. One possible explanation is that traditional masculine norms may encourage greater self-confidence when reporting health promotion behaviours while simultaneously discouraging acknowledgement of mental health needs or help-seeking (45). Importantly, this item-specific DIF did not translate into overall sex differences in health literacy, supporting the conclusion that the observed non-invariance is localized rather than indicative of systematic measurement bias across the scale.

Known-groups validity

Known-groups analyses provided additional evidence supporting the construct validity of the HLS-EU-Q6. As expected, educational attainment showed the strongest association with health literacy, with participants holding university or postgraduate qualifications reporting significantly higher scores than those with technical or elementary education. Although the observed effect size was small, this educational gradient has been consistently reported in previous validation studies conducted in Ecuador, Peru, and Venezuela (46–48), supporting the instrument's ability to discriminate between groups with theoretically expected differences.

Age was also associated with health literacy, although the direction of this association differed from that reported in the European Health Literacy Survey (HLS-EU), in which younger adults generally demonstrated higher levels of health literacy (8). One possible explanation is the composition of the present sample. Online convenience sampling may preferentially recruit older adults with greater educational attainment and digital engagement than the general population, thereby attenuating the negative age gradient commonly observed in population-based surveys. Notably, Peru represented an exception to the pooled findings, with younger participants reporting higher health literacy than older adults within the national subsample, suggesting that contextual differences in digital health information access and educational opportunities may shape age-related patterns differently across countries.

In contrast, no significant differences were observed according to chronic disease status. Although individuals living with chronic conditions are often expected to develop greater health literacy through repeated interactions with healthcare services, the absence of significant differences in the present study may reflect the relatively small effect sizes involved, cross-country heterogeneity, or the limited sensitivity of a six-item screening instrument to detect subtle differences in health literacy.

Public health implications

The present findings have several implications for public health research and practice in Latin America. First, the HLS-EU-Q6 provides a brief, reliable, and psychometrically supported instrument for rapidly assessing health literacy within individual countries. Its brevity makes it particularly suitable for incorporation into population health surveys, community screening initiatives, and primary healthcare services, where the rapid identification of individuals with limited health literacy may facilitate the implementation of targeted educational and preventive interventions.

Second, the lack of scalar invariance across countries indicates that health literacy should not be monitored using a single regional reference or common cut-off score. Instead, country-specific normative values are needed to ensure that screening, surveillance, and programme evaluation accurately reflect the characteristics of each population. Without such contextual interpretation, observed differences between countries may lead to inappropriate prioritisation of health policies or inefficient allocation of public health resources.

Third, improving health literacy has implications beyond individual knowledge. Individuals with limited health literacy are less likely to understand health information, navigate healthcare systems, engage in preventive behaviours, adhere to treatment recommendations, and participate effectively in health promotion programmes. Consequently, strengthening health literacy represents an important strategy for supporting disease prevention, improving access to primary healthcare, promoting informed health decision-making, and reducing health inequalities across Latin America.

Finally, the present findings highlight the importance of routinely evaluating measurement invariance before using health literacy instruments in multinational studies. Instruments may demonstrate satisfactory reliability and structural validity while still failing to support meaningful comparisons across countries if response processes differ systematically across cultural contexts. Incorporating measurement invariance into cross-cultural validation studies should therefore become standard practice in international public health research.

Strengths and limitations

This study has several strengths. It represents the first cross-national psychometric evaluation of the HLS-EU-Q6 in Latin America and, to our knowledge, the first study of this instrument to comprehensively assess measurement invariance across the three participating countries and major sociodemographic groups. The large multicountry sample and the combination of exploratory, confirmatory, and invariance analyses provide a robust evaluation of the instrument's psychometric performance.

Nevertheless, several limitations should be acknowledged. Participants were recruited through convenience sampling, limiting the generalisability of the findings to the adult populations of Venezuela, Ecuador, and Peru. Online recruitment may also have resulted in the overrepresentation of individuals with higher educational attainment and greater digital literacy, potentially inflating health literacy estimates. In addition, the brief six-item format necessarily limits the complexity of the latent structure that can be modelled and precludes reliable assessment of the theoretical health literacy domains. Finally, the absence of national normative data limits the clinical and public health interpretation of individual scores, highlighting an important direction for future research.

5. Conclusions

The HLS-EU-Q6 demonstrated satisfactory reliability, convergent validity, and a stable unidimensional structure for assessing global health literacy in adults from Venezuela, Ecuador, and Peru. The instrument supported meaningful comparisons across educational groups and, at the latent level, across age groups. However, the absence of scalar invariance across countries indicates that observed score differences should not be interpreted as equivalent differences in underlying health literacy, limiting the comparability of scores across national contexts.

These findings support the use of the HLS-EU-Q6 as a brief screening instrument for assessing health literacy within individual countries. However, multinational applications should rely on country-specific normative values and avoid direct comparisons of observed scores across countries. More broadly, the present findings highlight that establishing reliability and factor structure alone may be insufficient for instruments intended for cross-national use, underscoring the importance of evaluating measurement invariance before health literacy measures are applied in multinational surveillance, programme evaluation, or international comparative research.

List of abbreviations

ANOVA: Analysis of variance
AVE: Average variance extracted
BIC: Bayesian Information Criterion
CFA: Confirmatory factor analysis
CFI: Comparative fit index
CI: Confidence interval
CR: Composite reliability
DIF: Differential item functioning
EFA: Exploratory factor analysis
HLS: Health literacy
HLS-EU: European Health Literacy Survey
HLS-EU-Q: European Health Literacy Survey Questionnaire
HLS-EU-Q6: Health Literacy Survey European Questionnaire–Short Form
HLS-EU-Q16: European Health Literacy Survey Questionnaire 16-item short form
HLS-EU-Q47: European Health Literacy Survey Questionnaire 47-item version
HLS-EU-Q12M: Mexican version of the European Health Literacy Survey Questionnaire 12-item version
HLS-EU-SQ10-IDN: Indonesian 10-item short form of the European Health Literacy Survey Questionnaire
HLS-Q12: 12-item Health Literacy Survey Questionnaire
ICC: Intraclass correlation coefficient
KMO: Kaiser–Meyer–Olkin
MAP: Minimum average partial
MLR: Robust maximum likelihood estimation
RMSEA: Root mean square error of approximation
SD: Standard deviation
SRMR: Standardised root mean square residual
TLI: Tucker–Lewis index
VSS: Very Simple Structure
WLSMV: Weighted least squares mean- and variance-adjusted
α: Cronbach's alpha
ω: McDonald's omega
ε^2: Epsilon-squared
η^2: Eta-squared

Declarations:

- **Ethics approval and consent to participate:** This study involved adult human participants and was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained independently in each participating country: Universidad Peruana Cayetano Heredia, Peru (CIEI-389-36-23); Pontificia Universidad Católica del Ecuador, Ecuador (CEISH-860-2023); and Universidad Centroccidental Lisandro Alvarado, Venezuela (CBDCS-4-2023). All participants provided informed consent before completing the survey.
- **Consent for publication:** Not applicable. The manuscript does not contain any individual person's data, images, videos, or identifiable personal information.
- **Availability of data and materials:** The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.
- **Competing interests:** The authors declare that they have no competing interests.
- **Funding:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
- **Authors' contributions:** JFP contributed to the conception and design of the study. AMRF coordinated data collection in Venezuela. JFP coordinated data collection in Ecuador. DCC coordinated data collection in Peru. GOF performed the statistical analyses. VLG contributed to the interpretation and discussion of the results. JFP, DCC and AMRF drafted the manuscript. All authors critically reviewed the manuscript, approved the final version, and agreed to be accountable for all aspects of the work.
- **Acknowledgements:** The authors thank all participants who voluntarily contributed to this study. The authors also acknowledge the collaborating institutions and local research teams in Venezuela, Ecuador, and Peru for their support during data collection.
- **Authors' information:** Not applicable.

References

1. Abdi Almachavan S. The role of health literacy in enhancing preventive healthcare: A comprehensive review of challenges, interventions, and future directions. J Res Clin Med. 2024 Dec 15;12:36. http://dx.doi.org/10.34172/jrcm.35126

2. Marchant E. Re-energizing health literacy in Wales: a testbed for health, education and prosperity for all. Health Promot Int. 2024;39(3). https://doi.org/10.1093/heapro/daae055

3. Liu Y, Guo Y, Yan X, Ding R, Tan H, Wang Y, et al. Assessment of health literacy in patients with polycystic ovary syndrome and its relationship with health behaviours: a cross-sectional study. BMJ Open. 2023 Nov 24;13(11):e071051. https://doi.org/10.1136/bmjopen-2022-071051

4. Li Y, Lv X, Liang J, Dong H, Chen C. The development and progress of health literacy in China. Front Public Health. 2022;10. https://doi.org/10.3389/fpubh.2022.1034907

5. Palumbo R. Examining the impacts of health literacy on healthcare costs. An evidence synthesis. Health Serv Manage Res. 2017;30(4):197–212. https://doi.org/10.1177/0951484817733366

6. Ahmann E, Leikin SC, Smith K, Ellington L, Pille RO. Exploring health literacy and its relationship to health and wellness coaching. IJEBCM. 2020;18(2):83–100. https://doi.org/10.24384/9qz4-w404

7. Lega F, Kreutzer P. Building and Improving Health Literacy in the 'New Normal' of Health Care: Frameworks and Actions [Internet]. Emerald Publishing Limited; 2023 [citado 3 de septiembre de 2026]. https://doi.org/10.1108/9781837533367

8. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, et al. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). The European Journal of Public Health. 2015 Dec;25(6):1053–8. https://doi.org/10.1093/eurpub/ckv043

9. Lima ACP, Maximiano-Barreto MA, Martins TCR, Luchesi BM. Factors associated with poor health literacy in older adults: A systematic review. Geriatr Nurs (Minneap). 2024;55:242–54. https://doi.org/10.1016/j.gerinurse.2023.11.016

10. Alsalemi N, Kilpatrick K, Sadowski CA. Health and medication literacy in older adults. In: Health Literacy in Medicines Use and Pharmacy: a Definitive Guide [Internet]. 2024. p. 193–204. https://doi.org/10.1016/B978-0-12-824407-4.00001-5

11. Griebler R, Link T, Schütze D, Straßmayr C. Gesundheitskompetenz messen: Methoden und Instrumente zur Erfassung der allgemeinen Gesundheitskompetenz bei Erwachsenen. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2025 Feb 13. https://doi.org/10.1007/s00103-025-04010-y

12. Lorini C, Lastrucci V, Mantwill S, Vettori V, Bonaccorsi G, Literacy FH, et al. Measuring health literacy in Italy: a validation study of the HLS-EU-Q16 and of the HLS-EU-Q6 in Italian language, conducted in Florence and its surroundings. Ann Ist Super Sanità. 2019;55(1):10–8. https://doi.org/10.4415/ann_19_01_04

13. Gustafsdottir SS, Sigurdardottir AK, Arnadottir SA, Heimisson GT, Mårtensson L. Translation and cross-cultural adaptation of the European Health Literacy Survey Questionnaire, HLS-EU-Q16: The Icelandic version. BMC Public Health. 2020;20(1):1–11. https://doi.org/10.1186/s12889-020-8162-6

14. Kwon DH, Kwon YD. Patterns of health literacy and influencing factors differ by age: a cross-sectional study. BMC Public Health. 2025 Apr 26;25(1):1556. https://doi.org/10.1186/s12889-025-22838-6

15. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: A systematic review and integration of definitions and models. BMC Public Health. 2012 Dec 25;12(1):80. https://doi.org/10.1186/1471-2458-12-80

16. Bergman L, Nilsson U, Dahlberg K, Jaensson M, Wångdahl J. Validity and reliability of the arabic version of the HLS-EU-Q16 and HLS-EU-Q6 questionnaires. BMC Public Health. 2023 Feb 10;23(1):304. https://doi.org/10.1186/s12889-023-15226-5

17. Bergman L, Nilsson U, Dahlberg K, Jaensson M, Wångdahl J. Validity and reliability of the swedish versions of the HLS-EU-Q16 and HLS-EU-Q6 questionnaires. BMC Public Health. 2023 Apr 20;23(1):724. https://doi.org/10.1186/s12889-023-15519-9

18. Rouquette A, Nadot T, Labitrie P, Van den Broucke S, Mancini J, Rigal L, et al. Validity and measurement invariance across sex, age, and education level of the French short versions of the European Health Literacy Survey Questionnaire. PLoS One. 2018 Dec 6;13(12):e0208091. https://doi.org/10.1371/journal.pone.0208091

19. Pedro AR, Raposo B, Luís L, Amaral O, Escoval A, Simões Dias S. Portuguese Version of the HLS-EU-Q6 and HLS-EU-Q16 Questionnaire: Psychometric Properties. Int J Environ Res Public Health. 2023 Feb 7;20(4):2892. https://doi.org/10.3390/ijerph20042892

20. Mialhe FL, Moraes KL, Bado FMR, Brasil VV, Sampaio HA de C, Rebustini F. Psychometric properties of the adapted instrument European Health Literacy Survey Questionnaire short-short form. Rev Lat Am Enfermagem. 2021;29. https://doi.org/10.1590/1518-8345.4362.3436

21. Mousum S, Salwa M, Fatema K, Haque MA. Psychometric Validation of the Translated and Adapted Bengali Version of the HLS-EU-Q16/Q6 for Adults. INQUIRY: The Journal of Health Care Organization, Provision, and Financing. 2024 Jan 15;61. https://doi.org/10.1177/00469580241248132

22. Yeşildal M. Validity and reliability of the Turkish version of the HLS-EU-Q6 questionnaire. BMC Public Health. 2025 Aug 2;25(1):2633. https://doi.org/10.1186/s12889-025-24081-5

23. Rachmani E, Hsu CY, Nurjanah N, Chang PW, Shidik GF, Noersasongko E, et al. Developing an Indonesia's health literacy short-form survey questionnaire (HLS-EU-SQ10-IDN) using the feature selection and genetic algorithm. Comput Methods Programs Biomed. 2019;182. https://doi.org/10.1016/j.cmpb.2019.105047

24. 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 Dec 28;18(1):506. https://doi.org/10.1186/s12913-018-3275-7

25. García-Vera EM, Doubova S V., Sánchez-Arenas R, Monroy-Guzmán A. Validation of the health literacy scale in patients with hypertension in Mexico. Gac Med Mex. 2023 Sep 21;159(4). https://doi.org/10.24875/gmm.m23000788

26. Spencer M, Kemp N, Cruickshank V, Nash R. Examining the relationship between demographic variables and perceived health literacy challenges in Tasmania, Australia. Health Promot J Austral. 2025 Jan 29;36(1). https://doi.org/10.1002/hpja.905

27. Sørensen K, Van Den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, et al. 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). https://doi.org/10.1186/1471-2458-13-948

28. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing [Internet]. [cited 2026 Aug 17]. Available from: https://www.r-project.org/

29. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research [Internet]. Evanston, Illinois: Northwestern University; 2026 [cited 2026 Aug 25]. Available from: https://CRAN.R-project.org/package=psych

30. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw. 2012;48(2). https://doi.org/10.18637/jss.v048.i02

31. Jorgensen TD, Pornprasertmanit S, Schoemann AM, Rosseel Y. semTools: Useful tools for structural equation modeling. [Internet]. 2022 [cited 2026 Aug 25]. Available from: https://CRAN.R-project.org/package=semTools

32. Horn JL. A Rationale and Test for the Number of Factors in Factor Analysis. Psychometrika. 1965 Jun 1;30(2):179–85. https://doi.org/10.1007/BF02289447

33. Velicer WF. Determining the Number of Components from the Matrix of Partial Correlations. Psychometrika. 1976 Sep 1;41(3):321–7. https://doi.org/10.1007/BF02293557

34. Revelle W, Rocklin T. Very Simple Structure: An Alternative Procedure For Estimating The Optimal Number Of Interpretable Factors. Multivariate Behav Res. 1979 Oct;14(4):403–14. https://doi.org/10.1207/s15327906mbr1404_2

35. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling. 1999 Jan;6(1):1–55. https://doi.org/10.1080/10705519909540118

36. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research. 1981 Feb;18(1):39. https://doi.org/10.2307/3151312

37. Hair JF, Babin BJ, Anderson RE, Black WC. Multivariate data analysis. 8th ed. Cengage Learning; 2022.

38. Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Struct Equ Modeling. 2002 Apr;9(2):233–55. https://doi.org/10.1207/S15328007SEM0902_5

39. Chen FF. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. Struct Equ Modeling. 2007 Jul 31;14(3):464–504. https://doi.org/10.1080/10705510701301834

40. Byrne B, Shavelson R, Muthén B. Testing for the Equivalence of Factor Covariance and Mean Structures: The Issue of Partial Measurement Invariance. Psychol Bull. 1989;105(3):456–66. https://psycnet.apa.org/doi/10.1037/0033-2909.105.3.456

41. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review. 2016 Sep;41:71–90. https://doi.org/10.1016/j.dr.2016.06.004

42. Vandenberg RJ, Lance CE. A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. Organ Res Methods. 2000 Jan 1;3(1):4–70. https://doi.org/10.1177/109442810031002

43. Guo N, Guo Z, Zhao S, Ho SY, Fong DYT, Lai AYK, et al. Digital inequalities in health information seeking behaviors and experiences in the age of web 2.0: A population-based study in Hong Kong. PLoS One. 2021 Mar 30;16(3):e0249400. https://doi.org/10.1371/journal.pone.0249400

44. Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. Cogent Soc Sci. 2017 Jan 1;3(1). https://doi.org/10.1080/23311886.2017.1302785

45. Möller-Leimkühler AM. Barriers to help-seeking by men: a review of sociocultural and clinical literature with particular reference to depression. J Affect Disord. 2002 Sep;71(1–3):1–9. https://doi.org/10.1016/s0165-0327(01)00379-2

46. Ortiz G, Francisco-Pérez J, López-Guerra V, Vuele-Duma DM, Ayora-Apolo DC, Rojas A. Psychometric validation of the Spanish HLS-EU-Q16 in Ecuador: evidence for health literacy assessment and public health education. Front Public Health. 2026 Feb 11;14. https://doi.org/10.3389/fpubh.2026.1717946

47. Condor-Camara DF, Francisco-Pérez J, Ortiz-Francisco G, Leiton-Espinoza ZE, Jurado Boza JR, Coaquira Machaca NA, et al. Health literacy: psychometric properties of the HLS-EU-Q16 questionnaire in a Peruvian adult population. Front Public Health. 2026 Feb 5;14. https://doi.org/10.3389/fpubh.2026.1762195

48. Francisco-Pérez J, López-Guerra V, Ortíz G, Rojas A, Martínez C, Serrano R. Psychometric properties and factorial structure of the Spanish version of the HLS-EU-Q16 questionary in Venezuelan adults. PLoS One. 2025 Jun 25;20(6):e0324782. https://doi.org/10.1371/journal.pone.0324782