



## Artigos

### 0194/2025 - A global measurement tool to assess population health literacy fits all? A critical paper

Um instrumento para avaliar o letramento em saúde, em nível global, serve para todos? Um artigo crítico

[Facebook](#)[Twitter](#)[Whatsapp](#)

#### Autor:

• Fábio Luiz Mialhe - Mialhe, FL - <[mialhe@unicamp.br](mailto:mialhe@unicamp.br)>

ORCID: <https://orcid.org/0000-0001-6465-0959>

#### Coautor(es):

• Katarinne Lima Moraes - Moraes, KL - <[katarinnemoraes@gmail.com](mailto:katarinnemoraes@gmail.com)>

ORCID: <https://orcid.org/0000-0001-6169-0461>

• Vanessa da Silva Carvalho Vila - da Silva Carvalho Vila, Vanessa - <[vscvila@uol.com.br](mailto:vscvila@uol.com.br)>

ORCID: <https://orcid.org/0000-0002-1785-8682>

• Helena Alves de Carvalho Sampaio - Sampaio, HAC - <[dr.hard2@gmail.com](mailto:dr.hard2@gmail.com)>

ORCID: <https://orcid.org/0000-0001-5353-8259>

• Virginia Visconde Brasil - Brasil, VV - <[viscondebrasil@gmail.com](mailto:viscondebrasil@gmail.com)>

ORCID: <https://orcid.org/0000-0002-0279-9878>

• Flávio Rebustini - Rebustini, F - <[frebustini@gmail.com](mailto:frebustini@gmail.com)>

ORCID: <https://orcid.org/0000-0002-3746-3266>

#### Resumo:

Health Literacy (HL) is considered by the World Health Organization (WHO) as an important determinant of health and several instruments have been elaborated and adapted to measure this construct. In 2022, a group of researchers started a project aimed at developing a HL assessment instrument with global applicability, based on different versions of the European Health Literacy Survey Questionnaire (HLS-EU-Q). This article critically analyses the characteristics of those instruments to help improve health assessment practices and avoiding omitting measurement of locally important information, therefore committing an epistemic injustice..Analysis detected several flaws in the instruments' aforementioned aspects and the conceptual model proposed by WHO which limits their application as global instruments. Countries worldwide should develop specific instruments to measure the general HL of their population based on its sociocultural characteristics and specificities of their health care systems thus avoiding wrong decision-making, which can compromise the effectiveness of policies resulting from these surveys

#### Palavras-chave:

Health literacy; Psychometrics; Social determinants of health.

## Abstract:

O letramento em saúde (LS) tem sido considerado, pela Organização Mundial da Saúde (OMS), como importante determinante da saúde e diversos instrumentos têm sido elaborados e adaptados para mensurar esse construto. Em 2022, um grupo de pesquisadores iniciou um projeto voltado ao desenvolvimento de um instrumento de avaliação do LS com aplicabilidade global, baseado em diferentes versões do European Health Literacy Survey Questionnaire (HLS-EU-Q). O objetivo deste artigo é analisar criticamente as características desses instrumentos, com o intuito de contribuir para o aprimoramento das práticas de avaliação em saúde e evitar a omissão de informações localmente relevantes, o que configura uma injustiça epistêmica. A análise identificou diversas problemas nos instrumentos e no modelo conceitual proposto pela OMS, os quais limitam sua aplicabilidade como ferramenta global. Recomenda-se que os países desenvolvam instrumentos específicos para avaliar o LS de suas populações, considerando suas características socioculturais e particularidades dos respectivos sistemas de saúde, a fim de evitar decisões equivocadas que possam comprometer a efetividade das políticas públicas resultantes dessas avaliações.

## Keywords:

Letramento em Saúde; Psicometria; Determinantes Sociais da Saúde.

## Conteúdo:

### Introduction

Health Literacy (HL) is currently regarded as an important determinant of population health by the World Health Organization (WHO)<sup>1</sup>. According to the 2021 Health Promotion Glossary of Terms, health literacy “represents the personal knowledge and competencies that accumulate through daily activities, social interactions and across generations. Personal knowledge and competencies are mediated by the organizational structures and availability of resources that enable people to access, understand, appraise and use information and services in ways that promote and maintain good health and well-being for themselves and those around them”<sup>1</sup>.

Evidence has shown that HL levels are associated with several health outcomes, thus measuring HL population levels has been an object of interest for several researchers, managers and politicians over the years<sup>2,3,4</sup>.

The US was the first country to measure HL population levels by surveys conducted since the 1990s, with greater emphasis on instruments focusing on HL’s functional dimension, i.e., reading, writing and numeracy skills in the health context. In this perspective, instruments like the Test of Functional Health Literacy in Adults (TOFHLA) and the Rapid Estimate of Adult Literacy in Medicine (REALM) stand out<sup>5</sup>. TOFHLA is a 50-item reading comprehension and 17-item numeracy test that requires up to 22 minutes to administer. The reading comprehension items present sentences with missing words and individuals select one out of four response options to fill in the gaps. For the numeracy items, questions related to medical prescriptions are presented and the individual is asked to respond verbally<sup>5</sup>. REALM is a reading recognition test consisting of 125 words, in its original form, or 66 words, in the revised version, arranged in three columns in ascending order of number of syllables and difficulty. It measures patients’ ability to pronounce medical words and lay terms related to body parts and illnesses<sup>5</sup>.

In the 2010s, the scientific community was proposed two new instruments which considered HL as a multidimensional construct: the European Health Literacy Survey Questionnaire (HLS-EU-Q) and the Health Literacy Questionnaire (HLQ)<sup>6-9</sup>. Both aimed to assess HL at an individual, but the HLQ additionally took a grounded approach and focused on what community members saw as important<sup>8</sup>. HLQ is a 44-item multidimensional instrument that covers 9 conceptually distinct health literacy areas and can be self-administered or applied via interviews. Areas one to five have a four-point Likert scale ranging from “totally disagree (1)” to “totally agree (4),” whereas in areas six to nine the answers range from “always difficult (1)” to “always easy (5)”<sup>8,9</sup>.

HLS-EU-Q was developed by HLS-EU Consortium researchers and used in the HL European survey conducted in 2011 with citizens aged 15 years and older from eight EU member states<sup>7,10</sup>. For this purpose, researchers elaborated a health literacy model based on a systematic literature review and content analysis of peer-reviewed HL definitions and conceptual frameworks published so far<sup>6,7</sup>. Based on this model, the authors elaborated the European Health Literacy Survey Questionnaire (HLS-EU-Q47) consisting of 47 questions that intent to measure one’s perceived difficulty related to accessing, understanding, appraising and applying health-related information within health care, disease prevention and health promotion contexts<sup>7,10,11</sup>.

Two other short versions were created based on HLS-EU-Q47: HLS-EU-Q16 and HLS-EU-Q610,12. All three versions of HLS-EU-Q have already been adapted to different countries and languages<sup>5,10,12,13,14</sup>.

WHO’s “Health literacy: the solid facts” report recommended regular, standardized HL measurement in the general population<sup>11</sup>. In 2018, WHO-Europe founded the Action Network on Measuring Population and Organizational Health Literacy (M-POHL) with 28 participating WHO-Europe member countries to “support the availability of high-quality internationally comparative data on population and organizational HL” (p.14)<sup>15</sup>. M-POHL conducted a multinational standardized study - the Health Literacy Population Survey 2019–2021 (HLS19) - with 17 WHO-Europe member states using three instruments based on previous HLS-EU survey to collect data, namely the HLS19-Q47 (with 47 items) and two short forms, the HLS19-Q16 and the HLS19-Q1215.

In addition to the survey, the HLS19 project presented specific objectives such as “provide empirical data for evidence-based HL policies in the HLS19 countries, other countries in M-POHL (with observer status), and the larger WHO European Region,” “use the findings to inform policymakers from different sectors and levels about the relevance of HL in their field of decision making,” and “provide evidence for knowledge-based recommendations for HL interventions in the fields of health promotion, disease prevention, and healthcare”, among others<sup>15</sup>.

Based on the HL survey’s experiences in Europe and WHO demands for evaluating HL levels at a global level, a committee of specialists including M-POHL participants oversaw the discussions aimed at developing an instrument for that purpose.

On May 18-19th, 2021, an Informal Expert Consultation on the Development of a Measurement Tool to Assess Population Health Literacy virtual meeting led by European researchers was held. In December 2022, the WHO-Europe committee invited representatives and researchers from WHO regions to participate in an informal technical consultation virtual meeting on the development of a WHO global measurement tool to assess population health literacy via Zoom platform. Prior to the virtual meeting, the committee shared a draft document with country representatives presenting the HL instruments used in the last WHO European survey and requested comments on the HLS19-Q47 for use as a global tool to meet population and country needs within each region<sup>16</sup>. Although the meeting discussions were productive, no consensus was reached. The authors of this article participated in that meeting and, since then, have been discussing this matter, its potentialities and limitations.

Although having a global instrument to measure HL is an interesting idea from an academic standpoint, researchers must be careful in doing so since adopting a global HL measurement without critically reflecting about its validity to the sociocultural context in which it will be applied can generate measurement biases and produce epistemic injustice, i.e., unfair discrimination, marginalization and exclusion of certain population groups in their capacity as knowers or holders of knowledge<sup>4</sup>. In the health care context, this occurs when vulnerable individuals "are regarded as lacking credibility or authority to speak about their experience of their illness or their preferences and interests when making medical decisions"<sup>17</sup>. Additionally, inadequate measurements based on biased models will lead to wrong decision-making, thereby compromising the effectiveness of policies resulting from these surveys not only in Europe, but also in Latin America, Africa and Oceania, with their different cultures, and most Indigenous peoples across the world. It is therefore an ethical imperative to carefully assess the cultural validity and other psychometric properties of HL instruments before claiming them as global tools.

The aim of this article is to critically analyse the characteristics of the WHO instruments proposed as global tools to measure HL from the perspective of psychometrics and social sciences, in order to collaborate with the improvement of health assessment practices and avoid the implementation of public policies based on biased information. For this purpose, the paper was divided into four interconnected subsections that clarify the reader about the chronological and methodological processes involved in developing the HLS-EU-Q instruments: HL definition and conceptual model adopted by M-POHL to propose a global assessment instrument; the HLS-EU-Q47 instrument and its limitations; content of the HLS19-Q12 instrument, proposed as an initial tool to measure HL at the global level by WHO-Europe group, and its limitations; finally, considerations about the HLS19-Q12 psychometric properties.

HL Definition and conceptual model adopted by M-POHL to propose a global assessment instrument

The WHO-Europe Technical Committee has been suggesting to base a global HL instrument on an adapted and revised version of one of the instruments used in the 2019–2021 European Health Literacy Population Survey, i.e., the HLS19-Q47 or one of its short forms - the HLS19-Q16 or HLS19-Q12<sup>16</sup>. All these three new tools were based on previous HLS-EU-Q instruments, including their HL conceptual models and definition<sup>18</sup>. Thus, before discussing the questionnaire suggested by WHO-Europe for developing a new global instrument, we will initially focus on the HL definition and model adopted by the committee.

The HL definition used by the European committee for the 2019-2021 European survey was built based on a systematic review of studies published in Medline, PubMed and Web of Science databases that identified HL definitions and conceptual models<sup>7</sup>.

For that, authors adopted the following inclusion criteria: "(1) written in English; (2) concerned with health literacy in a developed country; and (3) offering relevant content with regard to the definition or conceptualization of health literacy, or a combination of these issues". The review identified 17 definitions published between 1998 and 2009 which were then used by researchers to develop an 'all-inclusive' HL definition as follows: "Health literacy is linked to literacy and encompasses people's knowledge, motivation and competencies to access, understand, appraise and apply information to form judgments and take decisions in terms of healthcare, disease prevention and health promotion to improve quality of life during the life course"<sup>7</sup>, which is still widely used in HL research<sup>19</sup>. Additionally, researchers developed an integrated model of HL that captures the main dimensions of the existing conceptual models reviewed until 2009<sup>7</sup>. This model was mainly developed from a North American perspective, as this country accounted for most publications at that time, and some Western EU countries.

Undoubtedly, the European researchers developed a valuable work and presented a notable contribution to developing the field of HL. However, science is a field under ongoing transformation, and self-criticism of its findings is what ensures its evolution and betterment for humanity.

Despite systematic reviews being a means of generating definitions and instruments, the selected studies in the aforementioned review deal with 'opposite knowledge,' i.e., the viewpoint of researchers about a construct which limits the emergence and development of additional items. Reviewing literature derived from a limited and rather homogeneous sector (white North American and EU academics with the capacity to publish in academic journals) could lead to a very limited perspective and definitely exclude the voice those populations most at risk of marginalisation and stigma, including large portions of people in Asia and South America.

Moreover, this methodology limits the possibility of criticisms based on new field information about whether the proposed model, its concept and theory are effectively efficient and consistent. Clearly, using systematic reviews to elaborate both the concept and the theory is insufficient to generate precision regarding the concept of the latent variable, as well as the intention of creating uniform items. Another limitation of that review is that only studies written in English and published on the Medline, PubMed and Web of Sciences databases were selected, restricting its external validity for other non-European countries and cultures.

In relation to the conceptual model of HL adopted by the European Committee to construct a global instrument, it was based on 12 previous conceptual models developed up to 2009 and have been adjusted to develop an 'integrated HL conceptual model'<sup>6</sup>. According to the authors, people need to have competencies related to the process of accessing, understanding, appraising and applying health-related information to make decisions about maintaining and improving health<sup>6</sup>. Yet, according to authors, "this process generates knowledge and skills which enable a person to navigate three domains of the health continuum: being ill or as a patient in the healthcare setting, as a person at risk of disease in the disease prevention system, and as a citizen in relation to the health promotion efforts in the community"<sup>6</sup>.

Our first consideration about this model is: What is the scientific evidence that validates its assumption that people must necessarily

have these four skills (access, understand, appraise, and apply health information) and necessarily go through all of them to generate HL knowledge and competencies in their life contexts? Additionally, what is the strength of the evidence that people equipped with few of these claimed essential individual abilities are unsuccessful in managing aspects of their health even though they may be illiterate and lack understanding of prevention concepts or how the health system works?

Second: What is the scientific evidence that this process generates knowledge and skills which enable one to navigate the three domains of the health continuum proposed by the model? The Western-derived concept of what constitutes and promotes health, what treatment falls within the scope of healthcare, and what behaviours are considered acceptable or unacceptable within such context are often promoted by Western ideas or even imposed by former colonial perspectives. Consequently, beliefs, trust, or mistrust in Westernized medicine, also known as modern medicine, could also prevent people from modifying their health behaviours or accepting healthcare offered in Western-influenced healthcare systems.

Third: None of the HLS-EU-Q47 questions nor HLS19 instruments explicitly assess individuals' 'motivation,' as described in the definition. In fact, it constitutes a serious omission to include declared concepts in a model and then ignore it in an assessment tool claiming to measure the concept.

Fourth: Regarding the model's domain on disease prevention, which evaluates one's ability to access/understand/interpret and evaluate/make decisions based on risk factors or information on risk factors and derive meaning, what if the individual is uninterested in accessing information about health risks because they like to drink, smoke and eat ultra-processed foods? This fact will likely generate bias in the HL instrument scores. ↯↯

The HLS-EU-Q47 instrument and its limitations

As already described, the European committee's proposal for a global HL measurement instrument establishes that it should be based on an adapted and revised version of the HLS19-Q47 which, in turn, is an adapted version of the HLS-EU-Q47 instrument. Content validity of the HLS-EU-Q47 was evaluated through a focus group with a convenience sample of students and academic staff from three participant universities. However, the external validity of this group's opinions was not tested for the general non-university population.

Again, the adopted procedure makes clear one of the major issues in developing instruments, especially those with universal pretensions: a model established by specialists. In this case, the latent variable studied is highly 'contaminated' by cultural, social, economic and structural aspects from the researcher's context. It is necessary to pay attention to the use of culturally appropriate oral and written language, considering differences associated with urban, regional or cultural minority groups; establish the expected relations between services and the target audience, including synchrony between what is intended by both parties; know the existing metaphors regarding the meanings of symbols and concepts shared by a certain cultural group; learn about values, customs and traditions shared by ethnic and minority groups; evaluate the constructs of the theoretical model to be used; establish procedures that must be followed to achieve intended objectives; and consider broader social, economic and political interests of the target audience<sup>20,21,22</sup>. Thus, using material produced in another country goes far beyond simply translating and culturally adapting it<sup>23</sup>.

In short, since the conceptual model of the HLS-EU-Q47 instrument was not validated with the general population, its content may represent only the researchers' point of view of the field and its constructs without meaning or importance in people's daily lives. This fact is reiterated by Osborne et al. when describing that "The development of a measurement instrument is necessarily influenced by the context and values of the instrument's authors and by the measurement and construct definitions on which the instrument is based. To date, health literacy instruments are based on Western definitions of health literacy and measurement, and these may or may not be in harmony with the worldviews of all participants in the studies using these instruments" <sup>4</sup>.

Hence, future HL definitions and conceptual models should consider the different cultural and social worldviews of people and use mixed methods in their development, with the objective of understanding the phenomenon more broadly in the context of the involved subjects<sup>24</sup>.

Content of the HLS19-Q12 instrument, proposed as an initial tool to measure HL at a global level by the WHO-Europe group, and its limitations

On the draft document shared by the WHO-Europe Technical Committee, it was described "It is strongly recommended that the global tool at minimum includes the HLS19-Q12 scale in its original form to facilitate comparison with existing data"<sup>22</sup>. On the same page, another paragraph states: "We strongly suggest that these items remain unchanged to ensure comparability among regions and countries, as a global population-based tool for measuring health literacy. Countries that would like to add more questions from the Q47 tool may make changes in the other items (questions) [...]"<sup>16</sup>. Despite the group's intention, an important reflection prior to setting "facilitate comparison with existing data" as a goal is whether the content of the instrument used for data collection is valid and relevant to the culture in which it will be applied.

Considering that all adapted versions of the HLS19-Q12 were based on items from the HLS19-Q47, the items of which were either chosen from existing examples in the literature or newly drafted by a Delphi procedure among consortium members or by expert consultation<sup>7</sup>, there is a great risk that the HLS19-Q12 items representing only a statistical adjustment of the HLS-EU-Q47 survey samples, making little or no sense to the respondents. Thus, the actual recommended HL theoretical framework used as basis for developing new HL instruments may not be socioculturally representative of what is valid and significant for populations regarding decision-making related to their health and health services according to their daily experiences; rather, they just represent what researchers 'think' is significant or should be. In social sciences, this aspect is considered an 'etic' approach to the phenomenon, i.e., an interpretation of a health phenomenon aspects from the categories of its observes<sup>25</sup>. Despite international guidelines developed to promote best practices in translating and cross-culturally adapting these tools, the cultural relevance of etic measures remains a potential bias<sup>26</sup>.

Inconsistencies between the content of the HLS-EU-Q47 items and its relevance in people's real lives were demonstrated by Domanska et al.<sup>27</sup> in a study with German adolescents. According to most respondents, the instrument questions had no relevance to their realities; moreover, they had limited or no experience in managing some of the health-related tasks in the HLS-EU-Q47<sup>27</sup>. Hence, the items proved to be culturally inappropriate or invalid for that population<sup>27</sup>. Unfortunately, this type of content validity assessment was

not conducted with the European population that answered the previous HLS-EU-Q47 or the recent HLS19 survey, characterizing an etic approach in developing HL tools.

Conversely, the emic approach intends to understand a phenomenon based on population references. Faced with such a perspective, researchers who assume an emic approach believe that factors such as language, cultural origin, social values, morals and lifestyles influence in developing an instrument to measure a health phenomenon that is triggered in the cultural context itself. Thus, different ways of assessment were developed by different cultural groups to understand and comprehend the cultural, social, political, economic, and natural environments in which they live<sup>28</sup>. Moreover, it is counterproductive to be ideological about the presence or absence of cultural differences and argue in favour of exclusively emic or etic methods<sup>29</sup>. Looking at the instrument based solely on the consensus of researchers imposes strong limitations. More worrisome, if we only examine 'etic' and ignore 'emic' approaches, we end up creating transnational images of societies that are more homogeneous than they actually are and generating a superficial portrait of each individual society<sup>30</sup>.

In critically reading the HLS19-Q12 items we identify several problems, both in relation to possible interpretation bias according to different populations and their relevance to measure HL.

Table 1 presents some critical considerations about HLS19-Q12 questions considering the Brazilian and Latin American context, as many groups in North America and non-European settings.

[insert Table 1]

In the English (original) version, the instrument presents an introductory question: "It is not always easy to get understandable, reliable, and useful information on health-related topics. With the following questions we would like to find out which tasks related to handling health information are more or less easy or difficult. On a scale from Very easy to Very difficult, how easy would you say it is..."

As observed in Table 1, some items appear to be irrelevant as most people already have knowledge on the subject; others were worded in a confusing and incomplete way for good comprehension. Others violate the premise that items should only evaluate one factor, or focused on a biased biomedical view of the disease, or even the word/content may not represent the same construct for people worldwide.

Considering the aforementioned criticisms, the uncritical use of the instrument in different contexts becomes problematic. Thus, the issues identified provide clear arguments that render the content validity of items highly problematic.

Considerations about the psychometric properties of the HLS19-Q12 instrument content

Not only the HLS19-Q12 instrument content is subject to criticism, but also its psychometric properties. According to authors who developed the HLS19-Q12, "The construct validity of the general and specific HL measures was established using methods from Classical Test Theory (e.g., Cronbach's Alpha for internal consistency and Confirmatory Factor Analysis for the factor structure) and Item Response Theory (e.g., the Rasch model)"<sup>32</sup>.

Firstly, research shows that Cronbach's Alpha coefficient can be an imprecise way to measure reliability, as it is highly influenced by the number of items in the instrument which can artificially increase its precision<sup>33,34,35</sup>. Cronbach's Alpha was used for developing previous HLS-EU-Q instruments (2009-12) and those for HLS-19 survey (2019-2021).

According to Pelikan and Straßmayr, "Inspired by the Asian and Norwegian short forms and to overcome the shortcomings of the HLS-EU-Q16 and HLS-EU-Q6, a new short form, the HLS19-Q12 with 12 items, one for each of the cells in the conceptual HLS-EU matrix, was constructed using the original HLS-EU data for eight countries and two additional countries based on Rasch analyses[...]"<sup>15</sup>. Importantly, Rasch analysis is limited to 1PL models, and applying just one technique also presents limitations in the concepts of validity evidence adopted currently<sup>36,37</sup>. Consequently, there is a leap within the evidence since before applying internal structure evidence techniques, it is essential to know whether the items are suitable for assessing health literacy in different populations. An adjustable internal structure does not eliminate possible biases established in the content phase, including the need for changes for adaptation to culture and target populations.

As Iliescu points out, test adaptation is a complex scientific process and therefore is guided by the principles of the scientific method, most prominent of all the need to provide evidence for the appropriateness of this linguistic transformation, in terms of both language and other psychometric characteristics<sup>38</sup>. It is not just the translation that will allow us to assess the suitability of the instrument for a given culture. Similarly, van de Vijver<sup>29</sup> points out that translating a test into another language requires different types of knowledge, such as linguistic, cultural and psychometric skills. To Cheung, cultural variations affect the linguistic, conceptual and measurement equivalence of psychological assessment when these measures are translated and applied to different cultural groups, posing challenges to their reliability and validity<sup>26</sup>.

The team of researchers also seem to poorly address content validity directly. HLS19-Q47 had 47 items; with a very much reduced number of items, how (and whether) was content validity actually preserved? On M-POHL's International Report on the Methodology of European Health Literacy Population Survey 2019-2021 (HLS19), authors stated that Pearson's correlation of the HLS19-Q12 (short form) and the HLS19-Q47 presented values ranging from 0.898 to 0.949 and therefore "the HLS19-Q12 behaves in a sufficiently equivalent manner to the HLS19-Q47 or the HLS19-Q16 so that it can be used as a short form substitute, and the results of publications using the different measures can be compared to a certain degree"<sup>32</sup>. However, high Pearson's correlation values do not guarantee equivalence of either instrument content or metrics. This hypothesis is confirmed when the authors state that "it should be noted though that, because the HLS19-Q47 is not proven to be a unidimensional scale, it is not guaranteed that the HLS19-Q47 and the HLS19-Q12 strictly measure the same latent construct"<sup>32</sup>. In this case, content and metrics equivalences are not possible because the short versions of the instrument behave differently from the long version, especially regarding theoretical dimensionality as the short versions do not follow the dimensionality of the original instrument. Thus, measurement equivalence can only exist if the two versions have proven similar interpretations regarding the established theory.

Additionally, many processes of the HLS-EU instruments have used only confirmatory factor analysis (CFA) to assess internal structure of the instrument and its dimensions. This can be problematic as it is "assumed" that there is already some conformity in the instrument's internal structure. Applying the CFA directly makes it impossible to find other instrument settings that are interpretable, reinforcing the

etic bias already discussed previously. It also limits the search for evidence of internal structure validity. More robust analyses should be performed to test the psychometric properties of the instruments, as seen in the studies using HLS-EU-Q6 and HLS-EU-Q16 in Brazilian adults<sup>13,14</sup>.

We also highlight aspects of scalar invariance of the HLS-EU-Q instruments. According to Osborne et al.<sup>4</sup>, "configural invariance for factor structures, metric invariance for factor loadings and scalar invariance for item intercepts" are important and must be achieved if national and minority group data are to be validly compared. Additionally, they state that "generation of such evidence would require every country to first demonstrate within-country utility and acceptability of the measure, and then to demonstrate, qualitatively and quantitatively, at the item and construct levels, that the questions are understood (and responded to) in the same way, despite different contexts, such as health systems, entitlements to services and cultural practices." In the HLS19 survey, authors used IRT models to evaluate item invariance across levels of different sociodemographic factors such as gender, schooling level, and health status<sup>40</sup>. To this end, they conducted DIF analyses using gender and the dichotomous criteria age (median split) and schooling level (< higher education entrance qualification vs. at least higher education entrance qualification). DIF results show that the items were affected by the variables analyzed<sup>40</sup>. According to authors, "the possible reasons are manifold and include, for instance, somewhat different meanings of the items due to translation, social and cultural context, and differences in the health systems"<sup>40</sup>. Clearly, the authors failed to establish scalar invariance, with considerable evidence suggesting that HLS tools are prone to providing users with inaccurate information.

Another criticism concerns the lack of evidence that the score categorization suggested by the authors to classify individual HL levels is valid for different populations. Wrongly classifying individuals' HL levels can have important implications for public health planning and expenditure. A study investigating normative data of the HLS-EU-Q6 instrument in Brazilian adults showed that the classification criteria to be used with that population should differ from those proposed by the original study to avoid biases in the measurement of health literacy levels<sup>39</sup>. Additionally, in the HLS19 survey, researchers observed a high ceiling effect and relative ease of the HLS19-Q12 suggested tasks regarding score distribution.<sup>15</sup> Thus, the applicability and sensitivity of these questions to identify individuals with different HL levels is doubtful.

#### Final Considerations

Developing a global instrument requires a complete understanding of the nuances concerning the latent variable in the most diverse countries, cultures, languages and sub-groups. However, the development process of the general HL instruments for the HLS-19 survey presents no qualitative exploratory studies in different contexts and countries that even subsidize the inclusion of items specific to the culture in which the instrument was applied. Several studies have shown that disregarding the balance between the etic and the emic approach has led to less accurate instruments.

Such imbalance will inexorably generate bias in score configuration/standardization and on the ability to validate interpretation of the instrument scores. Moreover, one should be aware of the probable infeasibility of developing a cross-cultural instrument. An effective instrument would require, before its construction, ensuring that the concept of HL is at least equivalent across cultures to guarantee possible equivalence of measurement in the most diverse countries. But this aspect seems absent in the HL definition, conceptual model and general instruments developed by WHO Europe researchers which limits the use of their instruments, including the HLS19-Q12 short version to facilitate comparison with existing data from the HLS-19 survey.

Thus, HL levels should only be compared between nations once the construct is solid in all countries and the measurement properties allow unbiased comparison. Countries (and the many cultures and marginalised groups within them) must first ensure that measurements generate ethical and reproducible data, then research teams can perhaps confirm if the assessment tool is invariant across countries for some of the subconstructs.

The suggestion to make such comparisons among countries and regions using the HLS19-Q12 violates the emic principle, treating the latent variable as universal. This fact can occur, but should include regional nuances which will oblige the modification, inclusion or removal of items if researchers desire a precise instrument.

Despite the invaluable contributions of the European Consortium in developing the field of HL, including its conceptual model, definitions and instruments for measuring HL, the HLS19-Q12 instrument should not be applied uncritically in other countries outside the HLS19 survey solely based on WHO's justification to "ensure comparability among regions and countries"<sup>16</sup>.

As the conceptual model has not been properly validated by qualitative research with populations around the world, its validity to explain health literacy is clearly a risk, potentially generating biases in its conceptualization and constructs at individual, population, and organizational levels, as well as contributing to epistemic injustice. To increase concept transferability across different contexts, researchers should provide comprehensive descriptions of the research context, use purposeful sampling to incorporate diverse perspectives, engage in reflexivity to recognize and address biases, leverage data triangulation to increase credibility, and contextualize their findings for broader applicability in diverse contexts, including family medicine and community health settings worldwide.

The present paper focused on HLS-EU-Q instruments, as their content were clearly proposed by the WHO-EU group for use as a global instrument. However, similar or equal criticisms can be applied to all instruments with universal claims.

In conclusion, the value of a one-size-fits all tool for measuring health literacy at a global level is highly questionable given its potential negative impact in increasing risk of marginalisation and stigma, especially in undeveloped countries and vulnerable populations, resulting in false reports and misuse of public resources.

#### REFERENCES

1. WHO - World Health Organization. Health promotion glossary of terms 2021 [Internet]. 2021. [cited 2024 jan 14]. Available from: <https://www.who.int/publications/i/item/9789240038349>.

2. Lorini C, Caini S, Ierardi F, Bachini L, Gemmi F, Bonaccorsi G. Health Literacy as a Shared Capacity: Does the Health Literacy of a Country Influence the Health Disparities among Immigrants? *Int J Environ Res Public Health* 2020;17(4):1149.
3. Sørensen K. Health literacy is a political choice: a health literacy guide for politicians [Internet]. 2017. [cited 2024 apr 20]. Available from: <https://www.globalhealthliteracyacademy.org/post/2016/12/06/health-literacy-is-a-political-choice-a-health-literacy-guide-for-politicians>.
4. Osborne RH, Cheng CC, Nolte S, Elmer S, Besancon S, Debussche X, Dias S, Kolarcik P, Loureiro MI, Maindal H, O DN, Smith JA, Wahl A, Elsworth GR. Health literacy measurement: embracing diversity in a strengths-based approach to promote health and equity, and avoid epistemic injustice. *BMJ Global Health* 2022;7:e009623.
5. Tavousi M, Mohammadi S, Sadighi J, Zarei F, Kermani RM, Rostami R, Montazeri A. Measuring health literacy: a systematic review and bibliometric analysis of instruments from 1993 to 2021. *PLoS One* 2022;17(7):e0271524.
6. 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; 12:80.
7. Sørensen K, Van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, Kondilis B, Stoffels V, Osborne RH, Brand H. HLS-EU Consortium. 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;13:948.
8. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013;13:658.
9. Batterham RW, Buchbinder R, Beauchamp A, Dodson S, Elsworth GR, Osborne RH. The OPTimising HEalth LIterAcY (Ophelia) process: study protocol for using health literacy profiling and community engagement to create and implement health reform. *BMC Public Health* 2014;14:694.
10. Pelikan JM, Ganahl K. Measuring health literacy in general populations: Primary findings from the HLS-EU Consortium's health literacy assessment effort. *Stud Health Technol Inform* 2017; 240: 34-59.
11. Kickbusch I, Pelikan JM, Apfel A, Tsouros AD. Health literacy: the solid facts [Internet]. 2013. [cited 2024 apr 14]. Available from: <https://apps.who.int/iris/handle/10665/326432>.
12. Pelikan JM, Roethlin F, Ganahl K. Measuring comprehensive health literacy in general populations - the HLS-EU instruments [Internet]. 2014 [cited 2024 apr 20]. Available from: <https://www.bumc.bu.edu/healthliteracyconference/files/2014/06/Pelikan-et-al-HARC-2014-fin.pdf>
13. Mialhe FL, Moraes KL, Bado FMR, Brasil VV, Sampaio HAC, Rebutini F. Psychometric properties of the adapted instrument European Health Literacy Survey Questionnaire short-short form. *Rev Lat Am Enfermagem* 2021;29:e3436.
14. Mialhe FL, Sampaio HAC, Moraes KL, Brasil VV, Rebutini F. Psychometric properties of the Brazilian version of the European Health Literacy Survey Questionnaire short form. *Health Promot Int* 2022;37(4):daac130.
15. Pelikan JM, Straßmayr C. Background of the HLS19. In: The HLS19 Consortium of the WHO Action Network M-POHL: International Report on the Methodology, Results, and Recommendations of the European Health Literacy Population Survey 2019–2021 (HLS19) of M-POHL [Internet]. 2021. [cited 20 apr 2024]. Available from: <https://m-pohl.net/NB1>
16. WHO - World Health Organization. WHO Global Survey of Population Health Literacy in countries across WHO regions. Instructions for commenting on the health literacy questionnaire. DRAFT, 2022.
17. Watson JC. Talking the talk: enhancing clinical ethics with health literacy best practices. *HEC Forum* 2019;31(3):177-199.
18. Straßmayr C, Pelikan JM, Link T. The HLS19 instruments. In: The HLS19 Consortium of the WHO Action Network M-POHL: International Report on the Methodology, Results, and Recommendations of the European Health Literacy Population Survey 2019–2021 (HLS19) of M-POHL [Internet]. 2021. [cited 20 apr 2024]. Available from: <https://m-pohl.net/NB1>
19. Larsen MH, Mengshoel AM, Andersen MH, Borge CR, Ahlsen B, Dahl KG, Eik H, Holmen H, Lerdal A, Mariussen KL, Thoresen L, Tschamper MK, Urstad KH, Vidnes TK, Wahl AK. "A bit of everything": Health literacy interventions in chronic conditions - a systematic review. *Patient Educ Couns* 2022;105(10):2999-3016
20. Bernal G, Jiménez-Chafey MI, Domenech Rodríguez MM. Cultural adaptation of treatments: a resource for considering culture in evidence-based practice. *Prof Psychol Res Pr* 2009; 40(4): 361–368.
21. Helms JE. An examination of the evidence in culturally adapted evidence-based or empirically supported interventions. *Transcult Psychiatry* 2015;52(2):174-97.
22. Salamanca-Sanabria A, Richards D, Timulak L, Castro-Camacho L, Mojica-Perilla M, Parra-Villa Y. Assessing the efficacy of a culturally adapted cognitive behavioural internet-delivered treatment for depression: protocol for a randomised controlled trial. *BMC Psychiatry* 2018; 18(1):53.
23. Menezes JCL, Nobre-Sandoval LA, Murta SG. Local Adaptations to Implement the Strengthening Families Program in Northeastern Brazil. *Int J Ment Health Addict* 2020; 18(2): 407-421.
24. Creswel J, Plano Clark VL (eds.) *Designing and conducting mixed methods research*, 3rd ed. Thousand Oaks: Sage Publications; 2018.
25. Galperin BL, Punnett BJ, Ford D, Lituchy TR. An emic-etic-emic research cycle for understanding context in under-researched countries. *Int J Cross Cultural Managem* 2022; 22(1): 7-35.
26. Cheung F. Culture and Personality Assessment. *Oxford Research Encyclopedia of Psychology* [Internet]. 2023. [cited 12 mar 2024]. Available from: <https://oxfordre.com/psychology/view/10.1093/acrefore/9780190236557.001.0001/acrefore-9780190236557-e-455>.
27. Domanska OM, Firnges C, Bollweg TM, Sørensen K, Holmberg C, Jordan S. Do adolescents understand the items of the European Health Literacy Survey Questionnaire (HLS-EU-Q47) - German version? Findings from cognitive interviews of the project "Measurement of Health Literacy Among Adolescents" (MOHLAA) in Germany. *Arch Public Health* 2018; 76:46.
28. Rosa M, Orey DC. The field of research in ethnomodeling: emic, ethical and dialectical approaches. *Educ Pesq* 2012; 38(4): 865-879.
29. van de Vijver PJR. Test adaptations. In: Leong FTL, Bartram D, Cheung KF, Iliescu D, organizers. *The ITC international handbook of testing and assessment*. New York: Oxford University Press, 2016. p. 364–376.

30. Smith TW. Improving multinational, multiregional, and multicultural (3MC) comparability using the total survey error (TSE) paradigm. In: Jonhson TP, Pennell BE, Stoop IAL, Dorer B, organizers. *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC)*. New York: John Wiley & Sons, 2018. p.13-43.
31. Papen U. Literacy, learning and health – a social practices view of health literacy. *Literacy & Numeracy Studies*, 2009; 16(2): 19-34
32. The HLS19 Consortium of the WHO Action Network M-POHL. International Report on the Methodology, Results and recommendations of the European Health Literacy Population Survey 2019-2021 (HLS19) of M-POHL. Austrian National Public Health Institute; Vienna, Austria: 2021. [Internet]. 2021. [cited 20 apr 2024]. Available from: <https://m-pohl.net/NB1>
33. Schmitt N. Uses and abuses of coefficient Alpha. *Psychol Assess* 1996; 8: 350.
34. Vaske JJ, Beaman J, Sponarski CC. Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences* 2016; 39: 163-173.
35. McDonald RP. *Test theory: a unified treatment*. Mahwah, NJ: Lawrence Erlbaum, 1999.
36. American Educational Research Association. American Psychological Association & National Council on measurement in education. *Standards for educational and psychological testing* [internet]. 2014. [cited 2024 dec 18]. Available from: <https://www.apa.org/science/programs/testing/standards>
37. American Educational Research Association. American Psychological Association & National Council on measurement in education. *Standards for educational and psychological testing* [internet]. 2019. [cited 2024 dec 18]. Available from: <https://www.aera.net/Publications/Books/Standards-1999-Ed>
38. Iliescu D. *Adapting tests in linguistic and cultural situations*. Cambridge University Press, 2017.
39. Mialhe FL, Rebustini F. Evaluation of normative data of the instrument European Health Literacy Survey Questionnaire short-short form (HLS-EU-Q6) in Brazilian adults. *Rev Eletr Acervo Saude* 2023; 23(6): e12305 Available from: <https://acervomais.com.br/index.php/saude/article/view/12305/7706>
40. Pelikan JM, Link T, Straßmayr C, Waldherr K, Alferts T, Bøggild H, Griebler R, Lopatina M, Mikšová D, Nielsen MG, Peer S, Vrdelja M; HLS19 Consortium of the WHO Action Network M-POHL. Measuring Comprehensive, General Health Literacy in the General Adult Population: The Development and Validation of the HLS19-Q12 Instrument in Seventeen Countries. *Int J Environ Res Public Health* 2022;19(21):14129.

## Outros idiomas:



Como

### Citar

Mialhe, FL, Moraes, KL, da Silva Carvalho Vila, Vanessa, Sampaio, HAC, Brasil, VV, Rebustini, F. A global measurement tool to assess population health literacy fits all? A critical paper. *Cien Saude Colet* [periódico na internet] (2025/jun). [Citado em 01/12/2025]. **Está disponível em:** <http://cienciaesaudecoletiva.com.br/artigos/a-global-measurement-tool-to-assess-population-health-literacy-fits-all-a-critical-paper/19670>

Últimos

## Artigos

[Análise das internações e óbitos por Acidente Vascular Cerebral em diferentes regiões do Brasil antes e durante a Pandemia de COVID-19.](#)

0399/2025

[QUANDO AS SOLUÇÕES DE TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO ENTRAM EM CENA: UMA PROPOSIÇÃO TEÓRICO METODOLÓGICO PARA ANÁLISE DE POLÍTICAS](#)

0398/2025

[Efeito positivo da segurança alimentar na diminuição do estado de ansiedade de mulheres durante a gestação e pós-parto.](#)

0397/2025

[ASSOCIAÇÃO ENTRE SEGREGAÇÃO SOCIOECONÔMICA E QUALIDADE DA DIETA EM PRÉ-ESCOLARES](#)

0396/2025

[Avaliação das capacidades institucionais para resiliência do SUS a partir dos Blocos de Construção definidos pela Organização Mundial de Saúde](#)

0395/2025

Realização



Patrocínio



FIOCRUZ



MINISTÉRIO DA  
EDUCAÇÃO

MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÃO

MINISTÉRIO DA  
SAÚDE



Organização  
Pan-Americana  
da Saúde



**Revista Ciência & Saúde Coletiva da Associação Brasileira de Saúde Coletiva**

Impressa ISSN 1413-8123 | Online ISSN 1678-4561

*Avenida Brasil, 4036 / sala 703 – Manguinhos – CEP: 21040-361, Rio de Janeiro/RJ*

*(21) 3882-9153 - Todos os direitos reservados para ABRASCO.*

