

Expert Evaluation of an Online Digital Health Literacy Tool

Corine OLDHOFF-NUIJSINK^{a,b}, Isabel ENGELEN^{a,b}, Thomas ENGELSMA^{a,b},
Marloes DERKSEN^{a,b}

^a *Amsterdam UMC location University of Amsterdam, Department of Medical Informatics,
eHealth Living and Learning Lab, Amsterdam, the Netherlands*

^b *Amsterdam Public Health, Digital Health, Amsterdam, The Netherlands*

ORCID ID: Corine Oldhoff-Nuijsink <https://orcid.org/0009-0002-5834-399X>

Abstract. This study aims to gain first insights into developing and using an interactive assessment tool with an automatic scoring method to measure digital health literacy with performance-based tasks. We developed a prototype assessment tool based on an informal rapid literature review. The prototype was then evaluated using a heuristic evaluation (n=4) and qualitative interviews (n=5) with domain experts. The expert evaluation of the developed prototype identified 27 usability problems and resulted in recommendations for future development and research. The use of a digital interactive assessment instrument with performance-based tasks has potential and could support digital health literacy assessment. However, further research is needed into the definition of digital health literacy and end-user evaluation of its usability.

Keywords. eHealth literacy, computer-based assessment, interactive assessment tool

1. Introduction

The digitalization of healthcare calls for a new skill: digital health literacy (DHL) [3]. The World Health Organization defines DHL as “the ability to seek, find, understand and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem” [4]. While digital health is often seen as a positive development—enhancing accessibility, efficiency, and patient empowerment [1, 2]—not everyone is equally equipped to benefit. People with lower socio-economic status, ethnic minorities, and older adults often lack the DHL skills needed to navigate digital health environments effectively [5–8]. This gap can lead to health inequities, as some individuals struggle to access essential health information or services [5, 9].

To address these disparities, it is important to assess DHL and tailor interventions accordingly [5, 9, 10]. Many existing tools rely on self-reported data, which are easy and cost-effective to collect [5, 9–11]. However, such methods are often unreliable, as individuals tend to over- or underestimate their abilities [1, 5, 9, 10, 12], resulting in weak correlations between perceived and actual skills [1, 10, 11].

Objective, performance-based assessments offer a more accurate alternative by requiring individuals to demonstrate their DHL through tasks or responses [9–13]. These approaches often involve simulated internet tasks or website evaluations, providing

insight into one's ability to search for and assess online health information [10, 14–16]. However, they are typically resource-intensive, requiring trained observers, time, and funding [10–12], making them impractical for routine use in healthcare or research settings.

To address this, we developed a prototype of an interactive, performance-based DHL assessment tool that is easy to use and does not require an evaluator. In this study, we examine the usability of this prototype.

2. Methods

To assess the usability of the prototyped DHL assessment tool, we conducted a heuristic evaluation with UX experts and semi-structured interviews with domain experts. The Medical Ethics Review Committee of Amsterdam UMC (location AMC) confirmed that the Medical Research Involving Human Subjects Act (“WMO”) did not apply; therefore, no further approval was required (#2024.0351). All participants gave informed consent after being briefed on the study.

2.1. Heuristic evaluation with UX experts

In May 2024, a heuristic evaluation was conducted using Nielsen's ten usability principles [17]. Three UX experts, recruited via the eHealth Living & Learning Lab at Amsterdam UMC, participated—meeting the recommended number for such evaluations [18]. Experts received a digital briefing and were instructed to explore the prototype by completing three predefined tasks. They identified usability flaws, noted their locations, and rated severity using Nielsen's scale [19]. Given the study's focus, only issues related to the performance-based tasks or the overall website were included (n=12). Evaluations were conducted independently at the participants' convenience. Findings were compiled into a master list, duplicates removed, and a final list of unique usability issues created, categorized by severity and frequency.

2.2. Interviews with domain experts

To explore the prototype's applicability, semi-structured interviews were held with experts in (digital) health literacy, inclusive eHealth design, or illiteracy. Participants were selected through convenience sampling using personal networks. Interviews were conducted individually in Dutch, either online or in person. An iterative topic list guided the discussions, focusing on the prototype, performance-based tasks, and potential applications. Topics were tailored to each expert's background. Interviews were recorded, transcribed verbatim, and analyzed in ATLAS.ti (version 24). Thematic analysis began with deductive coding using a predefined tree, followed by inductive, axial, and selective coding to identify themes and relationships.

3. Results

3.1. Heuristic evaluation with UX experts

Four UX experts participated in the heuristic evaluation, identifying 27 unique usability issues across four areas: home screen ($n=4$), self-report questions ($n=11$), performance-based tasks ($n=5$), and the overall website ($n=7$). All ten Nielsen heuristics were violated at least once [18]. The most frequently violated heuristics were ‘user control and freedom’ and ‘error prevention’ (3 out of 4 experts), both with the highest average severity scores. ‘Match between system and the real world’ and ‘Help’ were also violated three times but with lower severity. No issues were rated as catastrophic (severity 4).

Usability issues with the performance-based tasks mainly involved unsaved answers, complex language (not B1 level), and lack of clear instructions. The overall mean severity rating was 2.19 ($SD = 0.83$), suggesting that most problems were of minor severity.

3.2. Interviews with domain experts

Five domain experts were interviewed, including two health literacy researchers, one expert on illiteracy, one in inclusive eHealth design, and one DHL expert. All were enthusiastic about the prototype and expressed interest in using it in research, provided further development.

Experts agreed that the language used in tasks should be simplified and that the performance-based tasks offered good coverage of DHL, although views on which specific skills should be included varied. Several suggested adding self-report questions to provide context to the performance-based results.

Recommendations for future use included limiting test duration to 10 minutes, making it available in multiple languages, using the tool for research (rather than clinical) purposes, conducting it with an observer present, and clearly communicating that test results are non-judgmental.

4. Discussion

This study explored the usability of a digital prototype for objectively assessing digital health literacy (DHL) through performance-based tasks. A heuristic evaluation identified 27 usability issues, while expert interviews provided recommendations for future development and implementation.

4.1. Interpreting results

Experts saw potential in the tool to improve DHL assessment, which can support health equity [5]. However, they disagreed on which skills should be assessed—highlighting the ongoing challenge of defining DHL consistently [1, 9, 16]. A unified definition is crucial to guide the development and standardization of assessment tools [1].

Although most usability issues identified were minor [19], they may still pose barriers—especially for people with low DHL—potentially exacerbating health

inequities [5, 20]. Therefore, tools should be designed with accessibility for digitally vulnerable groups in mind.

One advantage of digital performance-based tools is the potential for home use with automatic scoring, reducing the need for trained observers [11]. However, experts recommended having an observer present to support participants and gather additional insights into their digital skills. While this limits the autonomy of remote testing, it can enhance understanding of a participant's overall DHL.

4.2. Strengths and limitations

A key strength of this study was the inclusion of experts from diverse backgrounds, leading to broad and valuable feedback. However, the small sample size limits the generalizability of the findings. Additionally, the prototype was not tested with end-users, who could identify different usability issues than experts.

4.3. Practical and research implications

The findings offer direction for improving DHL assessment tools. Future research should define DHL more clearly and develop validated, performance-based instruments accordingly. Including both experts and end-users in the development process will be essential to ensure the tool is both valid and accessible.

5. Conclusions

This study identified key considerations for developing a digital, performance-based assessment tool for digital health literacy (DHL). The findings offer a foundation for future development and research. While the prototype shows potential to improve DHL assessment and support healthcare equity, further steps are needed. These include establishing a clear, unified definition of DHL and conducting usability testing with end-users to ensure the tool is accessible, relevant, and effective for diverse populations.

References

- [1] Wang C, Wu X, Qi H. A comprehensive analysis of E-health literacy research focuses and trends. *Healthcare (Basel)*. 2021 Dec 30;10(1):66.
- [2] Majcherek D, Hegerty S. W, Kowalski A. M, Lewandowska M. S, Dikova D. Opportunities for healthcare digitalization in Europe: Comparative analysis of inequalities in access to medical services. *Health Policy*. 2024 Jan;139(104950):104950.
- [3] van der Vaart R, Drossaert C. Development of the digital health literacy instrument: Measuring a broad spectrum of health 1.0 and health 2.0 skills. *J Med Internet Res*. 2017;19(1):e27.
- [4] Seidel E, Cortes T, Chong C. Digital Health Literacy [Internet]. PSNet. 2023 Oct 31 [cited 2024 Apr 25]. Available from: <https://psnet.ahrq.gov/primer/digital-health-literacy>

- [5] Faux-Nightingale A, Philp F, Chadwick D, Singh B, Pandyan A. Available tools to evaluate digital health literacy and engagement with eHealth resources: A scoping review. *Heliyon*. 2022 Aug 23;8(8):e10380.
- [6] Neter, E., & Brainin, E. eHealth literacy: extending the digital divide to the realm of health information. *J Med Internet Res*. 2012 Jan 27;14(1):e19.
- [7] Kontos, E., Blake, K. D., Chou, W. Y. S., Prestin, A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. *J Med Internet Res*. 2014 Jul 16;16(7):e172.
- [8] Guo, Z., Zhao, S. Z., Guo, N., Wu, Y., Weng, X., Wong, J. Y. H., et al. Socioeconomic disparities in eHealth literacy and preventive behaviors during the COVID-19 pandemic in Hong Kong: cross-sectional study. *J Med Internet Res*. 2021 Apr 14;23(4):e24577
- [9] Arias López M. D. P, Ong B. A, Borrat Frigola X, Fernández A. L, Hicklent R. S, Obeles A. J. T, et al. Digital literacy as a new determinant of health: A scoping review. *PLOS Digit Health*. 2023 Oct 12;2(10):e0000279.
- [10] Crocker B, Feng O, Duncan LR. Performance-based measurement of eHealth literacy: Systematic scoping review. *J Med Internet Res*. 2023 Jun 06;25(1):e44602.
- [11] Neter E, Brainin E. Perceived and performed eHealth literacy: Survey and simulated performance test. *JMIR Hum Factors*. 2017 Jan 17;4(1):e2.
- [12] van der Vaart R, van Deursen AJ, Drossaert CHC, Taal E, van Dijk JA, van de Laar MA. Does the eHealth literacy scale (eHEALS) measure what it intends to measure? Validation of a dutch version of the eHEALS in two adult populations. *J Med Internet Res*. 2011 Nov 09;13(4):e86.
- [13] New York State Education Department. Performance-Based Learning and Assessment Networks (PLAN) Program. [Internet]. Available from: <https://www.nysed.gov/plan-pilot/program-description>. [Accessed 2024 May 20].
- [14] Faux-Nightingale A, Philp F, Chadwick D, Singh B, Pandyan A. Available tools to evaluate digital health literacy and engagement with eHealth resources: A scoping review. *Heliyon*. 2022;8:e10380.
- [15] Dijkman EM, ter Brake WWM, Drossaert CHC, Doggen CJM. Assessment Tools for Measuring Health Literacy and Digital Health Literacy in a Hospital Setting: A Scoping Review. *Healthcare*. 2023;12:11.
- [16] Lee J, Lee E-H, Chae D. eHealth Literacy Instruments: Systematic Review of Measurement Properties. *J Med Internet Res*. 2021;23:e30644.
- [17] Nielsen J. 10 Usability Heuristics for User Interface Design. [Internet]. Available from: <https://www.nngroup.com/articles/ten-usability-heuristics/>. [Accessed 2024 Jun 13].
- [18] Nielsen J. How Many Test Users in a Usability Study?. [Internet]. Available from: How Many Test Users in a Usability Study? - NN/<https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>. [Accessed 2025 May 12].
- [19] Nielsen J. Severity Ratings for Usability Problems. [Internet]. Available from: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>. [Accessed 2024 Jun 13].
- [20] Nielsen J. First rule of usability: Don't listen to users [Internet]. Fremont (CA): Nielsen Norman Group; 2001 Jan 28. Available from: <https://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/> [Accessed 2025 May 12].