Remote Photoplethysmography System for Vital Signs Estimation: Perceived Usability Evaluation by Elderly People

Anna Maria Carluccio, Andrea Manni, Andrea Caroppo, Pietro Aleardo Siciliano and Alessandro Leone
National Research Council of Italy - Institute for Microelectronics and Microsystems
Via Monteroni, c/o Campus Università del Salento, Palazzina A3, Lecce, Italy
email: annamaria.carluccio@imm.cnr.it, andrea.manni@cnr.it, andrea.caroppo@cnr.it, pietroaleardo.siciliano@cnr.it
alessandro.leone@cnr.it

***Abstract*— Remote photoplethysmography has emerged as a promising continuous, non-invasive vital signs monitoring technique. This technique provides real-time estimation of key vital signs, including blood oxygen saturation, breathing rate, and heart rate, by analyzing a video of the user's face. To assess perceived usability, elderly end-users and related caregivers completed the System Usability Scale and the short version of the User Experience Questionnaire, providing quantitative scores and qualitative feedback on usability, reliability, and satisfaction. The results demonstrate the robustness and user-friendliness of the system, particularly for caregivers, suggesting some refinement to make it more accessible to older users.**

Keywords- Remote Photoplethysmography (rPPG); Contactless Health Monitoring; Vital Sign Estimation; Usability Evaluation; System Usability Scale (SUS); User Experience Questionnaire (UEQ).

I. INTRODUCTION

Continuous monitoring of vital signs plays a key role in preventing various heart and respiratory system diseases. Heart Rate (HR), Breathing Rate (BR), and blood oxygen saturation (SpO_2) are critical indicators for assessing the state of human health. While wearable devices such as smartwatches have gained widespread popularity, contactless systems have seen increased adoption, particularly following the COVID-19 pandemic [1]. In recent years, researchers have investigated remote photoplethysmography (rPPG), a non-contact technique that analyses subtle variations in skin color caused by blood fluctuation in peripheral vessels [2]. These color variations are due to periodic changes in blood volume linked to the cardiac cycle, which can be extracted from Red-Green-Blue (RGB) video signals using signal processing and filtering techniques. The rPPG method typically involves face detection, Region-Of-Interest (ROI) selection, signal decomposition, and post-processing to estimate heart rate, breathing rate, and blood oxygen saturation with high accuracy [3]. This technology enables non-invasive vital signs monitoring by analyzing video data of the user's face acquired using consumer devices, such as RGB or smartphone-integrated cameras, which are low-cost and extremely diffused [4]. This innovation offers substantial promise for remote health monitoring, especially within telemedicine applications. Ensuring usability is essential to make health monitoring solutions practical, efficient, and accessible in real-world scenarios. Usability and acceptability are particularly critical for adopting systems, such as rPPG, especially among older adults, as they strongly influence engagement and sustained use [5]. This paper examines the experimental development phase of the proposed solution, focusing on usability outcomes, feedback on sustainability, and its potential integration into services and interventions designed for aging populations. The structure of the paper is as follows: Section II outlines the materials and methods used, detailing the device under usability investigation, as well as the protocols and questionnaires used for user experience data collection; Section III presents the findings related to usability and discusses the obtained results, while Section IV provides conclusion and future works.

II. MATERIALS AND METHODS

This section provides an overview of the rPPG system employed in the study followed by an introduction to the basic concepts of usability and a description of the questionnaires used to evaluate the user experience in this research.

A. Hardware and Software Description

Remote photoplethysmography (rPPG) allows the monitoring of vital signs using only a vision sensor and a processing unit. Most studies use consumer webcams or cameras connected to PCs to capture the video stream of the user's face. In our system, the NexiGo N960E webcam (Figure 1a) was selected for facial video acquisition since the built-in light ring ensures optimal signal quality even in low-light conditions (three adjustable brightness levels) and the Raspberry Pi 4 Model B was selected as the processing unit due to its efficiency and cost-performance ratio (Figure 1b), as evidenced in [6]. The input of the algorithmic pipeline for vital signs estimation is a video stream taken by the selected webcam. As detailed in [6], the pipeline consists of two main stages: (1) the pre-processing stage and (2) the feature extraction and vital signs estimation stage.

The system was tested in a controlled environment to assess its accuracy. The experiment involved measuring vital signs at various distances from the user to the camera, using data collected by certified devices for ground truth.

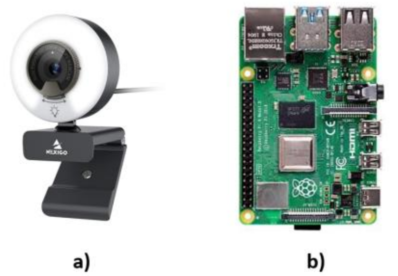

Figure 1. (a) NexiGo N960E Webcam and (b) Raspberry Pi 4 Model B.

Different kinds of metrics were proposed in this research area for evaluating vital signs measurement methods. Here, the commonly used Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics were utilized. At 0.5m, the system demonstrated accurate HR estimation with a MAE of 2.20 and an RMSE of 3.96. Similarly, the best results for BR were achieved at 0.5m, with a MAE of 1.80 and an RMSE of 2.15. For SpO_2 estimation, the average percentage difference from ground truth increased with distance, with the lowest error (0.85%) at 0.5m. Performance declined as distance increased, emphasizing optimal accuracy at closer proximity. The experiment was conducted in a controlled environment to ensure optimal lighting and positioning. However, real-world conditions may introduce factors such as variable lighting, background noise, and user movement, which could affect both the accuracy of rPPG measurements and overall usability.

The Graphical User Interface (GUI) was designed to be intuitive. Figure 2 illustrates the GUI that caregivers and elderly users interacted with. In the upper-left corner, a live feed from the webcam is displayed, assisting users in correctly positioning their faces for capturing. Once the acquisition is completed, the estimated vital signs are displayed in the upper right corner (green box). In addition, there is a section (black box) for manual entry of parameters from certified devices. Under this area (red box) any messages about data transmission or connection errors are displayed. Below the data transmission area, a countdown timer, set to 30 seconds, informs the user of the remaining acquisition time. Since the graphical user interface is entirely in Italian, Figure 3 shows for clarity a translated English version of the interface, created specifically for dissemination purposes and not presented to users.

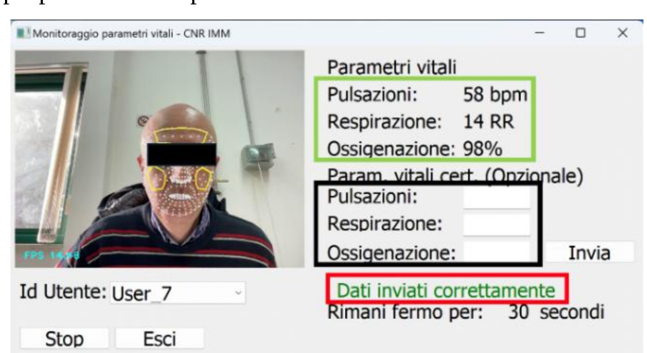

Figure 2. Graphical User Interface of the rPPG system, Italian version.

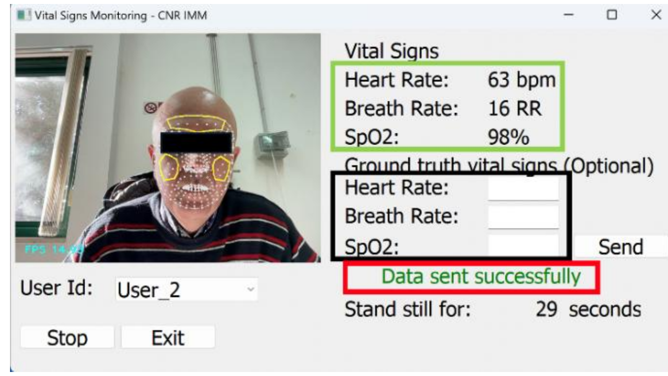

Figure 3. Graphical User Interface of the rPPG system, English version.

B. *Usability Rules and Protocols*

Usability measures how easy and intuitive a software product, website, application, or interactive system is for users. Two widely accepted definitions of usability come from Jakob Nielsen and ISO 9241-11. Nielsen describes usability as a quality attribute that evaluates ease of use and includes components such as learnability, efficiency, memorability, error reduction, and satisfaction. He proposed ten general heuristics to guide User Interface (UI) design, focusing on accessibility and intuitiveness [7]. Usability principles have been widely studied since the foundational works of Nielsen, and more recent studies have further explored their applications in digital health technologies [8].

The ISO 9241-11 standard defines usability as "the extent to which a system, product or service can be used by specific users to achieve specific goals with effectiveness, efficiency, and satisfaction in a defined context of use" [9]. Usability evaluation typically combines quantitative approaches, such as standardized questionnaires and metrics, with qualitative methods like interviews and observations to provide deeper insights into user behavior and preferences. For decades, practitioners and researchers in user-centered design and Human-Computer Interaction (HCI) have had a strong interest in the measurement of perceived usability [10].

A key tool for measuring perceived usability is the System Usability Scale (SUS), developed in the 1980s. The SUS is a 10-item questionnaire where participants rate each item on a 5-point Likert scale. The resulting score, ranging from 0 to 100, offers a quick and reliable assessment of usability and is especially useful for benchmarking systems [11]. Table 1 provides the full list of SUS items.

Another widely used tool is the User Experience Questionnaire (UEQ) which provides a more detailed assessment of usability dimensions. This questionnaire examines specific aspects such as reliability, intuitiveness, and satisfaction, offering a nuanced perspective on user perceptions. The standard UEQ includes 26 items, taking 3–5 minutes to complete, while its short version (UEQ-S) comprises only 8 items, making it suitable for constrained circumstances. The UEQ-S Questionnaire has been employed in this study to complement the SUS. It evaluates

 ISBN: 978-1-68558-255-5

two main dimensions: Pragmatic Quality (PQ), focused on usability and efficiency, and Hedonic Quality (HQ), related to attractiveness and emotional engagement [12]. The decision to limit the analysis to these dimensions aligns with the UEQ-S structure and ensures a focused assessment of the system's perceived usability and user experience. The UEQ-S uses a 7-point bipolar scale, ranging from -3 (extremely negative) to +3 (extremely positive), with 0 indicating neutrality. This balanced scale effectively captures both positive and negative feedback. Key pairs in the scale include “Confusing – Clear,” “Complicated – Easy” (PQ), and “Boring – Exciting,” “Uninteresting – Interesting” (HQ), which evaluate how users perceive the system's functionality and its emotional impact. Table 2 lists the items included in the UEQ-S.

TABLE I. SYSTEM USABILITY SCALE (SUS) ITEMS.

	SUS items
1.	I think that I would like to use this system frequently.
2.	I found the system unnecessarily complex.
3.	I thought the system was easy to use.
4.	I think that I would need the support of a technical person to be able to use this system.
5.	I found the various functions in the system were well integrated.
6.	I thought there was too much inconsistency in this system.
7.	I would imagine that most people would learn to use this system very quickly.
8.	I found the system very cumbersome to use.
9.	I felt very confident using the system.
10.	I needed to learn a lot of things before I could get going with this system.

TABLE II. SHORT VERSION OF THE USER EXPERIENCE QUESTIONNAIRE (UEQ-S) ITEMS.

	UEQ-S items	
Obstructive	-3 -2 -1 0 +1 +2 +3	**Supporting**
Complicated	-3 -2 -1 0 +1 +2 +3	**Easy**
Inefficient	-3 -2 -1 0 +1 +2 +3	**Efficient**
Confusing	-3 -2 -1 0 +1 +2 +3	**Clear**
Boring	-3 -2 -1 0 +1 +2 +3	**Exiting**
Not Interesting	-3 -2 -1 0 +1 +2 +3	**Interesting**
Conventional	-3 -2 -1 0 +1 +2 +3	**Inventive**
Usual	-3 -2 -1 0 +1 +2 +3	**Leading**

The short version of the UEQ (UEQ-S) was chosen over the more recent UEQ+ because it allows for a rapid yet reliable evaluation of user experience while minimizing cognitive load for elderly participants. Given the target population's limited familiarity with technology, a more extensive questionnaire could have impacted response quality and completion rates. The UEQ-S retains the core dimensions of usability and user engagement, making it well-suited for our study's goals.

III. RESULTS

The usability and acceptability of the rPPG system were tested in two elderly care facilities, involving 27 participants: 20 beneficiaries (age 65-85, with a mean age of 74.5 years) with varying levels of education and low to moderate familiarity with technology, and 7 staff members (age 30-55), mainly nurses and care assistants with greater technological proficiency. Training sessions were conducted to ensure the correct use of the device and accurate data collection procedures. During the experiment, the camera and computer were positioned in a controlled environment with optimal lighting and seating conditions. SUS and UEQ-S tests were completed by both the elderly beneficiaries and the nursing home staff. The results of the perceived usability evaluation are detailed below.

Table 3 provides the average scores for each item of the SUS questionnaire by users and staff during trials. The scores for each item are then transformed: for odd-numbered items, 1 is subtracted from the response, and for even-numbered items, the response is subtracted from 5. The transformed scores are then summed and multiplied by 2.5 to obtain a score ranging from 0 to 100.

TABLE III. SUS SCORES.

Item	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	Scores
Users	5	1	5	5	5	1	4	4	5	1	**80.0**
Staff	5	1	5	1	5	1	4	1	5	1	**97.5**

A SUS score of 68 is widely considered to be the benchmark for good usability, based on extensive research and studies of SUS interpretation. The SUS evaluation revealed a notable difference between the two user groups. Employees gave an average score of 97.5, reflecting their perception of the system as exceptionally easy to use and well-suited to their professional needs. This high score underscores the functionality, reliability, and user-friendly design of the rPPG system, which fits well with the workflow requirements of trained professionals.

In contrast, users gave a lower average score of 80. While this score still indicates good usability, it also indicates minor difficulties experienced by non-professional elderly users. These challenges are mainly related to specific items, such as the need for technical support and system comfort (items 4 and 8), maybe due to differences in technological familiarity and user expectations. It could also be due to the difficulty in terms of accessibility to interact with the user interface. This differentiation underlines the importance of user-centered design and emphasizes the need to develop health monitoring systems that meet the different needs of all user groups.

Table 4 shows the average scores given by both staff and beneficiaries for the items in the UEQ-S, while Table 5 presents the scores given by the two groups in terms of PQ and HQ. For each dimension (PQ and HQ), the transformed scores were summed and then divided by the number of items in the respective category. The overall quality score was

 ISBN: 978-1-68558-255-5

calculated as the average of the PQ and HQ scores. The results of the UEQ-S questionnaire show that users highly appreciate both the PQ and HQ of the system. Among users, the system achieved a PQ score of 2.250 and an HQ score of 3.000, resulting in an overall quality score of 2.625. These results suggest that the system is not only highly functional but also emotionally appealing. The higher hedonic score indicates that the interface design and user experience resonate strongly with users and evoke a positive emotional response. While the pragmatic quality score is strong, its slightly lower value suggests opportunities for further improvement in task-related usability. For employees, the system received a PQ score of 2.000 and an HQ score of 2.250, resulting in an overall quality score of 2.125. These slightly lower scores, compared to those of beneficiaries, suggest that while staff find the system competent and effective, they may experience minor functionality or emotional engagement challenges. These differences between users and staff are probably due to the different contexts in which each group interacts with the system. Users may approach the system with lower initial expectations and find the interactivity with the system particularly appealing, increasing their hedonic perception. Usually, staff familiar with professional tools may prioritize pragmatic aspects such as efficiency and precision, resulting in slightly lower hedonic ratings.

TABLE IV. UEQ-S SCORES.

	Users	Staff	
Obstructive	2	1	**Supporting**
Complicated	2	3	**Easy**
Inefficient	2	1	**Efficient**
Confusing	3	3	**Clear**
Boring	3	2	**Exiting**
Not Interesting	3	2	**Interesting**
Conventional	3	3	**Inventive**
Usual	3	2	**Leading**

TABLE V. PRAGMATIC, HEDONIC, AND OVERALL QUALITY SCORES.

	PQ	HQ	Overall
Users	2.250	3.000	2.625
Staff	2.000	2.250	2.125

This balance between functionality and pleasure is critical for health monitoring applications, as it promotes both short-term effectiveness and long-term adherence. However, the UEQ-S overall scores of 2.125 for the staff and 2.625 for the elderly users highlight the system's strong ability to effectively support user tasks while providing engaging and positive user experience.

The SUS and UEQ-S questionnaires highlight different but complementary aspects of the system's performance. The SUS focuses on usability, emphasizing functionality, efficiency, and ease of learning, making it ideal to assess the effectiveness of the system in completing tasks. This explains the higher SUS scores from staff, prioritizing seamless integration into professional workflows. In contrast, the UEQ-S evaluates both usability and overall user experience, capturing emotional engagement and aesthetic appeal through its Pragmatic and Hedonic Quality dimensions. Together, these tools provide a holistic view, combining functional reliability with user-centered design insights.

The findings indicate the reliability and effectiveness of the rPPG system, especially among professional personnel. However, to achieve a universally excellent user experience, it is essential to address the specific challenges faced by elderly users. Refinements such as larger fonts, high-contrast color schemes, and a more guided user experience could significantly enhance accessibility and satisfaction, encouraging broader adoption in diverse settings. Moreover, this study focuses on short-term usability assessment. While initial feedback is positive, long-term user engagement and system sustainability are crucial aspects of health monitoring applications, warranting further investigation. While the sample size (N=27) is appropriate for an initial usability study, future research should involve a larger and more diverse participant pool to improve generalizability. Additionally, participants had different levels of prior exposure to digital health technologies, which could influence their perceptions of usability. Usability evaluations are inherently subjective and may be influenced by participants' prior experience with technology. Staff members, being more technologically proficient, reported higher usability scores, while elderly users encountered minor difficulties. Future studies should account for this factor by stratifying participants based on their digital literacy levels. The system interface was initially developed in Italian to match the target population. An English version was created for dissemination purposes, but future research should explore how cultural and linguistic factors may affect usability in international contexts. The study was conducted in two elderly care facilities, where participants had access to structured assistance. These results may not be fully generalized for older adults living independently or in different cultural and socio-economic contexts. Future research should expand the evaluation to diverse settings.

IV. CONCLUSION

This study investigated the usability of an rPPG-based system for non-invasive vital sign monitoring, particularly for elderly users. A total of 27 participants (20 care facility residents and 7 staff members) evaluated the system through SUS and UEQ-S. Staff rated it highly (SUS: 97.5), reflecting professional suitability, while beneficiaries gave it a strong but lower score (SUS: 80.0), indicating room for improved accessibility. Similarly, the UEQ-S results highlighted a positive balance between functionality and emotional engagement, with overall scores of 2.125 for Pragmatic Quality and 2.625 for Hedonic Quality. While previous methods have focused primarily on accuracy, our approach emphasizes both technical performance and usability, making it a viable solution for real-world healthcare applications. These results underline the system's potential and highlight the necessity to improve it, ensuring wider acceptability and better user experience across different target groups. Future studies should explore long-term usability and effectiveness in diverse real-world scenarios, particularly focusing on iterative improvements based on user feedback from different demographic groups.

ACKNOWLEDGMENT

This paper was developed within the project funded by Next Generation EU-"Age-It-Ageing well in an ageing society" project (PE0000015), National Recovery and Resilience Plan (NRRP)-PE8-Mission 4, C2, Intervention 1.3. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

REFERENCES

[1] N. Mohammadzadeh, M. Gholamzadeh, S. Saeedi, and S. Rezayi, "The application of wearable smart sensors for monitoring the vital signs of patients in epidemics: a systematic literature review," J Ambient Intell Humaniz Comput. vol. 14, no. 5, pp. 6027-6041, 2023.

[2] P. V. Rouast, M.T.P. Adam, R. Chiong, D. Cornforth, and E. Lux, "Remote heart rate measurement using low-cost RGB face video: a technical literature review," Front. Comput. Sci. Front. Comput. Sci. vol. 12, pp. 858-872, 2018.

[3] W. Verkruysse, L.O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," Opt Express. vol. 16, no. 26, pp. 21434-21445, 2008.

[4] R. J. Lee, S. Sivakumar, and K. H. Lim, "Review on remote heart rate measurements using photoplethysmography," Multimedia Tools and Applications, vol. 83, pp. 44699–44728, 2024.

[5] E. Takano et al., "User Experience of Older People While Using Digital Health Technologies: A Systematic Review," Appl. Sci. vol. 13, pp. 12815, 2023.

[6] A. Caroppo, A. Manni, G. Rescio, P. A. Siciliano and A. Leone, "Vital signs estimation in elderly using camera-based photoplethysmography," Multimedia Tools and Applications, vol. 83, pp. 65363–65386, 2024.

[7] J. Nielsen, Usability Engineering. Morgan Kaufmann Publishers, 1993.

[8] I. Maramba, A. Chatterjee, and C. Newman, "Methods of usability testing in the development of eHealth applications: a scoping review," International journal of medical informatics, vol. 126, pp. 95-104, 2019.

[9] International Organization for Standardization, ISO 9241-11:2018 Ergonomics of Human-System Interaction – Part 11: Usability: Definitions and Concepts, 2018.

[10] J. Lewis, B. Utesch, and D. Maher, "Measuring perceived usability: The SUS, UMUX-LITE, and AltUsability," Int. J. Hum.-Comput. Interact., vol. 31, no. 8, pp. 496-505, 2015.

[11] J. Brooke, "SUS: A quick and dirty usability scale," Usability Evaluation in Industry, vol. 189, no. 194, pp. 4–7, 1996.

[12] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S)," Int. J. Interact. Multimedia Artif. Intell., vol. 4, pp. 103-108, 2017.