Usability evaluation of an integrative exergaming system for the senior population

Teresa Paulino
Madeira Interactive Technologies Institute
Faculdade de Ciências Exactas e da Engenharia
Universidade da Madeira
Funchal, Portugal
teresa.paulino@m-iti.org

Sergi Bermúdez i Badia
Madeira Interactive Technologies Institute
Faculdade de Ciências Exactas e da Engenharia
Universidade da Madeira
Funchal, Portugal
sergi.bermudez@m-iti.org

Mónica Cameirão
Madeira Interactive Technologies Institute
Faculdade de Ciências Exactas e da Engenharia
Universidade da Madeira
Funchal, Portugal
monica.cameirao@m-iti.org

***Abstract*—Cardiovascular diseases such as heart attack or stroke are one of the major causes of death worldwide. Physical activity plays a key role in the prevention and delay of such diseases. Configurable exergames have been suggested as a solution to promote physical exercise in the senior population. An integrative system was designed and implemented following human-centered design techniques involving the main prospective interactors with the system. The success of such a system can only be achieved if it is usable by the target audience. In this paper, we present the efforts towards evaluating the usability of an integrative system for exergames targeting the senior population. Results from the study revealed a high level of satisfaction and provided a list of improvements for future development iterations.**

Keywords—usability, user-centered design, software engineering, human-computer interaction, exergames

I. Introduction

The percentage of the senior population is rapidly increasing worldwide. Sedentarism in older adults has emerged as a new health risk which may lead to cardiovascular diseases such as heart attack or stroke [1]. Exergames have been proposed as a viable solution to promote physical activity in the senior population [2], [3]. The diversity of needs and limitations of older adults require that the exergames allow the configuration and adjustment of game parameters according to each profile. The versatility of such exergames enables them to be used in senior houses, nursing homes, senior gymnasiums or even rehabilitation centers. In this context, an integrative system was designed and developed to assist in the configuration and management of fitness training sessions and plans with exergames.

It is critical for the success of such complex systems that during its development life cycle it is ensured that they will be usable by the target audience. Usability is a multidimensional concept that has been studied and explored over time by different researchers, which added different aspects to its definition [4]–[6]. According to [6], usability is defined by five quality components: learnability, efficiency, memorability, errors, and satisfaction. More recently the international standard ISO 9241-11:2018 defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [7]. Effectiveness corresponds to what extent a task was completed, or an objective achieved. The number of errors or the error intensity can be used to define the level of effectiveness. Efficiency can be measured by considering the tasks time and productivity considering the time used to accomplish a task. Satisfaction considers measures such as user pleasure, physical comfort, and feature utilization [8].

Previous work focused on evaluating the usability of similar systems from the perspective of older adults. For instance, Billis and colleagues [9] evaluated the affective usability and satisfaction of an exergaming platform which was developed according to the seniors' needs and preferences. Their results revealed that their system stimulated positive emotions and overall was perceived as beneficial. A web version of this platform was evaluated by Konstantinidis and colleagues [10]. They applied the System Usability Scale (SUS) and a preference sorting technique to gather feedback from the target end-users: the senior population. Another study by Palestra and colleagues [11], performed an evaluation of an exergaming rehabilitation system based on customizable exergames. Answers from a qualitative questionnaire have shown very satisfactory results revealing high interest in the elderly. Similarly, Nawaz and colleagues [12] obtained a high score on the SUS when assessing the seniors' user experience of exergames for balance training. Like these authors, many others focused on assessing the usability of the exergaming systems from the perspective of older adults. However, the adoption of such systems in health and sports institutions highly depends on the acceptance by the healthcare professionals. Therefore, it is fundamental to run usability studies also focusing on the perspective of the professionals that employ them, since they are the main interactors with the management interface of the exergaming systems.

To evaluate our system, we invited the main prospective users of the software: sports professionals from senior gymnasiums.

II. Materials and Methods

A. Integrative system description

A set of configurable exergames was designed and developed in collaboration with the target population and sports professionals using user-centric techniques [13]. Each game targets mostly one or few fitness domains such as cardiorespiratory endurance, muscular strength, agility, balance or flexibility. The target population consists of healthy older adults aged over 60. A software system was envisioned to manage training sessions and plans and to intelligently automate the exergames selection and configuration process based on the end-user profile.

978-1-7281-3636-3/19/$31.00 ©2019 IEEE

The integrative system object of this study was designed and developed following user-centered design techniques involving sports and health professionals that are used to work with the senior population in their daily practice [14]. The purpose of this system is to universally integrate multiple independent exergames and user assessments. This means that it is not restricted to only the currently implemented games or assessments. It should also allow managing users and training sessions and plans offering individual personalization of the exergames parameters. The user fitness profile (Fig. 1) consists of the same physical aspects found on games and is built based on assessment results which are then compared with normative data. A database stores all the information required to keep track of the historical user progression. The interface of the system was highly simplified during the development cycle to keep it intuitive and easy to use.

Fig. 1. User profile displaying the different fitness domains under the profile picture. Names and personal data are fictional.

B. Setup

Setup of the usability tests consisted of a computer and a screen with touch capability. A touch screen was used because the software system was conceptualized to be installed on the PEPE mobile platform [15]. The user sat down in front of the screen performing the test of the software (Fig. 2). A video camera was installed to record audio and video during the sessions.

C. Participants

Ten sports professionals (7 female, 3 male, 32.7±8.8 years old, 8±6.9 years of experience) who did not have previous knowledge about the system were invited to participate in the usability evaluation sessions (Table I). All participants had previous experience in prescribing physical exercise for the senior population. Concerning the use of digital technologies in their daily practice, none used exergames or similar tools as training complement. The digital technologies they mostly used in their daily practice were mainly in the context of administrative functions such as organizing users' data and planning classes, for this, Microsoft Office tools were the preference among most participants.

D. Procedure

In order to run all usability evaluation sessions in a standardized way, a script was written and divided into three main stages:

1. Introduction to the system - in this stage, three videos of the exergames were shown to the participant to provide a general idea of the type of games and different physical domains trained. After the videos, a diagram representing an overview of the system was explained for the participant to understand the main objectives of the integrative system. This stage concluded by describing what would happen during the usability evaluation and asking the participant to sign informed consent.

TABLE I. DEMOGRAPHIC CHARACTERIZATION OF THE PARTICIPANTS.

Participant ID	Gender	Age	Years of Experience
1	Male	22	1
2	Male	30	8
3	Male	35	3
4	Female	41	19
5	Female	27	6
6	Female	21	1
7	Female	39	15
8	Female	27	2
9	Female	48	15
10	Female	37	10

2. Usability evaluation - in this stage, a fictional scenario was presented to the participant who was led through two sets of microtasks to accomplish two activities with the system. A microtask is considered each individual step of a given activity. For instance, the first activity consisted of creating an assessment session for a user with given details. To accomplish this, the participant was required to first login into the system in the authentication screen (Fig. 3a), then select an end-user from the users' list (Fig. 3b), create a session (Fig. 3c), select the general assessment routine from the session components gallery (Fig. 3d), edit session date, and finally save the session. In total, the first activity consisted of six microtasks as detailed in Table II. The two activities corresponded to the two of the most important functionalities of the integrative system, namely, create sessions and training plans. Hence, the first activity (A1) had as the primary goal to create an assessment session and was divided into six microtasks. The second activity (A2), was composed of three microtasks, and the primary goal was to create a training plan. Each microtask had a level of complexity which corresponds to the number of atomic steps, gestures or "clicks" that the participant had to perform to accomplish the microtask. For instance, microtask A1.1-Log had the complexity of 5 because consisted of at least five atomic steps: touch the username input field, type username, touch password input field, type password, and press login button. Table II shows a brief description of the microtasks and their level of complexity. The last one

Fig. 2. Participant performing the fictional scenario steps during the usability evaluation of the integrative system.

was the most complex of all, basically consisting of filling in a form with given details. During the session there was the constant presence of a moderator who was there to assist in case the participant could not perform an action. However, the moderator intervention was restricted to support the participant to find the solution by him/herself and never to explicitly tell the participant how to achieve some action with the system. The participant was invited to speak out loud while performing the tasks (Fig. 2).

3. Usability questionnaire – at the end of the evaluation session the participants were asked to fill in the Portuguese version of the Usefulness, Satisfaction, and Ease of use (USE) questionnaire [16]–[18]. The USE questionnaire aims to measure the subjective usability of a product or service, in this case, a software system. It is composed of 30 items and examines four dimensions of usability: usefulness, ease of use, ease of learning, and satisfaction. Participants rate each item on a 7-point Likert scale (1 = "Strongly Disagree"; 7 = "Strongly Agree") with an "N/A" option. In the end, participants may freely indicate the negative and positive aspects of the system. The scoring of the USE is found in literature as being conducted separately for each of the four dimensions by averaging the item scores on each dimension [18].

Each session lasted approximately from 25 to 30 minutes. All participants permitted to video record the sessions which enabled the posterior logging of the events and verification of task accomplishment.

Data analysis was performed using IBM SPSS Statistics version 25.0 for Windows (Armonk, NY: IBM Corp.).

TABLE II. LIST OF THE USABILITY TEST MICROTASKS AND THEIR COMPLEXITY

Microtask	Brief Description	Complexity
A1.1-Log	Log in the system	5
A1.2-Select	Select a specific user	2
A1.3-Create	Create a session for this user	1
A1.4-Add	Add the general assessment routine to the session	3
A1.5-Edit	Edit session date	2
A1.6-Save	Save session	1
A2.1-Select	Select another specific user	2
A2.2-Open	Open this user profile	1
A2.3-Create	Create a training plan with given details	22

Fig. 3 Some screenshots of the software system: a) authentication screen; b) users' list; c) session form; d) session components gallery. Users' pictures and names are fictional.

III. RESULTS

A. Usability test logging

From the video recording analysis, the logging of events was highly useful to identify the major usability problems and errors of the system. Table III shows the means and standard deviations of accomplishment, duration, and the number of detected issues on each microtask.

The accomplishment was measured in three ways: if the participant accomplished the microtask as requested it got a value of 2, if the microtask was accomplished but not precisely as requested it got a value of 1, and 0 if the microtask was not accomplished. All microtasks were generally accomplished, except microtasks A1.4-Add, A2.2-Open, and A2.3-Create. In the A1.4-Add, some participants did not add the assessment routine to the session. Instead, they added a single assessment and concluded the microtask considering they had done it

well. The microtask A2.2-Open, as mentioned above, opening the profile was not intuitive at all in the current implementation. Most participants pressed the username instead of the user picture to open the profile. The microtask A2.3-Create was accomplished by all participants. However, some did not input the correct details for the scenario sheet.

TABLE III. MICROTASKS MEANS AND STANDARD DEVIATIONS OF ACCOMPLISHMENT RATE, DURATION, AND DETECTED ISSUES.

Microtask	Accomplishment	Duration (MM:SS)	Number of issues
A1.1-Log	2±0	0:20±0:04	0.3±0.5
A1.2-Select	2±0	0:25±0:18	1.4±1.8
A1.3-Create	2±0	0:08±0:15	0.0±0
A1.4-Add	1.7±0.5	1:07±0:56	3.3±3.4
A1.5-Edit	2±0	0:14±0:12	0.7±1.9
A1.6-Save	2±0	0:08±0:13	0.6±1.9
A2.1-Select	2±0	0:18±0:15	0.5±1
A2.2-Open	0.7±0.9	0:29±0:41	2.1±1.7
A2.3-Create	1.8±0.4	2:26±0:43	4.7±2.9

Concerning microtask duration, which is presented in minutes and seconds, it is consistent with the complexity of the task. For instance, the fastest microtask to accomplish were A1.3-Create and A1.6-Save with an average of 0:08 seconds. The two tasks had a complexity of 1, and this means that each task is only composed of one single step, for instance, pressing a button. Interestingly, there was another microtask with the same complexity, which also only required to press one button, but had the lowest accomplishment rate which was the microtask of viewing the user profile (A2.2-Open). The longest task was the A2.3-Create, consistent with its complexity, and was the one that had the highest number of detected issues.

Issues correspond to events on the logging sheet and could be of the following types: assistance from the moderator (A), bug (B), general comment (G), miscellaneous event (M), opinion (0), and usability problem (X). Following, a description of the most critical items (registered by 5 or more participants):

• Wrong component selection in A1.4-Add (X) – as mentioned above, some participants added wrong components to the session. This may be because the scenario was requesting to add an assessment routine, and individual assessments were added – which appeared first on the list – instead of routines.

• Comment from moderator to assist in A1.4-Add (A) – assistance was given in two different circumstances on this microtask. First, after opening the session, the participant could not always find intuitively how to add the routine to the session. In this situation, the moderator led the participant to think about how he/she could add the session component. Second, since some users added wrong components to this microtask, sometimes when they realized, they did not know how to correct their mistake and needed assistance.

• Selecting the username to open the user profile in A2.2-Open (X) – participants intuitively pressed the user name instead of the user photo.

• Selecting the current month to change the month in A2.3-Create (X) – when opening calendar to select a date, participants pressed the current month label considering that it would enable them to change the month.

• Comment from moderator to assist in A2.3-Create (A) – this assistance was given mostly concerning the input type or by restarting the application when the participant closed it thinking it would close only the current window.

• Buttons were not responding at first try in A2.3-Create (B) – this happened when pressing dropdowns after selecting the plan's end date. Also, when pressing the "add" button when adding assessments to the plan. The software responsiveness of the buttons to touch was not entirely stable which didn't provide an immediate response in some circumstances.

B. USE questionnaire

The results of the USE questionnaire were analyzed by its domains: usefulness, ease of use, ease of learning and satisfaction. Boxplots for these domains depict results above the value of 5 of the Likert scale (Fig. 4) where 7 is the best usability score that one item can get. One user was identified as an outlier and displayed very bad acceptance and skepticism concerning the presented system. Overall, all domains display very high scores, being Usefulness the lowest with a mean of 5.3±1.4. The highest score was identified in the ease of learning domain which obtained a mean of 6±1.3. The domains ease of use and satisfaction both obtained identical results: 5.7±1.3.

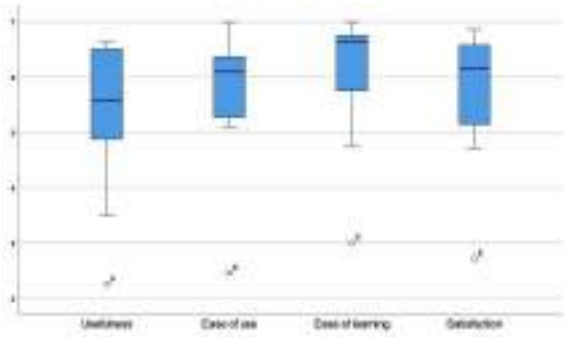

Fig. 4. USE questionnaire results averaged by domain. The Y-axis corresponds to the answer on the Likert scale from 1 to 7.

In the final part of the USE questionnaire, each participant was invited to mention the negative and positive aspects of the evaluated system. Among all participants, 27 positive and 10 negative aspects were referred, showing that this system generates more positive than negative impressions. On the positive side, the ease of use, the intuitive interface and the possibility for long-term planning were the most mentioned aspects. The most mentioned negative aspect was the fact that games are not playable by big groups at the same time (N=2), this may be because is something very different from the daily routine of sports professionals who typically prescribe physical exercise to groups and not to one individual at a time. Some participants made suggestions for further implementations, for instance, the actual system does not allow to categorize end-users by groups or classes. This is important when considering institutions that have multiple groups of end-users. Others suggested towards addressing issues such as some tabs are not very intuitive, too much information at the beginning, some concern about

vulnerability to technical failures, and the system should consider user's pathologies and precedents.

IV. Discussion and Future Work

Planning and systematically run the interviews assisted in gathering both quantitative and qualitative data during the evaluation tests. Concerning the event logging, a better solution would have been to have implemented it on the system itself. Nonetheless, that would only allow to keep track of the accurate duration of each microtask and not registering the user comments and difficulties. Therefore, in future evaluations, both the video recording and system's data logging should be considered.

Results from the USE questionnaire were high, and in general, the impressions on the system were very positive, which supports the value and applicability of our proposed integrative system. Based on the observation and registry of issues during the usability sessions, the following implementations are considered for the next iteration:

1. Adding components to a session was not so intuitive as it should be. The users needed clarifications from the moderator to accomplish the task. Perhaps changing the icon or adding extra information will result in a more intuitive interaction.
2. Pressing the user photo to open the profile was not intuitive since most of the participants pressed on the username instead, which resulted in ignoring this task possibly because by pressing the username nothing was changing. Therefore, both the username and photo should be used as a button to open the profile.
3. Calendar interaction should be redesigned because it was not intuitive for some participants. There is a need to distinguish better which elements are buttons from those that are not, and to implement the functionality as participants were expecting (i.e., enabling changing month or year when touching it).
4. Swipe gesture was many times detected as touch, while other times touching buttons or drop-down selection did not respond as they should, and users were required to press more than once. Hence, the touch detection and interaction must be improved to avoid such usability problems.
5. When pressing the login button, sometimes the system took longer to respond than expected, which led participants to press twice or more. Therefore, a wrong user was selected on the next screen. The existence of a loading screen for a few seconds may prevent the issue.
6. Session components and routines should be better distinguished or presented separately.
7. The Quit button was sometimes interpreted as a button to close the view. The icon should be changed.
8. The input of text in numeric fields resulted in a system error. Hence, adding restrictions on numeric input fields, or changing the input method for the frequency, moments of assessments and session duration fields will overcome this problem.

V. Conclusion

This paper presented a novel Exergaming integrative platform, and the methodology applied to its usability evaluation. Results were positive and data gathered show the importance of running these types of tests with the prospective target users. This study enabled us to find errors and non-intuitive functionalities of the system, while it also enabled to assess how easy, satisfactory and useful users find working with our system. In particular, the high scores in the USE questionnaire show that this integrative system is a promising tool that sports professionals may be willing to use on their daily practice. Therefore, further efforts and improvements will be considered to bring this system to life.

Acknowledgments

This work was supported by the Portuguese Foundation for Science and Technology through the Augmented Human Assistance project (CMUPERI/HCI/0046/2013), LARSyS (UID/EEA/50009/2019), and MACBIOIDI project (INTERREG program MAC/1.1.b/098).

Special thanks to all sports professionals that participated in the usability study.

References

[1] "United Nations - World Population Ageing 2017 Report.pdf." [Online]. Available: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Report.pdf. [Accessed: 09-Nov-2018].

[2] N. Skjæret, A. Nawaz, T. Morat, D. Schoene, J. L. Helbostad, and B. Vereijken, "Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies, safety and efficacy," *Int. J. Med. Inf.*, vol. 85, no. 1, pp. 1–16, Jan. 2016.

[3] D. L. Kappen, P. Mirza-Babaei, and L. E. Nacke, "Older Adults' Physical Activity and Exergames: A Systematic Review," *Int. J. Human–Computer Interact.*, vol. 35, no. 2, pp. 140–167, Jan. 2019.

[4] K. D. Eason, "Towards the experimental study of usability," *Behav. Inf. Technol.*, vol. 3, no. 2, pp. 133–143, 1984.

[5] B. Shackel and S. J. Richardson, *Human Factors for Informatics Usability*. Cambridge University Press, 1991.

[6] J. Nielsen, *Usability Engineering*. Elsevier, 1994.

[7] 14:00-17:00, "ISO 9241-11:2018," *ISO*. [Online]. Available: http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/35/63500.html. [Accessed: 20-Nov-2018].

[8] N. Bevan, J. Carter, J. Earthy, T. Geis, and S. Harker, "New ISO Standards for Usability, Usability Reports and Usability Measures," in *Human-Computer Interaction. Theory, Design, Development and Practice*, 2016, pp. 268–278.

[9] A. S. Billis, E. I. Konstantinidis, A. I. Ladas, M. N. Tsolaki, C. Pappas, and P. D. Bamidis, "Evaluating affective usability experiences of an exergaming platform for seniors," in *2011 10th International Workshop on Biomedical Engineering*, 2011, pp. 1–4.

[10] E. I. Konstantinidis, G. Bamparopoulos, and P. D. Bamidis, "Moving Real Exergaming Engines on the Web: The webFitForAll Case Study in an Active and Healthy Ageing Living Lab Environment," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 3, pp. 859–866, May 2017.

[11] G. Palestra, M. Rebiai, E. Courtial, and D. Koutsouris, "Evaluation of a Rehabilitation System for the Elderly in a Day Care Center," *Information*, vol. 10, no. 1, p. 3, Jan. 2019.

[12] A. Nawaz, N. Skj\a eret, K. Ystmark, J. L. Helbostad, B. Vereijken, and D. Svan\a es, "Assessing Seniors' User Experience (UX) of Exergames for Balance Training," in *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, New York, NY, USA, 2014, pp. 578–587.

[13] A. Gonçalves, J. Muñoz, E. Gouveia, M. Cameirão, and S. Bermúdez i Badia, "Portuguese Tradition Inspired Exergames for Older People Strategic Tools to Promote Functional Fitness," presented at the icSports 2017, Funchal, Portugal, 2017.

[14] T. Paulino, J. Muñoz, S. Bermudez i Badia, and M. S. Cameirao, "Design of an Integrative System for Configurable Exergames

Targeting the Senior Population," in *Human Systems Engineering and Design*, 2018.

[15] H. Simão and A. Bernardino, "User Centered Design of an Augmented Reality Gaming Platform for Active Aging in Elderly Institutions," presented at the Special Session on Exergaming as a Strategic Tool to Promote Functional Fitness in Older People, 2017, pp. 151–162.

[16] A. M. Lund, "Measuring usability with the use questionnaire," *Usability Interface*, vol. 8, no. 2, pp. 3–6, 2001.

[17] C. Dantas, A. L. Jegundo, J. Quintas, A. I. Martins, A. Queirós, and N. P. Rocha, "European Portuguese Validation of Usefulness, Satisfaction and Ease of Use Questionnaire (USE)," in *Recent Advances in Information Systems and Technologies*, 2017, pp. 561–570.

[18] M. Gao, P. Kortum, and F. Oswald, "Psychometric Evaluation of the USE (Usefulness, Satisfaction, and Ease of use) Questionnaire for Reliability and Validity," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 62, no. 1, pp. 1414–1418, Sep. 2018.