

Usability evaluation methods for mobile serious games applied to health: a systematic review

Jaison Dairon Ebertz Schmidt¹ · Ana Carolina Bertoletti De Marchi¹

Published online: 22 December 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract This paper aims to present the results of a systematic review focused on usability evaluation methods for serious games (SG) of mobile devices applicable to health care. The research questioned which usability evaluation methods have been available for mobile serious games. The research was conducted into four databases (ACM, IEEE, Science Direct and Springer) in two periods (23–30 March, 2015, and 01–07 June, 2016). After evaluating 2191 papers, the researchers considered that 9 met the eligibility criteria. As a result, similarities between some methodologies used have been found, however a specific methodology for SGs usability evaluation applicable to health has not been encountered.

Keywords Serious games · Usability · Evaluation · Mobile devices · Smartphone · Tablet · Health

1 Introduction

With an increasing processing power and storage, the use of smartphones has become constant in people's lives, often something indispensable. Its applicability covers the most varied day-by-day activities, being present in the workplace, education, communication and entertainment.

Among applications for entertainment, there are digital games which attract users in a wide age range, from children to elderly. For Wiemeyer and Kliem [1], digital games

divide opinions within the scientific field, being, on the one hand pointed out as the cause of sedentary lifestyle, dependence, aggressiveness and other social risk factors and, on the other, defended as a mechanism for improving cognitive, sensory, motor and social interaction. However, when the game goal goes beyond simple entertainment, encouraging mental exercise and/or physical education, health, training and knowledge acquisition, such games are called serious games (SG) [2].

By using mobile serious games (MSGs), users can transpose the length of space of the game screen [3]. This characteristic makes SGs integrate into the real world through the use of resources from the mobile devices, such as accelerometers, GPS tracking systems and messaging and gives the user a differentiated interaction experience. Due to the characteristics of the target group addressed in this paper—patients under recovery—and the various applications of SGs, it is necessary to be extremely cautious when using these applications properly.

The normative ISO 9241-11 [4] defines usability as “the measure to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction.” Usability is a key element in the digital game development process and directly influences the overall user experience [5]. Serious games pose unique challenges considering their usability efficiency. Because of their specific audience, SG users are generally expected to focus their attention on carrying out the proposed activities. An application with low level of usability would make the user divide the attention between technical use and the achievement of the proposed tasks and would lead to an undesirable output [5].

Consecutively, usability tests are tools aimed at providing a heterogeneous population with the use of many different technologies [6] and the developer with an

✉ Jaison Dairon Ebertz Schmidt
63227@upf.br

Ana Carolina Bertoletti De Marchi
carolina@upf.br

¹ Universidade de Passo Fundo, Passo Fundo, Brazil

assessment of how effectively an application is usable [5]. Yet, according to Moreno-Ger et al. [5] many of the existing usability evaluation tools are not currently applicable to SGs due to the fact that traditional digital games focus on competition, errors and attempts, and these factors are not always desirable or even applicable for SGs.

In the above context, it is mandatory to use usability evaluation methods to ensure that the desirable characteristics of serious games have been accomplished. Rubin [08] highlights that usability testing techniques are used to evaluate any type of product or system and are also considered research tools, though in turn, they are specific to the type of application under evaluation. Therefore, it is necessary to study specific methods applicable to assess the usability of SGs within the context in which they operate. In this context, our aim is to systematically review the literature looking forward to identify which are the usability evaluation methods available for serious games applied to health care for mobile devices.

2 Method

The work presented in the present paper is a systematic review (SR). According to Campbell [7], an SR is intended to “summarize the best available research on a specific issue through the synthesis of the results of several studies.” Campbell [7] emphasized that a SR should use transparent procedures to find, evaluate and synthesize relevant search results. These procedures must be previously defined in order to ensure that the study can be replicated. Thus, an SR is guided by a specific research question and must also follow clear eligibility guidelines to make the research relevant. The SR consists of identifying, screening and setting eligibility criteria, and each of the stages performed aims to filter the selected studies adequately [7].

In the identification stage, researchers perform searches on databases and catalog all studies which meet the search criteria. In the screening process, the studies previously identified are filtered, by reading each title and abstract, to verify which match the scope of the research problem. If it is unclear whether the study fulfills the requirements or not, or if there is doubt, it should be shortlisted for the eligibility stage. The eligibility phase consists in fully reading selected articles in order to verify in depth whether they fulfill the established criteria for inclusion. At the end of these three phases, the final set of studies are synthesized aiming to answer the research problem. Thus, the research question posed in the present study is the following: which are the available usability evaluation methods for serious games applied to health care on mobile devices?

2.1 Eligibility criteria

The studies included into this SR had to meet the following eligibility criteria: (a) evaluate Serious games; (b) evaluate games compatible with mobile devices (e.g., smartphones or tablets); (c) evaluate SGs applicable to health care; and (d) make use of any kind of usability evaluation technique/method. All studies that did not meet these eligibility criteria were excluded from this SR.

2.2 Search strategy

The searches were conducted in four electronic databases: Association for Computing Machinery (ACM), Institute of Electrical and Electronics Engineers (IEEE), Springer International Publisher Science and Science Direct. They took place from 24 to 30 March, 2015, and from 01 to 07 June, 2016, and attempted to enrich this paper with the maximum number of material to maximize its review. For defining the search terms, the researchers performed result tests and verified the impact of those terms into the results. They also conducted an etymological research to determine which terms would bring the expected results from English to Portuguese. After initial testing and researching for relevant terms, the search was defined based on the following expression: **game AND usability AND evaluation AND (mobile OR smartphone OR tablet) AND health**.

2.3 Studies selection and data extraction

As a selection criterion, only papers published in related scientific journals were considered, disregarding book chapters, books, conference abstracts or seminars. The titles and abstracts of all papers identified by search strategy were evaluated. All abstracts that did not provide sufficient information on the eligibility criteria were evaluated entirely. It should be observed that the studies were evaluated by two researchers independently. If there was no agreement, or if they were in doubt about the inclusion of the paper into the study, a third researcher would be consulted in order to ensure a correct evaluation. This process has generated a document containing the job title, internet address and a brief evaluation report from each researcher. This document justifies the inclusion, or not, of particular studies in the final search results. The data extraction followed Sampaio and Mancini [8] guidelines, and the following pieces of information have been identified: authors, publication year, goals, methodological design, number of participants and main results.

3 Results and discussion

The researchers identified 2191 studies in the databases mentioned above. After the initial selection, 87 appeared to be potential papers for complete analysis. The papers were fully read as the title and abstract did not bring enough information whether they matched the eligibility criteria proposed or not. After reading the entire pieces, 78 did not fit the eligibility criteria, and thus, only 9 papers were considered suitable for the final analysis. Figure 1 shows the flowchart of identified studies, and Table 1 shows in detail all the studies included. It should be noted that the studies were selected based on the eligibility criteria that had been previously established. Therefore, for the identification and screening, only the title and abstract of each result found were read.

The paper “Dancing in the streets: The design and evaluation of a wearable health game” by Clawson et al. [9] aims to propose a game to fight teenage obesity. This application was named “Dancing in the streets” (DITS) and uses two wireless sensors and an accelerometer in addition to the mobile devices.

DITS can be described as a game that uses body movement through dance. It uses music and sensors to detect movement, and a mobile device to control the game and visually interact with the user. To play, the user holds a sensor on each ankle. Thus, the movement of the user’s feet becomes monitored and should coincide with the guidance provided by the application screen. The feedback over the activity progress is made through the player scores.

DITS evaluation was conducted in two schools and was composed of three phases: user training, game play and a Likert questionnaire. In total, 50 students were selected, including 28 males and 22 females, aged between 16 and 17. For the game phase, 4 groups of students were created and the students played the game twice using different songs. Each group took about 10 min to perform the activities.

Participants completed a questionnaire in order to measure their satisfaction level with the game. This questionnaire comprised topics such as: “Evaluate your overall experience while playing DITS,” “Initial impressions when using the application” and “Tell us what you liked playing the game.” As the game makes use of sensors, questions were designed to evaluate these mechanisms in particular, aiming to measure the user satisfaction with the gestural interaction and sensors’ performance and to learn about what improvements users indicated.

After the tests, the results showed that the DITS received a positive evaluation from almost all users. The necessity of adequately training the participants to use the sensors before starting the game and introducing

competitive modes to encourage students to continue playing was also verified. Another relevant point was the inclusion of collaborative resources and training tutoring.

The paper “Mobile games and design requirements to increase teenagers’ physical activity” by Arteaga et al. [10] aims to explore the games design requirements targeted to encourage physical activities (PA), in order to improve health and prevent cardiovascular problems or diabetes. The target audience were Hispanic teens.

The study aimed to evaluate five characteristics of the game development focusing on physical activity: (a) exploring the design requirements for motion-based games; (b) evaluating current games for mobile devices which encourage physical activity; (c) evaluating motivational phrases; (d) analyzing teenage preferences over games; and (e) analyzing differences between games aimed at teenagers with normal weight or overweight.

The study included 51 teenagers aged between 15 and 18, including 38 female and 13 male subjects. Each participant had an hour to perform the activities. They answered three questionnaires and played nine games in the iPhone/iPod Touch devices at random, each game being 2 min long. At the end of the experience, the participants were also interviewed. They were asked questions in random order, and the questionnaire aimed to evaluate motivational phrases used in the game, as well as hold a personality test and conduct an overall evaluation of the game.

For evaluating the motivational phrases used in the game, a survey containing 10 sentences was prepared. Each survey used a rating scale ranging from 1 (not motivated) to 5 (very motivated). Thus, the idea was to evaluate which motivational phrases impacted the user the most. In the end, each participant would suggest three motivational phrases of their choice.

For personality evaluation, the Big 5 Model of Personality was used that consists of evaluating the effect caused by a particular phrase according to the user’s personality characteristics being evaluated. The game evaluation questionnaires asked the players to choose, in their opinion, what were the top and worst three games they have played and their reasons for those choices.

Finally, each participant was interviewed. The interview questions asked about their physical activities and experiences with games for mobile devices, their opinion about those games and the main obstacles for the development of physical activities (PA) using mobile devices.

According to the authors, the results obtained have shown that the players’ personality interferes in many issues, such as preferences for certain motivational phrases, features expected for a particular game or choosing particular types of games. For example, individuals with a higher degree of consciousness could be invited to set a

Table 1 Details of the studies included in the review

Authors	Year	Evaluation technique	Method focus	Evaluation steps	Participants	Age	Time spent	Results
Clawson et al. [9]	2010	Likert questionnaire	Verify game acceptance by the users	1) Training 2) Play the game 3) Questionnaire	50 Total 28 Males 22 Females	16–17 years old	10 min per game	Good game acceptance by participants
Arteaga et al. [10]	2012	Likert questionnaire and interview	Effectiveness of game phrases; game preferences and aspects of the game acceptance	1) Questionnaire 2) Play the games 3) Interview	51 Total 38 Males 13 Females	15–18 years old	01 h	Competitiveness increases the game acceptance; the player's personality affects phrases effectiveness
Barros et al. [11]	2014	Recording each game session	Application Ease of use by the elderly	3 gaming sessions with different groups	25 participants divided into three sessions, each session with distinct groups	65–96 years old	20 min per session	Several usability recommendations for seniors
Brown et al. [12]	2011	Think-aloud	Application ease of use by people with some kind of intellectual disability	1) Prototype evaluation 2) Experts evaluation 3) Users evaluation	8 users for prototype evaluation 12 specialists for usability evaluation 12 users for application evaluation	15–32 years old	N/I	Use of one field per screen; the use of touches should be rationalized to this audience in particular; scroll is not desirable; other related policies
Brown-Johnson et al. [13]	2015	Think-aloud	Verify the use of game for people with lung cancer	Structured interview	8 Total 7 Females 1 Males	20–50 years old	01 h	The application showed promising; several improvements noted
Sunwoo et al. [14]	2010	Recording game session	Verify which of 2 games is more easily played and which characteristics improve the game's usability	1 game session per user	21 users	11–37 50–63 years old	N/I	Games based on gestures seem to be more pleasant for users
Schmitz et al. [15]	2015	System Usability Scale	Verify elements for potential game improvement	1) Training 2) Play the game 3) Questionnaire	156 users 8 professionals 149 students	Students: 10–18 years old	N/I	Several guidelines and improvements must be implemented for this kind of application
Tong et al. [16]	2015	Health professionals observation	Verify elements for potential game usability improvement	1 game session per user	8 Occupational therapists 16 users	> = 18 years old	3 min and 35 s on average	Users usability improvements and interesting resources for specialists
Pirani et al. [17]	2016	Likert questionnaire (online)	Verify if their game is more easily played by elderly than existing ones	N/I	N/I	N/I	N/I	Their application is better in comparison with other existing ones

playing group or invite a friend to play, as they demonstrate a greater degree of organization.

Games with greater emphasis on competitiveness had a better acceptance by male users, whether they had a story behind it or not. The design requirements reached through qualitative analysis demonstrate that simply developing a game which contains a PA, and establishing goals, will not be sufficient to keep young people's interest and, in turn, the inclusion of competitive elements may stimulate the constant use of the application.

The paper "Design and Evaluation of a Mobile User Interface for Older Adults: Navigation, Interaction and Visual Design Recommendations" by Ana Correia de Barros, Roxanne Leitão and Jorge Ribeiro [9] aims to describe design and evaluation processes of a smartphone application focused on elderly, whose objective is to promote physical activity and prevent falls. For the study development, a game named "Dance! Don't Fall" (DDF) was designed based on the Windows Phone 7 (WP7) platform. In this game, users can dance alone or compete with friends. To play the game, the user should hold the mobile device on his back and follow the instructions on a Google TV feature.

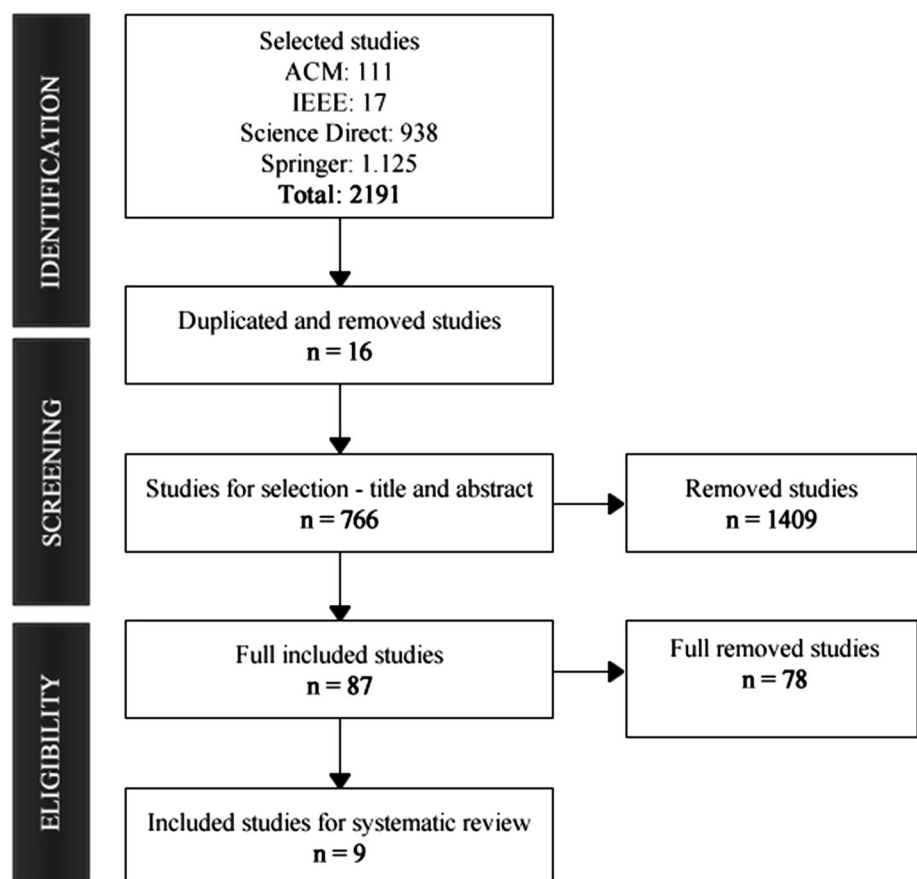
The tests were conducted in three 20-min sessions in the clinics where the recruited elderly lived. Each session was

assisted by two researchers and was recorded for future analysis. Nine subjects participated in the first session, two males and seven females, aged between 65 and 92. The purpose was to evaluate the application home screen, which was based on a WP7 navigation model. The game operations were submitted using the native WP7 model that did not reach its goal due to difficulties faced by participants who failed to perform the actions required efficiently.

Nine new subjects took part in the second session, five males and four females, aged between 68 and 89. In this session, the application home screen was changed resulting in better user acceptance. According to the authors, this improvement was due to the use of icons, not only text, to describe the actions which helped the users to understand the requested operations. Participants had difficulty using the virtual keyboard, not finding characters or failing to read them displayed. Another difficulty presented was the fact that the elderly do not relate dance with competition, and in their concept, dancing is not an activity you do alone, but something to be performed in pairs or groups.

Session three aimed to validate the changes made in the application interface, the terms used, icon legibility and the WP7 List Picker feature. This session was performed with seven seniors aged between 65 and 96. Participants had problems to scroll and click items, since apparently

Fig. 1 Flowchart of the identified studies



different gestures did not fit different actions. The element “Back to main menu” was not well assimilated, because the participants were not familiar with the concept. Finally, no participants could use the List Picker component without the help of the researchers.

As a result, the authors have listed a series of guidelines for developing interfaces for elderly, which are: be careful when using the Panorama and Pivot features; use the home screen as a central return point; use the back button as a way to undo the wrong operations; make use of the scroll only if there is time for teaching it in advance; minimize virtual keyboard use; use words that have association with elderly physical world; use a larger spacing between the screen items; use icons and text in the buttons; and be careful with elements positioned near screen border.

The paper “Designing location-based learning experiences for people with intellectual disabilities and additional sensory impairments” by Brown et al. [12] presents the development of an application that combines serious games with location-based systems to help people with intellectual disabilities and additional sensory impairment plan new routes to their work, or memorize these routes, leisure and learning opportunities.

The application, the so-called route mate (MRI), was developed for the Android operating system. For interface design evaluation, eight users were selected to evaluate the prototypes screens. The tests were divided into two separate sessions in which each participant had several intellectual disabilities such as Down’s syndrome and others. The implementation stage was divided into two stages, of which only the second was described by the authors.

The interface design followed principles of simplicity using icons and brief texts, without quick messages or elements which would hinder the understanding, and contained only the most relevant information needed to use the application, enabling the icons to be easily and directly accessed. For the interface evaluation, 12 experts in diverse applications for mobile devices (including serious games) were consulted and suggested improvements according to heuristics commonly used in this application type.

For the user evaluation phase, the think-aloud technique was applied, in which the end user uses the application together with the researcher to make suggestions in real time or even answer questions. According to the authors, this is a very effective evaluation technique for applications whose target audience has learning difficulties. For the tests, 12 members aged between 15 and 32 with various types of disabilities were selected and received several tasks to do such as: create a commuting route, set up a warning and send a distress message.

In general, the experts’ evaluation group considered the application to be extremely beneficial for the target audience. Some suggestions for improvement were collected,

such as symbols and language used in buttons, because some users with hearing impairment had problems with English terms. Some directives have also been mentioned, namely the placement of a single task per screen to avoid confusion by the user, and prevention from using scroll, as for this target audience this is an unwanted feature that hinders usability. This last finding is a relevant fact since to the public studied by Barros et al. [11], the scroll use was not a limitation factor in application usage, after learning how to use it properly.

The paper “Development and usability evaluation of the mHealth Tool for Lung Cancer (mHealth TLC): A virtual world health game for lung cancer patients” by Brown-Johnson et al. [13] aims to test the viability and usability of a game called Mobile Health Tool for Lung Cancer (mHealth TLC). This game approaches the realization of virtual doctor visits for patients with lung cancer through an immersive 3D environment using an iPad platform.

Similarly to the technique utilized by Brown et al. [12], the TLC evaluation methodology was also performed by using the think-aloud technique through structured interviews. For usability tests, eight health professionals were recruited at the University of California, aged between 20 and 50, being 7 women and 1 man. Researchers interviewed the subjects about their experiences and impressions while using the application. These interviews were recorded for later analysis.

Usability evaluation results demonstrate that the application can be useful for patients with lung cancer. The search reported low-quality audio elements and the virtual trainer feature that presented a low performance as points to be improved though the virtual trainer proved to be a feature of users’ interest. Improvements in the application also involve the reduction of interactions with the application when the patient fails a task or yet participative tasks generation and objectives to be reached.

The paper “Mobile Games for Elderly Healthcare” by Sunwoo et al. [14] aims to evaluate 2 games for iPhone and iPod Touch devices, i.e., Bowling Game and Penguin Toss. The games selected allowed elderly users to practice physical activities involving muscle movement to prevent arthritis and osteoporosis.

The usability tests were conducted with 5 subjects aged between 50 and 63 and 16 users aged between 11 and 37. The researchers recorded the time required by the users to understand how to play each game, with 24% of the subjects being able to learn how to play Bowling in less than 3 min, and 57% to play Penguin Toss in the same time. All participants were capable of understanding both games in maximum 10 min. This statistics remark that Penguin Toss is a game extremely easy to understand. One of the factors considered by the researchers regarding the games’ ease of use was their playing types. Penguin Toss can be played

almost solely by gesture-based moves, while Bowling requires touch screen interaction. By the end of each experiment, the users had to scale their level of physical exhaustion, although they did not evaluate degrees of the game usability.

The paper “Attuning a mobile simulation game for school children using a design-based research approach” by Schmitz et al. [15] aimed to evaluate an application named HeartRun. The application aims to teach kids and teenagers first aid Cardio-Pulmonary Resuscitation (CPR), as well as to discover which design principles support the development of this kind of applications, by testing its usability. The researchers did not detail much regarding the type of the device used, only mentioning it was a smartphone.

For usability tests, the subjects received orientation about the application use and researchers provided them with video recordings and notes. Participants had to fill in a form after playing the game. At this stage, we highlight their use of system usability scale (SUS) as a tool to elaborate the application usability assessment questionnaire. This was the first paper to justify its usability assessment tool. The tests were performed in three sessions with a total of 157 participants aged between 12 and 18, in which each session evaluated users of similar age. The evaluation process consisted of training in the use of the application, recording each session and administering the questionnaire previously mentioned. As a result, many guidelines have been listed for the development of applications with the same or similar purposes.

The paper “Case Study: A Serious Game for Neurorehabilitation Assessment” by Tong et al. [16] consisted in evaluating the efficiency and usability of an application focused on patients rehabilitation with any kind of cognitive disability. As the target audience has cognitive disabilities, the application evaluation was obtained from their occupational therapists (OT) similarly to the study Brown et al. [12] had performed.

A total of 5 OTs and 16 patients participated in the tests using a tablet. The researchers did not provide much information about the equipment used; however, through the pictures provided in the paper, it is possible to presume it is an Android device. A demographic survey (e.g., age, sex) was not conducted, though all participants were over the age of 18. The usability evaluation of this study showed peculiar data, due to being assessed over the OTs perception and notes instead of the patient’s direct usage feedback. Therefore, the search received only feedback over what the OTs expected the application to do and where it could be applied in, bringing up no reports regarding usability. The paper lists several desirable guidelines for this application, such as large elements and spaces to minimize cognitive difficulties faced by the users.

The paper “Android Based Assistive Toolkit For Alzheimer” by Pirani et al. [17] aimed to present the development and evaluation of an application named Alzheimer Application System (AAS) for the Android platform to support people with Alzheimer in their daily activities. This study does cover not only a Serious Game, but also a toolkit containing a number of sub-applications, such as a game similar to a Quiz addressed to improve the user’s memory.

The tests compared AAS with existing applications and evaluated usability among other features. The tests were based on an online survey conducted using the Survey Monkey tool. The users were able to evaluate matters regarding usability in a Likert scale of 5 points. This application received a good evaluation of usability reaching an average of 4 points. We highlight that this paper did not explain its questionnaire and did not report the number of respondents or their age range. On the other hand, this was the only study that used online tools for the survey.

The papers by Arteaga et al. [10] and Clawson et al. [9] used similar user groups in their evaluations, both with respect to age and sample size. Therefore, similarities could be identified regarding the evaluation method, in which both used the Likert scale, and questionnaire evaluation. A relevant point to be noted in these studies is that although the evaluation methods are similar, the goals were different. While the first study aimed to evaluate specific application questions (e.g., effectiveness of used phrases), the second evaluated the application globally.

The study presented by Barros et al. [11] was the only one focused on serious games for seniors. Both Barros et al. [11] and Sunwoo et al. [14] used recording mechanisms for further analyzing usability in their sessions, though a relevant point of Barros paper is that it does not explicitly list indicators of how the recordings were used and/or analyzed, leaving a subjective opinion to the enquiry that the recommendations cited were extracted from them.

The think-aloud evaluation technique was used by Brown et al. [12] and Brown-Johnson et al. [13]. Both studies presented the same number of participants, but used different groups to evaluate the application. The first study used healthcare professionals and listed usability recommendations for their target audience, and the second paper aimed at patients with cognitive disabilities in order to verify whether the application was promising for the area it was intended to.

Another aspect observed in the studies refers to the use of different techniques to obtain similar results. This can be observed in the works of Clawson et al. [9] and Brown-Johnson et al. [13], where different techniques were used to evaluate the application acceptability. Finally, the studies from Arteaga et al. [10], Barros et al. [11] and Brown et al.

[12] used different evaluation techniques, though all indicated guidelines for the application development.

An element to be considered refers to the time used for conducting those studies. Each technique stipulated without further details the time at which the user should be subjected to evaluation, not making clear the reasons that led the researchers to stipulate this variable. It may be noted that Arteaga et al. [10] and Brown-Johnson et al. [13] used the same study time, but different evaluation techniques. Thus, their results have shown some similarity, since both obtained several usability recommendations to the public to which the work was intended to. The paper by Schmitz et al. [15] was the only work which used an acknowledged usability evaluation instrument, even though not aiming to evaluate MSG usability. The other papers did not justify their usability evaluation methods, providing evidence that there is no standardized mechanism to evaluate this kind of applications.

4 Conclusions

This study aimed to seek methods for evaluating usability in serious games applied to health for mobile devices, the so-called mobile serious games (MSGs). After analyzing the results, it was found that there are different techniques for this type of evaluation. However, as much as these methods have similarities (e.g., think-aloud, Likert questionnaire), it was not possible to detect a standard way to evaluate mobile serious games. For future studies, it is necessary to detect which variables are relevant for usability evaluation in serious games, namely what are the main factors that influence the usability and acceptability of such application. Furthermore, the authors intend to propose a usability evaluation method for these applications related to health.

Acknowledgements The authors are grateful to Brazilian National Council for Scientific and Technological Development for their support (Process numbers 444088/2014-8 e 303618/2015-9).

References

1. Wiemeyer, J., Kliem, A.: Serious games in prevention and rehabilitation—a new panacea for elderly people? *Eur. Rev. Aging Phys. Act.* **9**, 41–50 (2012)

2. Carvalho, S., Ishitani, L. Fatores motivacionais para desenvolvimento de mobile serious games com foco no público da terceira idade: uma revisão de literatura. *Educação Temática Digital* (2013)
3. Sánchez, J., Olivares, R.: Problem solving and collaboration using mobile serious games. *Comput. Educ.* **57**, 1943–1952 (2011)
4. ABNT Catálogo, (2015). <http://www.abntcatalogo.com.br/norma.aspx?ID=86090>
5. Moreno-Ger, P., Torrente, J., Hsieh, Y.G., Lester, W.T.: Usability testing for serious games: making informed design decisions with user data. *Adv. Hum. Comp. Int.* **2012**, 13 (2012). doi:10.1155/2012/369637
6. Procci, K., Chao, A., Bohnsack, J., Olsen, T., Bowers, C.: Usability in serious games: a model for small development teams. *Comput. Technol. Appl.* **4**, 315–329 (2012)
7. What is a systematic review? The Campbell Collaboration, (2015). http://www.campbellcollaboration.org/what_is_a_systematic_review/
8. Sampaio, R. F., Mancini, M. C. Estudos De Revisão Sistemática: Um Guia Para Síntese Críteriosa Da Evidência Científica. *Revista Brasileira de Fisioterapia.* (2007)
9. Clawson, J., Patel, N., Starner, T.: Dancing in the streets: the design and evaluation of a wearable health game. In: *International Symposium on Wearable Computers*, Oct 2010
10. Arteaga, S.M., González, V.M., Kurniawan, S., Benavides, R.A.: Mobile games and design requirements to increase teenagers' physical activity. *Pervasive Mob. Comput.* **8**, 900–908 (2012)
11. Barros, A.C., Leitão, R., Ribeiro, J.: Design and evaluation of a mobile user interface for older adults: navigation, interaction and visual design recommendations. *Procedia Comput. Sci.* **27**, 369–378 (2014)
12. Brown, D.J. et al.: Designing location-based learning experiences for people with intellectual disabilities and additional sensory impairments. *Comput. Educ.* **56**, 11–20 (2011)
13. Brown-Johnson, C.G., Berrean, B., Cataldo, J.K.: Development and usability evaluation of the mHealth Tool for Lung Cancer (mHealth TLC): a virtual world health game for lung cancer patients. *Patient Educ. Couns.* **98**, 506–511 (2015)
14. Sunwoo, J., Yuen, W., Lutteroth, C., Wünsche, B.: Mobile games for elderly healthcare. in *Proceedings of the 11th International Conference of the NZ Chapter of the ACM Special Interest Group on Human-Computer Interaction 73–76* (ACM, 2010)
15. Schmitz, B., Klemke, R., Walhout, J., Specht, M. Attuning a mobile simulation game for school children using a design-based research approach. *Comput. Educ.* **81**, 35–48 (2015)
16. Tong, T., Chignell, M., Sieminowski, T.: Case study: a serious game for neurorehabilitation assessment. *Procedia Comput. Sci.* **69**, 125–131 (2015)
17. Pirani, E.Z., Bulakiwala, F., Kagalwala, M., Kalolwala, M., Raina, S.: Android based assistive toolkit for alzheimer. *Procedia Comput. Sci.* **79**, 143–151 (2016)