User Experience Evaluation Methods in Games for Children with Cognitive Disabilities: A Systematic Review

Carolina Rico-Olarte
Telematics Engineering Research Group
Universidad del Cauca
Popayán, Colombia
carolinarico@unicauca.edu.co

Diego M. López
Telematics Engineering Research Group
Universidad del Cauca
Popayán, Colombia
dmlopez@unicauca.edu.co

Sara Kepplinger
Human-centered Media Technologies (HMT)
Fraunhofer Institute for Digital Media Technology IDMT
Ilmenau, Germany
kpl@idmt.fraunhofer.de

***Abstract*— Background: In recent years, the interest in the field of the User Experience methods for assessing technology solutions has increased. Objective: Conduct a Systematic Review to analyze the existing evidence in the field of user experience methods used in games for children with special needs. Methods: The systematic review was a three-stage process. The first two stages were based on the Kitchenham's guidelines for Systematic Reviews in Software Engineering; as for the third stage, the Preferred Reporting Items for Systematic Reviews and Meta-analyses statement used in healthcare was adapted to report the systematic review transparently. Results: This review identified 143 papers, 10 of them meeting the eligibility criteria. Main findings include the identification of the most used evaluation methods in user experience, the detection of the research outcomes from each study and finding the main characteristics of the population of the studies. Quality of Experience methods for the addressed topic, were not found. Conclusions: There is no consensus as how the methods should be applied in the evaluation of the user experience or quality of experience. This systematic review specifies the research gaps identified for future research works in the area.**

Keywords—systematic review, evaluation methods, user experience, children, games.

I. INTRODUCTION

In recent years, the interest in the topic of User eXperience (UX) from the Human-Computer Interaction (HCI) area and its evaluation in games for children has increased [1], especially because children are considered as the main consumers of these types of technological solutions [2]. Therefore, developers are attempting to design software that meets the abilities, interests and developmental needs of the children [3]. The physical and cognitive abilities of every child are different, especially when dealing with children with cognitive impairments [4]. Despite this growth, it has not been found a comprehensive overview about UX or Quality of Experience (QoE) evaluation methods, particularly those for evaluating health technologies for children with special needs. In order to identify, evaluate and interpret all the available research relevant to this particular matter, a systematic review (SR) of the literature was conducted. This process was carried out from the current scenario of an on-going project: the system (consisting in a mat as input device and a computer game) supports the rehabilitation of children with intellectual and cognitive disabilities, focusing on memory and concentration therapies. Through conducting this review, it was possible to identify the most used methods for evaluating UX in children, the elements used in the evaluation and the participants assessed. The implications from the type of methods found are discussed and finally, some conclusions regarding the review and the challenges it presented for future work are drawn.

II. METHODS

A systematic literature review was conducted to summarize the existing evidence concerning the topic at hand. This review is a three-stage process where, first, it is important to plan the review to lately conduct it through a review protocol, and finally report the obtained information in a synthetic way. The first two stages were based on the Kitchenham's guidelines for performing a SR from the Software Engineering area [5]; as for the third stage, the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement consists in a 27-item checklist and a four-phase flow diagram used in healthcare [6]; the statement was adapted to report the SR transparently.

A. Planning the review

1) Purpose for reviewing

For undertaking this SR, it was necessary to confirm the need for one. The purpose of this review came from the known fact that there are several evaluation methods for UX [7]. Nevertheless, regarding the information about the evaluation of UX in games for children with cognitive disabilities, there was not enough knowledge available. Due to the UX relationship

with the definition of QoE [8], its evaluation methods were also considered in the review. QoE describes a user's subjective assessment of their experience when using a particular application [9]. Traditionally, QoE has focused on multimedia applications like video and audio streaming by measuring performance indicators such as throughput, delay, and packet loss. However, these indicators do not reveal much about the experience of users.

2) Research questions

The research questions were defined bearing in mind that most of the known methods for UX or QoE evaluation are subjective, i.e., there are based on the assumption that users are able to reflect their experience and communicate it through questionnaires [10]. However, applying subjective evaluation methods is a difficult task when talking about children with cognitive impairments. Given this reason and stated the purpose, three research questions (RQ) were raised:

- Which are the most used evaluation methods for UX and QoE in games for children?
- Which are the elements considered in the evaluation of UX and QoE in games for children?
- Is there any substantial difference between the UX and QoE evaluation methods?

3) The review protocol

This review protocol was based on the steps from the conducting stage of the SR and the process for performing systematic mapping studies [11].

B. Conducting the review

1) Search string

Some keywords, their synonyms, acronyms or combinations were identified and used to build the search string in order to answer the RQ. The search string was constructed using the Boolean operators:

- ("user experience" OR UX OR QoE OR "quality of experience") AND (children OR kids) AND (games OR exergames OR "serious applications")

Although the evaluation of UX or QoE was a relevant issue for searching information and answering the RQ, the evaluation was established as an exclusion criterion. The searching criteria used in the databases were limited to seek for the keywords in the sections of the paper: article title, abstract and key-words.

2) Information sources

The studies to be included in the review were identified by searching in the databases: (i) Scopus, (ii) IEEE Xplore, (iii) ACM Digital Library, and (iv) PubMed. In addition, some of the studies selected to be screened were identified in a previous systematic mapping study performed to scope the RQ stated.

3) Information extraction

The papers selected to be screened by title and abstract met the exclusion search criteria defined by year of the publication (from 2008 onwards), along with the removal of the duplicated papers. This process was performed for double-checking of the selected papers taking into account: (i) paper title containing substantial information to continue reading the abstract, (ii) the language of the paper (English), (iii) if papers were secondary studies, and (iv) if papers were available for downloading. Finally, 20 studies were selected to be full-text assessed in order to select the final studies to include in this SR. The PRISMA flow diagram (Fig. 1) provides the information regarding the selection process of studies, which comprises four phases: identification of studies, screening of studies, eligibility of studies, and studies included.

C. Reporting the review

Some of the 27 items from the PRISMA statement were followed in order to give clarity over the results from the SR process that was performed. This statement was chosen due to its focus on ways in which authors can ensure the transparent and complete reporting of SRs [6].

III. Results

This review identified 143 papers after the initial search in the electronic databases: in Scopus *n = 92*, in IEEE Xplore *n = 12*, in ACM DL *n= 36*, and PubMed *n = 3*; and 12 papers were retrieved from a list of references contained in a systematic mapping study performed. The double-screening phase involved the examination of titles and abstracts of all studies, resulting in 85 papers being excluded, as they were deemed not suitable for this SR. Consequently, 20 studies were selected for the eligibility phase. Out of these, ten studies were excluded mainly for (i) not being addressed to children (*n = 5*), (ii) not UX nor QoE evaluation mentioned (*n = 2*), (iii) not HCI systems (*n = 1*), (iv) discussion paper (*n = 1*), and (v) had a learning purpose (*n = 1*). Only 10 studies fully met the stipulated eligibility criteria for inclusion in the SR process.

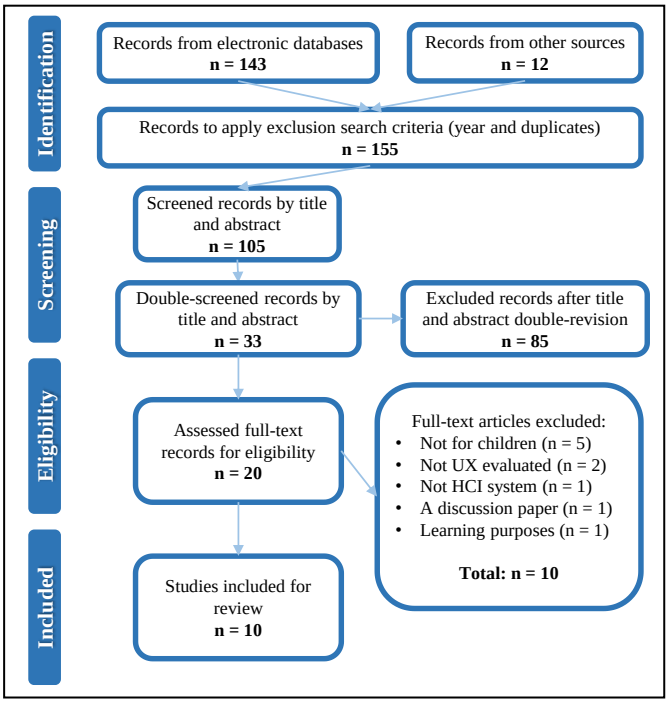

Fig. 1. Flow diagram of the study selection process

A. Study characteristics

In accordance with the RQ stated, some characteristics of the studies are selected in order to provide more information.

1) UX and QoE methods

The term QoE did not have any results: in the in-depth review of the selected articles, no result was found despite the inclusion of the term in the search string. Perhaps, the reason for this result is that QoE has been addressed towards multimedia applications, and maybe, video games are not considered within this classification. Accordingly, the question about the differences in the UX and QoE methods remains unanswered, at most of the knowledge we found in this SR. In consequence, the analysis presented below covers only UX evaluation methods. Table I presents in detail the evaluation methods used to determine the UX according to the system developed and presented by the authors.

Among the UX evaluation methods, the subjective ones such as questionnaires, direct observation and self-report are the more frequent with the 38.4% of the studies reporting these methods of evaluation. In addition, objective methods, such as recording and analyzing the log of the system or taking physiological measures like the electroencephalography (EEG) are frequently used in 30.8% of the studies. Finally, 30.8% of the studies employ video recording, which, despite of being considered by authors as an objective method, its content analysis is done by an expert, hence, this does not guarantee the complete objectivity in the results. It is important to mention that 40% of the used methods in the UX evaluation are subjective and only 20% of the studies used purely objective methods, the remaining 40% showed a combination between subjective and objective methods.

2) Research outcomes

The studies included in this SR evaluate the UX to fulfill different purposes, according to the field of expertise of the researchers or the population to which the research is addressed for. The research outcomes of the review studies are, among others:

TABLE I. EVALUATION METHODS FROM UX

Reference	UX methods
[12]	Drawing intervention, direct observation, survey method, likert scale, inquiry test, usability based on the QUIS, USE, GEQ, and UEQ questionnaires.
[15]	MemoLine, UX Curve, direct observations.
[17]	Direct observation, video recording, EEG measures, capture motion.
[13]	Direct observation, video recording, automatic-logging, Wizard of Oz.
[18]	Proved the attention improvement in children through metrics game; however, no UX evaluation methods were mentioned.
[19]	User Experience Questionnaire adapted, direct observation.
[16]	EMODIANA, video-recording, REP (Raiting of Perceived Exertion), pulse measures.
[10]	Video-observation
[20]	Fun toolkit, smiley-o-meter, the fun sorter, the again-again table, the this or that method.
[14]	Video-observation, skin conductance, self-reporting

- For designing: (i) a serious game prototype meeting the criteria of being educational and entertainment [12]; (ii) the NIKVision tabletop and a farm game for improving little children's cognitive development [13]; and (iii) the GraPM: an educational game about project management [14].
- For proposing a UX method: (i) the MemoLine tool for measuring the UX over time in children, through recalling methods and the UX curve [15], and (ii) a methodology for evaluating UX in children when interacting with active video games [16].
- For evaluating the actual UX: (i) KAPEAN is a tool for collecting data from the children with Attention-Deficit and Hiperactivity Disorder (ADHD) while they are playing with its games in order to improve its attention levels [17]; (ii) the Harvest challenge, a neuro-feedback serious video game that was evaluated with the assistance of a toolbox to obtain EEG signals from a brain-computer interface (BCI) device [18]; and (iii) the MFolktales, a mobile application prototype designed and developed for children with an educational purpose [19].

These research outcomes are important for QoE community, given that there is a gap in the research of other technological solutions such as video games.

3) Participants

Several screened papers were discarded because they were not addressed entirely to children, the subject of this study. Only four studies aimed at children with disabilities: one for children with hearing impairment, two for children suffering from Attention-Deficit Hyperactivity Disorder – ADHD, and one for children with Autism Spectrum Disorder – ASD.

4) Countries and years in which the data were collected

Mexico, UK, Spain and Colombia had two studies each [13], [15]–[18], [20], giving a special attention to a collaboration made between Mexico and Colombia [12]; the rest of the studies were from Belgium [15], Malaysia [19], Finland [10], and Poland [14].

B. Results of the individual studies

The study performed by [17], the KAPEAN tool and its components meet all the elements sought for when performing this SR. This tool gathers information in both subjective and objective ways from children suffering from ADHD, while the children are playing with mini-games that may improve their attention levels. The researchers were able to demonstrate the improvement in the attention level of the children; however, they fail to report the measurement of fun or some other indicator that identifies how the child's experience with the video game was in the experimental sessions. Thus, the purpose from the UX was not achieved completely.

C. Risk of bias across studies

An unbiased selection process is difficult to ensure; however, some precautions were taken: (i) the planning of the SR was crucial because it was possible to go through the protocol review several times in order to evaluate it and redesign it; and (ii) the risk of bias regarding the searching of the studies in the area was minimized by taking into account several important scientific databases.

IV. DISCUSSION

Therefore, several considerations about what was obtained in this study are made:

Regarding the evaluation methods from UX:

- There is no consensus as how the methods to be used in the evaluation of the UX should be selected, this selection process depends mainly on the purpose of the system/technology (in this case, games) developed, or the expertise of UX researchers and experts.
- Most used methods in the UX evaluation are subjective ones (40%), and only 20% of the studies used purely objective methods. The remaining 40% of the studies reviewed showed a combination between subjective and objective methods. From the review performed, the most commonly used method when evaluating UX in children is the Fun Toolkit [21], in this method, the authors established that "satisfaction" in adults can be considered as the equivalent of "fun" in children (regarding the game purposes due to the capacity for abstraction of thought in adults), the tool measured fun in three dimensions: expectations, engagement and endurability. This method is widely used; however, the reliability of this method is based on the children's responses and from earlier experiences, the method does not gather all the information that is necessary for performing a good evaluation of the experience of the children.
- QoE evaluation methods in games for children with cognitive disabilities were not found, despite two of the screened papers referred to it [22], [23].

Regarding the research outcomes from the studies:

- As in the evaluation methods, researchers do not find a consensus among the systems to be developed to implement the UX assessment, especially when it comes to rehabilitating children with cognitive disabilities. The options of systems-solution vary from serious games to video games, always with a generalized tendency, not only with respect to this subject, to reach towards the mobile solutions.
- Although mobile systems have been growing for some years now, regarding the UX evaluation for children with cognitive disabilities, there is not much evidence of this type of solution, because a system that meets these characteristics needs extra elements given its complexity.

Regarding the participants and the countries where the studies were performed, it is interesting to note that the studies performed in the countries of Latin America presented in this review (Mexico and Colombia) are aimed at children with a particular cognitive disability. Given the geographic proximity and the historical past, these two countries have a similar cultural context, which can lead to identify some common elements that can be used in different projects to enrich them.

Finally, regarding to the general purpose of the UX evaluation, the main objectives are related to (i) design solutions for a population, (ii) evaluate a solution, (iii) assess the user, (iv) improve a solution, and (v) improve the conditions of the user; it becomes important for the UX researcher to identify the main purpose of the evaluation and take a course of action leading to create standardized processes and methods.

V. Conclusions

The main purpose of the review was to summarize the existing evidence and provide an in-depth analysis of what has been investigated in the context of the UX evaluation methods in games for children with cognitive impairments. We followed the SR process by planning the review and conducting it based on a well-defined protocol to finally extract 10 studies that were analyzed; the reporting stage of the SR was based on the PRISMA statement, deemed as essential for transparent reporting of the results from a SR.

One of the contributions of this study was the identification of the evaluation elements and aspects of the UX methods used in the reviewed studies; these characteristics gave an overview about the most used methods for assessing children in a subjective way. Particularly, we were interested in identifying subjective and objective methods in the UX evaluation.

It is a concerning issue the fact that no results were obtained regarding the evaluation methods in QoE, confirming the gap that exists between this area and UX. One of the main objectives of this study, as stated in one of the research questions, was to obtain information in order to start building bridges and closing the gaps between these two areas, because they have some common elements that would be beneficial for the research communities associated to them. Further research is needed and it will be performed regarding this situation.

As future work, we will investigate the gaps identified in this study regarding the UX evaluation concerns, in order to have a conceptual and developmental standard framework as a validated approach for estimating the UX in games for children with cognitive disabilities.

References

[1] C. L. B. Maia and E. S. Furtado, "A Systematic Review About User Experience Evaluation," in *Design, User Experience, and Usability: Design Thinking and Methods*, A. Marcus, Ed. Springer International Publishing, 2016, pp. 445–455.

[2] Entertainment Software Association, "Essential Facts About the Computer and Video Game Industry," Apr. 2017.

[3] S. Yarosh and P. Markopoulos, "Design of an Instrument for the Evaluation of Communication Technologies with Children," in *Proceedings of the 9th International Conference on Interaction Design and Children*, New York, NY, USA, 2010, pp. 266–269.

[4] A. Alhussayen, W. Alrashed, and E. I. Mansor, "Evaluating the User Experience of Playful Interactive Learning Interfaces with Children," *Procedia Manuf.*, vol. 3, pp. 2318–2324, Jan. 2015.

[5] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," in *Technical report, Ver. 2.3 EBSE Technical Report. EBSE*, sn, 2007.

[6] A. Liberati *et al.*, "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *PLoS Med*, vol. 6, no. 7, p. e1000100, 2009.

[7] V. Roto, A. Vermeeren, K. Väänänen-Vainio-Mattila, and E. Law, "User Experience Evaluation – Which Method to Choose?," in *Human-Computer Interaction – INTERACT 2011*, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, Eds. Springer Berlin Heidelberg, 2011, pp. 714–715.

[8] K. Brunnström *et al.*, "Qualinet white paper on definitions of quality of experience," 2013.

[9] F. E. Bustamante, D. Clark, and N. Feamster, "Workshop on Tracking Quality of Experience in the Internet: Summary and Outcomes," *SIGCOMM Comput Commun Rev*, vol. 47, no. 1, pp. 55–60, Jan. 2017.

[10] S. Mäkelä, R. Bednarik, and M. Tukiainen, "Evaluating User Experience of Autistic Children Through Video Observation," in *CHI*

'13 Extended Abstracts on Human Factors in Computing Systems, New York, NY, USA, 2013, pp. 463–468.

[11] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic Mapping Studies in Software Engineering," in *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, Swinton, UK, UK, 2008, pp. 68–77.

[12] S. Cano, D. M. Alghazzawi, J. M. Arteaga, H. M. Fardoun, C. A. Collazos, and V. B. Amador, "Applying the information search process model to analyze aspects in the design of serious games for children with hearing impairment," *Univers. Access Inf. Soc.*, pp. 1–13, Jan. 2017.

[13] J. Marco, S. Baldassarri, and E. Cerezo, "NIKVision: Developing a Tangible Application for and with Children.," *J UCS*, vol. 19, no. 15, pp. 2266–2291, 2013.

[14] A. Landowska and J. Miler, "Limitations of emotion recognition in software user experience evaluation context," in *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2016, pp. 1631–1640.

[15] G. Sim, M. Nouwen, J. Vissers, M. Horton, K. Slegers, and B. Zaman, "Using the MemoLine to capture changes in user experience over time with children," *Int. J. Child-Comput. Interact.*, 2016.

[16] C. S. González-González and V. Navarro-Adelantado, "Methods and Techniques for Evaluating the Emotional Experiences of Children with Active Videogames," in *Proceedings of the XVI International Conference on Human Computer Interaction*, New York, NY, USA, 2015, p. 16:1–16:2.

[17] F. Martínez, "KAPEAN: Understanding Affective States of Children with ADHD," *Educ. Technol. Soc.*, vol. 19, no. 2, pp. 18–28, 2016.

[18] D. Z. Blandón, J. E. Muñoz, D. S. Lopez, and O. H. Gallo, "Influence of a BCI neurofeedback videogame in children with ADHD. Quantifying the brain activity through an EEG signal processing dedicated toolbox," in *2016 IEEE 11th Colombian Computing Conference (CCC)*, 2016, pp. 1–8.

[19] N. Ibrahim, W. Fatimah, W. Ahmad, and A. Shafie, "User Experience Study on Folktales Mobile Application for Children's Education," in *2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies*, 2015, pp. 353–358.

[20] G. Sim and M. Horton, "Investigating Children's Opinions of Games: Fun Toolkit vs. This or That," in *Proceedings of the 11th International Conference on Interaction Design and Children*, New York, NY, USA, 2012, pp. 70–77.

[21] J. Read and S. Macfarlane, "Endurability, Engagement and Expectations: Measuring Children's Fun," in *Interaction Design and Children, Shaker Publishing*, 2002, pp. 1–23.

[22] M. Wijnants *et al.*, "Quality of Service and Quality of Experience Correlations in a Location-Based Mobile Multiplayer Role-Playing Game," in *Entertainment Computing – ICEC 2011*, 2011, pp. 101–112.

[23] M. Claypool, R. Eg, and K. Raaen, "The Effects of Delay on Game Actions: Moving Target Selection with a Mouse," in *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, New York, NY, USA, 2016, pp. 117–123.