Usability and User Experience Evaluation of Learning Management Systems
A Systematic Mapping Study

Walter Takashi Nakamura, Elaine Harada Teixeira de Oliveira and Tayana Conte
Computer Institute, IComp, Federal University of Amazonas, UFAM, Manaus, Amazonas, Brazil

Keywords: Usability Evaluation, User Experience Evaluation, Learning Management Systems, Systematic Mapping Study.

Abstract: *Background:* Advances in technology made possible the development of powerful platforms called Learning Management Systems (LMSs), designed to help the teaching and learning process. Studies show that usability and User Experience (UX) of such platforms may influence in this process. Although several studies had been conducted in this area, most of them are at initial stages and need improvements or deeper empirical studies. *Aim:* This work aims to analyze scientific publications in order to characterize the usability and UX evaluation techniques in the context of LMSs. *Method:* We performed a systematic mapping study regarding the usability and UX evaluation techniques in the context of LMSs. *Results:* A total of 62 publications were accepted in this mapping, which helped identifying the techniques used to evaluate the usability and UX of LMSs and their characteristics such as its origin, type, performing method, learning factors, restriction and availability. *Conclusion:* Several studies were conducted regarding the evaluation of LMSs. However, there are still some gaps such as the lack of techniques with some features, e.g., feedback with suggestions to correct the identified problems. Besides, there is no sufficient evidence of which of them is best suited for this context.

1 INTRODUCTION

The advances of the technology made possible to the e-learning evolving in a complex manner regarding the educational content, technological resources and interaction possibilities (Brusilovsky, 2004). Such an event led to the development of powerful tools designed to help the teaching and learning process called Learning Management Systems (LMSs).

Learning Management Systems are specialized platforms that provide educational content by digital means (Dubost *et al.*, 2004). A LMS can offer various functionalities such as the management of educational content, accomplishment of assessment activities and communication between the students and teachers (Freire *et al.*, 2012; Kakasevski *et al.*, 2008; Hijon-Neira *et al.*, 2014).

The adoption of LMSs is not limited to learning institutions such as universities and schools. Corporations and government bodies also have been implementing such platforms to promote the employees' education and training (Oztekin et al., 2010). Thus, there is a need to use adequate techniques to evaluate these platforms so that is possible to improve their quality and, consequently, the learning and teaching process through them.

Usability and User Experience (UX) play an important role on the quality of the LMSs and in the learning process. While usability is focused on the pragmatic aspects such as user's tasks and their accomplishment, UX augments the subjective, focusing on hedonic aspects, such as user's emotions and stimulations while interacting with a product (Hassenzahl *et al.*, 2006). When the learning process occurs through a LMS, besides learning the content of the disciplines, the learner also needs to learn how to use the platform (De Carvalho and Anacleto, 2008). If a LMS do not provide a good usability, the learner will spend more time trying to understand how to use it rather than learning the educational content (Lanzilotti *et al.*, 2006). Similarly, a good UX is essential to make the platform more pleasuring and satisfactory to the learner. Being usable and interesting at the same time, the platform would be much more attractive to the user (Hassenzahl *et al.*, 2000).

Nakamura, W., Oliveira, E. and Conte, T.
Usability and User Experience Evaluation of Learning Management Systems - A Systematic Mapping Study.
DOI: 10.5220/0006363100970108
In *Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017) - Volume 3*, pages 97-108
ISBN: 978-989-758-249-3

Due to the difficulty in evaluating the usability of educational systems, many techniques tried to consolidate pedagogical aspects with heuristics related to interfaces (Mtebe and Kissaka, 2015). Nevertheless, there is still a lack of a widespread culture of usability in the e-learning field, besides most of the studies are at initial stages and needs improvements or deeper empirical studies (Granić and Ćukušić, 2011; Mtebe and Kissaka, 2015). This need for further studies and the importance of usability and UX in the e-learning process motivated us to perform this systematic mapping to determine what usability and UX evaluation techniques have been applied on LMSs.

The goal of this paper is to describe a Systematic Mapping of studies related to usability and UX evaluation techniques in the context of LMSs. The objective was to identify the techniques employed in the published studies. Moreover, the intention was to identify the characteristics of these techniques, such as its origin, type and performing method. With this work we present conclusions regarding the state of the art in this field and contribute to the development/improvement of the usability and UX evaluation techniques in the context of LMSs.

The remainder of this paper is organized as follows: Section 2 presents related work. Section 3 describes the methodology applied to conduct this Systematic Mapping. Section 4 presents the results of the mapping study. Section 5 presents the threats to validity. Section 6 presents the discussion of our results. Finally, Section 7 concludes the paper.

2 RELATED WORK

This section presents the literature reviews that were conducted concerning to the usability and/or UX evaluation of Learning Management Systems.

Freire *et al.* (2012) presented a review about the relationship between ergonomics and usability in e-learning context. There are no details whether the research was conducted in a systematic way or in a narrative way. The analysis of the results allowed the authors to identify three main differences among these usability evaluation methods. Most of the methods were proposed from general (non-educational) contexts. Others were adapted from already consolidated evaluation methods (such as usability tests) with semi-structured questionnaires and interviews. There were also methods that were composed by a mix between methods, methodologies and techniques from the fields of Participatory Design, Interaction Design and User-Centered Design. Methods such as Cognitive Walkthrough were oriented to the socio-cultural aspects of the system and the user's profile. The authors concluded that knowing how to integrate the most pertinent methods to each type of evaluation and each kind of stakeholder is the most important issue to evaluate the usability of LMSs.

Cota *et al.* (2014) conducted a systematic mapping of the literature, aiming to evaluate and interpret all relevant publications related to mobile usability and m-learning (mobile learning), with emphasis on mobile devices like smartphones and tablets. The authors divided these publications in four categories: (i) m-learning applications, (ii) guidelines and frameworks, (iii) specific aspects of m-learning, and (iv) analysis and tendencies of m-learning. Based in this classification, the authors concluded that there were no guidelines, frameworks or tools that evaluate educational factors and usability in m-learning applications. Therefore, they proposed an initial model to develop and evaluate m-learning applications, considering learning aspects and the student experience.

The research conducted by Navarro *et al.* (2016) was an update of the systematic mapping mentioned previously. Only two frameworks developed to evaluate mobile devices regarding m-learning were found. The authors refined the proposed framework, detailing its factors and subdivisions. The framework is divided in two categories: pedagogical usability and user interface usability. The pedagogical usability is concerned to factors that ease and support teaching and learning activities, and is divided in five subcategories: (i) content, (ii) multimedia, (iii) tasks or activities, (iv) social interaction, and (v) personalization. The user interface usability is concerned to the easiness of the interface to be learned, recognized and remembered by the users. This category is also divided in five subcategories: (i) design, (ii) navigation, (iii) customization, (iv) feedback, and (v) motivation. The authors intend to develop a web tool to test the m-learning application through a questionnaire, aiming to score, compare and detect the possible improvements of the application.

In summary, we did not find a systematic review regarding the evaluation of both usability and UX of LMSs. Therefore, our contribution with this systematic mapping is to report, classify and describe the features of the techniques used to evaluate the usability and UX of LMSs for both desktop/Web and mobile context.

3 RESEARCH METHOD

A systematic mapping is a method used to identify, evaluate and interpret all relevant publications on a particular research question, under a rigorous and well defined methodology. It allows obtaining less biased results and broader information about a variety of empirical methods, requiring, on the other hand, more effort from the researcher (Kitchenham and Charters, 2007). The following subsections detail the activities concerning the planning and conducting stages of this systematic mapping defined in our review protocol.

3.1 Research Question

The goal of this research was to analyze the publications regarding usability and UX evaluation techniques in the context of LMSs. Our main research question was: *"Which usability and UX evaluation techniques were applied on Learning Management Systems and how have they been used?"* Sub-questions were defined in order to answer specific questions about each study and technique (see Table 1).

3.2 Search Strategy

In order to construct the search string, we defined the search terms based on the procedure described by Kitchenham and Charters (2007), who suggested defining the parameters for Population, Intervention, Comparison, Outcome and Context (PICOC). The population were Learning Management Systems; the Intervention was composed by techniques used to evaluate the usability and User Experience of LMSs; the Comparison was not applicable, since our goal was to characterize these techniques; the Outcome was the evaluation of usability and UX of LMSs; the Context was not applicable, since there is no comparison to determine the context.

The search terms were divided in two groups. The first group was related to the different spellings and synonyms of LMSs, while the second group was related to the different types of terms used for usability and UX evaluation. These terms were identified based on the reference set of relevant articles and the knowledge of experts in this context. The reference set of articles was composed by the following publications: Freire *et al.* (2012), Lanzilotti *et al.* (2011), Theng and Sin (2012), Kakasevski *et al.* (2008), Zaharias and Koutsabasis (2012) and Medina-Flores and Morales-Gamboa (2015).

Table 1: Systematic mapping sub-questions.

#	Description
SQ1	What is the origin of the technique? If new, what is the difference from the other existing techniques?
SQ2	What is the type of the technique?
SQ3	How is the technique performed?
SQ4	Does the technique consider learning specific factors?
SQ5	Does the technique consider usability, UX or usability and UX?
SQ6	Does the technique provide some kind of feedback to the evaluator?
SQ7	Was the technique empirically evaluated? If affirmative, what studies were performed?
SQ8	Does the technique have any kind of restriction/condition to perform?
SQ9	Is the technique available to download/consultation? Where?
SQ10	In what kind of platform was the study performed?
SQ11	Does the study perform a comparison between techniques?

The string was constructed using the Boolean OR between the alternative spellings and synonyms, and the Boolean AND to join these two groups. The string was tested in several runs to reduce the amount of noise in the results, while ensuring that the reference set was returned.

The search string is presented in Table 2 and was used on the Scopus[1] and Engineering Village[22] digital libraries. These two are meta-libraries that index publications from several well-known publishers such as ACM, IEEE, Springer and Elsevier, besides allowing defining filters such as document type, language and knowledge area.

3.3 Publication Selection Process

Two steps called filters composed the publication selection process. In the first filter, the researchers read only the title and abstract to select the publications related to the evaluation of usability and/or UX in the context of LMSs, applying the inclusion and exclusion criteria (see Table 3). In the second filter, we did the complete reading of the selected publications. The publications were selected according to the same criteria used in the first filter.

[1]http://www.scopus.com
[2]http://www.engineeringvillage.com

Table 2: Search string used in the systematic mapping.

("learning management system" **OR** "LMS" **OR** "online education platform*" **OR** "online education system*" **OR** "online education environment*" **OR** "e-learning environment*" **OR** "e-learning system*" **OR** "e-learning course*" **OR** "e-learning platform*" **OR** "e-learning application*" **OR** "e-learning course*" **OR** "distance learning system*" **OR** "distance learning platform*" **OR** "distance learning environment*" **OR** "distance education platform*" **OR** "distance education system*" **OR** "distance education environment*" **OR** "online learning platform*" **OR** "online learning system*" **OR** "online learning environment*" **OR** "virtual learning environment*" **OR** "LMS" **OR** "managed learning environment*" **OR** "MLE" **OR** "course management system*" **OR** "web-based learning system*" **OR** "remote learning system*" **OR** "m-learning system*" **OR** "m-learning platform*" **OR** "m-learning environment*" **OR** "mobile learning system*" **OR** "mobile learning platform*" **OR** "mobile learning environment*" **OR** "b-learning" **OR** "blended learning" **OR** "learning content management system*" **OR** "LCMS")*
AND
("usability evaluat" **OR** "usability assessment" **OR** "usability inspection" **OR** "usability improvement*" **OR** "usability test*" **OR** "usability technique*" **OR** "usability guideline*" **OR** "UX evaluat*" **OR** "UX assessment" **OR** "UX improvement*" **OR** "UX technique*" **OR** "UX guideline*" **OR** "user experience evaluat*" **OR** "user experience assessment" **OR** "user experience improvement*" **OR** "user experience technique*" **OR** "user experience guideline*")*

Regarding the exclusion criteria, some considerations must be done. We did not consider publications related to the evaluation of MOOCs - Massive Open Online Courses (EC2), since they are not LMSs, but online courses aimed for massiveness and openness (Pireva *et al.*, 2015). Publications related to augmented reality (EC3) or specifically to accessibility aspects (EC4) were also unconsidered. The former is out of the scope of this work, while the last does not consider the usability as a whole.

Table 3: Inclusion and exclusion set of criteria.

#	**Inclusion Criteria**
IC1	Publications that describes the utilization of the usability or UX evaluation techniques in LMSs
#	**Exclusion Criteria**
EC1	Publications not related to the evaluation of usability or UX in LMSs
EC2	Publications related to the evaluation of MOOCs
EC3	Publications related to the evaluation of augmented reality systems
EC4	Publications related specifically to accessibility questions
EC5	Publications in which the language is different from English and Portuguese
EC6	Publications that are not available for reading or data collection (publications that are only accessible through payment or are not provided by the search engine)
EC7	Duplicated publications

The publications that presented an empirical study were assessed regarding its quality. This type of assessment is important, especially for studies that present contradictory results (Kitchenham and Charters, 2007). The studies were assessed through a questionnaire composed by three questions regarding: (i) the detailing of the employed techniques; (ii) the detailing of the process regarding its application, and (iii) the clarity of the empirical results. Each question was composed by three answers: (i) Disagree; (ii) Partially agree, and (iii) Totally agree. The overall quality of the publications ranks from 1.0 (low quality) to 3.0 (high quality).

3.4 Data Extraction Strategy

This section details the data extraction process. The following information was extracted from each of the selected publications:

Regarding SQ1 (*Technique Origin*), the technique was classified as: **a) New**, if it was developed or adapted specifically to the e-learning context or **b) Existing**, if it already exists from other HCI contexts and is used as it is.

SQ2 (*Technique Type*) is related to the classification of the techniques, according to the taxonomy proposed by Ivory and Hearst (2001): **a) Inspection**, when the evaluator uses a set of criteria to identify potential usability problems; **b) Testing**, when the evaluator observes a participant interacting with an interface in order to identify usability problems; **c) Inquiry**, when the user provides a feedback about an interface through interviews, surveys and the like; **d) Analytical Modeling**, when the evaluator employs different kind of models to generate predictions of usability problems; **e) Simulation**, when the evaluator simulates an user interaction within an interface by employing any kind of simulation algorithm.

Concerning SQ3 (*Performing Method*), the technique can be classified as: **a) Manual**, when the technique is performed manually; **b) Semi-automatic**, when part of the evaluation is performed manually and another is performed through automatized tools; **c) Automatic**, when almost all

the evaluation process is performed through an automatized tool, requiring the evaluator to just interpret the results.

Regarding SQ4 (*Learning Factors*), the goal is to identify whether the technique considers specific learning factors, for example, content relevance, learner control, instructional feedback and so forth.

SQ5 (*Evaluation Focus*) aims to identify whether the technique considers only usability factors (e.g.: navigation and feedback), only UX factors (e.g.: user's emotions/feelings regarding the interaction with the platform) or both usability and UX factors.

Concerning SQ6 (*Feedback*), the goal is to identify whether the technique provides a feedback about how to correct the identified usability problems or just identifies them.

SQ7 (*Investigation Type*) verifies whether or not the study was empirically evaluated. The studies can be classified according to three main investigation strategies described by Wohlin *et al.* (2012): **a) Survey**, if the study is performed in retrospect, when a tool or technique has been used for a while; **b) Case Study**, if the study is observational, aiming to track a specific attribute or establish relations between different attributes; **c) Controlled Experiment**, if the study performs an empirical investigation that manipulates one or more variables or factors in the studied context, verifying the effects of this manipulation.

Regarding SQ8 (*Restriction*), we want to identify whether or not the technique have any kind of restriction or condition for being applied, for example, a specific tool or equipment or a need for specific knowledge.

Concerning SQ9 (*Availability*), the goal is to identify whether or not the complete technique (questionnaires, processes, etc.) is available for download or consultation in the publication itself or in some external resource (e.g.: websites).

SQ10 (*Platform Used*) aims to identify whether the study was conducted in Desktop/Web, Mobile or both Desktop/Web and Mobile context.

Regarding SQ11 (*Techniques Comparison*), the goal is to identify whether the study conducted a comparison between techniques.

4 RESULTS

In order to avoid the bias of a single researcher, the systematic mapping involved two researchers. One researcher specified the review protocol, which was reviewed by the second researcher.

For the first step, the researchers independently classified a sample of 17 randomly selected publications based on the selection criteria. The agreement between the researchers was evaluated by the Kappa statistical test (Cohen, 1960). The result of this evaluation showed a substantial strength of agreement between the two researchers (*kappa = 0.610*) according to the range described by Landis and Koch (1977).

Details of this systematic mapping can be found on technical report (Nakamura *et al.*, 2017).

4.1 Selected Publications

The search string returned a total of 177 publications in the Scopus library and 13 in the Engineering Village library (see Figure 1). Duplicated publications were found during the process. In these cases, the publications were accounted for only once, selecting the most complete of them. After removing the duplicated publications, the number of selected publications for the first filter was 175. Out of these 175 publications, 33 were rejected in the first filter, since they did not meet the inclusion criteria. The remaining 142 publications were fully read and classified in the second filter, according to the criteria. At the end of the process, 62 publications were accepted and extracted.

The selected publications were published between 2004 and August 2016. The graph presented in Figure 2 shows a variation of the number of publications related to the evaluation of usability/UX in the context of Learning Management Systems during this period, with peaks of publications in 2008 and 2009 with 7 publications, 2011 with 10 publications and 2015, the year with the highest number of publications, with 11 publications. Since this systematic mapping was done in September 2016, the data referring to this year are incomplete, which can possibly explain

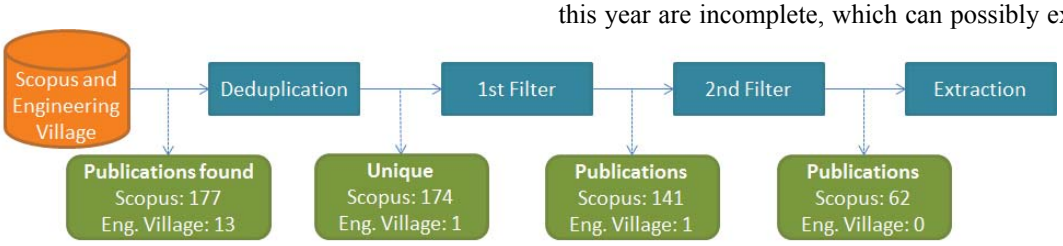

Figure 1: Publications selection process.

the low rate of publications for this year.

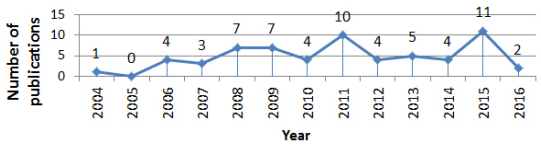

Figure 2: Frequency of publications by year.

4.2 Results Overview

An overview of the results is presented in Table 4. Although 62 publications were selected, the sub-questions regarding the studies counted only 58 publications. The studies from Freire *et al.* (2012) and Cota *et al.* (2014) were not considered in this counting, since they presented literature reviews, not being possible to extract their data as the techniques presented in other publications. For these cases a different extraction form was used, which can be found on technical report (Nakamura *et al.*, 2017). Although the publication of Navarro *et al.* (2016) presented a systematic review, this study was not excluded from the counting, since it proposed a technique based on the results of the systematic mapping. The publication from Zaharias (2006) and Zaharias (2008) were not included, considering that they presented an initial evolution stage of an already classified technique or a partial analysis of a study already done.

Regarding the classified techniques, some publications presented more than one technique, each of them being counted separately. Thus, the total number of techniques presented in the 58 publications was 104.

In SQ7, SQ10 and SQ11, given that they are related to the investigation strategy and platform of the study respectively, the publications were considered as a whole and not each technique individually.

Next sections will present the analysis of the results obtained through each research sub-question.

4.3 Technique Origin

The results of the SQ1 show that most of the techniques used in the e-learning context (67,31%) were originated from other areas of HCI. For example, Blecken *et al.* (2010) employed the Think-Aloud Protocol and SUS (System Usability Scale) questionnaire to perform a usability evaluation of a LMS called koaLA.

Table 4: Overall results for each research sub-question.

Sub-question	Answer	Qty.	%
SQ1. Technique origin	New	34	32,69%
	Existent	70	67,31%
SQ2. Technique type	Inspection	29	27,88%
	Testing	35	33,65%
	Inquiry	54	51,92%
	Analytical Modeling	1	0,96%
	Simulation	1	0,96%
SQ3. Performing method	Manual	94	90,38%
	Semi-automatic	3	2,88%
	Automatic	7	6,73%
SQ4. Learning factors	Yes	29	85,29%
	Not specified	5	14,71%
SQ5. Evaluation focus	Usability	72	69,23%
	UX	2	1,92%
	Usability and UX	30	28,85%
SQ6. Feedback	Yes	0	0
	No	104	100%
SQ7. Investigation type	Survey	15	25,86%
	Case Study	27	46,55%
	Controlled Experiment	10	17,24%
	No	7	12,07%
SQ8. Restriction	Yes	12	11,54%
	No	92	88,46%
SQ9. Availability	Yes	69	66,35%
	No	35	33,65%
SQ10. Platform used	Desktop/Web	46	79,31%
	Mobile	10	17,24%
	Desktop/Web and Mobile	2	3,45%
SQ11. Techniques comparison	Yes	6	10,34%
	No	52	89,66%

Around 32,69% of the techniques are new, specific to the context of e-learning. Most of them use some kind of questionnaire, such as proposed by Ssekakubo *et al.* (2014). The authors proposed a technique that involves instructional usability and motivation factors, besides the conventional usability factors.

4.4 Technique Type

The results of the SQ2 reveal that *Inquiry* was the most employed type of technique, with 51,92% of the techniques. Some of the most applied techniques were questionnaires, focus groups and interviews.

- *Questionnaires:* users provide answers to specific questions. Zaharias and Poylymenakou (2009) developed and empirically evaluated a questionnaire that considers the usability of the

platform and the instructional design, focusing on motivation to learn;

- *Focus Groups:* multiple users attend a discussion session coordinated by a moderator. An example of this technique can be found in the study conducted by Tee (2013), who employed open questions to get opinions from two groups (students and lecturers) about the interface usability of Moodle platform;
- *Interviews:* one or more users attend to a discussion session, where specific questions are asked to the participants. Santoso *et al.* (2014) conducted semi-structured interviews based on the Shneiderman's Eight Golden Rules (Shneiderman and Plaisant, 2010) to get the perceptions of the students about the usability of an e-learning platform called SCELE (Student Centered E-Learning Environment).

The second most employed type of technique was *Testing*, in other words, about 33,65% of the techniques employed some type of test involving users. Some of the most representative techniques were: performance measure, think-aloud protocol and log file analysis.

- *Performance Measure:* quantitative data, such as task completing time and number of errors, are collected during the test. Stickel *et al.* (2008) conducted a test using a technique called NPL Performance Measurement Method to calculate the metrics related to task efficacy, user efficacy and user relative efficacy;
- *Think-Aloud Protocol:* participants are encouraged to verbalize their thoughts during the test. Gordillo *et al.* (2014) applied the think-aloud protocol together with PrEmo, a self-evaluation instrument with a scale of emotions, composed by 7 pleasant emotions and 7 unpleasant emotions;
- *Log File Analysis:* evaluator analyses the logs containing collected and recorded usage data. De Kock *et al.* (2009) conducted a usability test with users supported by eye tracking, in order to detect participant's visual fixation points and identify possible usability problems.

Around 27,88% of the techniques were of *Inspection* type. The most representative techniques were: heuristic evaluation, pattern-based evaluation and checklist-based evaluation.

- *Heuristic Evaluation:* evaluators verify the conformity of the application according to a set of guidelines. Ssemugabi and De Villiers (2007) proposed a heuristic evaluation technique that considers the learning and teaching factors, divided in three categories. The first one is composed by the Nielsen's 10 heuristics (Nielsen, 1994) adapted to e-learning context. The second is composed by criteria for educational websites. The third contains learner-centred instructional design criteria;
- *Pattern-Based Evaluation:* inspectors perform a usability evaluation according to a set of predefined patterns that indicates the places to look for the problems and the actions that have to be taken to analyze these aspects. Ardito *et al.* (2006) applied a methodology called SUE (Systematic Usability Evaluation) to evaluate the usability of e-learning applications. This methodology consists in the definition of Abstract Tasks through a predefined template, which will be used to conduct a usability inspection and a task-based user testing;
- *Checklist-Based Evaluation:* heuristics are simplified in a questionnaire with elements that must be scored by an evaluator. Oztekin *et al.* (2010) proposed a technique called *UseLearn*, which is composed by a checklist related to 12 usability dimensions, such as visibility, aesthetics, and flexibility. The technique allows generating an overall usability index from the calculation of the weight of each attribute.

Regarding the *Analytic Modeling*, there was only one technique identified (0,96%), which was classified as of Design Analysis type.

- *Design Analysis:* allows the evaluator to represent a user interface in multiple levels of abstraction and evaluate its representation, being typically used to specify the UI design before its implementation. Rodrigues *et al.* (2011) utilized Markov Models to calculate the probability of change between states, allowing verifying whether the user can go to another state or whether will be stuck in it.

Since Markov Models simulates the user's interaction process between different navigation pages, this technique was also classified as *Simulation* type, being the unique technique (0,96%) classified in this category.

4.5 Performing Method

The results of the SQ3 show that the majority of the techniques are performed manually, representing 90,38% of the techniques, followed by automatic techniques, that represent about 6,73% and semi-automatic 2,88%.

Regarding the automatically performed techniques, the study conducted by Stickel *et al.* (2008) aimed to evaluate the usability and UX of an

LMS using psychophysiological methods to detect user-hostile systems, collecting data through Electroencephalograms, Heart Rate and Skin Conductance Level.

An example of semi-automatically performed technique can be found in the study conducted by Oztekin *et al.* (2013). First, the usability evaluation is conducted manually through a form using the *UseLearn* technique. Then, an analysis of this evaluation is performed automatically through a machine-learning program using artificial neural networks. The program calculates the usability problems that produce the highest impact on overall system usability and presents them to the evaluator.

4.6 Learning Factors

The results of the SQ4 reveal that about 85,29% of the new techniques consider learning factors. There was no consensus, however, about which learning factors should be evaluated. We identified 30 different learning factors. The most evaluated factor was "Content Relevance", with 50,00%, followed by "Interaction between participants", with 46,15%. "Feedback and Orientation" and "Instructional Assessment" were considered by 42,31% of the techniques employed. "Content organization and structure" accounted for 38,46%, followed by "Motivation" and "Support for Significant Learning Approach" with 30,77%. "Media Use" and "Collaborative Learning" were considered by 26,92% of the techniques. Other factors were considered by less than 25% of the techniques.

About 14,71% of the techniques did not specify whether they consider learning specific factors or not.

4.7 Evaluation Focus

The results of SQ5 show that about 28,85% of the techniques addressed usability and UX factors in the evaluation process. Navarro *et al.* (2016) proposed a framework to evaluate m-learning applications, considering pedagogical factors (e.g.: content, organization and objectives) and the motivation related to the affective factor.

Techniques that only evaluate usability factors accounted for about 69,23% of the techniques. Yusoff and Zin (2014) developed a usability evaluation questionnaire composed by 10 questions related to aspects such as ease of use, ease of navigation and class learning support.

Only 1,92% of the techniques were specific to evaluate the UX. However, all of them were generic techniques, not specific for the learning context. Santoso *et al.* (2014) used the UEQ (User Experience Questionnaire) to evaluate the User Experience of a platform called SCELE. The UEQ consists of 26 items that measures factors such as attractiveness, stimulation and novelty.

4.8 Feedback

The results of SQ6 show that none of the techniques provides a feedback with suggestions about how to correct the identified usability problems. All analyzed techniques just perform an identification of these problems.

4.9 Investigation Type

Concerning to the investigation type, we classified the empirical studies according to the three main investigation strategies described by Wohlin *et al.* (2012). The results of SQ7 show that the majority of the studies used case studies (according to the authors), representing 46,55% of them. Hamdi *et al.* (2011), for example, performed a case study to evaluate the usability of a m-learning application at the Arab Academy for Science and Technology (AASTMT). The study was conducted through a task-based user testing using the Think-Aloud Protocol and USE (Usefulness, Satisfaction and Ease of Use) questionnaire to analyze the user's efficacy, efficiency and satisfaction.

Survey was the second most used type of investigation, with about 25,86% of publications. Alkhattabi (2015), for example, performed a survey through an own questionnaire, composed by a 5-point Likert scale (strongly agree, agree, neutral, disagree and strongly disagree) and a field for comments, in order to analyze the usability, practicality, pedagogical efficacy and overall design of an LMS called Tadarus.

About 17,24% performed controlled experiments. Zaharias and Koutsabasis (2012) performed an experiment to compare two usability inspection techniques specifically developed to e-learning context: the Mehlenbacher *et al.* (2005) heuristics and Reeves *et al.* (2002) heuristics. The results showed that both heuristics have good coverage of the usability problems, however, they did not present a satisfactory distribution of problems. Some heuristics identified several problems, while others identified just some of them or even none. The authors also stated that there is a need for better orientation and organization of the

heuristics regarding its level of abstraction.

The publications that did not present empirical studies represented about 12,07%.

4.10 Technique Restriction

The results of SQ8 show that 11,54% of the techniques have some type of restriction for its application. These restrictions can be, for example, a specific knowledge, software or equipment. The *UseLearn* technique proposed by Oztekin *et al.* (2010) needs that the usability analyst has knowledge in structural equation modeling-based criticality metric analysis, which limits its application.

The other techniques, which do not have restrictions for utilization, represent 88,46%.

4.11 Technique Availability

The results of SQ9 show that 66,35% of the techniques are available, that is, they provide the questionnaires and/or procedures for carrying out the usability/UX evaluation in the article itself or in external sources (other publications or websites, for example). Junus *et al.* (2015), for example, provided the questionnaires in the article itself.

The techniques that are not available for download/consultation accounted for 33,65%.

4.12 Platform Used

The results of SQ10 show that the majority of the studies were performed in the context of desktop platforms, representing about 79,31%. Orfanou *et al.* (2015) performed a usability evaluation of the Moodle platform and Eclass platform (based on LMS Claroline). The evaluation was performed through the application of the SUS questionnaire. The authors performed a validity evaluation of the questionnaire and its correlation with several factors like genre, age and frequency of use of LMSs.

Studies in the mobile context represented about 17,24% of the publications. Fetaji *et al.* (2008) evaluated a mobile application called *MobileView* by utilizing a usability evaluation framework developed by the authors themselves, called MLUAT (Mobile Learning Usability Attribute Testing). This framework uses heuristic evaluation, usability testing and questionnaires to evaluate the usability of m-learning applications.

About 3,45% of the studies were carried out both in mobile and desktop context, making comparisons between the usability problems related to the interaction with the interfaces. Sánchez-Chamochin *et al.* (2008), for example, performed a usability test of a mobile version for a legacy e-learning platform. The users performed the tasks both in the desktop and the mobile device, in order to get, for example, the relation between the number of steps needed to perform a task between the two type of devices.

4.13 Techniques Comparison

Although several techniques were identified, only 6 publications (10,34%) presented a comparison between them. Lanzilotti *et al.* (2011), for example, performed a comparison between the HE of Squires and Preece (1999), Think-Aloud Protocol and Pattern-Based Evaluation. The authors concluded that the Pattern-Based Evaluation identified a higher number of different types of problems, provided more consistent and reliable results and had a better cost/benefit. However, it tended to induce to an overestimation of the problem's severity and did not reach a higher perceived value compared to the other techniques.

The publications that did not perform comparison of techniques represented 89,66%.

5 THREATS TO VALIDITY

Although we conducted this research under a systematic mapping methodology by defining a research protocol, some threats to validity can be identified: (i) the researcher's bias regarding the analysis of the primary studies; (ii) the university's limited access to some scientific databases, which can prevent some publications to being accessed; (iii) the limitation of the scope of this research to the two selected databases.

These threats were minimized by taking some actions. For the first threat, we reviewed the review protocol and conducted the Kohen's Kappa statistical test in order to reduce the researcher's bias. Additionally, other experienced researcher reviewed the execution process. For the second threat, we asked for the full publication to the author whenever it was possible and included those that were made available. Regarding the third threat, although the research had been conducted on only two databases, these databases indexes publications from a large number of very known publishers, journals and conferences, which can possibly reduce the number of publications that was not addressed by this research.

6 DISCUSSION

This systematic mapping aimed to identify the usability and UX evaluation techniques in the context of Learning Management Systems. The results showed that there are several studies regarding the techniques used to evaluate the usability and UX of LMSs. However, there are still some gaps that can be explored by further studies:

- We did not find evidences of techniques that provide a feedback with suggestions to correct the usability/UX problems, being limited to just identify them;
- There was no consensus about which learning factors must be considered in the evaluation process;
- Few studies were conducted in the m-learning context. Among the 12 studies that were found, only 5 of them presented a new technique. Altogether, 4 different frameworks were identified among them and just one was empirically evaluated;
- Only 6 studies presented a comparison between the different techniques, which difficult the definition of the most adequate technique for the e-learning context. Another limiting factor is the unavailability of the techniques in a considerably number of studies, which makes it difficult to perform this kind of comparison;
- Few techniques performs the usability/UX evaluation process in an automated way;
- Among the new techniques, none of them considered the UX evaluation with greater depth. Some studies considered the motivation factor, however, they were limited to this scope. Other techniques involved aspects related to the user's satisfaction only. Although some of the UX aspects could be captured by these techniques, none of them is concerned about, for example, the user's feelings and its change over the time;
- There were no studies relating the influence of the usability/UX improvement in the learning process.

Thus, despite the large number of techniques employed, there is still room for more research in this area. These identified gaps may contribute to give an insight to further research in order to improve the quality of the evaluation techniques and of the LMSs.

7 CONCLUSIONS

In this systematic mapping, we analyzed the publications regarding the usability/UX evaluation of Learning Management Systems. From a starting set of 190 publications, a total of 62 publications were selected and a total of 104 employed techniques were identified.

Although several studies had been conducted regarding the evaluation of LMSs, the results of this systematic mapping revealed that there is still a need for more research in this area. Among the identified techniques, for example, none of them provided a feedback with suggestions to correct the identified problems. Besides, even though many techniques had been developed to evaluate the usability/UX of LMSs, there is no sufficient evidence of which of them is best suited for this context. Thus, the identified gaps in this systematic mapping may be a starting point to other researchers.

We hope that our findings may contribute to the development and improvement of the usability and UX evaluation techniques for the e-learning context, hence the improvement of the LMSs.

ACKNOWLEDGEMENTS

We would like to thank the Brazilian funding agency FAPEAM through process number 062.02646/2014 and the Usability and Software Engineering Group (USES) of the Federal University of Amazonas (UFAM).

REFERENCES

Alkhattabi, M., 2015. Usability evaluation of Tadarus: Student perceptions. *Procedia Computer Science Journal*, 62(1), pp. 631-636.

Ardito, C., Costabile, M. F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V., 2006. An approach to usability evaluation of e-learning applications. *Universal Access in the Information Society Journal*, 4(3), pp. 270-283.

Blecken, A., Brüggemann, D., Marx, W., 2010. Usability evaluation of a Learning Management System. In *Proceedings of the 43rd Hawaii International Conference on System Sciences*. Washington DC, USA: IEEE, pp. 1-9.

Brusilovsky, P., 2004. KnowledgeTree: A distributed architecture for adaptive e-learning. In *Proceedings of the 13th International World Wide Web conference on Alternate track papers & posters (WWW Alt. '04)*. New York, USA: ACM Press.

Cohen, J., 1960. A coefficient of agreement of nominal scales. *Educational and Psychological Measurement*, 20(1), pp.37–46.

Cota, C. X. N., Díaz, A. I. M., Duque, M. Á. R., 2014. Evaluation framework for m-learning systems: Current situation and proposal. In *Proceedings of the XV International Conference on Human Computer Interaction (INTERACCION 2014)*. New York, USA: ACM Press.

De Carvalho, A. F. P., Anacleto, J. C., 2008. The importance of usability criteria on Learning Management Systems: Lessons learned. In *Proceedings of the 10th International Conference on Enterprise Information Systems (ICEIS 2008)*. Barcelona, Spain: SciTePress, pp. 154-159.

De Kock, E., Van Biljon, J., Pretorius, M., 2009. Usability evaluation methods: Mind the gaps. In *Proceedings of the 2009 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*. Vaal River, South Africa: ACM Press, pp. 122-131.

Dubost, A., Klein, M. R., Dang, J., 2004. Building interoperability between LMS and Brokerage Platforms. In *Proceedings of the E-ACTIVITIES conference (WSEAS 2004)*. Crete, Greece.

Fetaji, M., Dika, Z., Fetaji, B., 2008. Usability testing and evaluation of a mobile software solution: A case study. In *Proceedings of the ITI 2008 30th International Conference on Information Technology Interfaces*. Cavtat, Croatia: IEEE, pp. 501-506.

Freire, L. L., Arezes, P. M., Campos, J. C., 2012. A literature review about usability evaluation methods for e-learning platforms. In *18th World congress on Ergonomics - Designing a sustainable future (IEA 2012)*. Recife, Brazil: IOS Press, pp. 1038-1044.

Gordillo, A., Barra, E., Aguirre, S., Quemada, J., 2014. The usefulness of Usability and User Experience evaluation methods on an e-Learning platform development from a developer's perspective: A case study. In *Frontiers in Education Conference (FIE)* 2014. Madrid, Spain: IEEE.

Granić, A., Ćukušić, M., 2011. Usability testing and expert inspections complemented by educational evaluation: A case study of an e-learning platform. *Journal of Educational Technology & Society*, 14(2), pp. 107-123.

Hamdi, N., Elbadrawy, R., Aziz, R. A., 2011. An experimental study to measure mobile learning system usability at the AASTMT. In *International Conference Mobile Learning (IADIS 2011)*. Avila, Spain: IADIS Press, pp. 167-174.

Hassenzahl, M., Platz, A., Burmester, M., Lehner, K., 2000. Hedonic and ergonomic quality aspects determine a software's appeal. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '00)*. The Hague, Holland: ACM Press, pp. 201-208.

Hassenzahl, M., Law, E. L-C., Hvannberg, E. T., 2006. User Experience – Towards a unified view. In: User Experience – Towards a Unified View: Second International COST294-MAUSE Open Workshop – NordICHI'06, pp. 1-3.

Hijon-Neira, R., Velázquez-Iturbide, Á., Pizarro-Romero, C., Carriço, L., 2014. Merlin-know, an interactive virtual teacher for improving learning in Moodle. In *Proceedings of Frontiers in Education Conference (FIE 2015)*. Madrid, Spain: IEEE.

Ivory, M. Y., Hearst, M. A., 2001. The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, 33(4), pp. 470-516.

Junus, I. S., Santoso, H. B., Isal, R. Y. K., Utomo, A. Y., 2015. Usability evaluation of the student centered e-learning environment. *International Review of Research in Open and Distributed Learning*, 16(4), pp. 62-82.

Kakasevski, G., Mihajlov, M., Arsenovski, S., Chungurski. S., 2008. Evaluating usability in Learning Management System Moodle. In *30th International Conference on Information Technology Interfaces (ITI 2008)*. Cavtat, Croatia: IEEE, pp. 613-618.

Kitchenham, B., Chartes, S., 2007. Guidelines for performing systematic literature reviews in software engineering. *Technical Report (EBSE 2007)*. Durham, UK: Keele.

Landis, J.R. & Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), pp. 159-174.

Lanzilotti, R., Ardito, C., Costabile, M. F., 2006. eLSE methodology: A systematic approach to the eLearning systems evaluation. *Journal of Educational Technology & Society*, 9(4), pp. 42-53.

Lanzilotti, R., Ardito, C., Costabile, M. F., De Angeli, A., 2011. Do patterns help novice evaluators? A comparative study. *International Journal of Human-Computer Studies*, 69(1), pp. 52-69.

Medina-Flores, R., Morales-Gamboa, R., 2015. Usability evaluation by experts of a Learning Management System. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 10(4), pp. 197-2013.

Mehlenbacher, B., Bennet, L., Bird, T., Ivey, M., Lucas, J., Morton, J., Whitman, L., 2005. Usable e-learning: A conceptual model for evaluation and design. In *Proceedings of HCI International 2005: 11th International Conference on Human-Computer Interaction*, 4(1), pp. 1-10.

Mtebe, J. S., Kissaka, M. M., 2015. Heuristics for evaluating usability of Learning Management Systems in Africa. In *IST-Africa 2015 Conference Proceedings*. Lilongwe, Malawi: IEEE.

Nakamura, W. T., Oliveira, E. H. T., Conte, T., 2017. A systematic mapping on usability and user experience evaluation of Learning Management Systems, TR-USES-2017-0006. Available online at: http://uses.icomp.ufam.edu.br/relatorios-tecnicos/

Navarro, C. X., Molina, A. I., Redondo, M. Á., Juárez-Ramírez, R., 2016. Framework to evaluate m-learning systems: A technological and pedagogical approach. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 11(1), pp. 33-40.

Nielsen, J., (1994). Heuristic evaluation. In *Usability Inspection Methods*. New York, USA: John Wiley & Sons.

Orfanou, K., Tselios, N., Katsanos, C., 2015. Perceived usability evaluation of Learning Management Systems: Empirical evaluation of the System Usability Scale. *International Review of Research in Open and Distributed Learning*, 16(2), pp. 227-246.

Oztekin, A., Kong, Z. J., Uysal, O., 2010. UseLearn: A novel checklist and usability evaluation method for eLearning systems by criticality metric analysis. *International Journal of Industrial Ergonomics*, 40(4), pp. 455-469.

Oztekin, A., Delen, D., Turkyilmaz, A., Zaim, S., 2013. A machine learning-based usability evaluation method for eLearning systems. *Decision Support Systems Journal*, 56(1), pp. 63-73.

Pireva, K., Imran, A. S., Dalipi, F., 2015. User behaviour analysis on LMS and MOOC. *IEEE Conference on e-Learning, e-Management and e-Services (IC3e 2015)*. Malacca, Malaysia: IEEE, pp. 21-26.

Reeves, T. C., Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, C. S., 2002. Usability and instructional design heuristics for e-learning evaluation. In *ED-MEDIA 2002 World Conference on Educational Multimedia, Hypermedia & Telecommunications*. Denver, USA: AACE, pp. 1615-1621.

Rodrigues, J., Diniz, M., Ferreira, S. B. L., Silveira, D. S., Capra, E., 2011. Evaluation of usability in a remote learning system utilizing markov models. In *Proceedings of the IADIS International Conference on WWW/Internet*. Rio de Janeiro, Brazil: IADIS Press.

Sánchez-Chamochin, E. L., Cambranes-Martínez, E., Menéndez-Domínguez, V., 2008. Adapting mobile access scheme for a legacy e-learning platform. In *Mexican International Conference on Computer Science*. Mexicali, Mexico: IEEE, pp. 96-103.

Santoso, H. B., Isal, R. Y. K., Basaruddin, T., Sadita, L., 2014. Research-in-progress: User experience evaluation of student centered e-learning environment for computer science program. In *3rd International Conference on User Science and Engineering (i-USEr)*. Sha Alam, Malaysia, pp. 52-55.

Shneiderman, B., Plaisant, C., 2010. Designing the user interface: Strategies for effective Human-Computer Interaction. Boston, MA, USA: Pearson.

Squires, D., Preece, J., 1999. Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers Journal*, 11(1), pp. 467-483.

Ssekakubo, G., Suleman, H., Marsden, G., 2014. A streamlined mobile user-interface for improved access to LMS Services. In *eLmL 2014: Sixth International Conference on Mobile, Hybrid, and On-line Learning*. Barcelona, pp. 92-101.

Ssemugabi, S., De Villiers, R., 2007. A comparative study of two usability evaluation methods using a web-based e-learning application. In *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists*. Port Elizabeth, South Africa: ACM Press, pp. 132-142.

Stickel, C., Scerbakov, A., Kaufmann, T., Ebner, M., 2008. Usability metrics of time and stress - Biological enhanced performance test of a university wide Learning Management System. *HCI and Usability for Education and Work*. Graz, Austria: Springer, pp. 173-184.

Tee, S. S., Wook, T. S. M. T., Zainudin, S., 2013. User testing for Moodle application. *International Journal of Software Engineering and Its Applications*, 7(5), pp. 243-252.

Theng, Y-L., Sin, J., 2012. Evaluating usability and efficaciousness of an e-learning system: A quantitative, model-driven approach. In *12th IEEE International Conference on Advanced Learning Technologies*. Rome, Italy: IEEE, pp. 303-307.

Vogten, H., Koper, R., 2014. Towards a new generation of Learning Management Systems. *Proceedings of the 6th International Conference on Computer Supported Education (CSEDU 2014)*. Barcelona, Spain: SciTePress, pp. 514-519.

Wohlin, C., Höst, M., Regnell, B., 2012. *Experimentation in Software Engineering*. Verlag: Springer.

Yusoff, S. R. M., Zin, N. A. M., 2014. Design and evaluation of Collaborative Learning Management System (CLMS) framework for teaching technical subject. In *Proceedings of Revised Selected Papers on New Horizons in Web Based Learning*, 7.697(1), pp. 79-89.

Zaharias, P., 2006. A usability evaluation method for e-learning: Focus on motivation to learn. In *Proceedings of CHI EA '06 Extended Abstracts on Human Factors in Computing Systems*. New York, USA: ACM Press, pp. 1571-1576.

Zaharias, P., 2008. Cross-cultural differences in perceptions of e-learning usability: An empirical investigation. *International Journal of Technology and Human Interaction*, 4(3), pp. 1-26.

Zaharias, P., Poylymenakou, A., 2009. Developing a usability evaluation method for e-learning applications: Beyond functional usability. *International Journal of Human-Computer Interaction*, 25(1), pp. 75-98.

Zaharias, P., Koutsabasis, P., 2012. Heuristic evaluation of e-learning courses: A comparative analysis of two e-learning heuristic sets. *Campus-Wide Information Systems*. 29(1), pp. 45-60.