PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

TUXEL:
A Technique for User eXperience Evaluation in e-Learning

WALTER TAKASHI NAKAMURA

Manaus
February, 2018

PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

WALTER TAKASHI NAKAMURA

TUXEL:
A Technique for User eXperience Evaluation in e-Learning

Master dissertation presented to the Informatics Postgraduate Program (*Programa de Pós-graduação em Informática*) – PPGI, at Universidade Federal do Amazonas (UFAM) as one of the requirements for achieving the degree of Master in Informatics.

Advisor: Ph.D. Tayana Uchôa Conte
Co-advisor: Ph.D. Elaine H. T. de Oliveira

Manaus
February, 2018

TUXEL:
A Technique for User eXperience Evaluation in e-Learning

Walter Takashi Nakamura

Master dissertation presented to the Informatics Postgraduate Program (Programa de Pós-graduação em Informática) – PPGI, at Federal University of Amazonas (UFAM) as one of the requirements for achieving the degree of Master in Informatics.
Dissertation evaluation on February 22th, 2018 in Manaus (Brazil), by the evaluation committee constituted by the following professors:

_____
Ph.D. Tayana Uchôa Conte

_____
Ph.D. Thaís Helena Chaves de Castro

_____
Ph.D. Isabela Gasparini

Manaus
February, 2018

***To my family**,*

for all the love, encouragement and understanding,

without which I would not be who I am today.

ACKNOWLEDGEMENTS

To God, for granting me the gift of life and for all that He have given me during my walk in this master's program.

To my mother, Cecilia Tokuyo Nakamura, my father, Hiromu Nakamura and my sister, Neiva Saori Nakamura, for all the support and encouragement in starting this journey. You have never been against my quest for this goal, even knowing of the absence, distance and longing during these years. Thank you for always cheering for me and sharing my joys. I owe you everything I am today. Love you!

To my boyfriend Matheus Henrique Nogueira, for the love, support, understanding and companionship throughout this period. Thank you for sharing with me the greatest moments and supporting me in the difficult ones. My journey in Manaus would not be the same without you. I love you!

To my dear advisor, Tayana Uchôa Conte, who from the very first time welcomed me promptly. Thank you for all the teachings and orientations. When I started in this master's program, I had never carried out a research before. I never thought I could be a researcher. However, you always believed in my potential, encouraged me and guided me through the paths that I needed to walk. Thank you for all the valuable knowledge you have taught and all the opportunities you gave me for my personal and professional growth. The researcher that I am today and all of my achievements were thanks to you.

To my dear co-advisor, Elaine Harada Teixeira de Oliveira, for all the support I received since the first contact before the beginning of the master's program and for all the collaboration in my research. Thank you for sharing my achievements and always cheering for me.

To Bruno Gadelha and Leandro Galvão, for all the valuable advices regarding my research since the master's program qualification and for always being willing to collaborate.

To Thaís Castro and Isabela Gasparini for agreeing to participate in my dissertation evaluation committee and taking the time to read my work. All your critics and suggestions certainly helped me to improve my research.

To all from the USES group, for having received and accompanied me during this walk. Especially Bia, Bruna, Carol, Davi, Drica, Jacy, Leo, Lucho and Will. Thank you for sharing knowledge, joys, achievements and experiences.

To Leandro Ramos, for letting me know about this master's program, which led me to start this journey.

To the Nakamura Dojo family, with which I had some of the best moments of my life, especially the senseis Davi Pereira and Wevertton dos Santos for leading the dojo throughout this period.

To all the friends and family who cheered for me, and to all those directly or indirectly collaborated in my research, especially the subjects who participated in the empirical studies.

Finally, to FAPEAM, CAPES and CNPq, for the financial support granted throughout the master's degree.

ABSTRACT

The increasing interest in new learning technologies has led learning institutions and organizations to adopt powerful platforms designed to support the teaching and learning process, called Learning Management Systems (LMSs). The widespread adoption of LMSs raises the need to evaluate the quality in use of these platforms through attributes such as usability and User eXperience (UX). For instance, if the LMS does not provide a good usability, the learner will spend more time trying to understand how to use the platform rather than learning the educational content. Similarly, a positive UX is essential to make the platform more pleasuring, appealing and satisfactory to the learner. Although several studies have been conducted to evaluate LMSs, few techniques encompass both usability and UX. This work presents the development of a technique to evaluate the usability and UX of LMSs, called TUXEL, by means of Design Science Research methodology. Additionally, we developed a tool to support the evaluation process. We conducted two empirical studies in order to assess the feasibility of the technique, while comparing it with an existing technique from the literature. Results indicated that TUXEL allowed identifying a higher number of usability problems in comparison to an adapted version of the Heuristic Evaluation, while requiring less time.

Keywords: usability, user experience, learning management systems, informatics in education, human-computer interaction.

FIGURES INDEX

TABLES INDEX

TABLE OF CONTENTS

CHAPTER 1– INTRODUCTION

This chapter presents the contextualization of this master's research, the motivation, the question and the objectives of the research, and the followed methodology.

1.1. Context

Usability is one of the most important characteristics of a system, directly impacting on how users interact with it (Mtebe and Kissaka, 2015). In a simplified way, usability is a quality attribute that evaluates the ease of use of an interface (Nielsen, 2012). This ease or difficulty of a system during user's interaction can determine its success or failure (Fernandez et al., 2011). In Web systems, for example, usability is one of the most important attributes. If a website is confusing, difficult to navigate and obtain information, the user will not waste his time trying to understand it and will end up looking for alternatives (Nielsen, 2012).

Although usability is an important feature, technological evolution has generated new paradigms of interaction, requiring a revision of its concept (Rusu et al., 2015) and giving rise to the term User eXperience (UX), popularized primarily by Don Norman (Norman et al., 1995). The industry often considers UX as a synonym or extension for usability. The academy, however, differentiates between these two terms (Rusu et al., 2015; Hassenzahl, 2008). Hassenzahl et al. (2006), for example, proposed an approach that differentiates UX from traditional usability, based on three characteristics: holistic (balance between pragmatic and hedonic aspects versus usability-oriented aspects), subjective (in contrast to the objective approach of usability) and positive (in contrast to the focus of usability in removing barriers, problems, and frustrations). Thus, a software application would become much more attractive being usable and interesting at the same time, promoting aspects such as pleasure and satisfaction (Hassenzahl et al., 2000), demonstrating the importance of the evaluation of both usability and UX.

Regarding education area, the widespread adoption of information and communication technologies, such as Learning Management Systems (LMSs), has attracted an interest in evaluating the quality in use of these platforms regarding quality attributes such as usability and UX (Harrati et al., 2016). LMSs are specialized platforms designed to provide educational content, activities and support the management of learning programs in a digital way (Dubost et al., 2004). The use of such technology in education gave rise to the term called e-learning

(electronic learning), which "encompasses a broad spectrum of activities, from supported learning and blended learning to completely online learning" (Ssekakubo et al., 2014).

Given that most of the learning process occurs through the interaction of the learner with the LMS, and that the lack of high quality systems that meet individual and group needs forms a barrier to the development of effective technological learning (Kakasevski et al., 2008), it is important to evaluate the quality in use of these platform regarding aspects such as usability and UX. For instance, a LMS with poor usability and negative UX may make the learner spend more time trying to understand how to use it rather than learning the educational content (Ardito et al., 2006), while not motivating them enough during the learning process. In this context, one of the major challenges of the Human-Computer Interaction (HCI) area is "to develop software tools able to engage novice learners and to support their learning even at a distance" (Lanzilotti et al., 2006). By achieving this, it would be possible to understand the different ways students learn and make the interaction process as natural and intuitive as possible (Ardito et al., 2006).

In fact, differently from other contexts, the user's main task in e-learning is to learn. When the users are learners, there is not only a need for a software that support them doing tasks, but also support them learning while doing tasks (Zaharias et al., 2002). Thus, interaction process goes beyond the paradigm of traditional task-based usability, being directly related to the pedagogical value (Zaharias, 2008).

1.2. Motivation

The possibility of studying at any time and in any place has helped both students and teachers living away from schools and universities (Lanzilotti et al., 2006). Overcoming the limitations of traditional teaching as space and time, in e-learning, learning materials and processes are distributed over the Internet (Ardito et al., 2006; Orfanou et al., 2015). To assist in the management of these materials and to support teaching and learning process, learning institutions use specific platforms for this purpose, called Learning Management Systems (LMSs).

A LMS allows to "store, manage or modify educational contents through the interaction between the participants in the platform, by their authorization to insert, modify or erase data" (Freire et al., 2012), as well as the accomplishment of tasks and interaction and communication among participants (Medina-Flores and Morales-Gamboa, 2015). Some of the most used LMSs include Moodle, Edmodo and Blackboard, the latter a proprietary platform.

Currently, the use of LMSs is not limited to institutions working with distance education. Studies have shown that schools, colleges, large corporations and government bodies have been using this type of platform to promote continuous learning, to contribute to their experience and knowledge, and to find solutions to problems related to their studies and work (Oztekin et al., 2010; Guo et al., 2010; Granić and Ćukušić, 2011; Violante and Vezzetti, 2015). Considering the widespread adoption of LMSs in different contexts, it is important to evaluate these platforms regarding quality in use attributes, such as usability and UX, by means of evaluation techniques.

Although several usability and UX evaluation techniques exist in the literature, the difficulty in evaluating some specific aspects of LMSs, such as instructional feedback and flexibility, led many authors to develop specific techniques to evaluate these platforms, consolidating pedagogical aspects with usability principles (Reeves et al., 2012). However, most of these studies are still at initial stages and needs further improvements and empirical evidence (Zaharias and Koutsabasis, 2012), which indicates a room for improvement or development of new techniques to evaluate usability and UX of LMSs. By using these techniques, it may be possible to improve not only the usability and user experience, but also to contribute to the teaching and learning process by means of these platforms.

1.3. Problem definition

Technology made the evolution of e-learning possible in a complex way regarding educational contents, technological resources and possibilities of interaction through LMSs (Freire et al., 2012). In this context, a LMS should “take into account the different ways students learn and ensure that student’s interactions are as natural and intuitive as possible” (Ardito et al., 2006). Therefore, it is necessary to consider, during the evaluation, the way usability and learning interact, in order to obtain a synergy between learning process and student’s interaction with the software (Squires and Preece, 1996; Ardito et al., 2006).

According to Mtebe and Kissaka (2015), due to the difficulties in evaluating the usability of educational supporting systems, many authors have attempted to consolidate pedagogical aspects with interface-related heuristics. However, most of these studies are in early stages, with a relatively small number of heuristics, often vaguely defined, requiring further improvement or experimental studies.

Therefore, our research question is "How to evaluate the usability and UX of Learning Management Systems aiming to improve the quality in use of these platforms?” In order to

answer this question, we needed to obtain the state of the art about the techniques used to evaluate usability and UX of LMSs. Thus, we carried out a systematic mapping of the literature (Chapter 3).

The results revealed some opportunities for research: (i) few studies performed comparisons among techniques, which make it difficult to identify the most appropriate ones for evaluating LMSs; (ii) few studies evaluated the UX of LMSs; (iii) there was no evidence of techniques that evaluates both usability and UX in the context of LMSs. These results indicate that there is a need for more empirical evidence about the existing techniques in order to make it possible to researchers select those that are more adequate to their needs. There is also a lack of studies evaluating the UX of LMSs, an important quality attribute but frequently neglected in this context (Zaharias and Pappas, 2016). Finally, considering the importance of both usability and UX in LMSs, there is a need to develop techniques that evaluate both these quality attributes in order to make it possible to improve the teaching and learning process by means of these platforms.

Aiming to fill these gaps and answer our main research question, presented before, we developed a technique to evaluate both usability and UX of LMSs called TUXEL (Technique for User eXperience Evaluation in e-Learning). By integrating the usability and UX evaluation process, the technique may allow a more complete view regarding the quality in use of these platforms. The use of a technique that covers both aspects may provide the benefit of evaluating them in a more focused way, differently from using two independent techniques, which may evaluate aspects that may overlap between them. By doing this, it may also require less time and a fewer number of participants, since the evaluation process will be carried out at once. The empirical studies carried out to verify the feasibility of the technique during the development of TUXEL contributes to fill the gap regarding the lack of comparative studies between techniques. Finally, they also contribute to fill the gap related to the lack of studies carried out to evaluate the UX of LMSs

1.4. Research Goal

The main goal of this research is to develop a technique that is specific to evaluate LMSs composed by a set of usability and UX attributes. Thus, we hope to contribute improving the quality in use of LMSs by evaluating them through the proposed technique.

1.4.1. Specific Goals

The specific goals of our research are:

- Provide a body of knowledge regarding techniques to evaluate the usability and UX of Learning Management Systems;
- Obtain a set of usability and UX factors that may impact the teaching and learning process;
- Make possible the improvement of the quality of LMSs by evaluating them with the technique proposed.

Since students are the end users of these environments, the purpose of this work is to develop a technique to evaluate the Usability and UX of LMSs, while also being applicable by students. Managers will be able, through the evaluation process, to identify the possibilities to improve the quality in use of LMSs, improving learners' interaction with the platform and providing a better experience of use.

1.5. Methodology

We based our research methodology in the Design Science Research (DSR) approach, which consists in the design and investigation of artifacts (something created for some practical purpose), aiming to interact with a problem context and improve something in that context (Wieringa, 2014). In DSR, we change or improve something according to the desires of the stakeholders, in order to solve a problem by iterating over the activities of design and investigation, i.e., design cycle (Wieringa, 2009; Wieringa, 2014).

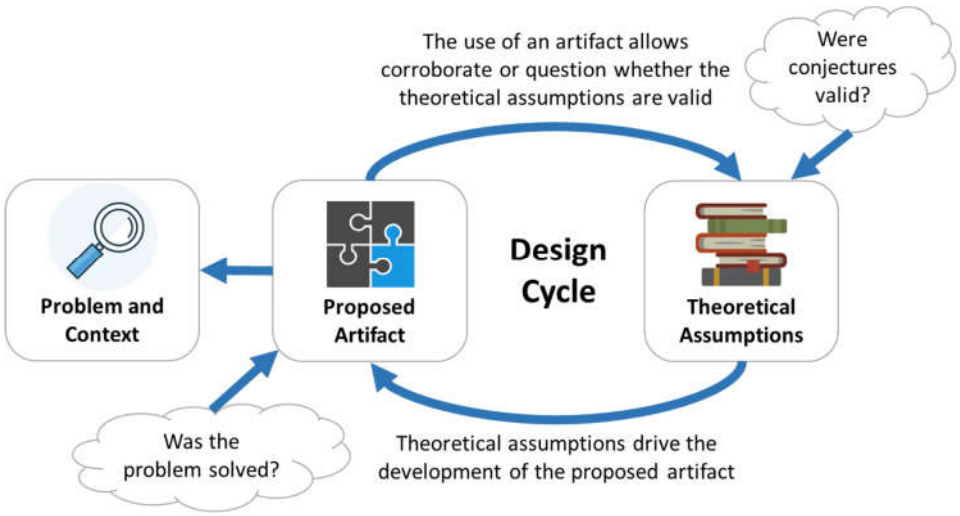


Fig. 1. Design science research map based on Hevner and Chatterjee (2010).

Figure 1 illustrates the main elements that compose DSR, based on the works of Wieringa (2014) and Hevner (2007). The DSR starts with the **definition of the problem** (problem and context) and/or opportunities of improvement in a given context (Hevner 2007). After defining the problem, the design science researchers then propose one or more artifacts to interact with that context.

The **artifact design** consists in the development and specification of an artifact that will interact with the problem context. Past knowledge and theoretical assumptions may drive its development, denoting the rigor of the DSR approach (Hevner 2007).

Finally, the **artifact is evaluated** by applying it to the original problem context, verifying whether it produces the desired effects (Wieringa, 2014). This step may lead to a new iteration over the design cycle.

Besides the experiences gained for performing the research, the DSR may provide additions to knowledge, such as extensions to theories and methods made during the conduction of the research (Hevner 2007). By combining theory and practice, the outcomes of the research may be useful for both academic and practitioner audiences.

Regarding problem investigation, we conducted a systematic mapping (Chapter 3) in order to obtain knowledge regarding the research topic, identify the problem and delimit the research scope. The results revealed that there is no evidence of a consolidated technique to evaluate the usability and UX of LMSs. It also revealed that few studies evaluated the UX of these platforms. Furthermore, among these studies, all of them employed generic UX evaluation techniques. Based on these results, we followed to the artifact design step.

In the artifact design step, we carried out a comparative study aiming to obtain useful knowledge to develop TUXEL. Since we found only few studies regarding the UX evaluation of LMSs, we carried out a comparative study with existing UX evaluation techniques (Chapter 4). The results from the systematic mapping and from the comparative study allowed us to develop the first version of TUXEL (Chapter 5).

In the next step, we carried out a study (Chapter 6) in order to verify whether the proposed artifact (TUXEL) produces useful results (artifact evaluation step). The results indicated the need for improvements, which led us to the refinement of the technique and development of TUXEL 2.0. Finally, we conducted a second study in order to evaluate the refined artifact (also in Chapter 6).

1.6. Organization

The next chapters are organized as follows: Chapter 2 provides definitions of terms used in the context of LMSs evaluation. Chapter 3 presents the methodology used to carry out the systematic mapping and its results. In Chapter 4, we describe the preliminary study carried out to evaluate two UX evaluation techniques. Chapter 5 presents our proposed technique. In Chapter 6, we present two empirical studies and the improvement of the technique. Finally, Chapter 7 concludes this dissertation.

CHAPTER 2 – USABILITY AND USER EXPERIENCE EVALUATION TECHNIQUES IN LEARNING MANAGEMENT SYSTEMS CONTEXT

This chapter provides the concepts, definitions and examples of techniques related to the usability and UX evaluation of Learning Management Systems. We also present the importance of evaluating both usability and UX to improve the quality in use of these platforms.

2.1. Learning Management Systems

Learning Management Systems (LMSs) are specialized platforms that provide educational content, activities and support for the management of learning programs by incorporating technologies to extend the learning in different environments (Dubost et al., 2004; Tee *et al.*, 2013). In the last decade, many universities and schools have adopted LMSs to complement regular teaching and learning (Vogten and Koper, 2014; Granić and Ćukušić, 2011). Moreover, corporations and government bodies also have been implementing such platforms to promote employee education and training (Oztekin *et al.*, 2010).

Differently from an academic system, which intends to manage student-related information such as personal data, attendance, enrollment, and grades, a LMS provides tools that expand teaching possibilities. A LMS can, for example, offer a number of functionalities and resources (not necessarily mandatory), such as storage, management and modification of educational content (Freire et al., 2012), assessments, tasks and mechanisms for tracking students' progress (Kakasevski et al., 2008). These platforms also may provide a set of tools that facilitate the communication with teachers and other students (Hijon-Neira et al., 2014), either synchronously (real-time, such as a videoconference) or asynchronously (e.g., discussion forums and emails).

Some publications differentiate learning platforms according to specific functionalities, dividing them into categories such as Learning Management Systems, Course Management Systems, Virtual Learning Environments or even Knowledge Management Systems. However, a study by Moore et al. (2011), which carried out a research on the different terminologies used in learning environments context, revealed that there is no consensus in the definition of these terms, having variations even between countries and continents. In their research, some authors referred to these terms as synonyms, while others alternated between them.

Regardless of its definition, learning support platforms have been introduced not only in universities, but also in schools, colleges, large corporations and government agencies for education and training of staff (Oztekin et al., 2010, Guo et al. 2010; Granić and Ćukušić, 2011). This widespread adoption reinforces the need to use appropriate techniques to improve the quality in use of these platforms in order to provide a better usability and a pleasurable experience of use.

2.1.1. Benefits and Difficulties in Using LMSs

One of the greatest benefits of using LMSs is that it can be accessed anytime and anywhere, making it possible, e.g., reducing the overcrowding of educational institutions and supporting students and teachers living in more distant places (Lanzilotti et al., 2006). By using LMSs, educational institutions can offer, for instance, courses in the form of distance education. According to a publication by EXAME.com[1] in 2014, about 25% of enrollments were in the distance education modality, with an expectation of increasing to 40% or 50% by 2019, leading institutions to invest in the expansion of poles and the provision of new courses.

In addition to distance education, face-to-face teaching can also benefit from the use of these platforms. Teachers, for instance, can use LMSs as an alternative means to motivate students by encouraging them in their activities and guiding them to study on their own (Hijon-Neira et al., 2014). Moreover, educational institutions are realizing that traditional education is not enough to cope with the growing demand for information, infrastructure limitations and lack of time needed for students (Sánchez-Chamochin et al., 2008). Thus, the adoption of LMSs to complement the traditional face-to-face classes has become increasingly common (Mtebe and Kissaka, 2015).

Although LMSs extend education possibilities, these platforms present difficulties in their use, such as technical limitations of computers and limited internet access (Tee et al., 2013). In countries where the Internet is more accessible and there is a better technological infrastructure, for instance, LMSs have been successfully deployed and used. On the other hand, emerging countries have not been able to realize the full potential of these platforms, mainly due to the high cost of Internet access, reduced number of computers and limited technological infrastructure (Ssekakubo et al., 2014).

In addition to the technological and infrastructure issues, there is also the lack of technical experience of instructors to develop and use multimedia software, the discomfort in

[1] https://exame.abril.com.br/brasil/ensino-a-distancia-no-brasil-pode-dobrar-em-5-anos/

the use of technological solutions and usability problems of LMSs (Tee et al., 2013, Ssekakubo et al., 2014). In this way, the challenge today is not to implement these platforms, but to provide means to use them effectively and efficiently (Aslan, 2011).

2.1.2. Examples of LMSs

A large number of LMSs have been developed for different contexts, such as academic, certification management, corporate training, among others. Next subsections present a brief description of two popular LMSs, which were objects of our empirical studies: Edmodo and Moodle. We selected Edmodo given that it is one of the most popular LMS, with more than 90 million users around the world[2]. Additionally, the author of this work identified some difficulties from learners during a course where he acted as manager of the Edmodo. Regarding Moodle, we selected it because it is one of the most used LMSs, with more than 127 million users in 232 countries[3]. Moreover, the evaluated LMS of our institution bases on this platform.

2.1.2.1. Edmodo

Edmodo (Figure 2.1) is an educational platform founded in 2008 designed to fill the gap between students' personal lives and school[2], enabling students and teachers to communicate with each other and connect, share ideas, problems and tips. The platform is accessed directly through a website, without the need to perform any local installation.

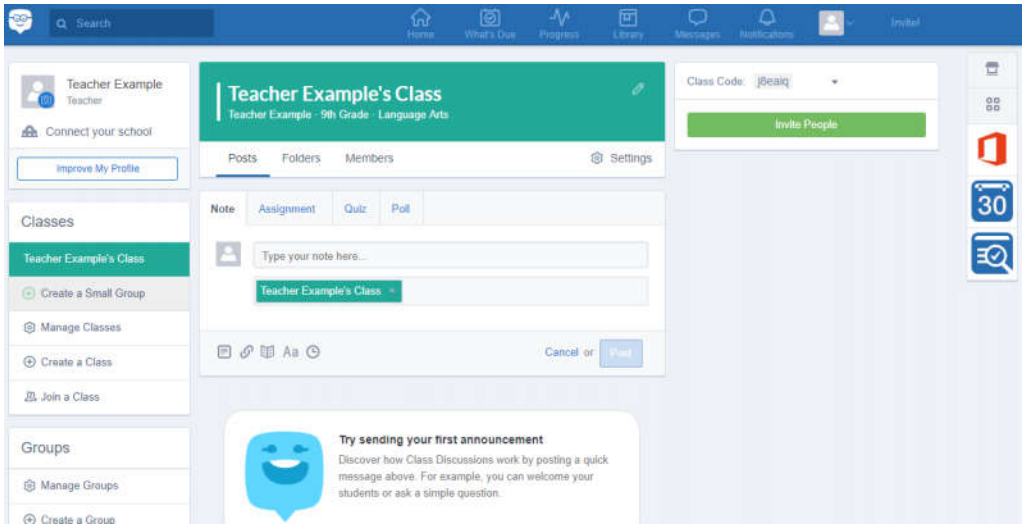


Figure 2.1 - Main page of the Edmodo LMS.

[2] https://www.edmodo.com/about?language=en
[3] https://moodle.net/stats/

Edmodo includes several functionalities, such as tasks definition, assessments, material distribution, note assignment and group creation. Each group has a unique code that makes it possible to restrict its access only to students who have that code (Kongchan, 2012). In addition, the platform presents features and aspects of a social network, such as posting messages, and options to follow, like and comment on a publication.

2.1.2.2. Moodle

Moodle (Modular Object-Oriented Dynamic Learning Environment) was developed as a PhD research project conducted by Dougiamas and Taylor (2003) at Curtin University of Technology. They developed the platform from the perspective of the social-constructivist and social-constructionism theory, whose epistemological positions focus on the collaborative discourse and knowledge development through social interaction. The authors' goals are to facilitate distance learning, improve teachers' pedagogical skills by providing free open source software and support a community of software contributors.

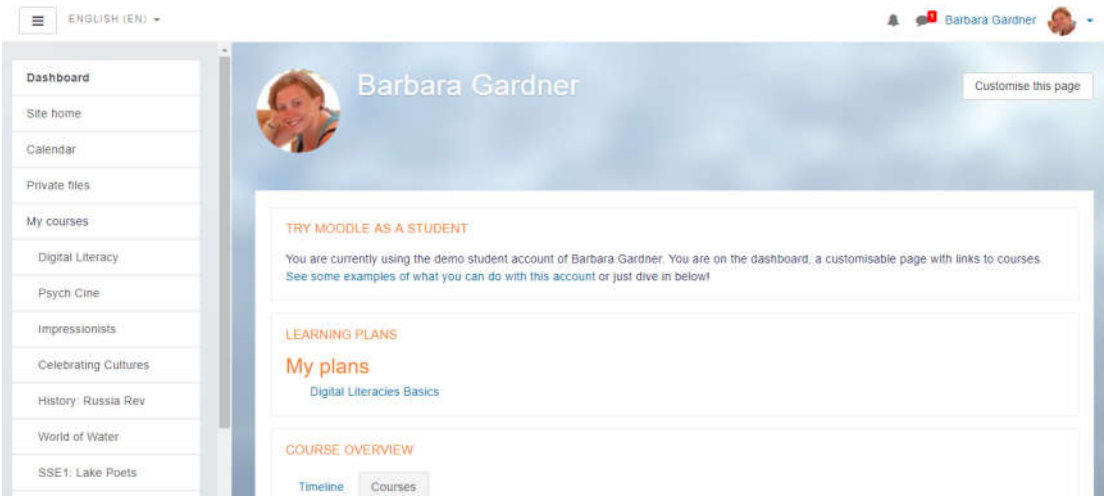


Figure 2.2 - Main page of Moodle demo website.

Moodle (Figure 2.2) integrates various functionalities, such as course management, assessments, collaboration tools, e.g., forums, Wikis, chats and blogs, as well as the possibility of installing and creating additional modules to expand its functionality[4]. According to Ssekakubo et al. (2014), universities with limited resources have been adopting open source platforms, such as Moodle, in order to avoid the concern with the use and renewal of licenses.

[4] https://docs.moodle.org/34/en/About_Moodle

Additionally, it provides scalability, supporting a great number of users, and flexibility, making it possible to change and improve its code to meet the needs of the institution.

2.2. Usability

There are many definitions for usability. According to ISO 9241-11 (2017), usability is "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". ISO/IEC 9126-1 (2001) defines usability as the "capacity of a software product to be understood, learned, used and be made attractive for the user when it is used under certain conditions". Nielsen (2012), on the other hand, considers usability as a quality attribute that evaluates the ease of use of an interface through five quality components: learnability, efficiency, memorability, errors and satisfaction.

Although there is no standard definition, usability is considered one of the most important factors in a Web application, such as LMSs, directly impacting users' quality in use regarding satisfaction, performance in the accomplishment of tasks and their acceptance (Bevan, 1995). A more usable application, for example, can quickly replace one that does not provide good usability (Whitten, 1990). In this way, it is essential to evaluate these applications regarding their usability, in order to identify opportunities for improvements and to enable a better quality in use.

2.2.1. Usability evaluation techniques

A variety of techniques has been proposed for usability assessment. Ivory and Hearst (2001) classified them into five types: inspection, testing, investigation, analytical modeling and simulation.

In **usability inspection**, the inspector, usually a usability expert, evaluates the platform by using a set of criteria to identify potential usability problems (Conte et al., 2007). Compared to usability testing, inspection may be more cost-effective, since it requires only a few experienced evaluators, while not needing any special equipment or laboratory and allowing evaluating the product even in its initial stage of development process (Matera et al., 2006). On the other hand, it depends on experts to be efficient, while the identified problems are generally less severe than those identified by usability testing (Hollingsed and Novick, 2007). In addition, this type of method relates to its subjectivity, i.e., different evaluators may produce different outcomes, being heavily depending on evaluator’s skills and experience (Conte et al., 2007).

Examples of inspection techniques include Heuristic Evaluation, Checklist-based Evaluation and Web Design Perspectives.

- **Heuristic Evaluation (HE):** its main goal is to identify the greatest number of usability problems possible with the lowest cost (WILSON, 2013). It consists in the evaluation of an interface, in which an expert verifies the violation of certain rules or simplified directives, called heuristics, to identify positive and negative aspects of an interface (Nielsen and Molich, 1990; Ivory and Hearst, 2001);
- **Checklist-based Evaluation:** in a checklist-based evaluation, the technique provides inspector with a list of questions to answer, giving him advices about what to look for during the inspection process (Laitenberger and Atkinson, 1999);
- **Web Design Perspectives (WDP):** it consists in a Web-based inspection technique proposed by Conte et al. (2009), developed with the aim of improving the efficiency of Nielsen's heuristic evaluation (Nielsen, 1994). Through perspective, it is possible to focus on specific aspects, allowing the identification of a greater number of usability problems. The technique associates the set of Nielsen heuristics with 04 Web perspectives: conceptual, presentation, navigation and structural. For each pair of heuristics-perspective, suggestions are provided to guide the inspectors in the use of the heuristics from the perspective's point of view.

Usability testing consists in observing the interaction of a user with an interface, usually involving the accomplishment of tasks, in order to identify usability problems (Ivory and Hearst, 2001). The main advantage of testing is that it provides evaluations that are more reliable, since it involves samples of real users. However, there are some drawbacks, such as the difficulty to select a proper sample for testing, the limited amount of time to reproduce the usage situations and the need for a prepared environment to conduct the evaluation (Matera et al., 2006). Additionally unlike heuristic evaluation, it covers only small parts of a product or service (Wilson, 2013). Some examples of usability testing techniques are the Think-Aloud Protocol and Log File Analysis.

- **Think-Aloud protocol:** in this technique, the user verbalizes his actions and thoughts during the accomplishment of the tasks (Fernandez et al., 2011);
- **Analysis of log files:** it consists in recording user interaction activities with the system in log files that are analyzed by evaluators or tools (Fernandez et al., 2011);

In **inquiry**, participants provide feedback about an interface through techniques such as questionnaires and interviews, in order to obtain subjective data from the participants

(Fernandez et al., 2011; Ivory and Hearst, 2001). In questionnaires, participants answer specific questions, while in interviews, an evaluator conducts a discussion session with a participant about his/her attitudes towards the artifacts being evaluated (Fernandez et al., 2011).

Analytical modeling techniques seek to predict usability problems through the modeling of aspects such as interface, task environments or user performance. An example of this type of technique is GOMS.

- **GOMS:** an acronym for Goals, Operators, Methods and Selection rules, is a technique that allows users to model interaction behavior with a software product, making it possible to predict, for example, the time required by an experienced user to be able to perform goal-directed tasks on a given interface (Schrepp, 2010). The goals describe what the user wants to accomplish, while the operators refer to the physical or cognitive processes that must be performed to achieve the goal. The methods are sequences of these operators, executed to achieve the goal. Finally, the selection rules decide which of these methods should be chosen.

Simulation consists in using simulation algorithms or using data analysis to simulate the interaction of a user with an interface, reporting the results of this interaction (Fernandez et al., 2011; Ivory and Hearst, 2001). As an example, we can mention the Monte Carlo simulation technique, which performs a simulation of a sample of a population to predict, for instance, the performance of a studied process or system, through mathematical models. Thus, a large amount of simulated data is generated, avoiding the need to conduct a large number of experiments or use a large number of samples (Davids, 2015).

2.2.2. Usability in LMSs

Usability, in LMSs, is a factor that has a great impact on the teaching and learning process. According to Lanzilotti et al. (2006), iff the platform does not provide a good usability, the learner will spend more time trying to learn how to use it rather than learning the educational content. Usability is also a key factor for acceptance, satisfaction and efficiency of academic institutions, regardless of the student's knowledge, experience or guidance (Harrati et al., 2016).

Traditional usability evaluation techniques have been applied to evaluate LMSs. However, their scope is limited. Despite identifying important usability factors such as ease of use and accessibility, such techniques do not cover specific aspects that can make an LMS usable, such as its pedagogical nature (Hovde, 2015).

Differently from other contexts, LMSs, in addition to offering good usability, should be effective in fulfilling the pedagogical objectives of the instructor, being necessary to integrate

the educational qualities to the usability evaluation process of the platform. Thus, it is not possible to obtain conclusions based on a simple failure or success analysis in the accomplishment of a task as in traditional evaluations, making this one of the major challenges of the HCI area in e-learning context (Lanzilotti et al., 2006 , Freire et al., 2012).

According to Mtebe and Kissaka (2015), due to the difficulty in evaluating the usability of LMSs and their specificities, many authors consolidated general interface usability with pedagogical aspects. Specific techniques have been developed for the evaluation of LMSs, considering, among other evaluated dimensions, the pedagogical usability.

Pedagogical usability relates to how the platform facilitates the learning of the provided material (Nokelainen, 2006). In the model proposed by Nokelainen (2006), pedagogical usability is a sub-concept of utility, defined by Nielsen (1990), which consists of 10 dimensions: (i) learner control, (ii) learner activity, (iii) collaborative/cooperative learning, (iv) goal orientation, (v) applicability, (vi) added value, (vii) motivation, (viii) evaluation of previous knowledge, (ix) flexibility and (x) feedback. Each dimension is composed of subitems, totalizing 51 elements.

An example of a technique that considers pedagogical usability is the one proposed by Zaharias and Poylymenakou (2009). The authors developed a questionnaire applicable with users of educational platforms in order to evaluate both Web usability and pedagogical usability and motivation to learn. The technique is based on techniques of other authors, such as the heuristics of Reeves et al. (2002) and Quinn's framework (1996).

Although specific techniques for evaluating the usability of LMSs have been proposed, most of them are in early stages (Mtebe and Kissaka, 2015). Moreover, there is still no widely accepted set of principles for the development of these techniques (Hovde, 2015), which reinforces the need for further studies in this context.

2.3. User Experience

As mentioned before, usability is an objective approach, focused on the accomplishment of tasks and on the removal of the barriers, problems and potential dissatisfaction (Hassenzahl et al., 2006). Although removing the negative is important, it does not necessarily imply that it would put a smile on user’s faces (Hassenzahl et al., 2006). With the development of new technologies, users are not seeking only to achieve a task, but also to amuse and entertain themselves (Petrie and Bevan, 2009). Thus, usability by itself is not enough to define the quality

of software. Aspects such as user's emotions and feelings when interacting with software should also be taken into account at the time of evaluation (Nascimento, 2016).

As well as usability, User Experience (UX) has many definitions. ISO 9241-210 (2010), for instance, defines UX as "perceptions and responses of a person resulting from the use and/or anticipated use of a product, system or service". Hassenzahl (2008), on the other hand, proposes an approach involving pragmatic (related to the accomplishment of tasks) and hedonic (related to emotions) aspects. In this way, the hedonic aspects contribute directly to a positive experience, while the pragmatic aspects facilitate the potential to achieve goals such as "being autonomous", "being competent", among others.

As technology evolves and interactive environments become increasingly ubiquitous in all aspects of life, including education, users expect such environments to provide more than just ease of use (Petrie and Bevan, 2009). Hassenzahl et al. (2000) suggest that instead of making software merely usable, aspects such as pleasure and satisfaction should be the main goal of the project. Being usable and at the same time interesting, the system would become much more attractive and consequently, it would improve the UX.

2.3.1. UX evaluation techniques

There are several UX evaluation techniques proposed in the literature using different approaches. Kujala et al. (2011), for instance, proposed UX Curve. It consists in a technique designed to evaluate UX over time. According to the authors, the momentary evaluation of UX is not, in most cases, very reliable in real situations, because they do not capture the long-term user experience, from initial learning to the enthusiasm in making it part of user's routine. In UX Curve, the user draws a curve and annotates these variations in a template (see Figure 2.3) to express their experience of using the software regarding UX dimensions, such as attractiveness, ease of use, stimulation, utility, among others. The authors concluded that, in this way, it is possible to obtain more qualitative data than only numerical results.

Laugwitz et al. (2008) developed a technique called UEQ (User Experience Questionnaire), aiming to carry out a simple and quick evaluation. The technique is composed by 26 pairs of items related to two classes: items that directly measure perceived attractiveness and items that measure product quality in the relevant aspects. Each pair corresponds to an adjective and its antonym, consisting of a 7-point semantic differential scale between them, where the evaluator should mark the point that is closest to the adjective that better describes his/her UX.

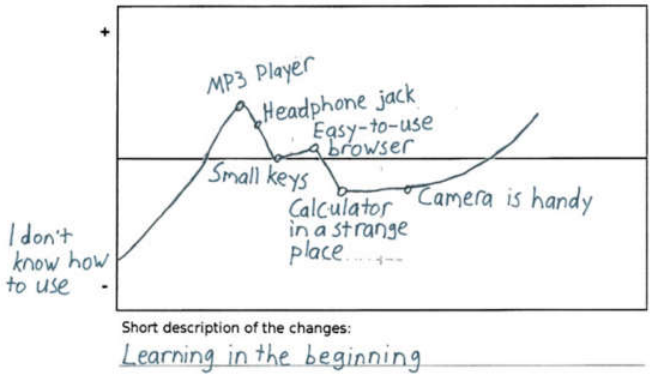


Figure 2.3 - A drawing from a UX Curve user with annotations of the reasons for each change in his/her UX. Source: Kujala et al. (2011).

Nascimento et al. (2016) proposed a technique to evaluate both usability and UX, called Userbility. The goal is to enable novice evaluators, not HCI experts, to evaluate his/her UX when using mobile applications. The technique bases on Nielsen's 10 heuristics (1994), which was used to generate the aspects related to usability, and the technique 3E (Expressing Emotions and Experiences), developed by Tähti and Niemelä (2006), to obtain emotional answers on the experience of use by the user. For each aspect of usability, the evaluator must perform four activities: 1) verify that the application meets this aspect; 2) report the feelings he/she had about this aspect; 3) suggest improvements to this aspect; and 4) choose a representation on a scale of five emotions to represent their degree of satisfaction with the application about this aspect.

2.3.2. User eXperience in LMSs

With technological innovations and new patterns of interaction, UX has been gaining importance and changing the expectations and demands of users. In e-learning context, it is important that the LMS provides good usability and a positive user experience. Failure to comply with these quality aspects may result in dissatisfaction, resistance, incomprehension or misuse of LMSs, leading to criticism and low acceptance of these platforms (Van Der Linden and Van De Leemput, 2015). In this way, it is necessary that the LMS provides, besides usability, a positive UX for the students, improving their interaction with the platform and enabling a greater engagement in the learning process.

Despite the importance of UX, few studies were carried out to evaluate this quality attribute in LMSs (Nakamura et al. 2017a). Moreover, there was no evidence of techniques designed specifically to evaluate the UX of these platforms. Thus, researchers have been using generic UX evaluation techniques (i.e., techniques designed to evaluate the UX of software in general, not specific for evaluating LMSs) to evaluate these platforms.

According to Zaharias and Pappas (2016), although it is very important, the UX evaluation has been neglected in the e-learning field, which may explain the lack of specific techniques for evaluating the UX of these platforms. Although some studies consider the importance of usability and human factors in the evaluation of LMSs, these works do not have a holistic vision oriented to UX. Thus, it reinforces the need for improvements of the existing techniques or the development of new techniques that cover subjective aspects related to users' emotions when using these platforms.

2.4. Conclusion

Usability and User eXperience play an important role in LMSs. If a LMS does not provide a good usability, the learner may focus his/her cognitive effort in understanding how to use the platform instead of learning the educational material (Ardito et al., 2006). Similarly, a negative UX may influence the attitude of the learner towards the use of the LMS, affecting the acceptance, satisfaction and efficiency of academic institutions (Harrati et al., 2016).

This chapter presented the concepts of usability and UX and the importance of evaluating these attributes in LMSs. It also provided an overview of the different types of evaluation techniques and examples of their applications.

Due to the specificities of LMSs, many authors developed specific techniques to evaluate these platforms. However, most of these techniques are at initial stages, while there is still no widely accepted set of principles for the development of these techniques (Mtebe and Kissaka, 2015). Thus, in order to gather further information about the existing usability and UX evaluation techniques in the context of LMSs, we carried out a systematic mapping of the literature, presented in the next chapter.

CHAPTER 3 – SYSTEMATIC MAPPING ABOUT TECHNIQUES TO EVALUATE USABILITY AND UX OF LEARNING MANAGEMENT SYSTEMS

This chapter presents a systematic mapping about techniques to evaluate usability and UX of Learning Management Systems. It aims to address the following research question: "Which usability and UX evaluation techniques have been applied in Learning Management Systems and how have they been used?" This section also describes how we carried out the systematic mapping and its results.

3.1. Introduction

This chapter presents the first step of Design Science Research (DSR) methodology employed in our research, which consists in the problem definition. In order to define the problem, we carried out a systematic mapping of the literature, which was published in Nakamura et al. (2017a).

A systematic mapping is a type of Systematic Literature Review (SLR). Unlike a conventional review, a SLR has a rigorous and well-defined methodology, which allows obtaining less skewed results and ample information on several experimental methods, however, requiring a greater effort from the researcher. According to Kitchenham and Charters (2007), a SLR is a "way to identify, evaluate and interpret all relevant research to a particular research question, area of a subject or phenomenon of interest." Systematic mapping, in turn, is a more comprehensive type of review that seeks to identify evidence or gaps related to a particular topic of research, making it possible to direct the focus of future systematic reviews. Given this, the goal of our systematic mapping was to obtain a broader view on the evidences and gaps about usability and UX evaluation techniques that have been used in the context of Learning Management Systems (LMSs).

This chapter will present the methodology used to conduct the systematic mapping, a discussion of the results found and a summary of the usability and UX evaluation techniques used to evaluate LMSs.

3.2. Review Protocol

The review protocol defines the procedures used to conduct a systematic mapping or systematic review of the literature, being important for the correct conduction and validity of the

review/mapping (Wohlin et al., 2012). The following subsections detail the revision protocol used to conduct this systematic mapping.

3.2.1. Goal

We defined the goal of the systematic mapping according to Basili's GQM (Goal-Question-Metric) Paradigm (Basili and Rombach, 1988), as presented in the table below.

Table 3.1 - Goal of the research according to the GQM Paradigm (Basili and Rombach, 1988).

Analyze	Scientific publications
For the purpose of	Characterize
With respect to	Usability and UX evaluation techniques
From the point of view of	Researcher
In the context of	Learning Management Systems

3.2.2. Research Question

This systematic mapping aimed to answer the following research question: **"What usability and UX evaluation techniques have been applied in Learning Management Systems and how have they been used?"** In addition to this general research question, we defined sub-questions with the purpose of answering specific questions about the usability and UX evaluation techniques in the context of LMSs, presented in Table 3.2.

3.2.2.1. Research Scope

We carried out this systematic mapping in the Scopus[5] and Engineering Village[6] digital libraries. The two are meta-libraries that index publications from several well-known publishers, such as ACM, IEEE, Springer and Elsevier. Its search engine makes it possible to define filters such as document type, language and area of knowledge. In addition, studies by Meho and Rogers (2008) indicate that Scopus has a greater coverage of journals and conferences in the HCI area.

[5] http://www.scopus.com
[6] http://www.engineeringvillage.com

Table 3.2 - Research sub-questions.

Sub-question	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 (inspection, testing, inquiry, analytical modeling or simulation)?
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 carried out?
SQ11	Does the study perform a comparison between techniques?

3.2.2.2. Language

The selected languages were English and Portuguese. We chose English given that it is adopted by the great majority of international conferences and periodicals. In turn, Portuguese was selected because it is the language used in national conferences, such as those promoted by the Brazilian Computer Society (SBC).

3.2.2.3. Terms Used in Search

To define the terms of the search, we used the procedure described by Kitchenham and Charters (2007), which suggests determining population, intervention, comparison, outcome and context parameters.

- **Population:** Learning Management Systems;
- **Intervention:** techniques, tools, processes;
- **Comparison:** it does not apply, given that the goal is to characterize the techniques;
- **Outcomes:** evaluation of usability or UX of Learning Management Systems;
- **Context:** does not apply given that there is no comparison to determine the context.

We divided the search terms into two groups (see Table 3.3). The first one relates to the different forms of writing and synonyms for LMSs, while the second group relates to the different types of terms used for usability and UX evaluation. We identified these terms based

on a reference set of relevant articles and expert knowledge in that context. The set of reference publications was composed by the following: Freire et al. (2012), Lanzilotti et al. (2011), Theng and Sin (2012), Kakasevski et al. (2008), Zaharias and Koutsabasis (2011) and Medina-Flores and Morales-Gamboa (2015).

Table 3.3 - Groups of terms used to compose the search string.

GROUP 1	GROUP 2
LMS	usability evaluat*
online education platform*	usability assessment
online education system*	usability inspection
online education environment*	usability improvement*
e-learning environment*	usability test*
e-learning system*	usability technique*
e-learning course*	usability guideline*
e-learning platform*	UX evaluat*
e-learning application*	UX assessment
e-learning course*	UX improvement*
distance learning system*	UX technique*
distance learning platform*	UX guideline*
distance learning environment*	user experience evaluat*
distance education platform*	user experience assessment
distance education system*	user experience improvement*
distance education environment*	user experience technique*
online learning platform*	user experience guideline*
online learning system*	
online learning environment*	
virtual learning environment*	
VLE	
managed learning environment*	
MLE	
course management system*	
web-based learning system*	
remote learning system*	
m-learning system*	
m-learning platform*	
m-learning environment*	
mobile learning system*	
mobile learning platform*	
mobile learning environment*	
b-learning	
blended learning	
learning management system*	
learning content management system*	
LCMS	

The search string was developed using the Boolean operator "OR" between the typing alternatives and the Boolean operator "AND" to join these two groups (Table 3.4). We tested the string in multiple runs in order to reduce the number of publications that are not related to the research topic, while ensuring that the set of reference publications is returned.

Table 3.4 - Search string used.

("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*")*

3.2.2.4. Selection Criteria

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 3.5). In the second filter, we did the complete reading of the selected publications. We selected the publications according to the same criteria used in the first filter

Table 3.5 - Inclusion and exclusion criteria used in the systematic mapping.

#	**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

Some considerations must be made regarding the exclusion criteria. We did not consider publications related to the evaluation of MOOCs - Massive Open Online Courses (EC2). Although they may be embedded in learning platforms, MOOCs are not LMSs, but online courses aiming at unrestricted massiveness (Pireva et al., 2015). Additionally, there is a limitation regarding access to the administrative part of the MOOCs, which prevents the validation of usability problems. For instance, it is not possible to identify if a resource is not present in a course due to limitations in the platform or due to the teacher's choice in not using it. We also did not consider publications related to augmented reality (EC3) or specifically designed to evaluate aspects of accessibility (EC4). The former, for being out of the scope of this work, while the latter, for not considering usability as a whole.

3.2.3. Data Extraction Strategy

After carrying out the selection process, we started the data extraction process by means of the complete reading of each of the publications selected in the second filter. To do so, we used the strategy of providing a set of possible responses. According to Fernandes et al. (2011), this ensures that the same data extraction criteria will be used, thus facilitating their classification. We extracted the information according to each sub-question.

With respect to **SQ1 (Origin of the technique)**, we classified the technique in:

a) **New:** when the technique used was developed or adapted specifically for the context of e-learning;
   - In this case, it is necessary to describe what differentiates the technique from other already existing techniques.

b) **Existing:** when the technique already exists in the areas of HCI and was applied in its original form, without adaptations for e-learning.

In relation to **SQ2 (Technique type),** we classified the technique according to the taxonomy proposed by Ivory and Hearst (2001):

a) **Inspection:** when an evaluator uses a set of criteria to identify potential usability problems. Example: Heuristic Evaluation, Cognitive Walkthrough, etc.;

b) **Testing:** when an evaluator observes a participant interacting with an interface in order to identify usability problems. Examples: Think Aloud Protocol, Analysis of log files, etc.;

c) **Inquiry:** when a user provides feedback about an interface through interviews, surveys, focus groups, etc.;

d) **Analytical modeling:** when an evaluator uses different user models and interfaces to predict usability problems. Examples: GOMS, Design analysis, Analysis of cognitive tasks, etc.;
e) **Simulation:** when the evaluator simulates the interaction of a user with the system using some kind of simulation algorithm. Examples: Petri Net Modeling, Information Processing Modeling, etc.

As for **SQ3 (Execution Form)**, we classified the technique as:

a) **Manual:** when the technique is executed manually, that is, the usability/UX evaluation is performed by a human evaluator;
b) **Semiautomatic:** when part of the usability/UX evaluation is executed manually and the other part using automated tools;
c) **Automatic:** when an automated tool performs almost all the evaluation process, requiring only the interpretation of the data by the evaluator.

With respect to **SQ4 (Learning factors)**, the possible answers are:

a) **Yes:** if the technique considers specific learning factors, such as content relevance, learning strategies, instructional feedback, instructional assessment, learner support, etc.;
b) **No:** if the technique considers only usability/UX factors.

Regarding **SQ5 (Evaluation Focus)**, we classified the technique as:

a) **Usability:** when the technique considers only general usability factors, such as navigation, error recovery and visibility of the system status;
b) **UX:** when the technique considers only specific factors of UX, that is, subjective factors, related to the emotions and feelings of the user when interacting with the platform;
c) **Usability and UX:** when the technique considers both usability and UX factors.

Regarding **SQ6 (feedback)**, the possible answers are:

a) **Yes:** if the technique provides suggestions for correction of identified usability problems;
b) **No:** if the technique only performs the identification of the usability problems.

As for **SQ7 (Type of research)**, we classified the studies (according to the conditions of the empirical investigation) into three main research strategies, as described by Wohlin et al. (2012):

a) **Survey:** if the study is performed retrospectively, when a tool or technique has been used for a while. In this type of study, the primary forms of obtaining qualitative and quantitative data are through interviews and questionnaires;
b) **Case study:** if the study is observational, aiming to track a specific attribute or establish relationships 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.
d) **No:** if the study provides only a proof of concept.

With regard to **SQ8 (Restriction)**, the possible answers are:

a) **Yes:** if the technique has some condition for application, such as the use of specific tools or equipment or the need for specific knowledge;
b) **No:** if the technique can be performed without restrictions by the evaluator.

Regarding **SQ9 (Availability)**, the possible answers are:

a) **Yes:** if the complete technique (questionnaires, procedures, etc.) is available in the article or in some external source, such as a website;
b) **No:** if the technique is not available.

As for **SQ10 (Platform used),** we classified the study as:

a) **Desktop/Web:** if the study was conducted on a desktop/Web;
b) **Mobile:** if the study was conducted on a mobile device;
c) **Desktop/Web and Mobile:** if the study was conducted both on desktop and mobile devices.

Regarding **SQ11 (Comparison between techniques)**, the possible answers are:

a) **Yes:** if the study carried out any comparison between different usability/UX evaluation techniques;
b) **No:** if the study did not carry out any type of comparison.

Details of the selected articles and the extraction forms are available in Appendix A and Appendix B and C, respectively.

3.3. Results

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

For the first step, the researchers independently classified a sample of 17 randomly selected publications based on the selection criteria. We evaluated the agreement between the researchers by applying 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).

3.3.1. Selected publications

Figure 3.1 shows the selection process of publications carried out in the conduction of the systematic mapping. The search string returned a total of 177 publications in the Scopus library and 13 in the Engineering Village library. We found duplicated publications during the process. In these cases, we accounted the publications 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. We fully read and classified the remaining 142 publications in the second filter, according to the criteria. At the end of the process, 62 publications were accepted and extracted.

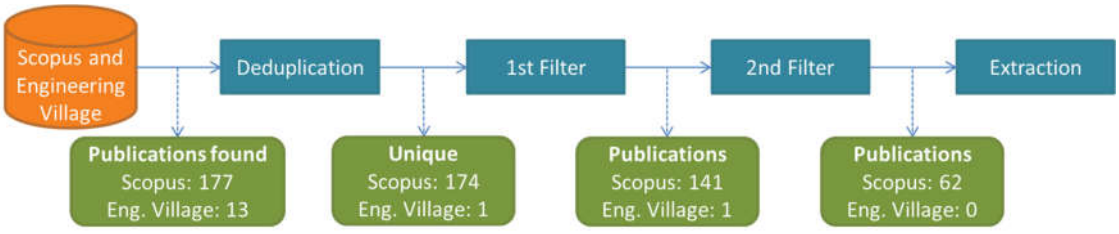


Figure 3.1 - Publications selection process.

3.3.2. Frequency of publications per year

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. Given that we carried out this systematic mapping in September 2016, the data referring to this year are incomplete, which can possibly explain the low rate of publications for this year.

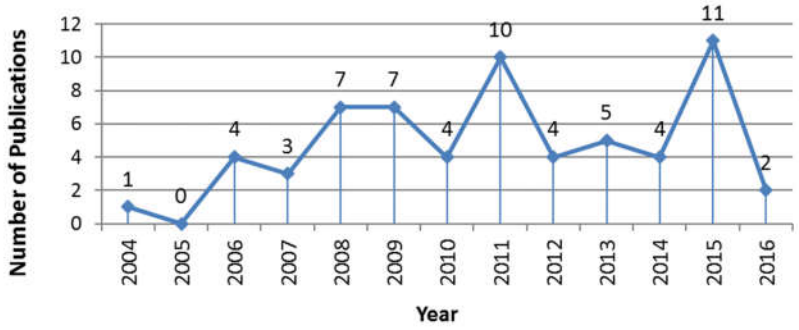


Figure 3.2 - Frequency of publications by year.

3.3.3. Results Overview

Table 3.6 presents an overview of the results. Sub-questions marked with a (*) are related to the techniques, while those marked with a (**) are related to publications. Although 62 publications were selected, the sub-questions regarding the studies counted only 58 publications. This is due to the fact that 02 publications presented only a literature review, and other two presented the initial stages of evolution of a technique that had already been classified.

We did not consider publications S37 and S49 given that they presented literature reviews, not being possible to extract their data in the same way as the other publications. Thus, we decided to keep them out of this table of results, adopting a specific data extraction form, which can be found in APPENDIX C. Although publication S61 presents an update of the systematic mapping from publication S49, we did not exclude it from the count by the fact that the authors proposed a technique based on the results of the systematic mapping. The remaining two publications (S04 and S10) were not considered given that they contained initial stages of evolution of an already classified technique or a partial analysis of a previous study.

In relation to the techniques classified, some publications used more than one technique for the evaluation of usability/UX, each being counted separately. Thus, the total number of techniques used in the 58 publications was 104.

The sub-questions SQ2 and SQ7 have data that is not unique. Therefore, the technique can be classified into one or more responses, generating a total percentage greater than 100%.

The results of each technique for each of the sub-questions are available in Appendix D. The following subsections present the analysis of the results obtained from each research sub-question.

Table 3.6 - Overall results for each research sub-question.

Sub-question	Answer	Results	
		Quantity	Percentage
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%

* Regarding techniques
** Regarding publications

3.3.4. Technique Origin

The results of the SQ1 (Figure 3.3) show that most of the techniques used in the e-learning context (67,31%) are 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.

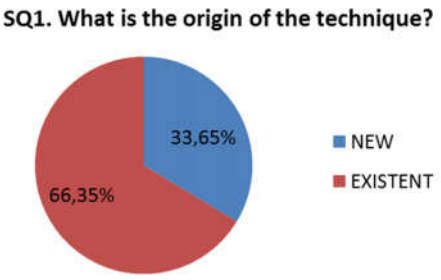


Figure 3.3 - Results of SQ1 (Technique Origin).

Around 33% of the techniques are new, specific to the context of e-learning. Most of them used some type of questionnaire, such as proposed by Ssekakubo et al. (2014). The authors proposed a technique that addresses instructional usability and motivation factors, besides the conventional usability factors. Others used some kind of Heuristic Evaluation, such as the HE proposed by Mtebe and Kissaka (2015), based on Nielsen's Heuristics (Nielsen, 1994) and other techniques specific to the e-learning context, such as the heuristics of Squires and Preece (1996) and Reeves et al. (2002).

3.3.5. Technique Type

The results of the SQ2 (Technique Type) reveal that **inquiry** was the most employed type of technique, with 51,92% of the techniques (Figure 3.4). Some of the most applied techniques were questionnaires, focus groups and interviews.

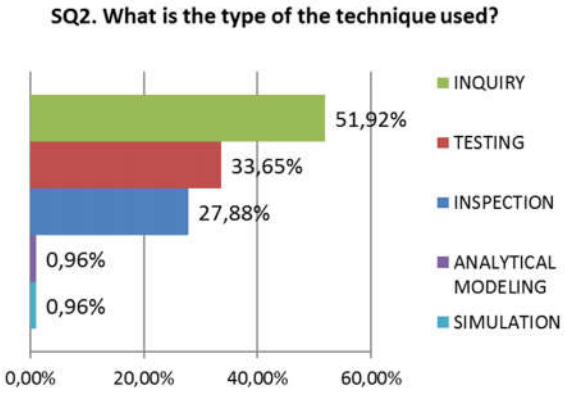


Figure 3.4 - Results of SQ2 (Technique Type).

- **Questionnaires:** users provide answers to specific questions. Zaharias and Poylymenakou (2009), for instance, 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 34% of the techniques employed some type of test involving users. Some of the most representative techniques were performance measurement, think-aloud protocol and log file analysis.

- **Performance Measurement:** 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 28% of the techniques were of **inspection** type. The most representative techniques were heuristic evaluation, pattern-based evaluation and checklist-based evaluation.

- **Heuristic Evaluation (HE):** 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-centered 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 by means of 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.

3.3.6. Execution Method

The results of the SQ3 (Execution Method) show that the majority of the techniques are performed manually, representing 90,38% of the techniques, followed by automatic techniques, that represent about 7% and semi-automatic about 3% (Figure 3.5).

Regarding the automatically performed techniques, the study conducted by Stickel et al. (2008) aimed to evaluate the usability and UX of an LMS by using psychophysiological methods to detect user-hostile systems, collecting data through Electroencephalograms, Heart Rate and Skin Conductance Level.

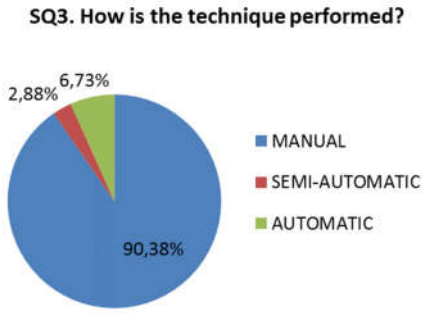


Figure 3.5 - Results of SQ3 (Execution Method).

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

3.3.7. Learning Factors

The results for the sub-question SQ4 (Learning factors) show that about 85% of the new techniques consider specific learning factors, while 14.71% of the techniques did not specify whether they considered these factors (Figure 3.6). Examples of techniques that consider specific learning factors are present in the works of Fetaji, M. and Fetaji, B. (2011) and Theng and Sin (2012).

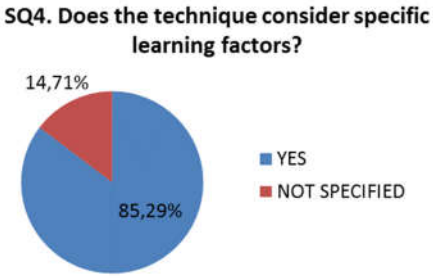


Figure 3.6 - Results of SQ4 (Learning Factors).

Fetaji, M. and Fetaji, B. (2011) proposed a specific framework for the context of m-learning, containing nine guidelines: (i) type of learner, skills and background; (ii) Easily perceptible representation of the m-learning system interface; (iii) Human factors; (iv) Technological factors; (v) Visibility of system status; (vi) Correspondence between the m-

learning system and the real world; (vii) Learner control and freedom; (viii) Recognition, diagnosis and recovery of errors; and (ix) Recognition rather than recall.

Theng and Sin (2012) developed a modified version of the TAM (Technology Acceptance Model) proposed by Davis et al. (1989). It is composed by a set of constructs and its variables: usability (Perceived Satisfaction, Perceived Usability and Perceived Usability), Engagement in E-learning (Learning By Interaction and Making Sense of Learning), Self-efficacy (E-learning Effectiveness and Efficacy of New Media), and E-learning Design and Support (Navigation Structure, User Interface and Personalization and Freedom of Control).

Although many techniques have considered specific learning factors, there has been no consensus, however, on which factors should be evaluated. We identified 30 different types of factors. Figure 3.7 shows the most evaluated learning factors by the evaluation techniques. The most evaluated factor was "Relevance of content" with 50%, followed by "Interaction among participants", with 46.15%. Around 42% considered "Feedback guidance" and "Instructional assessment". "Content organization and structure" accounted for 38.46%, followed by "Motivation" and "Support for meaningful learning approaches" with 30.77%. Around 27% considered "Media use" and "Collaborative learning". Other factors were evaluated by less than 25% of the techniques.

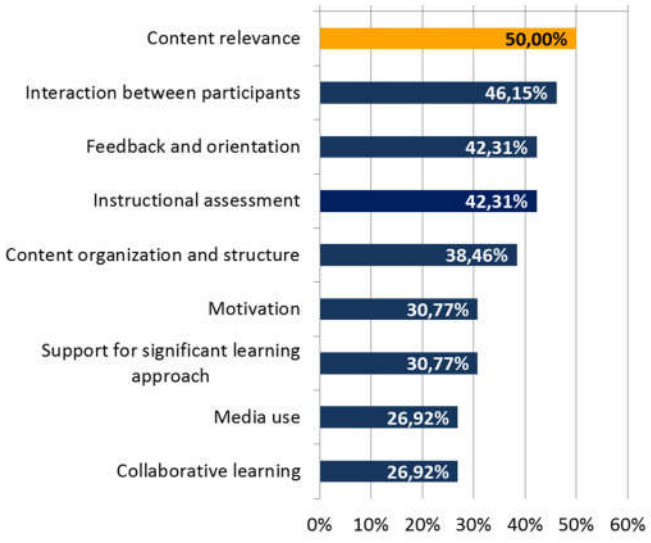


Figure 3.7- Most evaluated learning factors by the techniques.

3.3.8. Evaluation Focus

The results for the sub-question SQ5 (Evaluation Focus) show that about 29% of the techniques evaluate both usability and UX factors (Figure 3.8). Such techniques can be found in the works of Stickel et al. (2011) and Hijon-Neira et al. (2014). Stickel et al. (2011) proposed a technique

called SET (Shadow Expert Technique). In this technique, groups of evaluators identify usability problems by visualizing recordings of user interaction with the application. UX is also evaluated through visualization of video expressions and user audio verbalizations, allowing the evaluators to identify their expectations, intentions and emotions. Navarro et al. (2016) proposed a framework for evaluating m-learning applications, considering pedagogical factors such as content, organization and objectives, and motivation related to the affective factor.

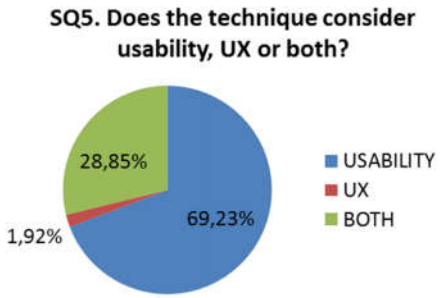


Figure 3.8 - Results of SQ5 (Evaluation Focus).

Techniques that only evaluate usability accounted for about 69% of the techniques. Examples of such techniques can be found in the works of Lanzilotti et al. (2006) and Yusoff and Mat Zin (2011). Lanzilotti et al. (2006) propose the eLSE methodology, a technique that uses standards-based evaluation to detect usability problems. Yusoff and Mat Zin (2011) used a questionnaire for usability evaluation, containing 10 questions related to ease of use, ease of navigation, support for classroom learning, among others.

Only 1.92% of the used techniques are specifically focused on the evaluation of UX. However, all are generic techniques, not specific to the e-learning context. Examples of these techniques can be found in the works of Santoso et al. (2014) who used the UEQ (User Experience Questionnaire) questionnaire to evaluate the UX of a platform called SCELE (Student Centered E-Learning Environment).

3.3.9. Feedback

The results for the SQ6 (Feedback) sub-question show that none of the techniques provide feedback with suggestions for correction of usability/UX problems for the evaluator (Figure 3.9). All used or proposed techniques only identify these problems.

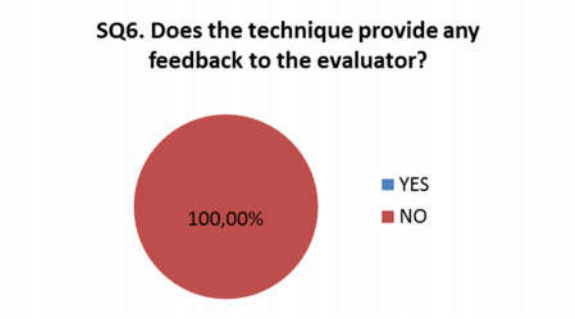


Figure 3.9 - Results of SQ6 (Feedback).

3.3.10. Investigation Type

Regarding type of investigation, we classified the experimental studies according to the three main research strategies described by Wohlin et al. (2012). The results for sub-question SQ7 (Investigation Type) showed that most of the studies carried out were Case Studies (according to the authors), with 46.55% (see Figure 3.10). Hamdi et al. (2011), for instance, conducted a case study to assess the usability of an m-learning application at the Arab Academy for Science and Technology (AASTMT). The study consisted of a user test by performing tasks using the Think Aloud protocol and applying the USE (Satisfaction and Ease of Use) questionnaire to verify user effectiveness, efficiency and satisfaction.

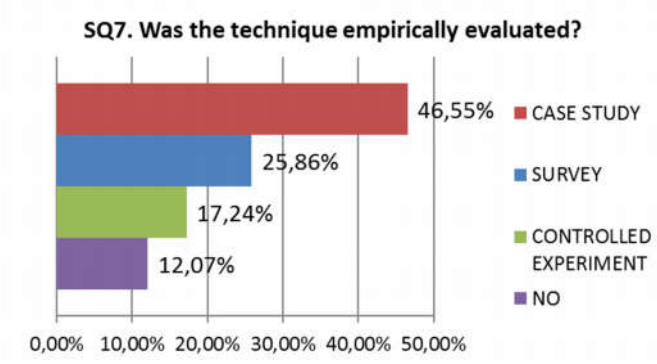


Figure 3.10 - Results of SQ7 (Investigation Type).

The Survey was the second most used type of research, with about 26% of the publications. Alkhattabi (2015), for instance, conducted a Survey using a self-administered questionnaire containing a Likert scale of 05 points (Strongly Agree, Strongly Agree, Neutral, Strongly Disagree and Strongly Disagree) and a feedback field to assess usability, practicality, pedagogical efficacy and overall design of an LMS called Tadarus.

About 17% performed some type of Controlled Experiment. Zaharias and Koutsabasis (2011) conducted an experiment to compare two specific usability inspection techniques for the

e-learning context: the Heuristics of Mehlenbacher et al. (2005) and the Heuristics of Reeves et al. (2002). The results showed that both heuristics have good coverage of usability problems, however, they do not present a satisfactory distribution of problems, with some heuristics that identify a high number of problems, while others identify few or none at all.

The publications that did not present experimental studies represented about 12%

3.3.11. Technique Restriction

The results for sub-question SQ8 (Technique Restriction) show that 11.54% of the techniques have some type of restriction for their use, such as specific knowledge, software or equipment (Figure 3.11). Examples can be found in the works by Stickel (2008) and Oztekin et al. (2010). The psychophysiological measurement method used in Stickel (2008) requires specific equipment for electroencephalograms, monitoring of skin conductivity and heart rate. The UseLearn technique, used in the work of Oztekin et al. (2010), requires the evaluator to have knowledge in the analysis of structural equations of critical metrics based on modeling to perform the usability index calculations.

The other techniques, which have no restrictions for use, represent 88.46%.

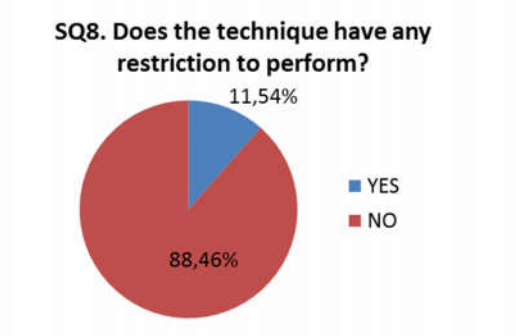


Figure 3.11 - Results of SQ8 (Technique Restriction).

3.3.12. Availability

Regarding the sub-question SQ9 (Technique Availability), the results showed that about 66% of the techniques are available (Figure 3.12), that is, the questionnaires and/or procedures needed to conduct the usability/UX evaluation are provided in the article itself or in external sources (e.g., other publications or Websites). Koohang (2004) and Junus et al. (2015), for instance, provided their questionnaires in the article itself.

Techniques that are not available for download/consultation accounted for 33.65%.

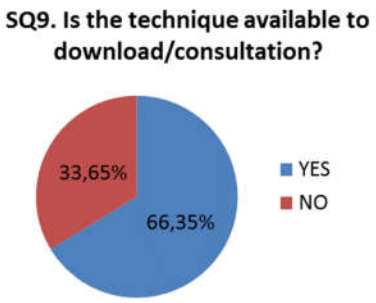


Figure 3.12 - Results of SQ9 (Technique Availability).

3.3.13. Platform Used

Regarding sub-question SQ10 (Platform Used), the results show that the majority of studies were carried out in the context of Desktop/Web platforms, representing 79.31% (see Figure 3.13). An example of a technique applied in a Desktop platform can be found in the work of Orfanou et al. (2015), who evaluated Moodle and Eclass LMS (the latter based on Claroline LMS), through the application of the SUS questionnaire. The authors assessed the validity of the questionnaire and the correlation of the questionnaire score with several factors such as gender, age and frequency of use of the LMS.

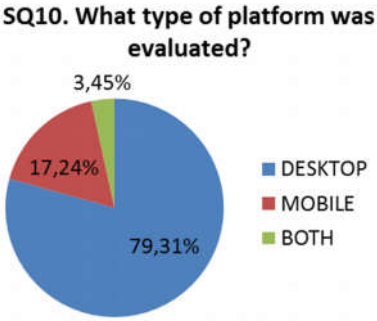


Figure 3.13 - Results of SQ10 (Platform Used).

Studies carried out in the context of mobile platforms represented 17.24% of the publications. An example of a technique applied in the context of mobile platforms can be found in the work of Fetaji et al. (2008), which evaluated a mobile application called MobileView. The authors carried out the evaluation by using a framework for usability evaluation developed by their own, called MLUAT (Mobile Learning Usability Attribute Testing), which uses heuristic evaluation, usability testing and questionnaires.

About 3% of the studies were performed in both mobile and desktop contexts, comparing usability problems in relation to the interaction. An example of a technique applied

in both contexts can be found in the work of Sánchez-Chamochin et al. (2008), who conducted usability tests of a mobile version for a legacy e-learning platform with tasks performed by both mobile device and desktop.

3.3.14. Techniques Comparison

Although several techniques were identified, only 6 publications (10,34%) presented a comparison between them. Lanzilotti et al. (2011), for instance, 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%.

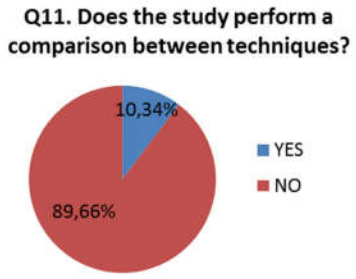


Figure 3.14 - Results of SQ11 (Techniques Comparison).

3.3.15. Results From the Identified Literature Reviews

The systematic mapping returned three reviews of the literature: Freire et al. (2012), Cota et al. (2014) and Navarro et al. (2016).

Freire et al. (2012) presented a review on the relationship between ergonomics and usability in the e-learning context. The authors identified three major differences among usability evaluation methods. Most methods have been proposed in general contexts. Others were adapted from already consolidated methods (such as usability tests) with questionnaires and semi-structured interviews. There were also methods composed by a mixture of methods, methodologies and techniques from the fields of Participatory Design, Interaction Design and User-Centered Design. The authors conclude that although usability studies have evolved considerably, new research on the most complete and quick forms of evaluation is still needed,

considering not only the integration of methods with stakeholders, but also the feasibility of such integration.

The study by Cota et al. (2014) was conducted through a systematic mapping of the literature, with the goal of evaluating and interpreting all articles relevant to mobile usability and m-learning. The authors analyzed a total of 68 publications and classified them according to four categories: (i) m-learning applications, (ii) guidelines and frameworks, (iii) specific aspects of m-learning, and (iv) analyzes and trends of m-learning. The results indicated a lack of guidelines, frameworks or tools designed to evaluate educational factors and usability in m-learning applications. Thus, the authors proposed an initial model of a framework to develop and evaluate m-learning applications, considering usability, learning aspects and student experience.

The work of Navarro et al. (2016) is an update of the systematic mapping conducted in 2014, described earlier. The study aimed to understand the trends and areas of opportunity in m-learning. The authors ranked the articles in: m-learning apps, guidelines/frameworks, specific aspects of m-learning, case studies, and m-learning trends and analyzes. With the classification of the techniques, the authors identified only two frameworks developed to evaluate mobile devices in relation to m-learning. Thus, they refined the framework proposed in Cota et al. (2014), dividing it into two categories: pedagogical usability and usability of the user interface.

3.4. Conclusion

This chapter presented the results of a systematic mapping of the literature to identify usability and UX evaluation techniques in the context of Learning Management Systems. From an initial set of 177 publications, 62 were selected after the first and second filters.

The results of this mapping were published in Nakamura et al. (2017a), contributing to the scientific community through an overview of the usability and UX evaluation techniques used in the context of LMSs. The identification of these techniques enables a synthesis of its main characteristics and applications, assisting in the improvement of those already existing or in the development of new techniques to promote a better usability and user experience.

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:

- No evidence has been found of techniques that provide the evaluator with feedback on suggestions for correcting UX/usability problems, but only to identify them;
- There was no consensus on which learning factors should be considered in the evalation process;
- Few studies were conducted in the m-learning context. Among the 12 studies found, only 05 presented new techniques (S12, S34, S35, S38 and S61). Among these studies, 04 different frameworks were presented and only 01 was evaluated experimentally (framework used in the S12 and S35 studies);
- Despite the large number of techniques identified, only six studies compared them (S06, S21, S27, S31, S32 and S35), not being possible to identify the most appropriate technique for the e-learning context. Another limiting factor is the lack of availability of the technique for consultation in most of the studies, which makes it difficult to perform this type of comparison;
- Few techniques perform the usability/UX evaluation process automatically. Most of the techniques are performed manually, which may require more effort by the researcher in the evaluation process;
- Among the new techniques, none considered the UX analysis in greater depth. Some studies considered motivational factors; however, they were limited to this scope (S06, S16, S17, S54, S58). Other techniques involved aspects related only to user satisfaction. Although some aspects of UX can be captured, none of them assess, for example, the user's feelings and their change over time;
- There were no studies relating the influence of usability and UX improvement in the learning process. Study S52 evaluated learning by comparing one group using a tutorial module developed for Moodle to perform content revisions and another group using the traditional (manual) method. The results showed that there was a significant difference in the learning of the first group, however, they are not related to the improvement of usability/UX of the platform Moodle, but to the use of the platform in conjunction with the module itself, compared to the manual method.

Therefore, although many studies have been conducted to evaluate LMSs, there are still several gaps to be filled. For this work, we carried out a preliminary study (Chapter 4) empirical studies between usability and UX evaluation techniques, aiming to contribute with empirical evidence regarding the existing techniques in the literature. With the results obtained from these studies and based on the body of knowledge acquired through the systematic mapping of the

literature, we started the development of a new technique to evaluate the usability and UX of LMSs (Chapter 5).

CHAPTER 4 – PRELIMINARY STUDY OF UX EVALUATION TECHNIQUES

In this chapter, we present a preliminary study carried out to evaluate two UX evaluation techniques. The results of this study made it possible to obtain an overview of the techniques and to identify opportunities for their improvement. Additionally, the results served as a starting point to the definition of the UX dimension of TUXEL.

4.1. Introduction

As the number of universities which use LMSs grows, researches in terms of principles related to human computer interaction, such as User eXperience (UX), have attracted considerable interest (Harrati et al., 2016). It is important that a LMS also provides a positive UX while being usable and facilitating learning. Failure to comply with these quality aspects may result in dissatisfaction, misunderstanding or improper use of these platforms, generating several criticisms and low acceptance (Van Der Linden and Van De Leemput, 2015).

Although there are several UX evaluation techniques, the results of our systematic mapping, presented in Chapter 3, indicated no evidence of techniques that consider the specifics for evaluating LMSs (Nakamura et al., 2017a). Researchers have been using generic techniques (i.e., techniques designed to evaluate the UX of software applications in general, not specific for evaluating LMSs) to evaluate these platforms. Consequently, there is a need for more empirical evaluations of existing UX evaluation techniques. By doing so, it may be possible to verify their suitability for understanding the experiences that are conveyed by LMSs, while allowing adapting them, if necessary.

This chapter presents the beginning of the Design Cycle of the Design Science Research (DSR). In Design Cycle, an artifact is proposed to interact with the problem context in order to improve something in that context (Wieringa, 2014). The development of this artifact may be driven by theoretical assumptions, while the results of the application of this artifact on the problem context may corroborate or question the validity of these conjectures (Hevner and Chatterjee, 2010). Regarding UX, we needed to identify whether these generic techniques are enough for students to convey their UX when using LMSs. To do so, we carried out an empirical study, aiming to corroborate or question the theoretical conjectures regarding the application of these techniques to evaluate the UX of software applications, specifically LMSs.

This chapter presents a study conducted with two UX evaluation techniques: User Experience Questionnaire – UEQ (Laugwitz et al., 2008) and Integrated Experience Acceptance Model – IEAM (Van Schaik and Ling, 2011), which was published in Nakamura et al. (2017b). We chose these techniques after a selection process, detailed in Section 4.3. Our goal was to compare these two techniques and verify whether they allow learners to fully convey their UX when using a LMS. To do so, we evaluated the UX of a LMS called Edmodo. We gathered participants' difficulties when performing the tasks in Edmodo and correlated them to learning issues. We obtained their perception about the technique they used regarding aspects such as ease of use, usefulness and intention to use the technique, which are some of the aspects that influence the acceptance of a technology (Venkatesh and Bala, 2008). Additionally, we identified which technique each participant preferred and the reasons for that choice. We also obtained their suggestions for improving the techniques and performed a qualitative analysis on their perceptions about the techniques. With this study, we provide researchers with information on these techniques, making it possible to improve them or create new techniques considering their strengths and weaknesses, while the identified UX problems faced by the students may contribute to the improvement of Edmodo's quality. Additionally, the results served as the basis for the development of the UX part of TUXEL.

4.2. Evaluated Techniques

In this preliminary study we compared two UX evaluation techniques: UEQ and IEAM, which were chosen after a selection process. These techniques aim to evaluate the UX of products regarding Pragmatic Quality (PQ) dimension (goal oriented) and Hedonic Quality (HQ) dimension (pleasure oriented). The HQ dimension is sub-divided in Hedonic Identification (HQ/I) and Hedonic Stimulation (HQ/S). The HQ/I relates to how user identifies with the product, i.e., how the product allows the user to express himself to others (Hassenzahl, 2003). People, for instance, may prefer products that communicate advantageous identities to others, such as products that are in fashion or developed by famous brands. The HQ/S, in turn, relates to how much the product stimulates the user with "novel, interesting or even exciting functionality, content, presentation or interaction style" (Hassenzahl, 2003).

UEQ is composed by a 7-point semantic differential scale where the users should mark the point that is closest to the adjective that better describes their UX. The technique uses 26 adjectives to evaluate 6 factors: (i) attractiveness, (ii) perspicuity, (iii) efficiency, (iv) dependability, (v) stimulation and (vi) novelty. Perspicuity, efficiency and dependability

attributes evaluates the PQ dimension, while stimulation and novelty evaluates the HQ dimension. UEQ also evaluates the Attractiveness (ATT), which relates to the general impression towards a product, evaluated by the attractiveness attribute. The HQ dimension evaluated by UEQ relates only to HQ/S dimension.

IEAM is composed by two parts. The first part evaluates the PQ, HQ/I and HQ/S dimensions, besides the Beauty and Goodness of a product. Pairs of adjectives and a 7-point semantic differential scale compose each dimension, similar to the UEQ. The second part is based on the Technology Acceptance Model (TAM) developed by Davis et al. (1989) and is composed by a 7-point Likert scale, aiming to assess participant's level of accordance for each affirmative regarding: (i) Perceived Enjoyment (PE), (ii) Perceived Ease Of Use (PEOU) and (iii) Perceived Usefulness (PU).

We translated both techniques to Brazilian Portuguese and two researchers reviewed them. The Cronbach Alpha indicated high internal consistency values for both translated techniques, with $\alpha > 0.7$.

4.3. Technique Selection Process

In Chapter 3, we conducted a systematic mapping to identify the usability and UX evaluation techniques that have been applied to evaluate LMSs. Given that we did not identify specific techniques for evaluating UX of LMSs, we performed a research to identify the techniques available from other contexts. To do so, we used the work of Rivero and Conte (2017) as a starting point. It contains a list of 227 publications returned from a systematic mapping conducted to identify methods, techniques and tools that have been proposed to evaluate the UX of software applications. The authors classified the publications according to a set of 9 research sub-questions and their responses (see Table 4.1).

We formulated five Exclusion Criteria (EC) according to the purposes of our study (see Table 4.2). Due to the large number of participants and the restricted time to carry out the evaluation, also considering that we wanted to identify techniques that are easy to apply and with low cost, we did not consider techniques that need a moderator or some type of monitoring tool (EC1). Since the focus of our study is on the perceptions of the UX after the use of a functional LMS from the users' point of view, we did not consider techniques that do not obtain data from users (EC2), not evaluate functional prototypes or final applications (EC4) and not perform the UX evaluation after the use (EC6). Finally, we did not consider publications which techniques are designed to a specific context (EC3), e.g., medical, journalism, etc. or were not

available for consultation (EC5). We analyzed these techniques to verify whether we can apply them to evaluate a LMS or not.

Table 4.1 - Research sub-questions from the work of Rivero and Conte (2017).

Sub-question	Responses
SQ1 – Type of technology	(a) written report (b) verbal report (c) observation/monitoring
SQ2 – Data origin	(a) users (b) development team (c) UX specialists
SQ3 – Local	(a) controlled environment (b) field
SQ4 – Type of evaluated application	(a) generic (b) web application (c) mobile application (d) others
SQ5 – Type of evaluated artifact	(a) conceptual ideas (b) project models (c) functional prototypes or final applications
SQ6 – UX evaluation period	(a) before use (b) during use (single episode) (c) during use (long-term) (d) after use
SQ7 – Collected data	(a) qualitative (b) quantitative (c) both
SQ8 – Support for correcting identified problems	(a) yes (b) no
SQ9 – Availability	(a) available for free (b) available under licence (c) not available

We performed two refinement steps. In the first refinement, we applied the publications to all the criteria, except EC3, whose application depended on the reading of the entire publication to identify the context of the artifact evaluated. We applied the first refinement directly in the categorization presented in the work of Rivero and Conte (2017). From the 227 initial publications, 170 were excluded based on the exclusion criteria, resulting in 57 accepted publications and a total of 50 unique identified techniques. Details of this refinement is presented in Appendix E.

Table 4.2 - Exclusion criteria defined based on the sub-questions used by Rivero and Conte (2017).

Base Sub-Question	Criteria	Description
SQ1	**EC1**	Publications whose techniques are conducted exclusively verbal or through monitoring.
SQ2	**EC2**	Publications whose data obtained by the techniques are not originated by users.
SQ4	**EC3**	Publications whose techniques were context specific.
SQ5	**EC4**	Publications whose applied techniques were not applied for the evaluation of functional prototypes or final products.
SQ9	**EC5**	Publications whose techniques are not available for free.
SQ6	**EC6**	Publications whose techniques did not perform UX evaluation after use.

In the second refinement, we analyzed and evaluated these techniques regarding aspects such as viability, ease of use, validity of the instrument, evaluated factors, among others. Regarding feasibility, for instance, some techniques were not applicable to our study, since they were specific to evaluate educational games or needed specific equipment to perform the evaluation. With respect to availability, some techniques did not provide the final applicable questionnaire. At the end of the analysis, we selected two techniques: UEQ (Laugwitz et al., 2008) and IEAM (Van Schaik and Ling, 2011). Details of this second refinement are presented in Appendix F.

4.4. Definition of the Evaluated LMS

We chose Edmodo LMS given that it is one of the most popular LMS, with more than 90 million users around the world. The author of this dissertation also identified difficulties of some students while submitting or performing activities through the platform in a class where he acted as an administrator of the environment. Furthermore, there was no evidence of studies about the evaluation of Edmodo in our systematic mapping.

4.5. Participants and Materials

We carried out the study with 34 students from Federal University of Amazonas (UFAM), enrolled in Human-Computer Interaction class during the first semester of 2017. We used the following materials in this study: (i) an informed consent form, explaining the study, the participants' voluntariness and the confidentiality of their identities; (ii) a script with a set of tasks to be performed on Edmodo; (iii) the UX evaluation techniques (UEQ and IEAM); (iv) the Technology Acceptance Model 3 (TAM3) questionnaire proposed by Venkatesh and Bala

(2008), with additional questions; and (v) a preference questionnaire. The TAM3 questionnaire (see Table 4.3) consists of a set of items evaluated by a 7-point Likert scale to obtain participant's perceptions regarding Perceived Ease Of Use (PEOU), Perceived Usefulness (PU) and Behavioral Intention (BI). The preference questionnaire is composed by a single-choice question, with three options: (i) UEQ, (ii) IEAM and (iii) none of them, in addition to a field where the participant should explain the reason for that choice.

Table 4.3 - Overview of the TAM3 questionnaire applied in the preliminary study.

TAM3 Questionnaire		
Dimension	**ID**	**Question**
Perceived Usefulness	PU1	Using the technique improves my performance when evaluating the experience of Edmodo.
	PU2	Using the technique improves my productivity when evaluating the experience of Edmodo.
	PU3	Using the technique allows me to fully evaluate the experience of Edmodo.
	PU4	I find the technique useful for evaluating the experience of Edmodo.
Perceived Ease Of Use	PEOU1	The technique was clear and easy to understand.
	PEOU2	Using the technique did not require much mental effort.
	PEOU3	I find the technique easy to use.
	PEOU4	I find it easy to report the experience of Edmodo using the technique.
Intention to Use	IU1	Assuming that I have access to the technique, I plan to use it to evaluate the experience of a learning platform.
	IU2	Given that I have access to the technique, I predict that I would use it to evaluate the experience of a learning platform.
	IU3	I intend to use the technique to evaluate the experience of a learning platform next month.
Additional Open-ended Questions		
1- Did you have any pair of adjectives that you did not understand or considered not applicable in this context? Which ones?		
2- Would you add any pair of adjectives to better describe your experience?		
3- Did you feel able to fully evaluate your experience with Edmodo using the technique?		
4- What was easy when applying the technique?		
5- What was difficult when applying the technique?		
6- What would you change to improve the application of the technique?		

4.6. Execution

Two days before the study, we provided a brief introduction on Edmodo to participants and informed them about the study. All participants signed the informed consent form. We divided them in two groups. Considering that the participants may have different backgrounds that may cause undesired effects on the results, we provided a pre-test questionnaire in order to

characterize them (principle of balanced design). This questionnaire contains questions regarding: (i) prior use of Edmodo, (ii) frequency of use of LMSs and (iii) knowledge level about usability/UX evaluations. We divided the participants in blocks according to their experience. From each block, we randomly assigned them to each group (see Table 4.4). Each group used only one technique.

Table 4.4 - Participants division according to the pre-test questionnaire.

Question	Answers	Participants	
		Group 1 (IEAM)	Group 2 (UEQ)
Prior use of Edmodo	Already used	2	3
	Never used	15	14
Experience with LMSs (frequency of use)	Several times a week	8	7
	Once a week	6	7
	Once a month	2	2
	Never used a LMS	1	1
Knowledge about usability/UX evaluation	Already performed this type of evaluation	2	2
	Already learned about it and did some class exercises	2	2
	Already read about it but not in depth	9	10
	Never heard about it	4	3
Total of participants		17	17

A day before the study, participants received, by e-mail, a script with a set of tasks to be performed on Edmodo. Given that Edmodo is Internet dependent and that the internal network of the institution is instable due to the high number of users, we decided that each participant would carry out the activities in their own home in order to avoid connectivity problems that could interfere on their experience of use. Therefore, we could not record the time spent by the participants during the accomplishment of the tasks.

The participants performed the following tasks in Edmodo: (i) register on the platform; (ii) join the group of the discipline through provided access code; (iii) change profile photo; (iv) download and read two available content; (v) perform tasks related to each content; (vi) perform an assessment activity. We chose these tasks since they reflect the main activities performed by learners in Edmodo. The last question in the assessment activity was an open-ended question where participants could describe their difficulties when performing the tasks in Edmodo, given that the selected techniques do not identify the difficulties faced by the participants. UEQ, for instance, aims to obtain “feelings, impressions, and attitudes that arise when experiencing the product” (Laugwitz et al., 2008), while IEAM aims to gather user’s

interaction experience and technology acceptance (Van Shaik and Ling, 2011). None of them focuses on evaluating the accomplishment of the tasks.

On the study day, the participants went to a room according to the group they were assigned. Each group evaluated Edmodo using only one technique. We recorded the time spent by each participant in the evaluation process in order to measure the average time necessary to employ the techniques. After the evaluation, the participants received the TAM3 questionnaire.

A day after the study, we applied the preference questionnaire. The goal was to identify which technique the participants would prefer and the reasons for this preference. We decided to apply this questionnaire a day after the evaluation in order to allow the participants to better think about the technique they used. First, we performed a brief presentation about both techniques, in addition to explaining to the participants that we divided them into two groups, each group using only one technique. After, we provided them the techniques. The group that used UEQ received IEAM and vice versa. The participants analyzed the technique for a while and then answered the preference questionnaire.

4.7. Results

In this section, we present the results of this study. Questionnaires composed by open-ended questions were analyzed qualitatively in order to better interpret the results. In the qualitative analysis, we performed an open coding (Corbin and Strauss, 2014) by analyzing the participants' answers and creating codes that represents the concepts identified in their answers. We divided the results in three subsections: (i) difficulties in Edmodo, (ii) UX evaluation of Edmodo, and (iii) evaluation of the techniques.

4.7.1. Results Regarding Difficulties in Edmodo

Results from the open-ended question in the assessment activity provided in Edmodo revealed that, within the 34 participants, 15 faced some difficulties during the execution of the tasks on Edmodo, giving a total of 12 unique reported difficulties. We identified five codes during the qualitative analysis: (1) lack of instructions to perform the matching task; (2) not intuitive navigation; (3) confusing interface; (4) lack of proper feedback; and (5) confusing language.

The first code indicates that **Edmodo does not provide sufficient instructions to perform the matching task**. This issue was indicated by 8 out of the 15 participants that reported difficulties. The matching task consisted in dragging the options on the right column to order them according to the content of the left column. Many participants stated that they

wasted too much time trying to perform this task. There were even participants that submitted the quiz without finishing this task due to the lack of information about how to perform it. Participant P22, for instance, stated, “I could not match the answers from the first column in the second column, it does not explain how it should be done, I tried everything and I could not do it, so I submitted it without solving the question”. This may affect the learning process, since it could cause a cognitive load to the learner. According to Ardito et al. (2006), if the platform does not provide a good usability, the learner will spend more time trying to understand how to use the platform then learning the educational content. Furthermore, teachers may not identify whether the students did not answer correctly because they did not really know the right answer or did not understand how to perform the task.

The second code relates to the difficulty in going to some pages, indicating that **the navigation of Edmodo is not intuitive enough**. Some participants had to navigate through different pages until they get to the desired page. Participant P1, for instance, stated, “I had difficulty to go back to the home page of the group to look for the next tasks […] I had to click on the ‘home’ icon to go back to the home page of Edmodo, access the group of the course and look for the tasks I needed to accomplish”.

The third code relates to participants’ difficulty in finding resources in the platform, e.g., course materials and activities, indicating that **the interface of Edmodo is confusing**. Regarding course materials, for example, participants reported difficulty in finding these materials, making them waste time during the tasks. Participant P15, for instance, stated: “It took me a long time to figure out where was the material to perform the tasks”. This difficulty may also affect the learning process, since it may cause cognitive load to learners when searching for the materials. Moreover, the learners may look for alternative content on the Internet and find wrong or inaccurate information. Participant P21, for instance, stated, “I could not interpret what I found on the Internet to answer the questions correctly”. Indeed, this participant answered three questions incorrectly and one question partially correct, from a total of nine questions. Although it is not possible to infer that the participant could have answered the questions correctly if he had consulted the materials in the group, it would have been easier for him to answer them.

There were also some comments related to the lack of adequate feedback from the platform (fourth code), indicating that **Edmodo does not provide sufficient feedback to learners**. Participant P12, for instance, stated, “In tasks 1 and 2, I was not sure whether the attached document was really delivered to the teacher, since the task of one of the students was visible to me right below the wording of the task, and mine was not”. Indeed, participant P19

attached the task files in the comment field of tasks 1 and 2. However, this participant did not report any difficulty when performing the tasks, i.e., he did not know that he did not submit the task in the right place, which may indicate that Edmodo, in fact, does not provide adequate feedback about the tasks. Learners may complain to the teacher, stating that they accomplished the task, when, actually, they attached it in the wrong place.

Finally, some participants complained about the language used in the platform (fifth category), indicating that **the language used in Edmodo confuses learners**. Participant P24 commented, "Edmodo mixes Portuguese with English and ends up confusing the user". Indeed, there are some words and sentences not translated to Brazilian Portuguese, which may cause difficulties for users who are not familiar with English language in finding or understanding information.

4.7.2. Results regarding UX evaluation of Edmodo

In this subsection, we will describe the results of the UX evaluation of Edmodo. Since the techniques evaluate some different aspects from each other, firstly we will present the results related to commonly evaluated aspects by both techniques: Pragmatic Quality (PQ) and Hedonic Quality Stimulation (HQ/S).

Figure 4.1 shows the results for the dimensions evaluated by: (i) both techniques, (ii) IEAM only, and (iii) UEQ only. Since a set of items composes these dimensions in order to capture a single concept (the dimension itself), we calculated the mean score of these items for each evaluated dimension (Sullivan and Artino Jr, 2013). The result is a number ranging from -3 (the most negative result) to 3 (the most positive). Scores smaller than -1 indicate a negative perception of the participants regarding this dimension. Scores between -1 and +1 indicate that the perception was neither positive nor negative. Finally, scores higher than +1 indicate a positive perception of the participants.

Results of the dimensions evaluated by both techniques indicate that Group 1 showed neutral to HQ/S, while Group 2 perceived it as positive. These neutral-positive scores may indicate that Edmodo fulfills the participants' needs regarding this dimension. However, there is room for improvements. For example, more interesting or exciting resources/functionalities can be developed in order to increase the stimulation provided by the platform. The results for PQ were similar to HQ/S. The Group 1 showed neutral, while Group 2 perceived it as positive. It may indicate that, in general, participants found Edmodo relatively easy to use. Results are in accordance with the additional questions, in which, within the 34 participants, 19 (around

56%) did not face any difficulty with Edmodo. However, 15 participants complained about one or more issues, which may indicate that there are opportunities to improve the platform.

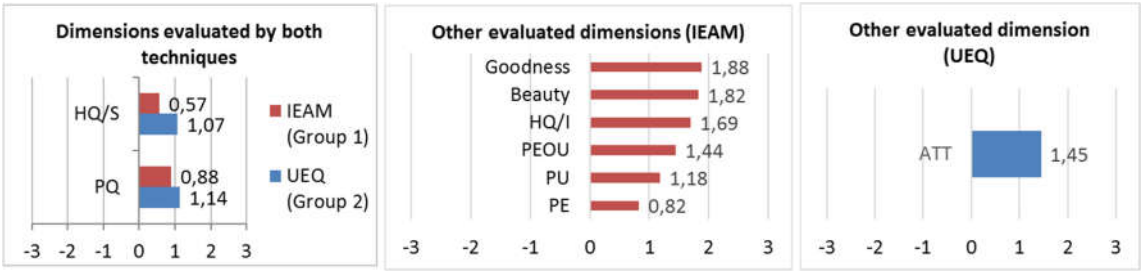


Figure 4.1 - Results from each evaluated dimension of the techniques.

Regarding the remaining dimensions evaluated by IEAM technique (Group 1), results showed that participants did not perceive Edmodo as either enjoyable or unenjoyable (PE = 0,82), which reflects the results from HQ/S dimension. On the other hand, they had a positive identification with the platform (HQ/I = 1,69). Despite their difficulties in performing some tasks, they found Edmodo easy to use (PEOU = 1,44) and useful (PU = 1,18). Additionally, the platform was also rated as good (Goodness = 1,88) and beautiful (Beauty = 1,82).

Regarding the other dimension evaluated by UEQ, the Attractiveness (ATT) was perceived as positive (ATT = 1,45). It means that although participants faced some difficulties, the platform was, in general, attractive to them.

In summary, the UX evaluation results showed that Edmodo provides a positive UX to users. However, there is still room for improvements, especially regarding task accomplishment aspects (PQ dimension) and stimulation aspects (HQ/S), in order to make the platform more intuitive and more interesting to learners.

4.7.3. Results regarding the evaluation of the techniques

In this subsection, we present the results of the evaluation of the techniques used in this study. We divided the results in two subsections: (i) TAM3 questionnaire; and (ii) preference questionnaire.

4.7.3.1. Results of the TAM3 questionnaire

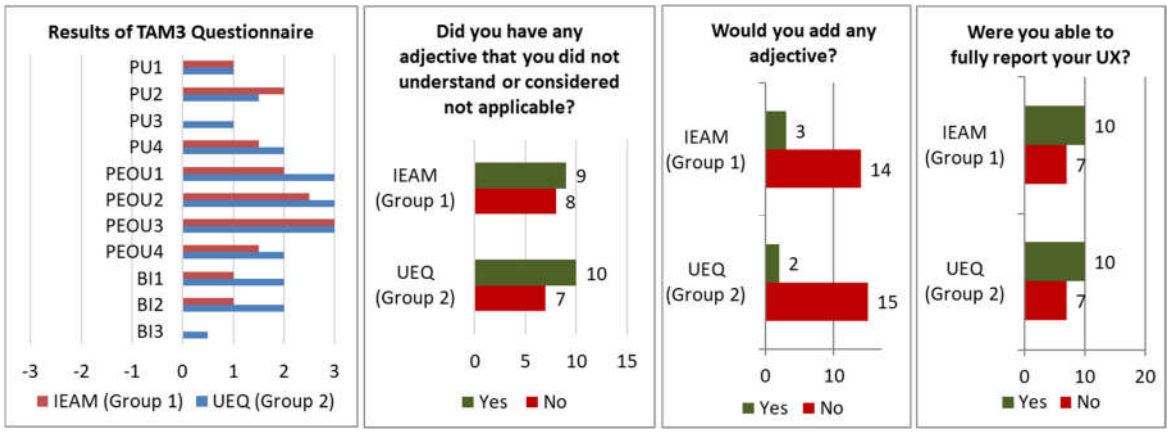
This section presents the results from the TAM3 questionnaire (Vankatesh and Bala, 2008), which aims to evaluate the level of acceptance of a given technology through attributes such as Perceived Ease Of Use (PEOU), Perceived Usefulness (PU) and Behavioral Intention (BI). We applied this questionnaire in order to evaluate the participants' level of acceptance regarding

the technique they used. We also added some questions in order to gather qualitative data about the techniques.

We calculated the median of each item of TAM3 questionnaire to compare the two evaluated techniques (Figure 4.3a). Regarding Perceived Usefulness (PU), participants considered that both techniques improve their performance a little when evaluating the UX of Edmodo (PU1). As for productivity (PU2), participants who used IEAM technique considered that it increases their productivity more than participants who used UEQ technique. On the other hand, participants using IEAM perceived it as neutral to fully evaluate the experience with Edmodo (PU3), while participants using UEQ perceived it as a little useful (PU3). Regarding overall usefulness to evaluate the UX of Edmodo, participants of both techniques perceived them as useful (PU4).

Regarding Perceived Ease Of Use (PEOU), both techniques were considered clear and easy to understand (PEOU1), with UEQ being perceived as clearer and easier than IEAM. The participants' responses also indicated that both techniques did not require much mental effort (PEOU2) and were very easy to use (PEOU3). Participants of both groups also perceived them as easy to report the experience with Edmodo (PEOU4).

Regarding Behavioral Intention, participants who used UEQ showed more intention to use it (BI1), while demonstrating a higher probability of using it (BI2) to evaluate the UX of a LMS than participants who used IEAM. However, participants of both groups did not express any intention to use the techniques so soon (IU3).



(a) (b) (c) (d)

Figure 4.2 - Results of TAM3 questionnaire and the additional open-ended questions.

Regarding the additional questions attached in TAM3 questionnaire (see Table 4.3), the results for question 1 (Figure 4.3b - "*Did you have any pair of adjectives that you did not understand or considered not applicable in this context?*") revealed that 9 participants from

IEAM and 10 participants from UEQ answered yes. Regarding IEAM, most participants complained that the pair "good/bad" is not applicable. Regarding UEQ, most participants complained about the similarity regarding some adjectives, e.g. "inventive-conventional" and "conservative-innovative", while other participants did not understand the meaning of "leading edge" adjective.

The results for question 2 (Figure 4.3c - "*Would you add any pair of adjectives to better describe your experience?*"), revealed that few participants would add more adjectives to the techniques. Three participants using IEAM and two using UEQ suggested the addition of pairs of adjectives. Regarding IEAM, the participants suggested adding: "easy to navigate/difficult to navigate", "intuitive/not intuitive" and "interactive/not interactive". Regarding UEQ, participants suggested adding: "intuitive/not intuitive" and "recommendable/not recommendable".

The results for question 3 (Figure 4.3d - "*Did you feel able to fully evaluate your experience with Edmodo using the technique?*") revealed that most of the participants was able to fully evaluate their UX. On the other hand, around 41% of the participants in both groups reported not being able to fully report their UX by using the techniques, indicating opportunities for improvements.

In order to better understand the reasons why some participants were able to fully report their UX and others not, in addition to identify their perceptions about the techniques (questions 3, 4, 5 and 6), we performed a qualitative analysis. We present the results of this analysis below. Regarding **participants' perceptions about UEQ technique**, we identified the following codes: (1) quick and easy; (2) adjectives convey the UX; (3) limited UX reporting; and (4) confusing adjectives.

The first code indicates that the semantic differential scale of UEQ makes the UX evaluation process quick and easy. Given this, participant P4 commented, "*applying rating on a scale eases reporting the emotion that I felt when using Edmodo*". Moreover, participant P14 stated, "*it was easy to choose the option of adjective and also, in the case of feeling neutral to a given pair of adjectives, there was an option for it*".

The second code reveals that the adjectives used by UEQ allowed participants to report their UX, covering the main feelings aroused in them when using Edmodo. Participant P12, for instance, stated, "*the adjectives reveal the feelings present during the use of the platform very well*". Participant P14 also highlighted that "*the pairs of adjectives address the most important features of the application, such as: ease of use, learnability, understanding, etc.*"

Although the above-mentioned codes reveal positive aspects of UEQ, there have also been some negative comments about the technique. Regarding third code, we identified that the UX evaluation of UEQ is limited due to the lack of a field for comments. Participant P8, for instance, stated, "*I could not describe clearly which tasks I had difficulty when using Edmodo*". Participant P5 also highlighted that *"[the technique does not allow me to describe] the reason for giving that answer. It is not possible to get the real cause for that answer*". Moreover, participant P13 commented, "*[the evaluation] seemed too broad; some functionalities were confusing, while others were not, for example*".

Finally, the fourth code reveals that some adjectives used by UEQ confuse learners during the evaluation. Participant P9 stated that he had difficulty with "*adjectives that he did not know the meaning (e.g. leading edge) […] [having to] try to infer [its meaning] by the opposite adjective*". Moreover, participant P12 commented, *"four pairs of adjectives [not understandable; unpredictable; usual; conservative] seemed to not fit very well in the context of use of the platform*".

Regarding **participants' perceptions about IEAM technique**, we identified the following codes: (1) quick and easy; (2) adjectives convey the UX; (3) covers the main topics; (4) difficulty in quantifying using scales; and (5) limited UX reporting.

The first code indicates that the scale-type approach used by IEAM makes the evaluation process quick and easy. Participant P28, for instance, stated, "*the adjectives used to evaluate the platform were already there, only needing to score each feature. The technique avoids the intellectual effort of having to think of a criterion to evaluate*".

The second code reveals that the adjectives used by IEAM allowed participants to express their UX. Participant P27, for instance, commented "*the adjectives, except those cited previously [Gaudy/Classy and Bad/Good], allowed me to immerse in the experience of use. While reading, I was able to associate them with my actions in Edmodo*".

In the third code, we identified that participants considered that IEAM covers the main topics regarding the UX evaluation of Edmodo. Participant P21 commented, *"[the technique] presented the main features of Edmodo [for being evaluated]*". Moreover, participant P30 stated, "*the questions were well elaborated and allowed me to evaluate the UX of the platform*".

Although some participants considered easy to evaluate the UX by using scales, others considered that quantifying the UX by using scales is difficult in IEAM (fourth code). Participant P30 commented, "*there are many divisions for the levels of satisfaction. There should be few levels and it would be interesting to express more right to the point opinions, such as 'liked' and 'disliked', objective opinions*". There were also some participants having

difficulty in differentiating "somewhat agree" from "agree". Participant P33, for instance, stated, "*it was difficult to decide between 'agree' and 'somewhat agree', both leaves a gap to detail the answer, have little difference [between them] and turned [the evaluation] difficult*".

The fifth code indicates that the UX evaluation of IEAM is limited due to the lack of a field for comments and the lack of questions more specific about the platform. Regarding the lack of a field for comments, participant P26 stated, "*Multiple choice questions without a field for descriptions leave out many things*". Participant P34 also highlighted, *"[I had difficulty in] evaluating an adjective in a broad way, since it can be applied to a given area and not in others, not allowing me to be specific*". Regarding the lack of specific questions about the platform, participant P19 stated, "*It could present more specific questions about the platform, questions that makes the experience of use clearer in this perspective*".

Regarding **participants' suggestions for improving the techniques**, the results were as follows. Suggestions from IEAM group: (i) reduce the scale to 5-point in order to avoid confusion within the terms "agree" and "strongly agree"; (ii) add a field to allow them to describe their difficulties and make comments; (iii) specify which feature is being evaluated by each adjective; (iv) split the neutral into "I do not want to answer" and "I do not know how to answer" options. Regarding UEQ, in general, participants suggested: (i) removing adjectives that are confusing or not suited for the context; (ii) removing similar adjectives; (iii) add a field for observations and comments.

Finally, in addition to these questions, we calculated the mean time needed to apply the techniques. The results indicated that the UX evaluation can be carried out very quickly through both techniques. However, participants using UEQ performed the evaluation faster (4 minutes) compared to participants using IEAM (7 minutes). It may be due to the second part of IEAM, which requires the participants to read the statements in order to answer the questions.

4.7.3.2. Results of the preference questionnaire

The preference questionnaire allowed us to identify which technique the participants would prefer to evaluate the UX of a LMS. Given that we applied it a day after the evaluation, not all the participants were present. Thus, two participants using UEQ and one participant using IEAM did not answer this questionnaire.

The results were as follows (Figure 3). Among the 15 participants who used UEQ, 4 stated that they would prefer to keep using UEQ to evaluate a LMS, while 11 demonstrated a preference to use IEAM. Regarding the group who used IEAM, among the 16 participants, 10

declared that they would prefer to keep using IEAM, 3 demonstrated their preference for UEQ, and 3 stated that they did not prefer any of the techniques.

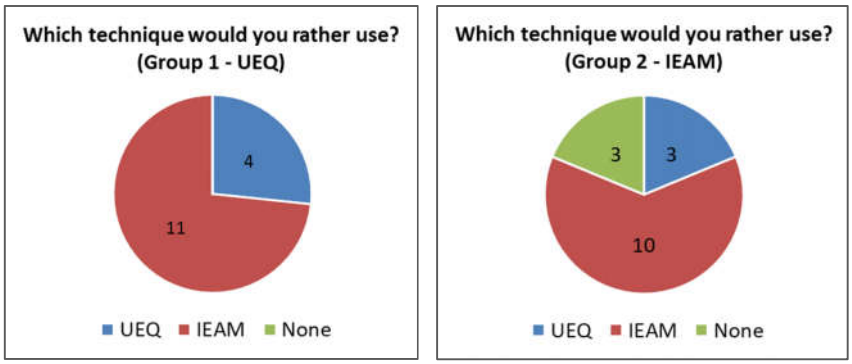


Figure 4.3 - Results of the preference questionnaire per group.

We identified the following **reasons for choosing UEQ**: (1) UEQ better qualifies the object being evaluated; (2) the highest number of adjectives in UEQ allows participants to better express their UX; (3) UEQ is more straightforward, practical and easy to understand. In order to exemplify these reasons, we present some quotations from the participants. For the first reason, participant P4 stated, "*IEAM is more succinct, but UEQ evaluates other aspects that IEAM does not address*". For the second reason, participant P6 commented, "*I think that the higher number of adjectives helps to better express the opinion*". Finally, the quotation from participant P13 exemplifies the third reason: "*In addition to being quick and practical, the additional pairs of adjectives make up the second part of IEAM*".

We identified the following **reasons for choosing IEAM**: (1) IEAM has a fewer number of adjectives; (2) IEAM is clearer and more objective; (3) the questions of IEAM evaluates the UX better. The following quotations exemplify these reasons. For the first reason, participant P1 commented, "*Straight away, it does not have excessive pairs of adjectives. In addition, it has specific questions related to application use*". For the second reason, we have the quotation from participant P11, "*IEAM is simpler and I consider that it allows me to answer what is needed (is more precise)*". Finally, for the third reason, participant P15 commented, "*the second section seems to better evaluate my experience and the first section has enough adjectives. UEQ extrapolates in adjectives*".

Among the participants, **three did not choose any of the techniques**. Participant P30 stated that "*both are unpleasant*" but did not specify why. Participant P33 commented that "*both have too much information*". Finally, participant P34 highlighted, "*both do not allow me to be specific in what exactly I am applying such an adjective, since it may be valid for one element [e.g. functionalities, resources, tasks, etc.] and not for another*".

With respect to the average time needed to employ the techniques, we calculated the average time spent by the participants per group and per experience with UX evaluation (Figure 4.5). The results indicated that the UX evaluation can be carried out very quickly through both techniques. However, participants using UEQ performed the evaluation faster (4 minutes) compared to participants using IEAM (7 minutes). It may be due to the second part of IEAM, which requires the participants to read the statements in order to answer the questions.

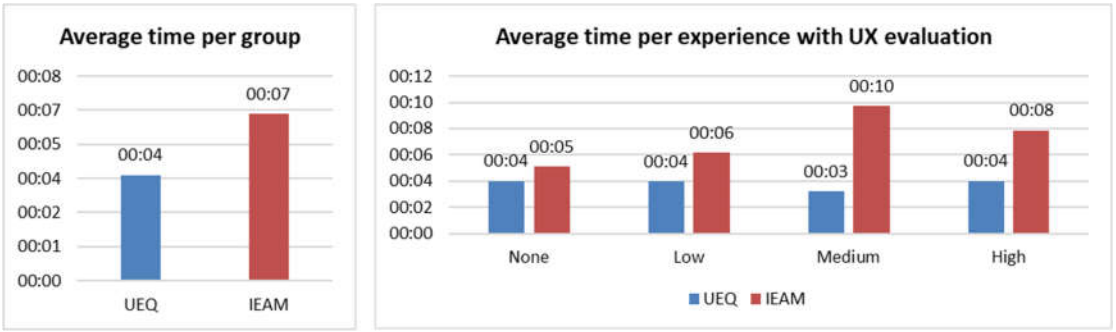


Figure 4.4 - Average time per group and knowledge level about UX evaluations.

The analysis per experience with UX evaluation revealed that the time spent by participants using UEQ remained stable. On the other hand, participants in IEAM group without any experience or low experience with UX evaluation spent less time than those with medium or high experience. It may also be due to the second part of IEAM. Participants with more experience may try to perform the evaluation more carefully, analyzing each statement in order to fully comprehend what is being evaluated by the technique.

4.8. Conclusion

This chapter presented an UX evaluation of a LMS called Edmodo in order to evaluate the adequacy of two generic UX evaluation techniques (UEQ and IEAM), chosen after a selection process. Our goal was to identify whether these techniques allows learners to fully evaluate their UX when using a LMS. We identified the participants' difficulties when performing the tasks in Edmodo, in addition to applying the TAM3 questionnaire with additional questions and the preference questionnaire in order to gather their perceptions on the techniques they used. Additionally, we performed qualitative analysis to better understand the results of the open-ended questions.

The qualitative analysis on the open-ended question regarding the difficulties faced by participants when performing the tasks on Edmodo revealed five codes. These codes address the main problems identified in Edmodo. As mentioned before, these problems may impair the teaching and learning process through the LMS. For instance, many students stated that they

spent too much time searching for the learning materials. Given this, one participant, who did not use the materials provided in the LMS, stated that he searched for the content on the Internet to answer the questions. In fact, this participant had his performance impaired, answering three questions incorrectly and one question partially correctly, from a total of nine questions. Although we cannot infer that he could have answered the questions correctly if he had consulted the materials on the group, it would have been easier for him to answer them. Similarly, many students did not understand how to perform the matching quiz correctly. This can directly impact the learning process, since it increases the cognitive load as students spend more time trying to understand how to use the LMS rather than learning the educational content (Ardito et al., 2006). Improvements may include positioning the files in locations that are more visible to the group and providing instructions in the matching quiz.

The lowest evaluation received by HQ/S dimension reveals that the platform does not motivate the students very much. Since "motivation and engagement are perhaps the most important elements of every form of learning experience" (Zaharias and Pappas, 2016), there is a need to provide more interesting features (e.g. gamification) in order to stimulate the users and make it possible to increase their engagement towards learning.

The results from TAM3 questionnaire indicated that there was no significant difference regarding their PEOU, PU and IU. Participants perceived both techniques as easy to use, however, not so useful and without intention to use them so soon.

Regarding additional questions, the results revealed that around 41% of the participants in both techniques were not able to fully evaluate the UX. It may explain the low score in PU3 item, related to how much the platform allows the learner to fully evaluate his UX. The most complained issue in both techniques was the lack of a field for comments. Participants commented that only using scales, without a field for comments, limits them reporting the UX conveyed by the LMS. They also reported that the techniques' evaluation is too broad, not allowing them to specify which features of the platform they are evaluating. Additionally, regarding UEQ, some participants considered some adjectives confusing, while others repetitive. Regarding IEAM, some participants reported difficulties in quantifying their experience in the Likert scale, e.g., between somewhat agree, agree and strongly agree. On the other hand, participants perceived the both techniques as easy and fast to perform the evaluation, specially the semantic differential scale.

When asked what technique they prefer, most participants chose IEAM. These participants stated that IEAM is more objective, has fewer adjectives and evaluates the UX better. This choice may be mostly due to the reduced number of adjectives and the second part

of the technique, which is composed by questions about the use of the platform, providing a more focused evaluation. On the other hand, some participants considered the evaluation method of UEQ easier and simpler, while providing a broader range of adjectives to express their UX.

According to these results, we identified some issues that researchers may consider when developing or improving UX evaluation techniques:

- **Let learners detail their experiences:** learners want to describe what they felt when using a product and what difficulties they had. Provide a field where participants can detail their experience;
- **Evaluate aspects related to the platform:** provide questions that focus on aspects related to the platform, for example, the level of instruction that the platform provides to learners during the tasks. Additionally, when using a semantic differential scale, provide more adjectives that relates to the use of the LMS, such as “interactive/not interactive”, “intuitive/not intuitive”;
- **Be specific or let learners specify:** when using semantic differential scales, specify which aspects are being evaluated by the adjectives or let learners explain their evaluation, since the LMS has many different features and an adjective may fit one, but not another;
- **Keep the evaluation quick and easy:** learners do not like to waste time. Use quick and straightforward evaluation methods, such as semantic differential scales;

Finally, the results of this study revealed limitations of these techniques when applied to the context of LMSs, contributing to add knowledge about the adequacy of the techniques. In addition, it served as a starting point to the development of the UX part of TUXEL, which we describe in next section. The findings also may contribute to the improvement of Edmodo and the development and improvement of UX evaluation techniques in the context of LMSs, making it possible to improve the quality of these platforms.

CHAPTER 5 – TUXEL: A TECHNIQUE FOR USER EXPERIENCE EVALUATION IN E-LEARNING

This chapter presents our proposal for the development of a technique to evaluate both usability and UX of Learning Management Systems, called TUXEL. Here we present the development process of the technique and the concepts and sources behind each evaluated dimension.

5.1. Introduction

In Chapter 3, we carried out a systematic mapping in order to define the context of the problem by identifying the techniques employed to evaluate usability and UX of LMSs. The results made it possible to identify the following gaps in this topic:

- **(G1) Lack of studies performing comparison among techniques:** due to this, there is no evidence of which technique is the most adequate to evaluate LMSs;
- **(G2) Few studies evaluated the UX of LMSs:** although some studies were classified as evaluating UX, most of them assessed only general aspects related to satisfaction, not measuring users' emotions and feelings. Furthermore, there was no evidence of studies that conducted UX evaluations by employing techniques that were specific to evaluate LMSs;
- **(G3) Lack of techniques that evaluates both usability and UX:** although we identified many usability evaluation techniques, there was no evidence of techniques evaluating both;

In order to fill the G2 and G3 gaps, we proposed a technique to evaluate both usability and UX from the learner's point of view. By integrating the usability and UX evaluation process under the learner's perspective, the technique may allow a more complete view regarding the quality of these platforms, while also filling the gap of lack of techniques to evaluate the UX of LMSs.

In Chapter 4, we started the Design Cycle of the Design Science Methodology (DSR) by carrying out a preliminary study to compare two UX evaluation techniques. We aimed to identify whether the UX evaluation techniques proposed in the literature are enough for learners to convey their UX when using a LMS. The results revealed some limitations of these questionnaire-based approach techniques using scales, indicating a need for improvements. This preliminary study and the two empirical studies that we will present in Chapter 6, allowed

us to gather further information about the existing techniques in order to obtain input for the development of TUXEL. The results of these studies may provide an overview of the techniques about attributes such as feasibility and effectiveness, as well as fill the gap of lack of studies that perform this type of comparison (G1).

Next section presents the artifact development phase of DSR, which consisted on the development of the first version of TUXEL.

5.2. TUXEL – Technique for User eXperience Evaluation in e-Learning

The Technique for User eXperience Evaluation in e-Learning (TUXEL) aims to evaluate the usability and UX of Learning Management Systems from the learner's perspective. Our main motivation to develop this technique relates to the increasing number of learning institutions and organizations that have been adopting LMSs to provide training and life-long education. As the number of educational institutions that use LMSs increases, research in LMSs have attracted considerable interest in order to evaluate how effective and how usable they are regarding principles of HCI, such as usability and UX (Harrati et al., 2016).

TUXEL

Name:_____

Start time: _____:_____

Labels
1- Totally disagree / 2- Partially disagree / 3- Neither agree nor disagree / 4- Partially agree / 5- Totally agree

ITEM	NAVIGATION					
N1	The options in the LMS's navigation menus are visible and easy to remember.	1	2	3	4	5
N2	I can easily know where I am and where I can go in LMS.	1	2	3	4	5
N3	I can easily go anywhere on LMS.	1	2	3	4	5
N4	The information is organized hierarchically to facilitate the navigation.	1	2	3	4	5
N5	I can easily identify the links I've visited.	1	2	3	4	5
N6	The LMS provides a search engine for finding information.	1	2	3	4	5
ITEM	**LEARNABILITY**					
L1	The LMS is intuitive enough that I can carry out the activities without difficulties.	1	2	3	4	5
L2	The instructions provided by the LMS are clear and objective.	1	2	3	4	5

Figure 5.1 - Part of the TUXEL 1.0 questionnaire.

Since learners are the final users of a LMS and that we wanted to evaluate its usability and UX from their point of view, we developed TUXEL to be used by learners who already use the LMS, regardless their level of knowledge about usability/UX evaluation. To achieve this, the technique needed to be quick and easy to use. Thus, we decided to develop TUXEL in form of a questionnaire (Figure 5.1). Although the results from the first empirical study have pointed out some drawbacks of questionnaires using scales, we decided to use this approach due to their

benefits of gathering the perceptions from the user's point of view, while being fast, easy and with low cost. Additionally, we wanted to verify whether these questions are enough to allow users to evaluate the usability of the LMS.

The use of questionnaires also makes it not necessary to perform any type of user testing prior to its application. The studies carried out by Zaharias and Poylymenakou (2009), Koohang (2004) and Alkhattabi (2015), for instance, employed questionnaire-based techniques with learners without performing any tasks before the evaluation. The only requirement is that the learners are using the platform. A drawback in this approach is that learners that are already using the platform for a long time may not have many difficulties as learners that are using the platform for a few days or weeks. However, TUXEL has some items that evaluate activities and/or elements that participants may not normally perform or notice throughout the course, for instance, the feedback provided by the LMS during assessments and the learner's progress. Therefore, carrying out the evaluation with a larger sample and with different profiles may provide better results. The next subsection describes the dimensions that compose TUXEL.

5.2.1. TUXEL Dimensions

The first step to develop TUXEL was defining the dimensions that it evaluates. As stated before, the development of the artifact in DSR may be driven by theoretical assumptions, while the application of this artifact on the problem context may produce new knowledge that can corroborate or question these conjectures (Hevner and Chatterjee, 2010). Given this, we searched for the dimensions evaluated by the techniques returned in our systematic mapping presented in Chapter 3.

We identified two main dimensions: general usability and pedagogical usability. According to Mtebe and Kissaka (2015), due to the difficulty in evaluating the usability of LMSs and their specificities, many authors consolidated general interface usability with pedagogical aspects. Some studies, as those conducted by Ssemugabi and De Villiers (2007), Zaharias and Koutsabasis (2011) and Mtebe and Kissaka (2015), have demonstrated that taking pedagogical aspects into consideration lead to the identification of several problems that were not addressed by general usability heuristics. Therefore, it is important that the technique encompass both general usability and pedagogical usability.

In addition to usability and pedagogical usability, we also included the UX dimension. According to Zaharias and Pappas (2016), technology innovations, combined with a heterogeneous user population has led to a set of new types of interaction, changing the user's expectations and demands. Users are not interested in an environment that is only easy to use,

but that amuses and entertains themselves (Petrie and Bevan, 2009). Regarding educational games, for instance, Ardito et al. (2007) identified that motivation and engagement are fundamental principles to provide a significant user experience. Zaharias and Pappas (2016) corroborates with this statement, bringing it to a broader context, stating "motivation and engagement are perhaps the most important elements of every form of learning experience" (Zaharias and Pappas, 2016). Given that, it is not only important to evaluate aspects oriented to task accomplishment, but also to address aspects related to feelings and emotions, such as motivation and engagement.

Therefore, we defined three dimensions for TUXEL: usability, pedagogical usability and UX. In the next subsections, we detail each of these dimensions.

5.2.1.1. General Usability

This dimension contains general usability criteria adapted for the context of LMSs. It is composed by nine categories, which were extracted from several works such as Mtebe and Kissaka (2015), Navarro et al. (2016), Nielsen (1994), Ogunbase (2014), Squires and Preece (1999), Ssemugabi and De Villiers (2006) and Zaharias (2009). One researcher categorized the items and two researchers, one of them usability expert and the other expert in informatics in education, reviewed the categorization.

Navigation (N): it has been an important aspect to be considered in LMSs (Zaharias, 2009). A LMS should be designed to facilitate the comprehension of content structure and its organization in order to support learner's navigation in the platform.

Learnability (L): relates to the ease with which a product enables the user to acquire essential skills to accomplish a task (Nokelainen, 2006). A LMS should be easy to be learned so that the student begins to study and perform activities as soon as possible (Ogunbase, 2014). Learners should be able, e.g., to accomplish certain tasks, locate information and remember how to use the platform after periods without using it (Zaharias, 2009).

Consistency (C): it is important to enhance learning efficiency (Zaharias, 2009). A consistent approach for layout and navigation allows learners to quickly recognize the design and predict the location of the information through different pages.

Visual Design (VD): a good design is important to enhance usability, minimize cognitive load and make the platform more appealing. The design should be structured and planned to allow learners to find information quickly and effectively (Shiratuddin et al., 2003).

Personalization (P): consists in the modification of the platform according to learner's preferences. Personalization is important to provide flexibility and maximize learner's efficiency (Navarro et al., 2016).

LMS Feedback (LF): it is important that learners know what is happening in the platform by providing enough feedback in order to prevent them from making mistakes (Navarro et al., 2016). The platform may, for instance, show a progress bar in time-consuming tasks or highlight mandatory fields that were left unfilled.

Help and Documentation (HD): the platform should provide online help and documentation that are easily retrievable and meet the learner's needs.

Learner Control (LC): considering that learners might make mistakes, it is important that the platform provide them facilities to get out of unwanted situations such as undo and redo operations (Nielsen 1994).

Error Prevention (EP): the platform should prevent users to make common mistakes and allow correcting them as well (Nielsen, 1994).

5.2.1.2. Pedagogical Usability

The concept of pedagogical usability is related to the assumption that the designers of the learning platform were "guided by either a conscious or subconscious idea of how the functions of the system facilitate the learning of the material it is delivering" (Nokelainen, 2006). Thus, the evaluation of pedagogical usability aims to verify how extent the platform was designed to facilitate the learning process.

In order to better understand the concept of pedagogical usability and its dimensions, we first identified, in our systematic mapping, the publications that considered pedagogical usability. After identifying these publications, we carried out a process known as backward snowballing, which consists in "using the reference list to identify new papers to include" (Wohlin, 2014). We identified the following publications: Silius and Tervakari (2003), Silius et al. (2003), and Nokelainen (2006).

The publication of Silius and Tervakari (2003) and Silius et al. (2003) defines pedagogical usability to denote whether the tools, content, interface and tasks of web-based learning environments support various learners to learn in a variety of learning contexts, according to the pedagogical objectives. The authors present a tool to evaluate the usefulness of web-based learning environments, proposing that usefulness is a combination of usability and utility, in which the last is also divided in two categories: pedagogical usability and value

added. However, they do not provide details regarding each of the sub-categories. Additionally, the tool is not available for consultation.

The work of Nokelainen (2006) was one of the most cited by the papers that considered pedagogical usability in our systematic mapping. The author provides a set of criteria to evaluate digital learning materials through 10 dimensions: (i) learner control; (ii) learner activity; (iii) cooperative/collaborative learning; (iv) goal orientation; (v) applicability; (vi) added value; (vii) motivation; (viii) valuation of previous knowledge; (iv) flexibility; and (x) feedback. The paper presents the concept behind each evaluated dimension, in addition to providing the questions that compose these dimensions. Thus, we decided to use this work as our starting point.

The pedagogical usability dimension of TUXEL is based on the 10 dimensions proposed by Nokelainen (2006), presented above. However, three of these dimensions (Learner activity, Applicability and Valuation of previous knowledge) were related to the evaluation of the learning material provided by the teacher. Since the learning material is dependent of the teacher and not of the LMS itself, we did not consider them. The selection and categorization process of the items for each category was the same of that performed for the general usability dimension.

Collaborative Learning (CL): learners are migrating from the perspective of acquiring knowledge to the perspective of participation. Instead of acquiring knowledge by their own, they are building it by practice as members of a community (Barab and Duffy, 2000; Lave and Wenger, 1991). In this way, it is important that the platform provides tools that allow communication and interaction among learners and teachers in order to build knowledge collaboratively (Nokelainen, 2006).

Goal Orientation (GO): as learning is a goal-oriented task, the platform should provide clear learning objectives (Nokelainen, 2006). The platform may provide, for instance, the learner's score related to a particular activity or his/her overall performance.

Instructional Assessment (IA): allows learners (and teachers) verifying whether the learning goals are being achieved or not (Zaharias, 2009). It is important that these platforms allow learners to perform this type of assessment in order to verify their learning level.

Added Value (AV): by introducing technological resources in learning, it is expected that they add value to the learning process compared to, for example, printed materials (Nokelainen, 2006).

Motivation to Learn (ML): e-learning usually has a higher dropout compared to traditional learning approach (Zaharias, 2009). One of the main factors is the motivation to

learn. Motivation affects the whole learning and making people behave the way they do (Nokelainen, 2006). It is important that the platform motivates students during the learning process.

LMS Flexibility (LFL): the LMS should be flexible to suit learners' interests, their knowledge level and their individual differences (Nokelainen, 2006).

Instructional Feedback (IF): one of the benefits of a human-computer interaction is the possibility to obtain immediate feedback. The LMS should provide feedback that supports the learner to become aware of the problematic part of his/her learning (Nokelainen, 2006).

5.2.1.3. User eXperience

This dimension aims to evaluate learner's perceptions regarding his/her experience using the LMS. Since there was little evidence regarding UX evaluation of LMSs and considering that generic techniques have been used to evaluate these platforms, we conducted a comparative study prior to the development of the technique.

In Chapter 4, we performed a comparison between two generic UX evaluation techniques, chosen after a selection process, in order to verify their applicability to evaluate a LMS called Edmodo: User Experience Questionnaire – UEQ (Laugwitz et al., 2008) and Integrated Experience Acceptance Model – IEAM (Van Schaik and Ling, 2011). The results indicated that around 41% of the participants in both groups could not fully evaluate their experience of use by using these techniques. Participants stated that the absence of a field to describe their difficulties and experience limited their evaluation. They also stated that some adjectives were not applicable to evaluating LMS, while others were confusing to understand or repetitive. On the other hand, both techniques were perceived as fast and easy to use, where participants using UEQ considered it a little easier than IEAM and had a little more intention to use the technique in the future. The results of this study provided us insights regarding how we could evaluate the UX of the learners.

In order not to cause a cognitive overload, we wanted learners to perform the UX evaluation in a quick and easy way. Since the results of the preliminary study indicated that semantic differential scales are fast and easy to be employed, we decided to use this approach.

Our semantic differential scale was based on the UEQ technique, which was perceived as a little easier and participants demonstrated more intention to use than IEAM in the comparative study. However, since UEQ has a large number of adjectives (26), we decided to use only two adjectives per dimension, selecting those that were not considered repetitive or difficulty to understand. We defined a total of 12 pairs of adjectives: pleasant/unlikable,

comfortable/uncomfortable, meets expectations/does not meet expectations, supportive/obstructive, efficient/inefficient, practical/unpractical, inventive/conventional, creative/without creativity, easy/complicated, evident/confusing, interesting/not interesting, motivate to learn/demotivate to learn. We also added two optional comment fields. In the first field, the participants can make criticisms considering their evaluation in the semantic differential scale. In the second field, the participants can make suggestions regarding the evaluated LMS.

5.3. Conclusion

The results of our systematic mapping (Chapter 3) reveled some opportunities of research, such as: (i) the need for more empirical studies regarding the existing techniques; (ii) the lack of studies evaluating the UX of LMSs; and (iii) the lack of techniques that evaluate both usability and UX of LMSs. Based on these results and considering the importance of evaluating these two quality attributes in LMSs (as described in Chapter 2), we developed TUXEL, a Technique for User eXperience Evaluation in e-Learning.

This chapter presented the development process of the proposed artifact (TUXEL), which was guided by theoretical assumptions regarding the dimensions that it evaluates. We designed TUXEL to be easily applied by students without the need for any user testing prior to its application. Its questionnaire-based form makes it easy to be applied at a low cost, while allowing gathering the perceptions from the user's point of view.

In order to verify the feasibility of the technique, we carried out two empirical studies by comparing TUXEL with other existing techniques, chosen after a selection process. The results of each empirical study provided important feedback that served as the basis to improve the technique. Next chapter details these empirical studies.

CHAPTER 6 – EMPIRICAL EVALUATIONS AND IMPROVEMENTS OF TUXEL

In this section, we present two empirical studies carried out to evaluate TUXEL and to provide empirical evidence about its feasibility. The results of the first study served as an input to the development of its second version, called TUXEL 2.0, and a tool to support the evaluation process. In the second empirical study, results indicated that TUXEL 2.0 enabled the evaluators to detect a greater number of usability problems in comparison to another technique, in addition to requiring less time.

6.1. Introduction

In the previous chapter, we proposed a new usability and UX evaluation technique designed to evaluate LMSs, called TUXEL. This chapter presents the application of the proposed artifact (TUXEL) to interact with the context of the problem (usability/UX evaluation of LMSs) in the Design Cycle of the Design Science Research (DSR) methodology. In Design Cycle, the artifact is applied to the context of the problem, verifying whether it solves a problem or improves something in the given context (Wieringa, 2014). This process can lead to a new iteration over the Design Cycle, in which the results of the application of the artifact produces new knowledge, in addition to providing feedback for improving the artifact. Our Design Cycle resulted in two iterations. The first iteration consisted in the development of the technique and its application on the problem context (first empirical study). The second iteration consisted in the improvement of the proposed artifact and its application on the problem context after the changes (second empirical study).

In the first empirical study, we compared TUXEL with a technique developed by Theng and Sin (2012), which consists in an adaptation of the TAM (Technology Acceptance Model) questionnaire proposed by Davis et al. (1989). The results of this first study served as the basis for improving TUXEL and developing its second version, called TUXEL 2.0. In the second empirical study, we carried out a comparison between TUXEL 2.0 and the Heuristic Evaluation proposed by Mtebe and Kissaka (2015).

By presenting the results of these studies, we intend to provide empirical evidence about the feasibility of TUXEL, while encouraging its adoption to evaluate LMSs. The remainder of this chapter are organized as follows: in Section 6.2, we describe the execution of the first empirical study; Section 6.3 presents the improvement of TUXEL and the development of its

second version, called TUXEL 2.0 and a tool to support the evaluation process; in Section 6.4 we describe the execution of the second empirical study; finally, Section 6.5 discusses and concludes this Chapter.

6.2. First Empirical Study

This first empirical evaluation aimed to verify the feasibility of TUXEL to identify usability and pedagogical usability problems, while being fast and easy to be applied by students. To do so, we performed a comparison between TUXEL and a technique that was chosen through a selection process.

Given that most of the studies identified in our systematic mapping employed techniques with a questionnaire-based approach using scales, we aimed to identify whether this approach is adequate for the identification of usability problems. In the next sections, we describe the technique selection process, the conduction and the results of the study.

6.2.1. Technique Selection Process

The technique selection process was carried out by defining a set of exclusion criteria (EC) according to the purposes of our study. Since our technique is specific to evaluate LMS, we did not consider generic techniques (EC1). We also did not consider techniques that were not available for download/consultation (EC2) or did not provide the final applicable questionnaire (EC3). Finally, since our technique aims to be applicable by students, we did not consider those that were applied only by expert evaluators (EC4).

Figure 6.1 illustrates the selection process. Firstly, we applied the EC1 and EC2 criteria. These criteria were directly applied on the SQ1 and SQ9 sub-questions from the systematic mapping of Rivero and Conte (2017), giving a subtotal of 22 techniques. After, we applied them to EC3 and EC4 criteria (Table 6.1).

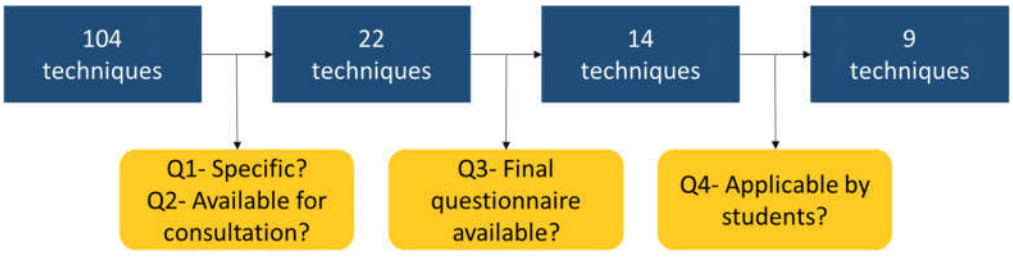


Figure 6.1 - Technique selection process workflow.

Table 6.1 - Mapping of the techniques according to the exclusion criteria of the first empirical study.

N.	ID	Technique	EC3	EC4
1	S01	Questionnaire proposed by Koohang (2004)		
2	S06	Framework proposed by Ssemugabi and De Villiers (2007)	X	
3	S16	Questionnaire proposed by Zaharias and Poylymenakou (2009)		
4	S17	Framework proposed by Zaharias, P. (2009)	X	
5	S21	Heuristic Evaluation proposed by De Kock et al. (2009)		X
6	S24	UseLearn technique proposed by Oztekin et al. (2010)		
7	S27	Heuristics of Squires & Preece (1999) applied by Lanzilotti et al. (2011)	X	X
8	S31	Heuristics of Squires & Preece (1999) applied by Granić and Ćukušić (2011)	X	X
9	S31	Heuristics of Quinn (1996) applied by Granić and Ćukušić (2011)	X	
10	S32	Heuristics of Reeves et al. (2005) applied by Zaharias and Koutsabasis (2011)		X
11	S32	Heuristics of Mehlenbacher et al. (2002) applied by Zaharias and Koutsabasis (2011)		X
12	S34	Framework proposed by Fetaji and Fetaji (2011)	X	
13	S38	Framework for usability testing proposed by Ivanc et al. (2012) to evaluate mobile interfaces	X	
14	S39	Adapted TAM, proposed by Theng and Sin (2012)		
15	S47	UX evaluation questionnaire proposed by Ssekakubo et al. (2014)		
16	S48	Usability evaluation questionnaire proposed by Yusoff and Mat Zin (2011)		
17	S51	Web-based Educational Environmental Attitude Scale applied by Torun and Tekedere (2015)		
18	S54	Heuristic Evaluation proposed by Mtebe and Kissaka (2015)		X
19	S55	Questionnaire proposed by Alkhattabi (2015)		
20	S58	Questionnaire proposed by Jusus, I. S. et al. (2015), adapted from Zaharias and Poylymenakou (2009)		
21	S59	Usability evaluation questionnaire proposed by Medina-Flores, R. and Morales-Gamboa, R. (2015)		X
22	S61	Framework proposed by Navarro, C.X. et al. (2016)	X	

Within the 22 techniques, 14 provided the final applicable questionnaire (EC3). Since the goal of our technique is to be applied by students (EC4), we selected, to the first empirical study, only the techniques that were applied by them. Within the 14 techniques, only 9 met this requirement (see Table 6.2).

We analyzed each technique regarding aspects such as the evaluated dimensions, scope and validity. Given that TUXEL consists of sets of items that evaluate different dimensions, we did not consider techniques that do not have separation by dimensions, in order to make the comparison possible. Regarding scope, we aimed to identify whether the technique can be

applied to LMSs in general. We did not consider, for instance, techniques developed to evaluate specific aspects of a given LMS. Finally, regarding validity, we aimed to identify whether the instrument is valid to be applied. We did not consider, for instance, techniques that were modified from the original version (e.g. by excluding or including questions) without explaining the modifications made. At the end of the analysis, we selected the technique developed by Theng and Sin (2012). Since the compared technique does not address UX, we did not consider this dimension during this evaluation. Details of this analysis are presented in Appendix G.

Table 6.2 - Overview of the selected techniques.

N.	ID	Technique
1	S01	Questionnaire proposed by Koohang (2004)
2	S16	Questionnaire proposed by Zaharias and Poylymenakou (2009)
3	S24	UseLearn technique proposed by Oztekin et al. (2010)
4	S39	Adapted TAM, proposed by Theng and Sin (2012)
5	S47	UX evaluation questionnaire proposed by Ssekakubo et al. (2014)
6	S48	Usability evaluation questionnaire proposed by Yusoff and Mat Zin (2011)
7	S51	Web-based Educational Environmental Attitude Scale applied by Torun and Tekedere (2015)
8	S55	Questionnaire proposed by Alkhattabi (2015)
9	S58	Questionnaire proposed by Jusus, I. S. et al. (2015), adapted from Zaharias and Poylymenakou (2009)

The technique developed by Theng and Sin (2012) consists in an adaptation of the Technology Acceptance Model (TAM) proposed by Davis (1989), which measures the degree of acceptance of a particular technology. This technique, which we will refer as Adapted TAM (AT), consists of a 5-point Likert questionnaire that valuates the usability and efficaciousness of LMSs through four constructs and their dimensions. The first construct is composed by ISO 9241-11 Standards for Usability dimensions: Perceived Satisfaction (PS), Perceived Usefulness (PU) and Perceived Ease Of Use (PEOU). The second construct evaluates the learners' E-learning Engagement by means of the following dimensions: Learning By Interaction (LBI) and Making Sense of Learning (MSL). The third construct, Self-efficacy, aims to evaluate their familiarity with e-learning technologies and it is formed by E-Learning Efficacy (ELE) and New Media Efficacy (NME) dimensions. Finally, the fourth construct evaluates general E-learning Design and Support by means of Navigation Structure (NS), User Interface (US) and Personalization and Freedom of Control (PFC) dimensions.

6.2.2. Participants and Materials

We carried out the study with 46 undergraduate students from Federal University of Amazonas (UFAM), enrolled in Technical Texts in Computer Science course. In this study, the participants evaluated the usability of a LMS called ColabWeb, which is based on version 3.1 of the Moodle LMS. We decided to evaluate this platform, since it is the main LMS used in the institution. Moreover, the platform had not yet been evaluated regarding its usability and UX. All the participants were already using the platform. Since both techniques can be applied at any time with learners that use the platform, we did not perform a user test prior to the application of the techniques. We present the details of the participants' profile and the groups balancing and distribution process in the next sub-section.

The following materials were used in this study: (i) an informed consent form, explaining the study and the subjects' voluntariness and confidentiality of their identities; (ii) a characterization questionnaire; (iii) the usability evaluation techniques (TUXEL and AT – See Appendix H and I for the questionnaires); and (iv) a feedback questionnaire composed by open-ended questions to gather qualitative data regarding the techniques used by the participants (see Table 6.3).

Table 6.3 - Open-ended questions used in the feedback questionnaire.

Did you have any difficulty in understand some item of the technique? Which ones?
Was there any item or any difficulty that you think should have been evaluated by the technique? Which ones?
Were you able to evaluate fully the usability of the LMS by using the technique?
Did you feel comfortable while using the technique?
What was easy when applying the technique?
What was difficult when applying the technique?
What would you change to improve the application of the technique?

6.2.3. Execution

Two days before the study, we presented an overview regarding usability concepts. We informed the students about the voluntariness of the study and the confidentiality of their identities. All participants signed the informed consent form and filled the characterization questionnaire. This questionnaire aimed to gather information regarding three attributes: (i) their frequency of use of LMSs, (ii) the number of disciplines in which they used a LMS, and (iii) their experience with usability evaluations.

We divided and balanced the participants in two groups, according to the results of the characterization questionnaire (Table 6.4). On the study day, participants went to the proper

room according to the group they were assigned. Each group used only one technique. We recorded the time spent by each participant in order to gather information about the mean time needed to apply these techniques. After the LMS evaluation process, the participants filled the feedback questionnaire.

Table 6.4 - Overview of the groups balancing and distribution of the participants (first empirical evaluation).

Question	Answers	Participants	
		Group 1 (TUXEL)	Group 2 (AT)
Number of disciplines which used the LMS	Four or more disciplines	4	3
	Two or three disciplines	17	18
	Only in this discipline	2	2
Experience with LMSs (frequency of use)	Several times a week	14	13
	Once a week	8	9
	Once a month	1	1
	Never used a LMS	0	0
Knowledge about usability /UX evaluation	Already performed this type of evaluation	1	1
	Already learned about it and did some class exercises	4	3
	Already read about it but not in depth	9	8
	Never heard about it	9	11
Total of participants		23	23

6.2.4. Results

This section presents the results regarding the first empirical evaluation. We divided the results in two subsections: (i) Results regarding the LMS's usability and (ii) Results regarding the techniques.

6.2.4.1. Results Regarding the LMS's Usability

This section presents the results from the usability evaluation of the LMS. We calculated the percentage of agreement, disagreement and neutral values for each evaluated item. Since the techniques use a 5-point Likert scale, the values regarding "agree" and "partially agree" options were consolidated in order to better interpret the results. The "disagree" and "partially disagree" options were consolidated as well.

Regarding the **results from TUXEL** (Figure 6.2), 7 out of 15 dimensions were positively evaluated: (i) Learnability (L); (ii) Consistency (C); (iii) LMS Feedback (LF); (iv) Learner Control (LC); (v) Goal Orientation (GO); (vi) Instructional Assessment (IA); and (vii) LMS Flexibility (LFL). It may indicate that the LMS is intuitive and easy to be learned and its interface elements are consistent across the pages. The LMS also provides an adequate level of control and feedback on student actions. Finally, it allows students to track their performance

and achievements, choose the path and pace of their study and carry out activities for their self-assessment.

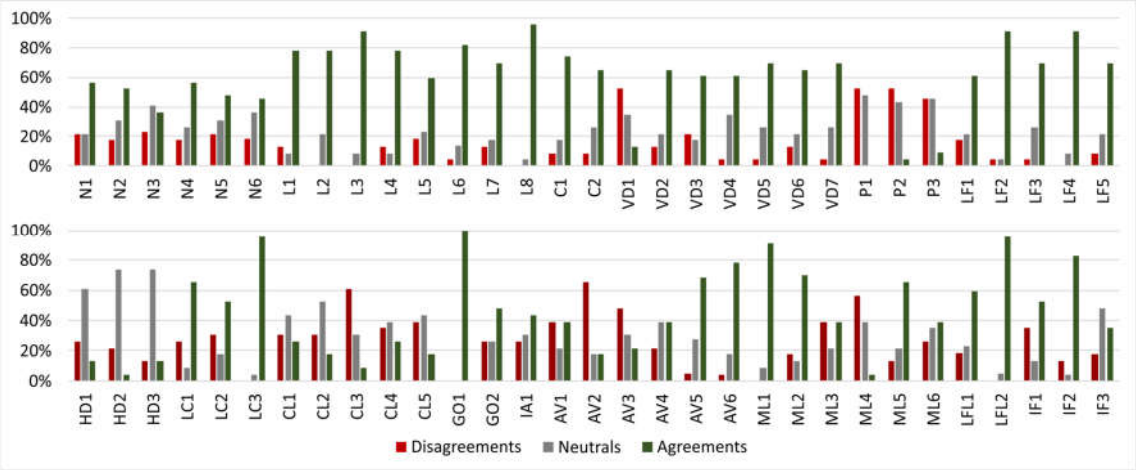


Figure 6.2 - Overview of the results from TUXEL.

On the other hand, Personalization was the most negatively evaluated dimension. Participants evaluated negatively all of the three items that compose this dimension. It may indicate that there is a need to provide features that allow learners to personalize the environment according to their preferences.

The Help and Documentation (HD) and Collaborative Learning (CL) received a great number of neutral evaluations. These results may indicate that: (i) the participants did not look for any kind of help/documentation or communication tools, thus being unable to state whether they exist in the platform, or (ii) the teacher does not provide any type of collaborative activity. The only item that received negative evalution was CL3, which reveals that the LMS does not allow the students to see what their colleages are doing in the platform, the most read contents, etc.

The remaining dimensions obtained positive evaluations, as well as some negative or neutral evaluations. Regarding Navigation (N), the participants neutrally evaluated the item N3, i.e., they were unsure whether they could go wherever they want in the platform or not.

In relation to the Visual Design (VD), the participants considered that the interface of the platform is not aesthetically attractive (VD1). It may indicate that there is a need to improve its overall design, in order to make it more appealing.

Regarding Added Value (AV), the results revealed that the participants consider the LMS useful to the learning process. However, it is not providing features that add value to the learning by means of the platform in comparison to the traditional learning. The low number of agreements in the items which makes a contrast between these two approaches (AV1, AV2 and

AV3) may indicate that there is a need do provide more interesting features to add more value to the use of the platform.

With respect to Motivation to Learn (ML), the participants stated that the platform, in general, is interesting and motivates them to learn. However, it does not provide simulations or gamification mechanics, neither a symbolic reward for their accomplishments (ML3 and ML4).

Regarding Instructional Feedback (IF), the results revealed that the LMS provides immediate and specific feedback, whenever possible, during the tasks. However, the participants were not sure whether the platform provides opportunities for additional feedback from the teacher (IF3). Since the participants are students from a face-to-face course, they would probably prefer to have their questions answered personally by the teacher rather than through the platform, thus not knowing the existence of such a feature.

Finally, in relation to the efficiency of the technique, the participants using TUXEL spent, on average, 11 minutes to fill the questionnaire. It demonstrates that the technique does not require too much time to be applied, being possible to employ it during part of a class.

Regarding the **results from Adapted TAM** (Figure 6.3), 05 out of 10 dimensions were positively evaluated: (i) Perceived Satisfaction (PS); (ii) Perceived Usefulness (PU); (iii) Perceived Ease of Use (PEOU); (iv) Navigation Structure (NS); and (v) Making Sense of Learning (MSL). These results indicate that the LMS was considered easy to be used and useful, while providing comprehensible navigation and being satisfactory to the learners. Additionaly, the participants considered that they usually understand well the messages in the discussion board and in the course topics.

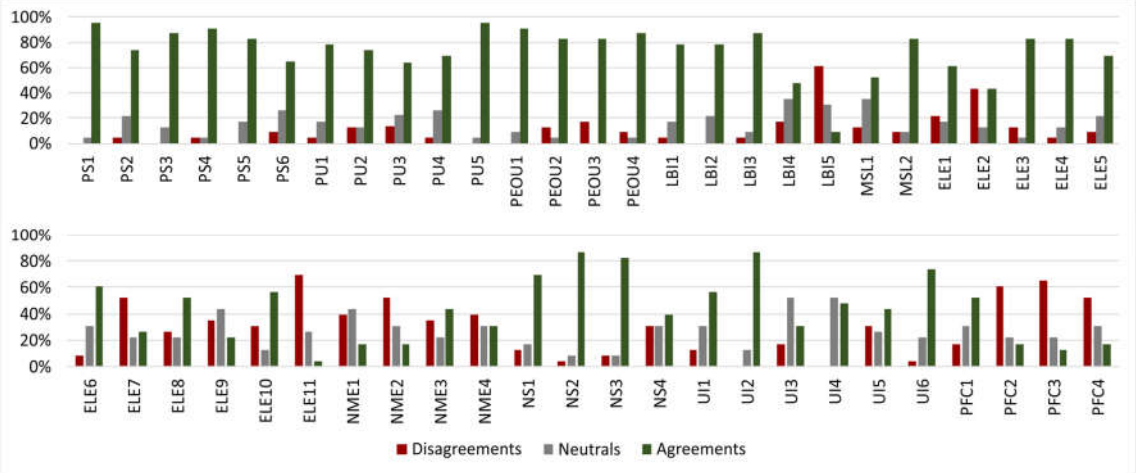


Figure 6.3 - Overview of the results from the AT technique.

On the other hand, two dimensions received a high number of negative evaluations: (i) New Media Efficacy (NME) and (ii) Personalization and Freedom of Control (PFC). Regarding

NME, it indicates that the participants did not consider them as experts in working with new media tools, such as wikis, discussion boards and blogs. With respect to PFC, the results revealed that the participants did not use some personal functionalities offered by the platform, such as the learner's panel and personal alerts.

The remaining dimensions were composed by positive evaluations, as well as some negative or neutral evaluations. Regarding Learning By Interaction (LBI), the results revealed that most of the participants stated that the LMS promotes the interdisciplinary and independent learning. However, they still prefer face-to-face interaction rather than interaction by means of the platform (LBI5), which is in accordance with the results from TUXEL.

In relation to E-learning Efficacy (ELE), the results indicated that the participants are aware of the functional capabilities of the LMS. It also revealed that the participants felt that they could accomplish the tasks without needing further help. On the other hand, the results demonstrated that they did not know all the functionalities of the platform (ELE2) and they did not consider themselves as computer experts (ELE7). Additionally, they stated that they did not used the help (ELE11).

With respect to User Interface (UI), the results demonstrated that the participants did not know whether the items of the main menu are the most used ones (UI3) and whether the platform provides a welcome message on the main page (UI4).

Finally, regarding the efficiency of the technique, the participants using Adapted TAM spent, on average, 10 minutes. It indicates that this technique can also be applied quickly during part of a class.

6.2.4.2. Results Regarding the Techniques

This section presents the results from the feedback questionnaire, composed by open-ended questions. The results provided some insight regarding the evaluated techniques. In relation to the question "did you have any difficulty to understand some item of the questionnaire?", all participants using TUXEL answered "no". On the other hand, seven participants using AT technique reported that they did not understand one or more items from the questionnaire. Most of these items were related to E-Learning Efficacy and New Media Efficacy dimensions.

Similarly, with respect to the question regarding the difficulties in using the techniques, 9 participants using TUXEL related some difficulty in contrast with 13 participants using AT technique. Most of the reported difficulties, in both techniques, were related to the unfamiliarity of the participants with some of the evaluated items.

When asked whether they could fully evaluate the usability of the LMS using the techniques, 5 and 11 participants using AT and TUXEL, respectively, answered “no”. With respect to AT, one participant related that there was a lack of questions to evaluate attributes such as frequency of use, ease of use and usefulness of some aspects of the LMS (he/she did not specify which aspect). Other participant stated that the technique did not evaluate the interface of the LMS, which in his/her opinion is confusing in the evaluated platform. One participant also reported “the technique has questions that has nothing to do with using the LMS”. Finally, two participants reported that they did not know about some of the evaluated functionalities. Regarding TUXEL, seven participants related some difficulty in remembering whether they used certain functionality or whether the functionality existed in the platform. Two participants stated the lack of a field to provide criticisms, while one participant considered the technique specific to learners that use the LMS in more than one course. Finally, one participant stated that the technique could evaluate aspects related to privacy.

About the question “What was easy when using the technique?”, most of the participants of both groups pointed the adoption of the 5-point scale and the simplicity and objectivity of the techniques. Similarly, when asked “What was difficult when using the technique?”, participants of both groups related the difficulty in remembering or evaluating functionalities that they do not have enough knowledge.

When asked “What would you change in the technique?”, most of the suggestions from participants using AT technique was to add a field for commentaries. There were also 3 participants who suggested removing some items that are not related to the LMS, such the New Media Efficacy dimension. Participants using TUXEL also suggested adding a field for commentaries. Two participants suggested removing the neutral point and other two suggested adding images of the evaluated LMS in order to help them remember its interface.

6.2.5. Conclusion from first empirical evaluation

The results of first empirical evaluation indicated that participants evaluated negatively some of the items presented by the techniques. Regarding the usability evaluation of the LMS, participants using TUXEL evaluated 10 items negatively, while participants using AT, 9. However, all of the negative evaluations in AT were in items related to participants’ self-assessments and preferences. Regarding TUXEL, within the 10 negative evaluated items, 4 (40%) were about problems related to personalization and collaborative learning (P1, P2, P3, CL3).

With respect to neutral evaluations, participants using AT technique evaluated 4 items as being neutral, while participants using TUXEL evaluated 10. The level of specificity of the questions may explain the higher number of neutral evaluations in TUXEL. In the open-ended questionnaire, 7 participants (30,43%) related that they had difficulty in remembering whether they used or whether the evaluated functionality existed or not in the platform.

Although both techniques provide some information regarding usability of the evaluated LMS, none of the techniques allowed identifying where the problems occur. Since the techniques used only a 5-point Likert scale, participants' answers were limited. Moreover, the use of a scale made it difficult to interpret the results when verifying whether the item would be considered an indicative of a problem or not. Some items, such as CL4 and ML3 of TUXEL, for instance, had few differences between the percentage of agreements, neutrals and disagreements. Therefore, TUXEL was viable from the point of view of being fast and easy to use. On the other hand, it was not so viable when it comes to identifying usability problems. Due to this, we decided to carry out a second iteration in the DRS methodology in order to refine the artifact (TUXEL), considering the drawbacks and the suggestions proposed by the participants.

6.3. Technique Improvement: Development of TUXEL 2.0

The first study provided important feedback about TUXEL. The results revealed that questionnaire-based techniques using scales do not provide enough information regarding the problems identified. Although this type of technique pinpoints the existence of problems, it does not allow the evaluator to identify where exactly they occur our whether it is really a problem or not, which makes it difficult to pinpoint solutions. Moreover, the participants reported difficulty in remembering the use or the existence of some of the evaluated items. Since these items evaluate specific aspects of usability that participants may not know or consider during their interaction with the LMS, the questionnaire-based approach using scales revealed to be inadequate. Considering these issues, we decided to perform a second iteration over the Design Cycle of the DSR methodology.

Our main concern was regarding the difficulty of the participants in remembering whether a functionality, aspect or element of the LMS exists or not, which may explain the variation in the answers. Since there are specific items of the LMS that the learner may not notice during his interaction with the platform, it would be important that he interact with the LMS while evaluating it by using the set of items. Given this, we analyzed the types of

technique presented in Section 2.2.1 in order to identify the approach that best fits this requirement. After the analysis, we selected the inspection-based approach.

As stated in Section 2, usability inspection is a type of technique in which an evaluator, usually a usability expert, inspects the interface of a product by means of checklists or heuristics (Conte et al., 2007). However, usually, experienced evaluators perform usability inspections. Since our goal is to make TUXEL applicable by students, which may never have carried out a usability inspection before, it must provide guidance during the evaluation process. Thus, we re-structured TUXEL and changed it to a guided inspection approach, giving rise to its second version, called TUXEL 2.0.

In TUXEL 2.0, we transformed all items in negative phrases in order to make it more intuitive for participants to find problems. We also included hints to items that evaluate specific aspects that novice inspectors may not know about.

Since there are many items to be evaluated, we removed those that did not provide useful results in the first study regarding usability evaluation of the LMS. The dimensions Added Value and Motivation were fully removed. The former because it just provided results related to participants' preferences, the latter because it could also be measured in the UX evaluation by using the semantic differential scale. We also grouped the items according to the page and/or activity being inspected, in order to present only the items that should be evaluated in each page/task. Thus, we created four categories: General Interface, User Login, Assignments and Assessments.

The General Interface category is composed by questions regarding interface usability, such as navigation structure, terminology used and organization of the information. The User Login is composed by questions related to login, such as login instructions and feedback. In Assignments we grouped questions aimed to evaluate specific aspects related to the tasks, such as the clarity of instructions, level of user control and feedback of the LMS. Finally, the Assessments category contains questions related to instructional feedback and goal orientation. Researchers may combine these categories or use them separately, according to the page/task being evaluated.

The technique also has a checklist that evaluates general aspects of a LMS, divided in the following categories: Help and Documentation, LMS Learnability, Learning Through the LMS and LMS Flexibility. The Help and Documentation category relates to items that evaluates whether the LMS has a help material and whether this material is comprehensive and straightforward. The LMS Learnability category relates to items that evaluate how the LMS is easy to learn. The Learning Through the LMS evaluates the extent the LMS allows learners to

learn by using the platform. Finally, the LMS Flexibility category evaluates the extent the LMS allows learners to have control of their learning process and of the platform. The complete set of items of TUXEL 2.0 can be found in APPENDIX J. Next subsection details how to conduct an inspection using TUXEL 2.0.

6.3.1. Inspection process

The researcher, i.e., the person who is interested in evaluating the LMS, must perform some actions prior to the beginning of the inspection process. Figure 6.4 illustrates this process.

Figure 6.4 - TUXEL 2.0 inspection process.

Firstly, the researcher must define the tasks that will be performed by the inspector, i.e., the student. Since there are many items to be evaluated, it is recommended to set just few tasks in order to not overload the inspector. After defining the tasks, it is necessary to match each task to one or more of the four categories described previously. If the task, for instance, consists in accomplishing an assignment, it should be evaluated by using the Assignment category checklist. The task may be evaluated, additionaly, by using the General Interface category.

The inspection process of TUXEL 2.0 is composed by three steps. In the first step, the inspector performs the tasks defined by the researcher while evaluates the usability of the LMS by using the corresponding category of checklist. For each item, the inspector should verify whether it occurs in the platform. When the problem is identified, the inspector must provide, in a specific form, further details about the problem (if necessary), where it occurs and the item ID used to identify it.

In the second step of TUXEL, after finishing the defined tasks, the inspector evaluates the LMS by using the General LMS Aspects category checklist. The inspector should mark the items that the evaluated LMS does not fulfill and describe the problems in detail in the comments field.

In the third and last step, the inspector evaluates the UX of the LMS through the semantic differential scale. After filling the scale, the inspector makes criticisms considering the responses and provides suggestions.

6.3.2. TUXEL Tool

Considering that some inspectors may have difficulty in reporting the problems in written form and to make the evaluation process more interactive, we developed a tool to support the inspection process. Since usually learners access the LMS by using a browser, and considering that Google Chrome is one of the most used one, we developed an extension for it, called TUXEL Tool.

The TUXEL Tool provides guidance to inspectors during the evaluation process, presenting the tasks that need to be performed and the items that should be evaluated. It encompasses the three steps defined in TUXEL 2.0 inspection process presented in the previous subsection. In the beginning, the tool asks the inspector for his/her name and then provides a brief introduction regarding its purpose and its use (Fig. 6.5-a). In the first step, the inspector is presented with a set of items (Fig. 6.5-b). Some of them have an icon that, when clicked, show further information about the item. When the inspector identifies a problem, he/she should click on the item that addresses this problem and then click and drag the cursor in the area where the problem occurs. After marking the area, the tool asks the inspector for confirmation. When confirmed, the tool presents a field where the inspector can provide details about the identified problem (Fig. 6.5-c). Finally, the tool prompts to save a screenshot of the screen with the marking. After evaluating all items, the inspector clicks in "Continue" button, where he/she is presented with other tasks and instructions.

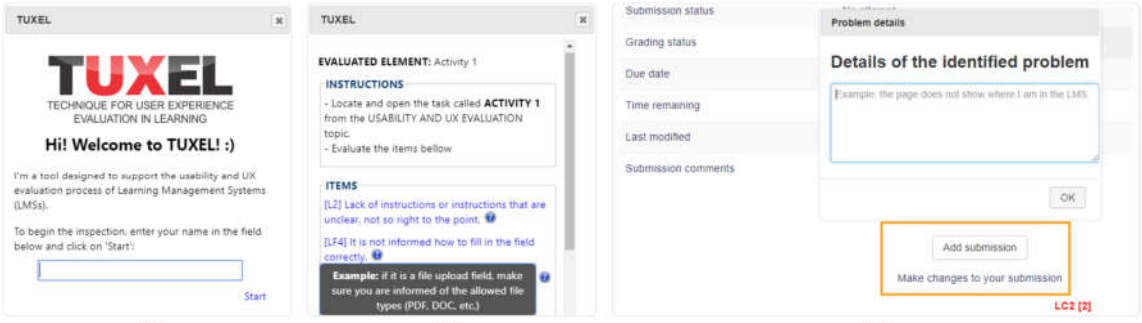


Figure 6.5 - TUXEL interface: (a) welcome screen; (b) list of evaluated items; (c) screen-marking feature usage.

In the second step, the tool presents a set of items that evaluate general usability aspects of LMSs, each item composed by a checkbox. The inspector should verify whether the problem

described by the item exists in the evaluated LMS or not and check the corresponding checkbox. The inspector can give further details about the checked items in the comment field.

The third step is composed by a semantic differential scale structured with a set of radio buttons between a pair of adjectives to evaluate the UX of the LMS. A quick guide about how to fill the scale is provided in the beginning of this step. After filling the scale, the inspector is presented with two comment fields. In the first field, the inspector should make critics regarding the LMS considering the scores given in the semantic differential scale. In the second field, the inspector can give suggestions of improvements for the evaluated LMS.

Finally, the tool generates a report in a CSV format file. This file contains all information related to the inspection process, such as the duration of the evaluation, the ID of each violated item and the respective URL of the page and commentaries. The tool can be found in the website[7] of our research group.

6.4. Second Empirical Evaluation

In order to evaluate our improved artifact (TUXEL 2.0), we carried out a second empirical evaluation, by applying it on the problem context. The second empirical evaluation aimed to evaluate TUXEL 2.0 in comparison with other technique. Since we changed TUXEL to an approach based on guided inspection, we needed to select a technique that uses the same or similar approach. Thus, we conducted a new technique selection process, still based on the results of the systematic mapping presented in Chapter 3.

6.4.1. Technique Selection Process

We used the same exclusion criteria defined in the first study (see Section 6.2.1) with some modifications. Given that we did not want to limit our results to techniques that were applied to students, we replaced the EC4 with the following criteria: “techniques that are not based on inspection approach”. A total of 4 techniques were obtained (rows 1, 10, 11 and 18 from Table 6.5).

We analyzed each technique in relation to their evaluation focus. The AHE (Adapted Heuristic Evaluation) proposed by De Kock et al. (2009) was focused on the evaluation of a Web service for assignments, which limits its scope. Regarding the AHE from Mehlenbacher et al. (2005) and Reeves et al. (2002), a study conducted by Zaharias and Koutsabasis (2011) demonstrated that both techniques could capture a good number of usability problems.

[7] http://uses.icomp.ufam.edu.br

However, the experienced inspectors faced difficulties to match some problems to the heuristics. Given that students, who usually do not have previous experience with inspection, compose the population of our study, we did not consider these techniques. At the end of the process, the AHE proposed by Mtebe and Kissaka (2015) was selected.

Table 6.5 - Mapping of the techniques according to the exclusion criteria of the second empirical study.

N.	ID	Technique	EC3	EC4
1	S01	Questionnaire proposed by Koohang (2004)		X
2	S06	Framework proposed by Ssemugabi and De Villiers (2007)	X	X
3	S16	Questionnaire proposed by Zaharias and Poylymenakou (2009)		X
4	S17	Framework proposed by Zaharias, P. (2009)	X	X
5	S21	Heuristic Evaluation proposed by De Kock et al. (2009)		
6	S24	UseLearn technique proposed by Oztekin et al. (2010)		X
7	S27	Heuristics of Squires & Preece (1999) applied by Lanzilotti et al. (2011)	X	
8	S31	Heuristics of Squires & Preece (1999) applied by Granić and Ćukušić (2011)	X	
9	S31	Heuristics of Quinn (1996) applied by Granić and Ćukušić (2011)	X	
10	S32	Heuristics of Reeves et al. (2005) applied by Zaharias and Koutsabasis (2011)		
11	S32	Heuristics of Mehlenbacher et al. (2002) applied by Zaharias and Koutsabasis (2011)		
12	S34	Framework proposed by Fetaji and Fetaji (2011)	X	X
13	S38	Framework for usability testing proposed by Ivanc et al. (2012) to evaluate mobile interfaces	X	X
14	S39	Adapted TAM, proposed by Theng and Sin (2012)		X
15	S47	UX evaluation questionnaire proposed by Ssekakubo et al. (2014)		X
16	S48	Usability evaluation questionnaire proposed by Yusoff and Mat Zin (2011)		X
17	S51	Web-based Educational Environmental Attitude Scale applied by Torun and Tekedere (2015)		X
18	S54	Heuristic Evaluation proposed by Mtebe and Kissaka (2015)		
19	S55	Questionnaire proposed by Alkhattabi (2015)		X
20	S58	Questionnaire proposed by Jusus, I. S. et al. (2015), adapted from Zaharias and Poylymenakou (2009)		X
21	S59	Usability evaluation questionnaire proposed by Medina-Flores, R. and Morales-Gamboa, R. (2015)		X
22	S61	Framework proposed by Navarro, C.X. et al. (2016)	X	X

The technique proposed by Mtebe and Kissaka (2015) is composed by two dimensions. The first dimension evaluates the general usability of LMSs and it is based on the Nielsen's ten heuristics (1994) adapted to e-learning context. The second dimension evaluates the didactic effectiveness of LMSs. This technique was chosen due to the following reasons: (i) it is newer

than the other techniques, which may result in a better coverage of the usability aspects; (ii) its heuristics are based on many works related to usability of LMSs, such as the works of Squires and Preece (1999), Nokelainen (2006) and also the work of Reeves et al. (2002); (iii) the authors suggested, as a future work, an evaluation performed by students using their technique.

6.4.2. Selected Technique Adaptation

Considering that TUXEL does not address aspects directly related to the content, since they are teacher dependent, we removed some questions from the AHE of Mtebe and Kissaka (2015), such as those related to "Instructional Materials" dimension. We also transformed the heuristics to negative sentences and adapted the tool to use the heuristics and the same resources provided in TUXEL tool in order to balance the evaluation process. However, we did not adapt the UX evaluation step in order not to bias the results in relation to the time spent in the evaluation, since it is not present in the original AHE.

6.4.3. Participants and Materials

We carried out a study with 22 students from two computer science courses at the Federal University of Amazonas. The study was conducted in the same LMS used in the first empirical study (ColabWeb LMS). All the participants were already using the LMS in one or more courses. Details of their profile are presented in Section 6.4.5.

The following materials were used in this study: (i) an informed consent form; (ii) a characterization questionnaire; (iii) the inspection tool for TUXEL 2.0 and AHE techniques; (iv) a video tutorial explaining how to install and use the tool; (v) the Technology Acceptance Model 3 (TAM3) questionnaire (Venkatesh and Bala, 2008) with the open-ended questions from the feedback questionnaire used in the empirical study 1. The TAM3 is a questionnaire based on TAM, originally proposed by Davis et al. (1989). It consists of a set of items evaluated by a 7-point Likert scale to obtain participant's perceptions about a technology regarding aspects such as Perceived Ease Of Use (PEOU), Perceived Usefulness (PU) and Behavioral Intention (BI). Table 6.6 presents an overview of the TAM3 questions.

6.4.4. Hyphotheses

We propose that TUXEL 2.0 is more efficient and effective than the AHE (Mtebe and Kissaka, 2015). We defined efficiency as the ratio between the number of identified problems and the time spent in the inspection process. We defined effectiveness as the ratio between the number

of identified problems and the total number of known problems. We tested the following hyphotheses (null and alternative respectively):

H_1: There is no difference in terms of efficiency between TUXEL 2.0 and AHE techniques.

H_{A1}: The efficacy of TUXEL 2.0 is greater than the efficiency of AHE technique.

H_2: There is no difference in terms of effectiveness between TUXEL 2.0 and AHE techniques.

H_{A2}: The effectiveness of TUXEL 2.0 is greater than the effectiveness of AHE technique.

Table 6.6 - Overview of the TAM3 Questionnaire applied in the second empirical study.

TAM3 Questionnaire		
Dimension	**ID**	**Question**
Perceived Usefulness	PU1	Using the technique improves my performance when evaluating the usability/UX of a LMS.
	PU2	Using the technique improves my productivity when evaluating the usability/UX of a LMS.
	PU3	Using the technique allows me to fully evaluate the usability/UX of a LMS.
	PU4	I find the technique useful for evaluating the usability/UX of a LMS.
Perceived Ease Of Use	PEOU1	The technique was clear and easy to understand.
	PEOU2	Using the technique did not require too much mental effort.
	PEOU3	I find the technique easy to use.
	PEOU4	I find it easy to report the usability/UX of the LMS using the technique.
Intention to Use	IU1	Assuming that I have access to the technique, I plan to use it to evaluate the experience of a learning platform.
	IU2	Given that I have access to the technique, I predict that I would use it to evaluate the experience of a learning platform.
	IU3	I intend to use the technique to evaluate the experience of a learning platform next month.

6.4.5. Execution

Prior to the execution of the study, we asked the teachers of the two disciplines to propose two types of tasks that students should accomplish during the usability evaluation. We instructed the teachers to provide simple and easy questions so that the students do not spend a lot of time answering the questions during the evaluation and do not bias the results. The first task was to send a file containing any interesting topic related to the content they are studying. The second task was a quiz composed by a drag and drop sentence completion and a multiple-choice question.

A day before the study we presented a brief introduction about usability inspection and the study, without giving too much details. We informed the participants about their voluntariness and confidentiality of their identities. All participants signed in the informed consent form and filled the characterization questionnaire, which was used to balance the participants regarding their experience with usability evaluations. This questionnaire aimed to gather information about two attributes: (i) the number of disciplines which they used a LMS, and (ii) their experience with usability/UX evaluations. We did not evaluate their frequency of use like in the first empirical evaluation since these participants use the LMS frequently to perform activities and group discussion in their disciplines.

We divided the participants in two groups, each group using one technique and both balanced according to the results of the characterization form (Table 6.7). Due to the limited time to conduct the evaluation, we decided that each participant should perform the inspection at home within one week and send the final report provided by the tool and the screenshots by e-mail. On the study day, we sent an email with instructions about the evaluation, the tool attached and the links to the video tutorial and the TAM3 questionnaire, according to the group that the participant was assigned to.

Table 6.7 - Overview of the group balancing and distribution of the participants (second empirical evaluation).

Question	Answers	Participants	
		Group 1 (TUXEL)	Group 2 (AT)
Number of disciplines which used the LMS	Four or more disciplines	8	9
	Two or three disciplines	2	2
	Only in this discipline	1	0
Knowledge about usability /UX evaluation	Already performed this type of evaluation	0	1
	Already learned about it and did some class exercises	4	4
	Already read about it but not in depth	4	5
	Never heard about it	3	1
Total of participants		11	11

6.4.6. Results

In this section we describe the results from the second empirical evaluation. We analyzed the data quantitatively, in order to make comparisons between the two techniques, and qualitatively, in order to gather further information about the techniques. Next sections detail the analysis process.

6.4.6.1. Quantitative Analysis

We tabulated the discrepancies reported by each participant in a spreadsheet with the ID of the heuristic/item and of the participant. An important task in usability evaluations when involving multiple evaluators is the consolidation of usability problems into one master list by searching for duplicates, filtering and merging them out (Hvannberg and Law, 2017). We carried out this consolidation process in order to enhance the validity of comparing different usability evaluation methods (Law and Hvannberg, 2008).

Discrepancies that had similar descriptions were consolidated in a single discrepancy, but accounted for each participant individually. We excluded duplicated discrepancies from the same participant in the counting, except if the participant pointed the same problem using different heuristics/items. Discrepancies that were reported by just one or two participants were analyzed in order to identify whether it is a false-positive or a problem.

Two researchers, experts in HCI, performed the analysis individually. We prepared a form with the description of the discrepancy, and a field where the researcher should indicate whether it is a problem or not and make comments if necessary. In the case of disagreement between the two researchers, a third researcher, also an expert, analyzed the discrepancy. At the end of the process, 36 were considered as false-positives, 3 as suggestions and 1 as a duplicate discrepancy, resulting in a total of 103 unique identified problems. Within these problems, 58 were identified only by TUXEL, 29 only by AHE and 16 by both techniques. In Table 6.8, we present the overall results of this study per participant and group.

Table 6.8 - Overall results per participant and group.

Group 1: Participants using TUXEL 2.0											
Inspector ID	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**
Experience with usability evaluation	M	L	L	None	M	None	M	M	L	None	L
Total problems	12	19	11	19	11	9	29	29	18	12	15
Time (min)	61	88	13	65	27	20	57	94	37	28	46
Efficiency (defects/hour)	11,8	13,0	50,8	17,5	24,4	27,0	30,5	18,5	29,2	25,7	19,6
Effectiveness (%)	13,8%	21,8%	12,6%	21,8%	12,6%	10,3%	33,3%	33,3%	20,7%	13,8%	17,2%
Group 2: Participants using AHE											
Inspector ID	**P12**	**P13**	**P14**	**P15**	**P16**	**P17**	**P18**	**P19**	**P20**	**P21**	**P22**
Experience with usability evaluation	H	M	L	L	L	L	None	M	M	L	M
Total problems	10	9	7	6	9	3	8	7	2	1	7
Time (min)	68	36	20	81	78	74	45	38	40	57	62
Efficiency (defects/hour)	8,8	15,0	21,0	4,4	6,9	2,4	10,7	11,1	3,0	1,1	6,8
Effectiveness (%)	11,5%	10,3%	8,0%	6,9%	10,3%	3,4%	9,2%	8,0%	2,3%	1,1%	8,0%

We analyzed and calculated the mean of the number of problems identified by each group according to the level of experience in usability evaluation (Figure 6.6). We considered only the total number of problems of each participant, disregarding the false-positives. The participants who used TUXEL that had no experience, low or medium experience in usability evaluation, identified a higher number of problems in comparison with participants using AHE. Since there was only one participant with high level of experience with usability evaluation, we could not make a comparison between the techniques in this case. The results also indicated that, regarding TUXEL, the number of problems identified increased according to the inspector's level of experience in usability. On the other hand, such pattern does not occur in the AHE group. The participant that had no experience with usability evaluation identified a higher number of problems than participants with low and medium experience. It may indicate that TUXEL provides a better guidance to inspectors in comparison to AHE technique.

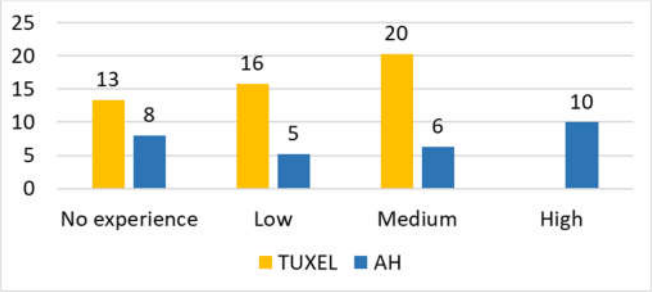


Figure 6.6 - Mean number of problems per usability level of experience.

Figure 6.7 shows the boxplots graphs for the efficiency and effectiveness indicators. We carried out a statistical analysis with the IBM SPSS Statistics version 24. Due to the low number of participants in the sample (Dybå et al., 2006), we applied the Mann-Whitney non-parametric statistical test, as suggested by Wohlin et al. (2012).

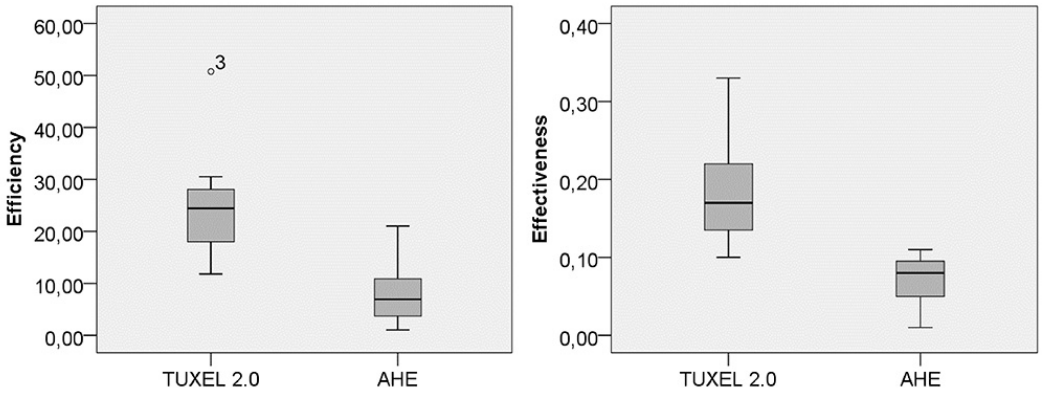


Figure 6.7 - Boxplot graphs evaluating the hypotheses indicators.

The results from the statistical analysis rejected both null hypotheses (H1 and H2), indicating that the effiency and effectiveness of TUXEL was significantly higher than the efficiency and effectiveness of AHE ($U = 7$, $p = .000$ and $U = 2$, $p = .000$ respectively). In fact, Fig. 9 shows that participants who used TUXEL 2.0 achieved higher efficiency and effectiveness scores than participants who used AHE.

With respect to the number of problems identified in each evaluated dimension/heuristic, the results were as follows. Regarding TUXEL, "Learnability of the LMS", "Feedback of the LMS" and "Navigation" were the dimensions that encompassed the highest number of problems, with 14, 13 and 12 respectively. These results indicate that the LMS does not provide enough instructions to learners accomplish the activities or it has elements that do not convey their functionality, which makes the platform not so easy to be learned and used. The platform also does not provide sufficient feedback about the learners' actions and its navigation is not so intuitive. An example of problem related to "Learnability of the LMS" is the lack of instructions to perform the second task, which consists on dragging and dropping some words in the blank spaces of a sentence.

Regarding AHE, "Learner control and Freedom", "Match between the system and the real world" and "Visibility of system status" accounted the highest number of problems, with 9, 7 and 6 respectively. These results are similar with the results from TUXEL, indicating that the LMS have elements that are not easily understood by learners and does not provide sufficient feedback about their actions. The results also indicate that the platform does not provide sufficient control to learners, such as options to exit whenever they need to. As an example of problem related to "Learner control and Freedom", we can cite the impossibility of go back to the topic of the activity by using the breadcrumbs.

During the analysis of the problems, we verified whether the participant assigned it to the right heuristic/item or not. We read the definition of each dimension of the heuristic/item in order to verify whether it addresses the problem identified by the participant. We identified that some participants assigned some problems to wrong items/heuristics of the techniques. Regarding TUXEL, within a total of 74 problems, 7 (9.46%) were identified using wrong items of the technique. With respect to AHE, within a total of 45 problems, 12 (26.67%) assigned to wrong heuristics. These results reveal that in addition to finding more problems, the participants that used TUXEL also made fewer mistakes in comparison to participants that used AHE.

With respect to the results from the UX evaluation step of TUXEL, we calculated the median for each evaluated item. We did not perform any comparison with AHE technique, since it does not evaluate the UX. The results revealed a neutral-negative perception about the UX of

the evaluated LMS (Figure 6.8). The unique positively evaluated aspect was the item “meets expectations”. This indicates that although the platform somewhat meets the needs of learners, they are not having a good experience of use. There is a need to provide more interesting features in order to stimulate and engage learners during their learning process through the LMS.

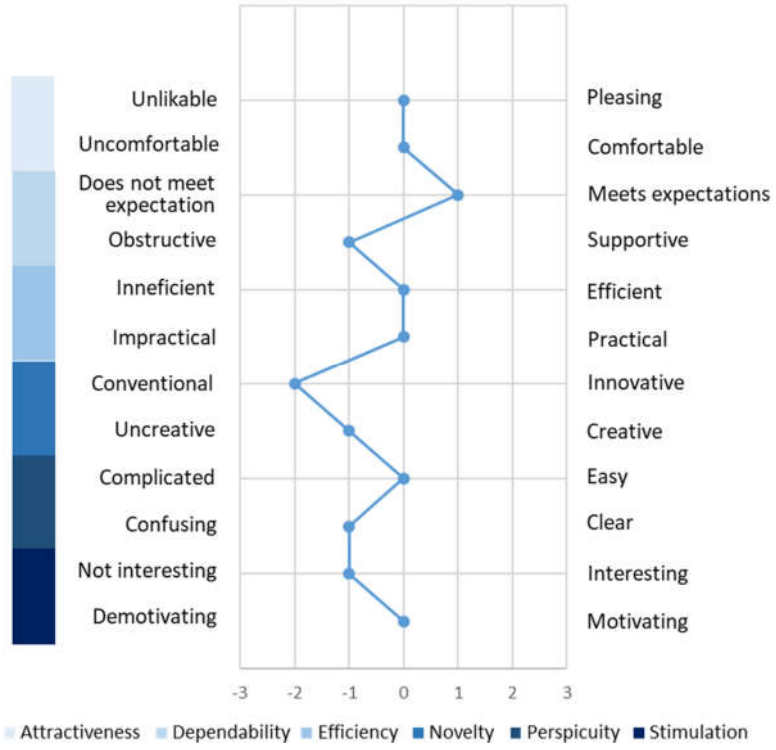


Figure 6.8 - Results of the UX evaluation part of TUXEL.

With respect to the results from the TAM3 questionnaire (Figure 6.9), we identified that, in general, participants showed a high level of acceptance of the techniques regarding Perceive Usefulness (PU). The item PU3 (effectiveness) was the lowest evaluated attribute in both techniques regarding PU, where participants using AHE evaluated the technique a little more negatively than participants using TUXEL. Since most participants had not performed this kind of evaluation before, they may have had difficulty in identifying whether the technique improves their effectiveness. The participants using AHE may have evaluated the PU3 a little more negatively due to the difficulty in addressing some problems with the heuristics.

With respect to ease of use (PEOU), both techniques reached similar levels of agreement from the participants. These results indicate that both techniques do not demand so much mental effort, while being easy and clear to be understood.

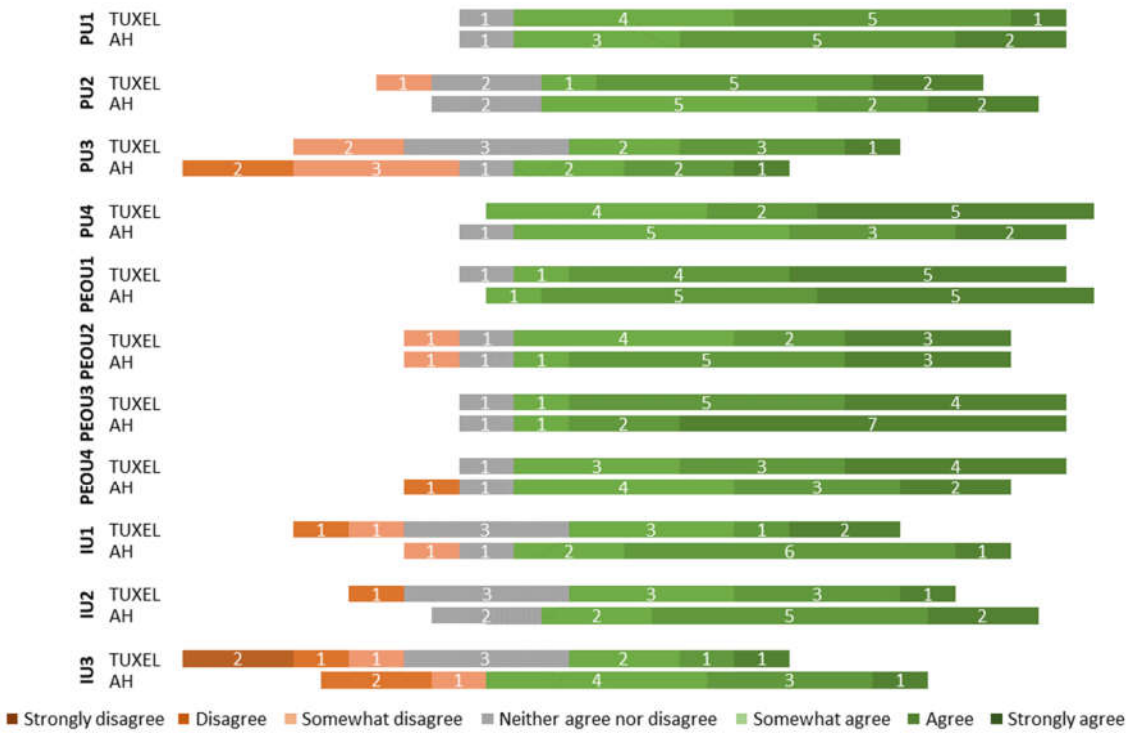


Figure 6.9 - Results of the TAM3 questionnaire (second empirical evaluation).

Regarding Behavioral Intention (BI), participants using AHE demonstrated more intention to use the technique than those using TUXEL. It may be due to the difficulties faced by the TUXEL group when using the screen capture feature of the tool. Since this group found a higher number of problems, they used this feature more frequently, which may have increased the probability of facing difficulties during the use of the tool. We will present further details of these difficulties in the next subsection.

6.4.7. Qualitative Analysis

The objective of the qualitative analysis was to gather further information with respect to participants' perception about the technique and the tool they used. We performed an open coding (Strauss and Corbin, 2014) by analyzing the participants' answers to the open-ended questions, creating codes according to the concepts identified in their quotations. Next, we analyzed and grouped the codes according to their aspects in order to abstract concepts that encompass categories and subcategories.

With respect to the **participant's perceptions regarding the UX of the LMS**, we analyzed the comments related to the UX part of TUXEL. We identified the following codes:

(1) somewhat intuitive; (2) confusing; (3) unorganized; (4) unattractive; and (5) do not meet the expectations.

Regarding codes 1, 2 and 3 (somewhat intuitive, confusing and unorganized), the results indicated that the platform is not so easy to be used and to find information. The participant P2, for instance, stated, “there is unnecessary information occupying a lot of space on the main screen, making navigation confusing and difficult to understand some tasks”. Moreover, the participant P11 stated the following: “many times I did not know how to go to a place I wanted, so I easily got lost and, for taking so long to find something, I gave up”. These quotations may explain the UX evaluation results, where the participants perceived the LMS as a little confusing and obstructive.

Regarding code 4, the participants considered the LMS as unattractive. According to participant P2, “the LMS sometimes becomes a festival of colors: black, blue, navy, words with bold, words without bold”. It is in accordance to the UX evaluation results, where the participants perceived it as neither attractive nor unattractive.

In relation to code 5, two participants stated that the platform does not provide some functionality, which may indicate that it does not meet their expectations. The participant P3 stated that the LMS “does not notify when a message is received”, while the participant P5 stated, “if it was similar to CodeBench [a type of online judge for learning programming] it would be better. […] It also leaves something to be desired regarding gamification”.

Finally, the participants provided suggestions for improvements. Most of the participants suggested an improvement in the communication through the LMS. Other suggestions include: (i) improve the interface in order to make it more attractive; (ii) have better search tools to facilitate locating contents and functionalities; (iii) remove unnecessary information on the interface; and (iv) allow personalization of the interface. Finally, the participant P8 stated that “the UX [of the platform] should be taken more seriously”.

Regarding the open-ended questions of the feedback questionnaire, we identified that the participants’ answers were sometimes composed by a mix between their perception about the technique and the tool. Since a tool supported the evaluation process, it is normal that the participants, who in general have low experience with usability evaluation, have difficulty in distinguishing these two concepts. Therefore, we divided the results in three sub-sections, according to the object of opinion, which will be presented next.

With respect to **participant’s perceptions regarding TUXEL**, we identified the following codes: (1) TUXEL was able to identify all participants’ daily use difficulties and/or

problems; (2) TUXEL helps inspectors identifying usability problems; (3) TUXEL is easy to use; and (4) TUXEL is easy to be understood.

With respect to the first code, the results indicated that the technique was able to evaluate all difficulties and/or problems faced by the participants during their daily use of the LMS. When the participants were asked whether the technique was not able to identify some problem or difficulty faced by them, all participants answered “no”. It may indicate that the set of items evaluated by TUXEL is adequate to evaluate LMSs, particularly the one in this scenario.

The second code reveals that “TUXEL helps inspectors identifying usability problems”. Some participants reported that TUXEL guides inspectors during the evaluation. According to participant P7, “it was easy to keep track of the items that should be inspected”. Participant P1 also highlighted that “it was easy to know where to begin searching for errors using the list of items that I should evaluate”. There were also some participants who stated that TUXEL allows finding defects more quickly. Participant P9, for instance, stated that “the method helps finding the errors in a quickly way”. Moreover, he also stated “TUXEL helped a lot with the tips and the preselected topics. Thus it was easier to find or not the errors”. Finally, there were also participants that stated that “TUXEL is objective”, which makes it easy to be understood.

The third and fourth codes indicate that TUXEL is easy to use . Participant P1, for instance, stated that he “had no difficulty with the method”, while participant P3 stated that “the objective questions described very well [the problems]”.

Although these results present some good indicators about some quality attributes of TUXEL, there were also participants that faced some difficulties using it. Participant P3 stated,“the part that is evaluated with radio buttons [UX evaluation] was a bit confusing to me”. Participant P9 stated that “[the technique] keeps [the evaluation] very open, there are times that you get lost”. Finally, the participant P6 stated that “[it was difficult] to look for where the cited problems can occur to check whether they happen or not”. These difficulties may be explained by the fact that participants P3 and P9 have low level of experience regarding usability evaluation, while participant P6 has no experience, although other 2 participants that also have no experience did not face or did not report any difficulty. However, these quotations indicate opportunities for improving the technique by making, for example, the items and the instructions clearer to novice inspectors.

Finally, two participants made suggestions to improve TUXEL. Participant P5, who has medium experience with usability/UX evaluations, suggested to “slightly reduce [the technique]”. It may indicate that he felt a slight cognitive overload throughout the evaluation process due to the number of evaluated items or the number of steps to carry out the evaluation,

indicating improvement opportunities. The participant P9, in turn, suggested to “have something more centered, following something less open”.

With respect to **participant’s perceptions regarding AHE**, we identified the following codes: (1) AHE is objective; (2) AHE is easy to be remembered; (3) AHE is difficult to be understand; and (4) AHE is difficult for evaluating.

Regarding the first and second codes, the results indicate that the items provided in AHE technique are objective and easy to remember. According to participant P20, “the method has exactly what should be performed and evaluated”, while participant P12 stated that “[it was easy] to remember the heuristics during the method”.

Even though these quotations indicate positive aspects of the AHE, several participants reported some difficulties in using the technique during the evaluation process. With respect to the third category, 2 participants reported some difficulty in understanding the heuristics. Participant P14, for instance, stated that his/her difficulty was regarding the “items without details of what it in fact is about”.

In fourth category, “AHE is difficult for evaluating”, the participants reported some difficulties regarding the use of the AHE technique. Some participants stated that the technique requires some familiarity from the inspector. According to participant P17, he/she “had to go back many times to each item to try to evaluate and check what could be useful, which turned out to be a bit annoying”. Other participants reported that AHE does not address some problems. As a result, the inspector may not report the problem or it may be assigned to a wrong heuristic, which is in accordance to the quantitative results (subsection 6.5.1), where participants using AHE made a higher number of mistakes in comparison to those using TUXEL. Finally, some participants stated that AHE has a high number of items to be verified at once. Since it does not select the heuristics according to the tasks/pages being evaluated, the inspector needs to check all the heuristics on every task/page.

Some participants presented suggestions to improve AHE. Participant P19 suggested detailing the items, “putting something more specific for a better evaluation”. In the same line, participants P12, P20 and P21 suggested putting examples in order to help the inspector during the evaluation. Finally, 2 participants stated that the techniques could vary according to the screen being evaluated. According to participant P17, we “could try to group the items that would appear one at a time and that would advance as inspector clicked. This would save effort in finding out what topics (among many) could be reported”.

Regarding participants’ perceptions about the tool, which were used to support the evaluation process in both techniques, we identified the following main codes: (1) easy to use;

(2) useful for the inspection process; (3) flexible; (4) makes the inspector feel comfortable; (5) has functionality problems; and (6) is difficulty to perform some tasks.

With respect to the first code, “the tool is easy to use”, 2 participants from TUXEL group and 6 from AHE stated that the tool is easy to use. The feature that was most mentioned as easy was reporting the problem by selecting and capturing the screen, cited by 5 participants. Other participants stated that the easiest part was to make comments about the problem identified. There was also one participant (P19) who stated that “in general, the tool is good and straight forward, very easy to use”.

In relation to the second code, 3 participants stated that the tool was useful to inspect the LMS. According to participant P2, “the tool facilitated/supported the evaluation”. Moreover, participant P19 stated, “the screen capture mode is very practical and helps a lot”. Regarding third code, “flexibility”, 3 participants considered the tool flexible. All these participants mentioned that the drag and drop feature of the window was interesting. The participant P12, for instance, stated that “[he/she felt comfortable with the inspection] in most of the time, because it is flexible, being possible to be dragged through the screen, closed when necessary, without losing the inspection process”.

With respect to the fourth code, one participant reported that the technique makes the inspector feels comfortable. According to participant P1, who has medium experience with usability, “did not feel forced with the inspection like in other experiences where he/she did not have the support from a tool”.

There were also statements that described some difficulties faced by the participants during the inspection process. Regarding functionality problems, the participants reported the following issues: (i) the tool occupies a lot of space on the screen; (ii) the tool repeats the filenames during the screen capture; and (iii) the tool does not save the folder. Regarding the second issue, although Google Chrome always renamed the duplicated files automatically in our tests, it seems that some participants faced problems with this feature, maybe due to different versions of the browser. Regarding the third issue, it is due to a limitation of Google Chrome regarding the permission of the extensions.

The sixieth code is composed by quotations related to “difficulties to perform some tasks” in the tool such as the use of the screen capture feature and the difficulty in cancelling an action and in reporting missing elements on the page. The participant P16, for instance, stated that he/she had difficulty in “selecting the adequate area to report the error (sometimes it was too big or it was not possible to select it, for example, the submenu that appears as we click in our username)”.

With respect to the suggestions, we identified 8 general suggestions to the tool and 3 specific ones to the AHE version. The general suggestions were the following: (1) to not save the files on the disk (locally); (2) to save the folder; (3) to rename the filenames automatically; (4) to change how the window appears on the screen; (5) to reduce the number of steps needed to identify the problem; (6) to add an image editor; (7) to add comments directly to elements on the page; and (8) to make the screen capture optional.

6.4.8. Conclusion from Second Empirical Evaluation

The results of the second empirical evaluation indicated that the inspection-based approach of TUXEL 2.0 provided useful results regarding usability of the evaluated LMS in comparison to the first questionnaire-based version. The two main limitations of the first version were: (i) the impossibility to identify where the problem occur; and (ii) the difficulty of participants in remembering whether a problem or functionality really existed in the platform.

The inspection-based approach of TUXEL 2.0 solved these two limitations. With this approach, the technique allowed identifying the location of the problem, while removing the issue related to the difficulty of the participants in remembering the functionalities or problems of the platform. On the other hand, it requires more time to perform the evaluation process, while it evaluates only the defined tasks, pages or functionalities.

Regarding AHE technique, some participants stated that they could not address some problems to any of the items provided by the heuristics. Moreover, participants who used the AHE matched a greater number of problems to incorrect heuristics in comparison to participants using TUXEL. It may be due to the generic nature of the heuristics. These results indicate that there is a need to provide more detailed items in order to make it easier for novice inspectors performing the evaluation.

In comparison to AHE technique, TUXEL 2.0 found a higher number of problems, generated fewer false-positives and demanded fewer time to be employed. Additionally, even participants without previous experience with inspection were able to find a high number of usability problems, which may indicate that the technique properly guides the inspectors. On the other hand, there is still room for some improvements such as: (i) redesign the UX evaluation step to make it clearer to the inspector; and (ii) provide further details regarding where the inspector may look for possible problems according to the evaluated item.

Regarding the tool, the participants perceived it, in general, as useful to the inspection process, specially the screen marking feature. It was also perceived as easy to use and flexible. Moreover, one participant stated that the tool did not make him/her feel forced like in traditional

inspection. These results may indicate that the tool provides useful features that supports the evaluator during the inspection process. However, some improvements are needed, especially with regard to the marking feature, where many participants reported some difficulty.

6.5. Discussions and Conclusion

In this chapter, we presented the two iterations we carried out during the Design Cycle of the Design Science Research (DSR) methodology. We presented the initial development of our proposed artifact, which consists of a technique to evaluate the usability and UX of Learning Management Systems (LMSs) called TUXEL. The development of TUXEL was driven by theoretical assumptions, from the definition of its approach to the definition its dimensions and items. We aimed to develop a technique that was fast, easy to use and applicable by students with a low cost. Considering these requisites, we decided to develop TUXEL by using a questionnaire-based approach. After developing its first version, we performed the first empirical evaluation to evaluate and compare the technique, by applying it on the problem context.

In the first empirical evaluation, we performed a comparison between TUXEL and Adapted TAM (AT) technique, developed by Theng and Sin (2012). The results showed both techniques were considered easy to use and useful, with a little advantage to TUXEL. However, both techniques did not provide, in general, sufficient feedback about the problems faced by the students. It was not possible, for example, to identify where the problem occurs on the LMS. Moreover, the use of a scale made it difficult to interpret the results when verifying whether the item would be considered an indicative of a problem or not, given that some items had few difference between the percentage of agreements, neutrals and disagreements. Finally, a considerable number of participants using TUXEL stated that they had difficulty in remembering whether the evaluated aspect was present or not in the platform. It may be explained by the fact that the technique evaluates aspects that students are not aware during their daily use. Considering these results and the feedback provided by the participants in the feedback questionnaire, we decided to perform a second iteration over the Design Cycle of the DSR methodology by improving our prosed artifact, giving rise to TUXEL 2.0.

To develop TUXEL 2.0 we considered the following: (i) the difficulty of participants in remembering or paying attention to specific aspects of the LMS; (ii) the need for a comment field to participants describe the problem.; (iii) the difficulty in identifying where the problem

occurs on the platform. By considering these issues, we decided to change the TUXEL to an inspection-based approach and also developed a tool to support the inspection process.

In the second empirical evaluation we evaluated TUXEL 2.0 by comparing it to an Adapted Heuristic Evaluation (AHE) proposed by Mtebe and Kissaka (2015). The results indicated that participants using TUXEL 2.0 identified a higher number of problems while performing the inspection, on average, in a fewer amount of time in comparison to participants using AHE. Additionally, the results revealed that TUXEL 2.0 can be successfully used to inspect a LMS, from novice evaluators, without knowledge of usability evaluation to expert evaluators. On the other hand, participants had difficulty in evaluating by using AHE. It may indicate that there is a need to provide items with higher level of granularity in order to make the technique applicable by novice inspectors.

The results of the second iteration corroborate to the theoretical assumptions regarding the use of inspection-based approach to evaluate LMSs. Additionally, our study provided new knowledge on the feasibility of using this approach by novice inspectors (i.e. learners), given that the studies identified in our systematic mapping applied this approach by using only expert evaluators, and not novice ones.

Regarding threats to validity, the main limitation with the first empirical evaluation relates to the sample of the population and the length of the questionnaires. Considering that students from a single institution composed the sample and that the questionnaires contained several questions, it may be possible that some students wanted to finish it quickly and did not fill the questions properly. However, considering that we did not interpret the results by using the mean, which is more sensible to variations (Manikandan, 2011), this issue may have been minimized.

With respect to the second empirical evaluation, we can identify three threats to validity regarding the employed instruments and one internal validity. First, the use of the tool may have affected, as stated before, the participants' perception about the techniques. However, the use of the open-ended questions allowed us to triangulate and better interpret the results. Second, since the participants carried out the evaluation at their own home, without supervision, the time spent by each participant may have not been accurately measured. To minimize this issue, we measured the time spent by each participant automatically by the tool. We also instructed them to take a time to carry out the evaluation and that, once started, they should carry it out until the end, without interruption. Third, since the participants performed the tasks during the inspection process, the time spent during the accomplishment of these tasks may have influenced the final inspection time. To minimize this threat, we asked teachers to provide

simple and easy questions, in order not to require them to search for answers. Finally, regarding internal validity, given that the author of this dissertation was one of the researchers who evaluated whether a discrepancy identified by the participants is a problem or not, the results may be biased. To minimize this threat, an external researcher carried out the evaluation. Additionally, a third researcher also performed the evaluation in the cases of divergence between the two researchers.

In conclusion, this chapter reports the development of a technique to evaluate the usability and UX of LMSs called TUXEL. The empirical studies conducted during the development process may provide some insights to researchers of this area about the techniques employed. Additionally, the applicability of TUXEL 2.0 indicates the possibility to employ inspection-based techniques by evaluators without previous experience with inspection, since the techniques provide adequate guidance to them. Moreover, the methodology employed in this research may serve as the basis to improve or develop new artifacts, such as usability/UX evaluation techniques.

CHAPTER 7 – CONCLUSIONS AND FUTURE WORK

In this chapter, we present our final remarks regarding TUXEL, a technique designed to evaluate the usability and UX of Learning Management Systems. We also highlight the contributions of this research and our future works.

7.1. Epilogue

This work presented the development process of TUXEL, a technique to evaluate the usability and UX of Learning Management Systems. To develop the technique, we followed the Design Science Research (DSR) approach.

Initially we carried out a systematic mapping to identify the usability and UX evaluation techniques applied to evaluate LMSs (see Chapter 3). The results allowed the identification of some gaps, such as the lack of comparative studies, the lack of studies evaluating the UX of LMSs, and the lack of a technique that evaluates both usability and UX of LMSs.

Given that few studies evaluated the UX of LMSs, and that generic UX evaluation techniques have been used in the evaluation process, we performed a preliminary study (Chapter 4) in order to obtain further information regarding the feasibility of these techniques. The results indicated that techniques based on questionnaires with scales limit users in conveying their UX, specifically due to the lack of a field for comments. The results served as the basis for the development of the UX part of TUXEL, while also filling the gap regarding the lack of studies conducted to evaluate the UX of LMSs.

Based on the results of the systematic mapping and of the preliminary study, we started the development of TUXEL, defining its dimensions and items. Given that the technique aimed to be applied by students, we designed it to be easy to use and to capture the usability and UX of LMSs from the student's point of view. Thus, we decided to use the questionnaire-based approach.

We conducted two empirical studies in order to verify its feasibility. The first empirical study revealed that the questionnaire-based approach used by the first version of TUXEL and the Adapted TAM (Theng and Sin, 2012) technique did not provide sufficient information about the identified problem. It was not possible, for instance, to identify where the problem occurs in the platform, which makes it difficult to develop solutions for it. Additionally, due to the specificity of some items of TUXEL, many students reported that they did not remember whether the platform has a given problem or lacks a given functionality.

The results of the first empirical study served as the basis for the development of TUXEL 2.0. In this version, we changed the questionnaire-based approach to a guided inspection approach. Additionally, we developed a tool to support the inspection process. In order to verify the feasibility of TUXEL 2.0, we carried out a second empirical study.

The results of the second empirical study revealed the success of the guided inspection approach used by TUXEL 2.0. The technique, for instance, allowed students that did not have previous experience with usability inspection to detect usability problems in the LMS. Additionally, TUXEL 2.0 identified a greater number of problems in comparison to the Heuristic Evaluation proposed by Mtebe and Kissaka (2015), while requiring less time. Thus, TUXEL 2.0 revealed to be adequate to evaluate the usability and UX of LMSs, answering our main research question: "How to evaluate the usability and UX of Learning Management Systems aiming to improve the quality in use of these platforms from the learners' point of view?" By using our technique, researchers may identify usability and UX issues, making it possible to develop solutions to improve the quality in use of these platforms and, consequently, the learning process through them. However, there are still some opportunities for improving the technique, which could be implemented in future works.

7.2. Contributions

The main contributions of this master's dissertation are:

- **A secondary study in the field of usability and UX evaluation of Learning Management Systems:** our systematic mapping may contribute to the scientific community by an overview of usability and UX evaluation techniques in the context of LMSs. The results may serve as a starting point for the beginning of new research in this area;
- **A methodology for the development of artifacts based on Design Science Research:** the methodology followed to develop TUXEL may serve as the basis for the development of new artifacts, especially usability/UX evaluation techniques;
- **A technique for evaluating usability and UX of LMSs:** researchers or professionals working with LMSs may use our technique (TUXEL) to evaluate these platforms regarding usability and UX. The outcomes of TUXEL may help them identify usability and UX problems and make it possible to improve the quality of the evaluated LMS;

- **A tool to support the evaluation process:** the tool we developed may be used to support the usability/UX evaluation, as well as serve as the basis for the development of new evaluation tools that makes the identification and reporting of problems easier and straightforward;
- **An overview of some usability and UX evaluation techniques:** the empirical studies carried out during the development of TUXEL may provide an overview of the compared techniques regarding their feasibility and effectiveness to evaluate LMSs. Additionally, the strengths and the weaknesses of each technique may provide an insight to the development of new evaluation techniques.

Research papers resulted from this research:

- Nakamura, W. T., de Oliveira, E. H. T., & Conte, T. (2017). Usability and User Experience evaluation of Learning Management Systems: A systematic mapping study. In Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS), 10(1), 97-108. (Best student paper in the HCI track)
- Nakamura, W., Marques, L., Rivero, L., Oliveira, E., & Conte, T. (2017). Are Generic UX Evaluation Techniques Enough? A study on the UX Evaluation of the Edmodo Learning Management System. In Brazilian Symposium on Computers in Education (Vol. 28, No. 1, p. 1007). (Best paper in the HCI track)
- Nakamura, W. T., Gadelha, B., de Oliveira, E. H. T., & Conte, T. (2018). TUXEL: a Technique for User eXperience Evaluation in e-Learning. In *Interacting With Computers Journal* (paper submitted).
- Nakamura, W., Marques, L., Rivero, L., Oliveira, E., & Conte, T. (2018). Are scale-based techniques enough for learners to convey their UX when using a Learning Management System? In Brazilian Journal of Informatics in Education (paper submitted)

7.3. Future Works

The results of our research provide some perspectives to be explored in future works:

- **Evolution of the proposed technique:** as shown in Chapter 6, although TUXEL 2.0 provided good results, there are still opportunities for improving the technique and the tool. An evolution of the technique may be the development of an approach to evaluate the UX on each task performed by the inspector;

- **Evolution of the proposed tool:** since our tool is an extension of Google Chrome, it has some limitations, such as the impossibility of saving the screenshots automatically without prompting the user, while also limiting its application to the aforementioned browser. The development of a tool that is not browser dependent may provide flexibility for researchers to carry out the evaluation in different browsers. Additionally, there is a need to change the approach used by the marking feature, given that it does not allow, for instance, marking menus that are activated by mouse hover. A possible solution for this problem is to capture the screen by pressing a given key and then perform the marking in a built-in image editor;
- **Further evaluations:** although the results of our second empirical study indicated the feasibility of TUXEL 2.0, further studies may be conducted, for instance, to evaluate different LMSs, as well as different student profiles, environments (mobile and Web) and courses. Additionally, further studies may evaluate functionalities different from those normally used or those that are specific of a given LMS in order to verify whether TUXEL 2.0 can evaluate them.

REFERENCES

Alkhattabi, M. (2015). Usability Evaluation of Tadarus: Student Perceptions. Procedia Computer Science, 62, 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, 4(3), 270-283.

Ardito, C., Costabile, M. F., Lanzilotti, R., & Montinaro, F. (2007). Towards the evaluation of UX. Towards a UX Manifesto, 6.

Aslan, S., & Reigeluth, C. M. (2011). A trip to the past and future of educational computing: Understanding its evolution. Contemporary Educational Technology, 2(1).

Barab, S. A., & Duffy, T. (2000). From practice fields to communities of practice. Theoretical foundations of learning environments, 1(1), 25-55.

Basili, V. R., & Rombach, H. D. (1988). The TAME project: Towards improvement-oriented software environments. IEEE Transactions on software engineering, 14(6), 758-773. BEVAN, N., 1995. Measuring usability as quality of use. Software Quality Journal, v. 4, p. 115-150.

Bevan, N. (1995). Measuring usability as quality of use. Software Quality Journal, 4(2), 115-130.

Blecken, A., Bruggemann, D., & Marx, W. (2010). Usability evaluation of a learning management system. In System Sciences (HICSS), 2010 43rd Hawaii International Conference on (pp. 1-9). IEEE.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), 37-46.

Conte, T., Massolar, J., Mendes, E., & Travassos, G. H. (2009). Web usability inspection technique based on design perspectives. IET software, 3(2), 106-123.

Conte, T., Massollar, J., Mendes, E., & Travassos, G. H. (2007). Usability evalua-tion based on Web design perspectives. In Empirical Software Engineering and Measurement, 2007. ESEM 2007. First International Symposium on (pp. 146-155). IEEE.

Corbin, J., & Strauss, A. (2014). Basics of qualitative research. Sage.

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 (p. 12). ACM.

Davids, M. R., Harvey, J., Halperin, M. L., & Chikte, U. M. (2015). Determining the number of participants needed for the usability evaluation of e-learning resources: A Monte Carlo simulation. British Journal of Educational Technology, 46(5), 1051-1055.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. Management science, 35(8), 982-1003.

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 (pp. 122-131). ACM.

Dougiamas, M. and Taylor, P. (2003). Moodle: Using learning communities to create an open source course management system.

Dubost, A., Klein, M. R., & Dang, J. (2004). Building interoperability between LMS and brokerage platforms. WSEAS Transactions on Information Science and Applications, 1(5), 1341-1347.

Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. Information and Software Technology, 53(8), 789-817.

Fetaji, M., & Fetaji, B. (2011). Devising m-learning usability framework. In Information Technology Interfaces (ITI), Proceedings of the ITI 2011 33rd International Conference on (pp. 275-280). IEEE.

Fetaji, M., Dika, Z., & Fetaji, B. (2008). Usability testing and evaluation of a mobile software solution: a case study. In Information Technology Interfaces, 2008. ITI 2008. 30th International Conference on (pp. 501-506). IEEE.

Freire, L. L., Arezes, P. M., & Campos, J. C. (2012). A literature review about usability evaluation methods for e-learning platforms. Work, 41(Supplement 1), 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 IEEE (pp. 1-8). IEEE.

Granic, A., & Cukusic, 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), 107.

Guo, Y., Qian, D., Guan, J., & Wang, J. (2010). Usability testing on a government training platform: A case study. In Education Technology and Computer (ICETC), 2010 2nd International Conference on (Vol. 2, pp. V2-211). IEEE.

Hamdi, N., Elbadrawy, R., & Aziz, R. A. (2011). An experimental study to measure mobile learning system usability at the AASTMT. mobile, 167.

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2016). Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. Computers in Human Behavior, 61, 463-471.

Hassenzahl, M. (2008, September). User experience (UX): towards an experiential perspective on product quality. In Proceedings of the 20th Conference on l'Interaction Homme-Machine (pp. 11-15). ACM.

Hassenzahl, M. (2003). The thing and I: understanding the relationship between user and product. In Funology (pp. 31-42). Springer Netherlands.

Hassenzahl, M., Law, E. L. C., & Hvannberg, E. T. (2006). User Experience-Towards a unified view. Ux Ws Nordichi, 6, 1-3.

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

Hevner, A. R. (2007). A three cycle view of design science research. Scandinavian journal of information systems, 19(2), 4.

Hevner, A. and Chatterjee, S. (2010). Design Research in Information Systems: Theory and Practice. In Integrated Series in Information Systems, v. 22. Springer.

Hijon-Neira, R., Velázquez-Iturbide, Á., Pizarro-Romero, C., & Carriço, L. (2014, October). Merlin-know, an interactive virtual teacher for improving learning in Moodle. In Frontiers in Education Conference (FIE), 2014 IEEE (pp. 1-8). IEEE.

Hollingsed, T., & Novick, D. G. (2007). Usability inspection methods after 15 years of research and practice. In Proceedings of the 25th annual ACM international conference on Design of communication (pp. 249-255). ACM.

Hovde, M. R. (2015, July). Effective user experience in online technical communication courses: employing multiple methods within organizational contexts to assess usability. In Proceedings of the 33rd Annual International Conference on the Design of Communication (p. 30). ACM.

ISO/IEC 9126-1 (2001). Software engineering - Product quality - Part 1: Quality model.

ISO 9241-11 (2017). Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts.

ISO 9241-210 (2010). Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems.

Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. ACM Computing Surveys (CSUR), 33(4), 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. The International Review of Research in Open and Distributed Learning, 16(4).

Kakasevski, G., Mihajlov, M., Arsenovski, S., & Chungurski, S. (2008, June). Evaluating usability in learning management system Moodle. In Information Technology Interfaces, 2008. ITI 2008. 30th International Conference on (pp. 613-618). IEEE.

Kitchenham, B., & Chartes, S. (2007). Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report – EBSE 2007, Durham.

Kongchan, C. (2008). How a non-digital native teacher makes use of Edmodo. In 5th Intenational Conference ICT for language learning.

Koohang, A. (2004). Expanding the concept of usability. Informing Science: International Journal of an Emerging Transdiscipline, 7, 129-141.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., & Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. Interacting with Computers, 23(5), 473-483.

Laitenberger, O., & Atkinson, C. (1999). Generalizing perspective-based inspection to handle object-oriented development artifacts. In Proceedings of the 21st international conference on Software engineering (pp. 494-503). ACM.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. biometrics, 159-174.

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), 52-69.

Lanzilotti, R., Ardito, C., Costabile, M. F., & De Angeli, A. (2006). eLSE methodology: a systematic approach to the e-learning systems evaluation. Journal of Educational Technology & Society, 9(4).

Laugwitz, B., Held, T., & Schrepp, M. (2008, November). Construction and evaluation of a user experience questionnaire. In Symposium of the Austrian HCI and Usability Engineering Group (pp. 63-76). Springer Berlin Heidelberg.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participa-tion. Cambridge university press.

Matera, M., Rizzo, F., & Carughi, G. T. (2006). Web usability: Principles and evaluation methods. In Web engineering (pp. 143-180). Springer, Berlin, Hei-delberg.

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), 197-203.

Mehlenbacher, B., Bennett, L., Bird, T., Ivey, M., Lucas, J., Morton, J., & Whitman, L. (2005, July). Usable e-learning: A conceptual model for evaluation and design. In Proceedings of HCI International (Vol. 2005, p. 11th).

Meho, L. I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: a comparison of Scopus and Web of Science. Journal of the Association for Information Science and Technology, 59(11), 1711-1726.

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same?. The Internet and Higher Education, 14(2), 129-135.

Mtebe, J. S., & Kissaka, M. M. (2015, May). Heuristics for evaluating usability of Learning Management Systems in Africa. In IST-Africa Conference, 2015 (pp. 1-13). IEEE.

Nakamura, W. T., de Oliveira, E. H. T., & Conte, T. (2017a). Usability and User Experience evaluation of Learning Management Systems: A systematic map-ping study. In Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS), 10(1), 97-108.

Nakamura, W., Marques, L., Rivero, L., Oliveira, E., & Conte, T. (2017b). Are Generic UX Evaluation Techniques Enough? A study on the UX Evaluation of the Edmodo Learning Management System. In Brazilian Symposium on Computers in Education (Vol. 28, No. 1, p. 1007).

Nascimento, I., Silva, W., Gadelha, B., & Conte, T. (2016, July). Userbility: A Technique for the Evaluation of User Experience and Usability on Mobile Applications. In International Conference on Human-Computer Interaction (pp. 372-383). Springer International Publishing.

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

Nielsen, J. (1990). Evaluating hypertext usability. Designing hypermedia for learning, 147-168.

Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R. L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, 25-64.

Nielsen, J. (2012). Usability 101: Introduction to Usability. Nielsen Norman Group. Available in: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>.

Nielsen, J., & Molich, R. (1990, March). Heuristic evaluation of user interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 249-256). ACM.

Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. Journal of Educational Technology & Society, 9(2).

Norman, D., Miller, J., & Henderson, A. (1995, May). What you see, some of what's in the future, and how we go about doing it: HI at Apple Computer. In Conference companion on Human factors in computing systems (p. 155). ACM.

Ogunbase, A. (2014). Pedagogical Design and Pedagogical Usability of Web-based Learning Environments: Comparative Cultural Implications between Africa and Europe. In J. Viteli & M. Leikomaa (Eds.), Proceedings of EdMedia 2014--World Conference on Educational Media and Technology (pp. 840-849). Tampere, Finland: Association for the Advancement of Computing in Education (AACE).

Orfanou, K., Tselios, N., & Katsanos, C. (2015). Perceived usability evaluation of learning management systems: Empirical evaluation of the System Usability Scale. The International Review of Research in Open and Distributed Learning, 16(2).

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), 455-469.

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

Petrie, H., & Bevan, N. (2009). The Evaluation of Accessibility, Usability, and User Experience.

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

Quinn, C. N. (1996). Pragmatic evaluation: lessons from usability. In 13th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (pp. 15-22).

Reeves, T. C., Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., ... & Loh, S. (2002). Usability and Instructional Design Heuristics for E-Learning Evaluation.

Rivero, L. and Conte, T. (2017). A systematic mapping study on research contributions on UX evaluation technologies. In Proceedings of the 16th Brazilian Symposium on Human Factors in Computing Systems (IHC 2017), Joinville, Brazil.

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.

Rusu, C., Rusu, V., Roncagliolo, S., & González, C. (2015). Usability and user experience: what should we care about?. International Journal of Information Technologies and Systems Approach (IJITSA), 8(2), 1-12.

Sánchez-Chamochín, E., Cambranes-Martínez, E., & Menéndez-Domínguez, V. (2008, October). Adapting Mobile Access Scheme for a Legacy e-Learning Platform. In Computer Science, 2008. ENC'08. Mexican International Conference on (pp. 96-103). IEEE.

Santoso, H. B., Isal, R. Y. K., Basaruddin, T., Sadira, L., & Schrepp, M. (2014, September). in-progress: User experience evaluation of Student Centered E-Learning Environment for computer science program. In User Science and Engineering (i-USEr), 2014 3rd International Conference on (pp. 52-55). IEEE.

Schrepp, M. (2010). GOMS analysis as a tool to investigate the usability of web units for disabled users. Universal Access in the Information Society, 9(1), 77-86.

Shiratuddin, N., Hassan, S., & Landoni, M. (2003). A usability study for promot-ing eContent in higher education. Educational Technology & Society, 6(4), 112-124.

Shneiderman, B., Plaisant, C., Cohen, M. S., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016). Designing the user interface: strategies for effective human-computer interaction. Pearson.

Silius, K., & Tervakari, A. M. (2003). An evaluation of the usefulness of web-based learning environments. The evaluation tool into the portal of Finnish virtual university. In International Conference of Network Universities and e-learning. Valencia, España.

Silius, K., Tervakari, A. M., & Pohjolainen, S. (2003). A multidisciplinary tool for the evaluation of usability, pedagogical usability, accessibility and informational quality of web-based courses. In The Eleventh International PEG Conference: Powerful ICT for Teaching and Learning (Vol. 28).

Squires, D., & Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interacting with computers, 11(5), 467-483.

Squires, D., & Preece, J. (1996). Usability and learning: evaluating the potential of educational software. Computers & Education, 27(1), 15-22.

Ssekakubo, G., Suleman, H., & Marsden, G. (2014). A streamlined mobile user-interface for improved access to LMS services. In The Sixth International Conference on Mobile, Hybrid, and On-line Learning.

Ssemugabi, S., & De Villiers, R. (2007, October). 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 on IT research in developing countries (pp. 132-142). ACM.

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, 173-184.

Stickel, C., Ebner, M., & Holzinger, A. (2011). Shadow expert technique (SET) for interaction analysis in educational systems. Universal Access in Human-Computer Interaction. Applications and Services, 642-651.

Sullivan, G. M. and Artino Jr, A. R. (2013). Analyzing and interpreting data from Likert-type scales. In Journal of Graduate Medical Education, v. 5, n. 4, p. 541-542.

Tähti, M., & Niemelä, M. (2006, February). 3e–expressing emotions and experiences. In WP9 Workshop on Innovative Approaches for Evaluating Affective Systems.

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), 243-252.

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

Van Der Linden, J., & Van De Leemput, C. (2015). Observatory of students' uses of computer-based tools. Psychologie française, 60(2), 145-157.

Van Schaik, P., & Ling, J. (2010). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. Interacting with Computers, 23(1), 18-32.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision sciences, 39(2), 273-315.

Violante, M. G., & Vezzetti, E. (2015). Virtual interactive e-learning application: An evaluation of the student satisfaction. Computer Applications in Engineering Education, 23(1), 72-91.

Vogten, H., & Koper, R. (2014). Towards a new generation of Learning Management Systems.

Whitten, N. (1990). Managing software development projects. Formula for success. New York: Wiley,| c1990.

Wieringa, R. (2014). Design Science Methodology for Information Systems and Software Engineering. Springer-Verlag Berlin Heidelberg.

Wieringa, R. (2009). Design science as nested problem solving. In Proceedings of the 4th international conference on design science research in information systems and technology (p. 8). ACM.

Wilson, C. (2013). User interface inspection methods: a user-centered design method. Newnes.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th international conference on evaluation and assessment in software engineering (p. 38). ACM.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.

Yusoff, S. R. M., & Zin, N. A. M. (2012, September). Design and evaluation of collaborative learning management system (clms) framework for teaching technical subject. In International Conference on Web-Based Learning (pp. 79-89). Springer, Berlin, Heidelberg.

Zaharias, P., Vasslopoulou, K., & Poulymenakou, A. (2002). Designing on-line learning courses: Implications for usability. Scientific Journal on Applied Information Technology, 1(1).

Zaharias, P. (2009). Usability in the Context of e-Learning: A Framework Augmenting 'Traditional'Usability Constructs with Instructional Design and Motivation to Learn. International Journal of Technology and Human Interaction (IJTHI), 5(4), 37-59.

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

Zaharias, P., & Koutsabasis, P. (2011). Heuristic evaluation of e-learning courses: a comparative analysis of two e-learning heuristic sets. Campus-Wide Information Systems, 29(1), 45-60.

Zaharias, P., & Pappas, C. (2016). Quality Management of Learning Management Systems: A User Experience Perspective. Current Issues in Emerging eLearning, 3(1), 5.

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

APPENDIX A – LIST OF SELECTED PUBLICATIONS IN THE SECOND FILTER IN SCOPUS DIGITAL LIBRARY

This appendix contains the list of articles selected in the 2nd filter of the Systematic Mapping presented in Chapter 3.

ID	Title	Authors
S01	*Expanding the concept of usability*	Koohang, A.
S02	*Determining effective distance learning designs through usability testing*	Miller-Cochran, S.K.
S03	*An approach to usability evaluation of e-learning applications*	Ardito, C.; Costabile, M.F.; De Marsico, M.; Lanzilotti, R.; Levialdi, S.; Roselli, T.; Rossano, V.
S04	*A usability evaluation method for e-learning: Focus on motivation to learn*	Zaharias, P.
S05	*eLSE methodology: A systematic approach to the e-learning systems evaluation*	Lanzilotti, R.; Ardito, C.; Costabile, M.F.; De Angeli, A.
S06	*A comparative study of two usability evaluation methods using a web-based e-learning application*	Ssemugabi, S.; De Villiers, R.b
S07	*Usability and accessibility evaluations along the eLearning cycle*	Martin, L.; Gutiérrez y Restrepo, E.; Barrera, C.; Ascaso, A.R.; Santos, O.C.; Boticario, J.G.
S08	*Accessible and adaptive e-Learning materials: Considerations for design and development*	Debevc, M.; Stjepanovič, Z.; Povalej, P.; Verlič, M.; Kokol, P.
S09	*Development of an E-learning system for occupational medicine: Usability issues*	Mazzoleni, M.C.; Rognoni, C.; Finozzi, E.; Giorgi, I.; Pugliese, F.; Pagani, M.; Pimbriani, M.
S10	*Cross-cultural differences in perceptions of E-learning usability: An empirical investigation*	Zaharias, P.
S11	*Evaluating usability in learning management system moodle*	Kakasevski, G.; Mihajlov, M.; Arsenovski, S.; Chungurski, S.
S12	*Usability testing and evaluation of a mobile software solution: A case study*	Fetaji, M.; Dika, Z.; Fetaji, B.
S13	*Adapting mobile access scheme for a legacy e-learning platform*	Sánchez-Chamochin, E.L.; Cambranes-Martínez, E.; Menéndez-Domínguez, V.
S14	*Pattern-based usability evaluation of e-learning systems*	Zub, D.; Eessaar, E.
S15	*Usability metrics of time and stress - Biological enhanced performance test of a university wide learning management system*	Stickel, C.; Scerbakov, A.; Kaufmann, T.; Ebner, M.

S16	*Developing a usability evaluation method for e-learning applications: Beyond functional usability*	Zaharias, P.; Poylymenakou, A.
S17	*Usability in the context of e-learning: A framework augmenting 'traditional' usability constructs with instructional design and motivation to learn*	Zaharias, P.
S18	*Application of VoiceXML in e-learning systems*	Azeta, A.A.; Ayo, C.K.; Atayero, A.A.; Ikhu-Omoregbe, N.A.
S19	*A case study of usability testing on an asynchronous e-Learning platform*	Guo, Y.; Wang, J.; Moore, J.; Liu, M.; Chen, H.-L.
S20	*Usability testing methods on e-learning environment*	Jókai, E.
S21	*Usability evaluation methods: Mind the gaps*	De Kock, E.; Van Biljon, J.; Pretorius, M.
S22	*Seeing the system through the end users' eyes: Shadow expert technique for evaluating the consistency of a learning management system*	Stickel, C.; Fassold, M.; Holzinger, A.; Ebner, M.
S23	*Usability evaluation of a learning management system*	Blecken, A.; Bruggemann, D.; Marx, W.
S24	*UseLearn: A novel checklist and usability evaluation method for eLearning systems by criticality metric analysis*	Oztekin, A.; Kong, Z.J.; Uysal, O.
S25	*Usability testing on a government training platform: A case study*	Guo, Y.; Qian, D.; Guan, J.; Wang, J.
S26	*Embedding interactivity into a university study portal by using social computing*	Todoran, I.
S27	*Do patterns help novice evaluators? A comparative study*	Lanzilotti, R.; Ardito, C.; Costabile, M.F.; De Angeli, A.
S28	*Evaluation of usability in a remote learning system utilizing Markov models*	Rodrigues, J.; Diniz, M.; Ferreira, S.B.L.; Silveira, D.S.; Capra, E.
S29	*An experimental study to measure mobile learning system usability at the aastmt*	Hamdi, N.; Elbadrawy, R.; Aziz, R.A.
S30	*Shadow Expert Technique (SET) for interaction analysis in educational systems*	Stickel, C.; Ebner, M.; Holzinger, A.
S31	*Usability testing and expert inspections complemented by educational evaluation: A case study of an e-learning platform*	Granić, A.; Ćukušić, M.
S32	*Heuristic evaluation of e-learning courses: A comparative analysis of two e-learning heuristic sets*	Zaharias, P.; Koutsabasis, P.
S33	*Establishment of the genetic/genomic competency center for education*	Calzone, K.A.; Jerome-D'Emilia, B.; Jenkins, J.; Goldgar, C.; Rackover, M.; Jackson, J.; Chen, Y.; Voss, J.; Feero, W.G.
S34	*Devising M-learning usability framework*	Fetaji, M.; Fetaji, B.
S35	*Comparative study of efficiency among the developed MLUAT methodology in comparison with*	Fetaji, B.; Fetaji, M.; Kaneko, K.

	qualitative user testing method and heuristics evaluation	
S36	*Work in progress - Open source usability evaluation: The case of Moodle*	Baytiyeh, H.
S37	*A literature review about usability evaluation methods for e-learning platforms*	Freire, L.L.; Arezes, P.M.; Campos, J.C.
S38	*Usability evaluation of a LMS mobile web interface*	Ivanc, D.; Vasiu, R.; Onita, M.
S39	*Evaluating usability and efficaciousness of an e-learning system: A quantitative, model-driven approach*	Theng, Y.-L.; Sin, J.
S40	*Perceived usability evaluation of learning management systems: A first step towards standardization of the system usability scale in Greek*	Katsanos, C.; Tselios, N.; Xenos, M.
S41	*Identifying cross-platform and cross-modality interaction problems in e-learning environments*	Da Silva, A.C.; Freire, F.M.P.; Da Rocha, H.V.
S42	*Interaction problems accessing e-learning environments in multi-touch mobile devices: A case study in teleduc*	Da Silva, A.C.; Freire, F.M.P.; De Arruda, A.V.P.; Da Rocha, H.V.
S43	*User testing for Moodle application*	Tee, S.S.; Wook, T.S.M.T.; Zainudin, S.
S44	*A machine learning-based usability evaluation method for eLearning systems*	Oztekin, A.; Delen, D.; Turkyilmaz, A.; Zaim, S.
S45	*A Voice-Enabled framework for recommender and adaptation systems in E-learning*	Azeta, A.A.; Ayo, C.K.; Ikhu-Omoregbe, N.A.
S46	*Evaluating E-learning platforms for schools: Use and usability, user acceptance, and impact on learning*	Stergioulas, L.; Abassi, M.; Xydopoulos, G.; Fakhimi, M.; Margineanu, R.; Rifon, L.A.; Iglesias, M.J.F.c
S47	*A streamlined mobile user-interface for improved access to LMS services*	Ssekakubo, G.; Suleman, H.; Marsden, G.
S48	*Design and evaluation of Collaborative Learning Management System (CLMS) framework for teaching technical subject*	Yusoff, S.R.M.; Mat Zin, N.A.
S49	*Evaluation framework for m-learning systems: Current situation and proposal*	Cota, C.X.N.; Díaz, A.I.M.; Duque, M.A.R.
S50	*The usefulness of usability and user experience evaluation methods on an e-Learning platform development from a developer's perspective: A case study*	Gordillo, A.; Barra, E.; Aguirre, S.; Quemada, J.
S51	*The usability analysis of an E-learning environment*	Torun, F.; Tekedere, H.
S52	*Merlin-know, an interactive virtual teacher for improving learning in Moodle*	Hijon-Neira, R.; Velazquez-Iturbide, A.; Pizarro-Romero, C.; Carrico, L.
S53	*Perceived usability evaluation of learning management systems: Empirical evaluation of the system usability scale*	Orfanou, K.; Tselios, N.; Katsanos, C.

S54	*Heuristics for evaluating usability of Learning Management Systems in Africa*	Mtebe, J.S.; Kissaka, M.M.
S55	*Usability evaluation of Tadarus: Student perceptions*	Alkhattabi, M.
S56	*Observatory of studentsg uses of computer-based tools*	Van Der Linden, J.; Van De Leemput, C.
S57	*Research-in-progress: User experience evaluation of Student Centered E-Learning Environment for computer science program*	Santoso, H.B.; Isal, R.Y.K.; Basaruddin, T.; Sadira, L.; Schrepp, M.
S58	*Usability evaluation of the student centered e-Learning environment*	Junus, I.S.; Santoso, H.B.; Isal, R.Y.K.; Utomo, A.Y.
S59	*Usability Evaluation by Experts of a Learning Management System*	Medina-Flores, R.; Morales-Gamboa, R.
S60	*Testing usability in Moodle: When and How to do it*	Ternauciuc, A.; Vasiu, R.
S61	*Framework to Evaluate M-Learning Systems: A Technological and Pedagogical Approach*	Navarro, C.X.; Molina, A.I.; Redondo, M.A.; Juárez-Ramírez, R.
S62	*Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis*	Harrati, N.; Bouchrika, I.; Tari, A.; Ladjailia, A.

APPENDIX B – EXTRACTION FORM FOR PRIMARY STUDIES

This appendix presents the extraction form used to extract the information from the selected primary studies in the systematic mapping.

TITLE: **AUTHORS:** **PUBLISHED IN:** **YEAR:**	
TABLE FOR DATA EXTRACTION	
Technique description	Brief description of the technique (What does it do? What are its characteristics? How it is applied?)
RESEARCH SUB-QUESTIONS	**ANSWERS**
Q1. What is the origin of the technique?	a) New b) Existing If it is new/adapted, in which it differs from existing techniques?
Q2. What kind of technique is used?	a) Inspection b) Testing c) Inquiry d) Analytical modeling e) Simulation
Q3. How is the technique performed?	a) Manual b) Semiautomatic c) Automatic Describe.
Q4. Does the technique consider specific factors to support learning?	a) Yes. Which ones? b) No
Q5. Does the technique consider only usability, UX, or both?	c) Usability d) UX e) Both
Q6. Does the technique provide any feedback to the evaluator (suggestions for correcting identified usability problems)?	a) Yes. How? b) No
Q7. Was the technique evaluated empirically?	a) Survey b) Case study c) Controlled experiment d) No If so, what were the results?

Q8. Does the technique have any restrictions/conditions to be applied?	a) Yes. Which one? b) No
Q9. Is the technique available for download/consultation? Where?	a) Yes. Where? b) No
Q10. On what kind of platform did the study take place?	a) Desktop b) Mobile c) Both
Q11. Did the study compare techniques?	a) Yes. Which techniques have been compared? b) No

APPENDIX C – EXTRACTION FORM FOR SECONDARY STUDIES

This appendix presents the extraction form used to extract data from selected secondary studies in systematic mapping.

TITLE: **AUTHORS:** **PUBLISHED IN:** **YEAR:**	
EXTRACTION TABLE FOR SYSTEMATIC REVIEWS AND MAPPINGS	
Q1. What is the purpose of the research?	Description of the research goals.
Q2. What are the research questions?	Description of the research questions that the SRL/SML sought to answer.
Q3. Which string was used?	String used in the search.
Q4. In what fields has the string been searched?	Description of the fields in which the string was searched, such as title, abstract, or full-text.
Q5. Which databases have been queried?	Listing of the databases in which the string was run.
Q6. What are the inclusion criteria?	Description of the inclusion criteria.
Q7. How many articles are included?	Total number of articles included after the 2nd filter.
Q8. What information is extracted from the articles?	Description of the fields used in the extraction form.
Q9. Describe the analysis of results.	Description of how the analysis was performed and its results.
Q10. What are the limitations of this SRL/SML?	Description of SRL/SML limitations.

APPENDIX D – MAPPING OF THE PRIMARY STUDIES

This appendix provides a mapping of the primary studies for each research sub-question of the systematic mapping presented in Chapter 3.

ID	Tech.	SQ1		SQ2					SQ3			SQ4		SQ5			SQ6		SQ7				SQ8		SQ9		SQ10			SQ11	
		a	b	a	b	c	d	e	a	b	c	a	b	a	b	c	a	b	a	b	c	d	a	b	a	b	a	b	c	a	b
S01	1	X				X			X			X				X		X	X					X	X		X				X
S02	1		X		X				X							X		X		X				X		X	X				X
S03	1		X	X	X				X					X				X				X		X		X	X				X
S05	1	X		X	X				X			X		X				X			X			X		X	X				X
S06	1	X		X		X			X			X				X		X						X	X						
	2	X				X			X			X				X		X			X			X		X	X			X	
	3	X		X					X			X				X		X						X		X					
S07	1		X	X					X					X				X		X				X		X	X				X
S08	1		X			X			X							X		X						X	X						
	2	X				X			X			X		X				X		X				X		X	X				X
S09	1		X		X				X					X				X		X				X	X						
	2		X			X			X							X		X						X	X		X				X
S11	1		X	X					X					X				X						X		X					
	2		X		X				X					X				X			X			X		X	X				X
	3	X				X			X				X	X				X						X		X					
S12	1	X		X	X	X			X				X	X				X		X				X		X		X			X
S13	1		X		X				X					X				X			X			X	X				X		X
S14	1	X		X					X			X		X				X		X				X		X	X				X
S15	1		X		X				X					X				X						X	X						
	2		X		X				X							X		X						X	X		X				X
	3		X			X			X							X		X						X	X						
	4		X		X						X				X			X					X		X						

ID	Tech.	SQ1		SQ2					SQ3			SQ4		SQ5			SQ6		SQ7				SQ8		SQ9		SQ10			SQ11	
		a	b	a	b	c	d	e	a	b	c	a	b	a	b	c	a	b	a	b	c	d	a	b	a	b	a	b	c	a	b
S16	1	X				X			X			X				X		X	X					X	X		X				X
S17	1	X				X			X			X				X		X				X		X	X		X				X
S18	1		X		X				X					X				X						X	X						
	2		X			X			X					X				X		X				X		X		X			X
S19	1		X		X						X			X				X					X			X					
	2	X				X			X			X		X				X		X				X		X	X				X
S20	1		X		X						X			X				X		X			X		X		X				X
S21	1	X		X					X				X	X				X						X	X						
	2		X		X						X			X				X			X		X		X		X			X	
	3		X			X			X					X				X						X		X					
S22	1		X	X	X	X			X							X		X		X			X		X		X				X
S23	1		X		X				X							X		X						X	X						
	2		X			X			X					X				X	X	X				X		X	X				X
	3		X			X			X							X		X						X		X					
S24	1	X		X					X			X		X				X		X			X		X		X				X
S25	1		X		X	X				X				X				X					X			X					
	2		X			X			X					X				X		X				X		X	X				X
S26	1		X		X	X			X					X				X	X					X	X		X				X
S27	1	X		X					X			X		X				X						X		X					
	2	X		X					X			X		X				X			X			X	X		X			X	
	3		X		X				X							X		X						X	X						
S28	1		X				X	X	X					X				X		X				X	X		X				X
S29	1		X		X				X							X		X						X	X						
	2		X			X			X							X		X		X				X	X			X			X
S30	1		X	X	X	X			X							X		X				X	X		X		X				X
S31	1		X		X				X							X		X						X	X						
	2		X	X					X					X				X		X				X		X	X			X	
	3	X		X					X			X		X				X						X	X						

ID	Tech.	SQ1		SQ2					SQ3			SQ4		SQ5			SQ6		SQ7				SQ8		SQ9		SQ10			SQ11	
		a	b	a	b	c	d	e	a	b	c	a	b	a	b	c	a	b	a	b	c	d	a	b	a	b	a	b	c	a	b
	4	X		X					X			X		X				X						X	X						
	5		X			X			X							X		X						X	X						
	6		X			X			X					X				X						X		X					
S32	1	X		X					X			X		X				X			X			X	X		X			X	
	2	X		X					X			X		X				X						X	X						
S33	1		X			X			X					X				X	X					X		X	X				X
S34	1	X		X					X			X		X				X				X		X	X			X			X
S35	1	X		X	X	X			X				X	X				X						X		X					
	2		X			X			X					X				X			X			X	X			X		X	
	3		X	X					X					X				X						X	X						
S36	1	X				X			X				X	X				X	X					X		X	X				X
S38	1	X			X				X			X		X				X				X		X	X			X			X
S39	1	X				X			X			X		X				X	X					X	X		X				X
S40	1		X			X			X							X		X	X					X	X		X				X
S41	1		X	X					X					X				X		X				X	X				X		X
S42	1		X		X				X					X				X		X				X	X						
	2		X			X			X					X				X						X		X		X			X
S43	1		X			X			X					X				X	X					X	X		X				X
S44	1	X		X						X		X		X				X		X			X			X	X				X
S45	1		X			X			X					X				X	X					X	X			X			X
S46	1		X		X				X					X				X		X				X	X						
	2		X			X			X					X				X						X	X		X				X
S47	1		X	X					X					X				X						X	X						
	2		X			X			X					X				X		X				X	X			X			X
	3	X				X			X			X				X		X						X	X						
S48	1		X		X				X					X				X						X	X						
	2	X				X			X			X		X				X		X				X	X		X				X
S50	1		X		X						X			X				X		X				X	X		X				X

ID	Tech.	SQ1		SQ2					SQ3			SQ4		SQ5			SQ6		SQ7				SQ8		SQ9		SQ10			SQ11	
		a	b	a	b	c	d	e	a	b	c	a	b	a	b	c	a	b	a	b	c	d	a	b	a	b	a	b	c	a	b
	2		X		X						X			X				X						X			X				
	3		X			X			X							X		X							X		X				
	4		X	X					X					X				X							X		X				
	5		X		X	X			X							X		X							X	X					
	1		X		X				X					X				X							X	X					
S51	2		X			X			X					X				X		X					X	X		X			X
	3		X			X			X					X				X							X	X					
S52	1		X			X			X							X		X			X				X	X		X			X
S53	1		X			X			X							X		X	X						X	X		X			X
S54	1	X		X					X			X		X				X			X				X	X		X			X
S55	1	X				X			X			X		X				X	X						X	X		X			X
	1		X			X			X					X				X							X	X					
S56	2		X		X				X					X				X	X						X	X		X			X
	3	X				X			X			X		X				X							X		X				
S57	1		X			X			X						X			X	X						X	X		X			X
	2		X			X			X					X				X							X	X					
S58	1	X				X			X			X				X		X	X						X	X		X			X
	2		X			X			X					X				X							X		X				
S59	1	X		X					X			X		X				X		X					X	X		X			X
	1		X			X			X					X				X							X		X				
	2		X			X			X							X		X							X	X					
S60	3		X		X				X					X				X					X		X		X	X			X
	4		X	X					X					X				X							X	X					
	5		X		X						X			X				X						X		X					
S61	1	X				X			X			X				X		X					X		X			X			X
S62	1		X		X	X				X						X		X		X				X		X		X			X

APPENDIX E – GENERAL CLASSIFICATION OF THE PUBLICATIONS IDENTIFIED BY RIVERO AND CONTE (2017)

This appendix presents the classification of the publications identified by Rivero and Conte (2017) according to the exclusion criteria defined in the selection process of publications related to UX evaluation.

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S001	X	X	X	X			X	X	X				X			X	X		X			X	X		X			EC1,EC4
S002		X		X			X		X						X				X		X			X	X		X	EC1
S003	X		X	X			X		X						X		X		X			X	X		X	X		
S004	X			X			X					X			X				X		X			X	X			EC3
S005			X	X				X			X				X		X	X			X			X			X	EC1,EC5,EC6
S006	X			X			X					X			X				X		X			X	X			EC3
S007			X	X			X					X			X				X		X			X			X	EC1,EC6
S008	X		X	X		X	X					X			X				X		X			X	X			EC3
S009	X	X	X	X			X					X			X		X		X			X			X	X			EC3
S010	X			X				X	X						X			X		X				X	X			EC5
S011		X	X	X				X			X				X		X		X	X				X				X	EC1,EC6
S012			X	X			X		X						X		X					X	X					X	EC1,EC5,EC6
S013	X	X	X	X			X					X			X		X		X			X		X		X	X		
S014		X	X	X			X		X						X		X		X			X		X		X	X		EC1
S015	X			X			X					X			X				X		X			X				X	EC6
S016			X	X			X		X						X		X					X	X					X	EC1,EC5,EC6
S017	X		X	X			X					X			X		X		X		X			X		X	X		EC3
S018	X			X			X	X	X						X				X		X			X		X			

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S019	X			X			X		X					X				X	X			X		X			X	EC6
S020	X		X	X				X				X		X			X		X			X		X			X	EC6
S021		X		X			X					X		X					X	X				X	X			EC1
S022	X		X	X			X	X		X				X			X		X			X		X			X	EC6
S023	X			X				X			X			X			X		X		X			X			X	EC6
S024	X			X				X	X					X				X		X				X	X			EC5
S025	X	X		X			X		X					X			X		X			X		X			X	EC6
S026			X	X			X		X					X			X					X	X				X	EC1,EC5,EC6
S027		X		X			X			X				X					X		X			X	X			EC1
S028	X		X	X				X			X			X				X				X		X			X	EC5,EC6
S029	X		X	X				X				X		X				X	X		X			X	X			EC3
S030			X	X			X		X					X			X			X				X	X			EC1,EC5
S031	X			X			X			X				X					X		X			X	X			
S032	X			X			X					X		X					X		X			X	X			EC3
S033	X	X		X			X					X		X					X			X		X	X			
S034	X			X			X					X		X					X		X			X	X			EC3
S035	X		X	X			X					X		X			X		X		X			X	X			EC3
S036		X		X		X	X				X			X					X		X			X	X			EC1
S037	X			X			X		X					X				X	X			X		X	X			
S038		X	X	X			X					X		X			X		X			X		X		X		EC1,EC6
S039	X			X			X					X		X					X		X			X			X	EC6
S040	X	X	X	X			X	X			X			X			X		X			X	X				X	EC6
S041	X			X			X					X		X		X			X		X			X	X			
S042	X			X			X					X		X					X		X			X			X	EC6
S043			X	X			X					X		X			X					X		X			X	EC1,EC5,EC6
S044			X	X			X					X		X			X				X			X			X	EC1,EC5,EC6

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S045		X	X	X				X			X				X				X	X				X	X			EC1
S046			X	X			X					X			X		X				X			X			X	EC1,EC5,EC6
S047	X			X				X				X			X				X		X			X	X			EC3
S048	X			X			X		X						X			X	X			X		X	X			
S049	X					X	X		X						X		X				X		X		X			EC2,EC5
S050			X	X							X				X		X			X				X				X	EC1,EC5,EC6
S051	X		X	X			X					X			X		X		X		X			X	X			EC3
S052		X	X	X			X			X					X		X		X			X		X	X			EC1
S053	X			X				X	X				X			X						X		X			X	EC4,EC5,EC6
S054	X	X	X	X				X	X						X		X		X			X		X	X			
S055	X	X	X	X			X			X			X	X	X		X		X			X	X					X	EC6
S056	X			X			X					X			X				X		X			X				X	EC6
S057			X	X				X		X					X				X		X			X	X			EC1
S058	X		X	X			X					X			X		X		X		X			X	X			EC3
S059	X			X				X		X					X				X		X			X	X			
S060	X			X				X				X			X				X		X			X				X	EC6
S061	X		X	X			X					X			X		X		X			X		X	X			EC3
S062	X			X			X				X				X				X		X			X	X			
S063	X			X				X	X						X				X			X		X				X	EC6
S064	X		X	X			X		X						X		X					X		X				X	EC5,EC6
S065	X			X			X			X					X				X			X		X				X	EC6
S066	X			X			X		X						X				X		X			X	X			
S067	X			X			X		X						X				X		X			X				X	EC6
S068	X			X				X	X						X				X	X				X	X			
S069	X			X				X	X				X		X	X			X	X				X	X			
S070			X	X			X		X						X		X					X	X					X	EC1,EC5,EC6

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S071	X			X			X		X						X				X			X		X	X			
S072			X	X			X		X						X		X				X			X		X		EC1,EC5,EC6
S073		X	X	X			X	X	X						X			X	X	X				X	X			EC1
S074		X	X	X			X					X			X		X				X			X			X	EC1,EC5,EC6
S075	X			X			X					X			X				X		X		X		X			
S076		X		X			X					X			X				X	X				X	X			EC1
S077	X			X			X					X			X				X		X			X	X			EC3
S078	X			X		X	X	X		X					X				X		X			X	X			
S079	X			X			X					X			X				X		X			X	X			
S080	X			X			X					X			X				X		X			X	X			
S081			X	X			X					X			X		X				X			X			X	EC1,EC5,EC6
S082	X			X			X					X			X				X	X				X			X	EC6
S083	X			X			X		X						X				X		X			X	X			
S084			X	X			X					X	X		X	X	X				X			X	X			EC1,EC5
S085	X		X	X			X			X					X		X		X		X			X	X			
S086	X				X	X	X		X						X				X		X			X			X	EC2,EC6
S087	X	X	X	X			X					X			X		X		X			X		X			X	EC6
S088	X			X			X					X			X	X					X			X			X	EC5,EC6
S089	X			X			X		X						X			X	X			X		X			X	EC6
S090	X			X			X					X			X				X		X			X	X			EC3
S091	X	X		X			X		X				X			X				X			X		X			EC4,EC5
S092	X	X			X		X					X			X		X		X			X		X			X	EC2,EC6
S093	X			X			X					X			X				X		X			X	X			EC3
S094	X	X		X			X		X				X			X				X				X	X			EC4,EC5
S095			X	X				X				X			X		X				X			X	X			EC1,EC5
S096			X	X			X		X						X		X				X			X	X			EC1,EC5

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S097		X		X				X				X			X				X			X		X	X			EC1
S098	X			X				X		X					X				X		X			X	X			
S099	X			X			X		X						X				X		X			X	X			
S100		X	X	X				X		X					X			X			X			X			X	EC1,EC5,EC6
S101			X	X			X					X			X		X				X			X		X		EC1,EC5,EC6
S102	X			X			X				X				X				X		X			X	X			
S103		X	X	X				X				X			X		X		X	X				X	X			EC1
S104			X	X			X			X					X		X				X			X	X			EC1,EC5
S105		X		X			X					X		X					X	X				X	X			EC1,EC4
S106			X	X			X					X			X		X				X			X			X	EC1,EC5,EC6
S107			X	X			X			X					X		X				X			X		X		EC1,EC5,EC6
S108			X	X			X					X			X		X				X			X		X		EC1,EC5,EC6
S109	X			X				X				X			X	X			X		X			X	X			
S110			X	X			X					X			X		X					X		X	X			EC1,EC5
S111	X			X				X			X				X			X			X			X			X	EC5,EC6
S112	X		X	X			X					X			X		X		X		X			X			X	EC6
S113	X	X		X			X				X				X		X		X			X		X			X	EC6
S114	X				X	X	X		X					X	X					X				X	X			EC2,EC5
S115	X			X			X					X			X				X		X			X	X			EC3
S116	X			X				X				X			X	X			X		X			X	X			
S117			X	X			X					X			X		X					X		X	X			EC1,EC5
S118	X					X	X			X	X			X			X				X			X	X			EC2,EC4,EC5
S119	X			X			X					X			X				X		X			X	X			EC3
S120	X	X	X	X			X		X						X		X		X			X		X	X			
S121			X	X			X		X						X		X				X			X			X	EC1,EC5,EC6
S122	X			X			X					X			X				X		X			X			X	EC6

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S123	X			X			X					X			X				X		X			X				X	EC6
S124	X			X			X			X					X				X		X			X		X			
S125	X			X			X					X			X				X		X			X		X			EC3
S126	X			X			X					X			X				X		X			X		X			
S127	X			X			X					X			X				X		X			X		X			
S128	X			X			X		X						X				X		X			X				X	EC6
S129	X		X	X			X		X				X	X		X	X					X		X		X			EC4,EC5
S130	X			X			X		X						X				X		X			X		X			
S131	X	X		X			X		X						X		X		X			X		X		X			
S132	X		X	X			X					X			X		X		X		X			X		X			
S133	X		X	X			X					X			X		X		X		X			X		X			
S134			X	X			X		X				X	X	X	X	X	X	X			X		X				X	EC1,EC6
S135			X	X			X		X				X				X				X			X				X	EC1,EC4,EC5,EC6
S136		X		X			X					X			X				X			X		X		X			EC1
S137			X	X				X		X					X		X				X			X		X			EC1,EC5
S138	X			X			X		X						X			X	X			X		X		X			
S139	X			X			X		X					X	X				X		X			X		X			
S140	X			X			X					X			X				X		X			X		X			EC3
S141	X			X			X					X			X				X			X		X		X			EC3
S142			X	X			X					X			X		X				X			X				X	EC1,EC5,EC6
S143	X	X		X			X			X					X				X			X		X		X			
S144			X	X			X					X			X		X				X			X			X		EC1,EC5,EC6
S145	X			X			X			X					X				X		X			X				X	EC6
S146			X	X			X		X						X		X				X			X				X	EC1,EC5,EC6
S147	X	X	X	X			X		X						X	X	X		X			X		X		X			

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S148			X	X			X		X				X	X	X		X				X			X	X			EC1,EC5
S149	X			X			X		X						X				X		X			X	X			
S150	X				X	X	X		X					X	X				X	X				X	X			EC2
S151	X			X			X		X						X				X		X			X	X			
S152	X			X			X					X			X				X		X			X			X	EC6
S153	X	X	X	X			X		X						X		X		X			X		X	X	X		
S154	X		X	X			X					X			X		X		X	X				X	X	X		EC3
S155	X			X			X		X						X				X		X			X	X			
S156	X			X			X			X					X				X		X			X			X	EC6
S157	X					X	X		X						X				X		X			X			X	EC2,EC6
S158		X	X	X			X					X			X		X		X			X		X			X	EC1,EC6
S159			X	X			X			X					X		X				X			X	X			EC1,EC5
S160	X	X	X	X			X		X						X		X		X		X			X	X			
S161		X		X			X		X						X				X			X		X			X	EC1,EC6
S162	X			X				X	X						X				X	X				X	X			
S163			X	X			X					X			X		X				X			X		X		EC1,EC5,EC6
S164			X	X			X					X			X		X				X			X			X	EC1,EC5,EC6
S165	X			X			X					X			X				X		X			X	X			
S166	X			X			X					X			X				X		X			X	X			EC3
S167	X		X	X			X				X				X		X		X		X			X	X	X		
S168	X			X			X		X					X		X			X			X		X	X			EC4
S169	X			X			X		X				X			X				X				X	X			EC4,EC5
S170			X	X			X					X			X		X		X		X			X			X	EC1,EC6
S171	X			X			X					X			X				X		X			X			X	EC6
S172	X			X			X		X						X				X		X			X	X			
S173	X			X				X				X			X		X				X			X			X	EC5,EC6

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S174	X	X	X	X			X		X				X	X	X	X	X		X		X			X			X	EC6
S175			X	X			X					X			X		X			X			X				X	EC1,EC5,EC6
S176	X		X	X		X	X			X					X				X		X			X			X	EC6
S177	X			X				X			X				X			X				X		X	X			EC5
S178	X		X	X			X					X			X		X				X		X		X			EC5
S179		X	X	X			X		X						X		X		X			X		X		X		EC1,EC6
S180	X			X				X		X					X				X		X			X	X			
S181	X		X	X			X		X						X		X		X			X		X			X	EC6
S182	X			X				X		X					X				X	X			X				X	EC6
S183			X	X			X					X			X		X				X			X		X		EC1,EC5,EC6
S184			X	X				X			X				X		X			X				X	X			EC1,EC5
S185	X			X		X	X			X					X	X			X		X			X	X			
S186			X	X			X					X			X		X		X		X		X				X	EC1,EC6
S187	X		X	X			X		X						X		X		X		X			X	X	X		
S188			X	X				X			X				X		X				X			X	X			EC1,EC5
S189	X		X	X			X				X				X		X					X		X			X	EC5,EC6
S190	X			X			X					X			X				X			X		X	X			
S191	X				X	X	X				X				X				X	X				X	X			EC2
S192	X	X	X	X			X					X			X		X		X	X				X			X	EC6
S193			X	X				X	X						X				X			X		X			X	EC1,EC6
S194	X	X		X			X		X						X				X	X				X			X	EC6
S195	X			X			X					X			X				X		X			X	X			EC3
S196	X		X	X			X	X	X						X				X		X		X				X	EC6
S197	X		X	X			X					X			X	X	X		X		X			X		X		EC6
S198	X	X	X	X			X			X					X		X		X			X		X	X	X		
S199	X			X			X					X			X				X			X		X	X			EC3

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S200	X			X			X		X						X				X		X			X			X	EC6
S201	X	X		X			X					X			X				X			X		X				EC3
S202	X		X	X				X				X			X				X		X			X			X	EC6
S203	X	X		X			X					X			X				X			X		X				EC3
S204		X		X			X					X			X				X	X				X			X	EC1,EC6
S205			X	X			X			X					X		X				X			X		X			EC1,EC5,EC6
S206	X			X			X					X			X				X		X			X				EC3
S207			X	X			X					X			X		X					X		X			X	EC1,EC5,EC6
S208	X			X			X			X					X				X		X			X				
S209	X		X	X			X		X						X		X		X		X			X			X	
S210			X	X			X		X						X		X				X			X			X	EC1,EC5,EC6
S211		X		X			X			X					X				X	X				X			X	EC1,EC6
S212	X			X			X			X					X				X		X			X				
S213	X	X		X			X			X			X						X			X		X				EC4
S214	X			X				X				X			X				X	X				X				EC3
S215	X			X			X					X			X				X		X			X				
S216	X			X			X		X						X				X		X			X				
S217			X	X			X					X			X		X				X			X		X	X	EC1,EC5,EC6
S218			X	X			X		X						X		X				X			X			X	EC1,EC5,EC6
S219	X			X				X	X						X				X			X		X			X	EC6
S220			X	X			X		X						X		X				X			X			X	EC1,EC5,EC6
S221		X	X	X			X				X				X		X		X			X		X				EC1
S222			X	X			X		X						X		X				X			X			X	EC1,EC5,EC6
S223	X			X			X					X			X				X		X			X			X	EC6
S224		X		X			X					X			X				X	X				X			X	EC1,EC6
S225	X			X			X					X			X				X		X			X	X			EC3

ID	SQ1			SQ2			SQ3		SQ4				SQ5			SQ6				SQ7			SQ8		SQ9			Applied Exclusion Criteria
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(a)	(b)	(a)	(b)	(c)	
S226			X	X			X		X						X		X				X			X		X	X	EC1,EC5,EC6
S227	X			X			X				X				X				X		X			X	X			EC3

APPENDIX F – ANALYSIS OF SELECTED TECHNIQUES FROM THE SYSTEMATIC MAPPING OF RIVERO AND CONTE (2017)

This appendix presents the result of the analysis of the identified and selected techniques of Rivero and Conte (2017) mapping. The table contains the original code of the publications where the technique was used, the name of the technique, if it was deleted, and the reason for the exclusion.

ID	Technique	Excluded?	Reason for Excluding
S068	10 UX Dimensions	Yes	The proposed technology needs an interviewer to conduct the study.
S139	Again Again Method	Yes	Questionnaire specific to evaluate the UX of childrens.
S033 S120 S131 S153	AttrakDiff	Yes	According to Laugwitz et al. (2008), the AttrakDiff questionnaire "lays a greater emphasis on the hedonic aspects of product quality than on the pragmatic aspects", which may not be appropriate to have a comprehensive evaluation of the platform being evaluated.
S102	Reduced AttrakDiff + NASA-TLX	Yes	The authors reduced the AttrakDiff questionnaire to two pairs of adjectives for each dimension, specifically to the conducted study (Os autores realizaram a redução do questionário AttrakDiff para 02 pares de palavras para cada dimensão, especificamente para o estudo realizado (evaluation of mapping software for mobile devices). They did not verify whether this reduced version captures the UX equivalent to the full version of AttrakDiff.
S133	Cognitive Absorption Scale (CAS)	Yes	Some questions of immersion dimension depend on the use of the platform for long periods of time, a situation more frequent in Distance Learning courses, which makes it difficult to apply.
S212	Custom UX Questionnaires	Yes	Questions related specifically to 3D Virtual Reality.
S080	Decomposed Expectation-Confirmation Model and Questionnaire	Yes	Questionnaire aimed at evaluating the continuous acceptance of a technology, addressing aspects related to usability and satisfaction. Hedonic aspects such as emotions and stimuli are not captured.

ID	Technique	Excluded?	Reason for Excluding
S131	EmoCards	Yes	It is necessary the presence of an interviewer to conduct the choice of cards with emotions, which makes their use unfeasible.
S185	Evaluation of User Experience and Interface Ergonomic Criteria	Yes	Questionnaire developed to evaluate the accessibility for the elderly.
S155	Experience Questionnaire	Yes	Three experiments were carried out with different types of questionnaires, without a standard questionnaire, which made its use unfeasible.
S013 S120 S153 S209	Eye Tracking	Yes	Monitoring techniques were not considered, since they are unfeasible due to lack of equipment.
S190	Framework of UX Questions	Yes	Questionnaire related only to aspects of usability, such as navigation, precision and layout. There is no addressing of hedonic aspects.
S075	Immersive Tendencies Questionnaire (ITQ)	Yes	Questionnaire specific to evaluate Virtual Environments.
S059	Integrated experience acceptance model	No	
S160	Integrative Multi-Dimensional Assessments of Usability Features	Yes	Final questionnaire not available.
S162	Intelligent User Experience Questionnaire	Yes	It uses a proprietary application (which is not available) to conduct personalized interview questions during interaction with the product.
S078	Interface Aesthetics Requirements Evaluation	Yes	Final questionnaire not available.
S127	iTV-UX questionnaire	Yes	It does not consider pragmatic aspects, important to evaluate aspects related to the tasks, frequently performed in the AVA.
S138	MemoLine	Yes	Questionnaire specific to the evaluation of long-term games for children.
S187	Panas	Yes	Questionnaire measuring the positives and negatives, without separation by factors, which hinders a deeper analysis and comparison with other techniques.
S209	Panas-X	Yes	Final questionnaire not available.
S062	Perceived Visual Aesthetics	Yes	Assessment related specifically to aesthetic aspects, limiting its scope.
S167	Physiological Measures	Yes	We did not consider monitoring techniques.

ID	Technique	Excluded?	Reason for Excluding
S071	Post-study Usability Questionnaire	Yes	Questionnaire evaluating attributes focused on the pragmatic aspects of UX.
S075	Presence Questionnaire	Yes	Questionnaire related to virtual environments.
S165	Adapted Quick-UX	Yes	The adapted technique was not validated by factors after its modification. The original technique also does not present, on the referenced website, studies related to the reliability of the technique.
S033 S062 S075	SAM	Yes	The figures used by SAM may not be easy to understand, and an explanation is necessary about the 03 dimensions evaluated. It is difficult, for example, to understand that the figure with a "burst" in the chest signifies a high state of arousal.
S126	Satisfaction Questionnaire	Yes	Questionnaire specific to the context of the evaluated mobile application.
S054	Schwartz's value model	Yes	Composed of observations and interviews, which we did not consider. The questionnaires were also specific to assess the opinion of the parents who left or did not let the children participate in the experiment, not being related to the evaluation of UX.
S079	Semantic Differentials Questionnaire	Yes	The developed technique presents 20 pairs of adjectives referring to 05 factors, being 02 factors related to quality and 03 related to the use. The pairs of adjectives used are not distributed among these factors. One of them has 08 of the 20 adjectives, while another factor has only 1 pair of adjectives, which may interfere with the results of the evaluation.
S085 S132	Semantic Web Exploration Tools Quality in Use Model (SWET-QUM)	Yes	The post-test questionnaire employed is specific to pragmatic aspects, specifically targeted to the tasks. There is no verification of the hedonic aspects. In addition, the techniqueology used employs the use of visual tracking to perform the tests, which makes it unfeasible to apply it.
S143	Semi-structured Interview	Yes	We did not consider verbally related techniques.
S075	Simulator Sickness Questionnaire	Yes	Questionnaire related to virtual environments.
S013 S139	Smileyometer	Yes	Questionnaire specific to the evaluation of children.

ID	Technique	Excluded?	Reason for Excluding
S003 S033 S062 S126 S133 S187	SUS	Yes	Questionnaire focused on aspects related to usability, not evaluating hedonic aspects.
S041 S116 S131	SUXES	Yes	The technique performs an evaluation of the user's expectation before the use of the system and an evaluation after the use, comparing the notes given by the users in the two stages. It may not be easy for users to answer questions that involve the evaluation of expectation, such as accuracy, speed, fatigue, and learning, only by displaying the general characteristics of the system.
S109	Modified SUXES + Experience Pyramid	Yes	The technique proposes to use the elements of the Experience Pyramid to evaluate the user's expectation before using the SUXES technique. Experience Pyramid's assertions are confusing, for example, the element "authenticity", whose statements are "The application is artificial and unreliable" and "The application is genuine and trustworthy." It becomes difficult for the user to know whether the application is reliable or not only with an overview of the platform's features.
S066 S099	The Fun Toolkit	Yes	Method specifc to measure the degree of fun of children.
S018 S216	UMUX	Yes	Questionnaire specific for the evaluation of perceived usability based on the SUS questionnaire, not involving hedonic aspects related to UX.
S216	UMUX-Lite	Yes	Questionnaire limited to two questions, specific to pragmatic aspects: "the capabilities of the system met my needs" and "the system is easy to use". Hedonic aspects, important for evaluating the UX, are not considered.
S031 S130	User Engagement Scale (UES)	Yes	Questionnaire originally developed for e-commerce applications. Some aspects such as immersion and the feeling of disconnection are obtained during the use of the system for long periods, a situation most commonly found in courses and disciplines carried out exclusively in the Distance Learning modality, which hinders its application.

ID	Technique	Excluded?	Reason for Excluding
S083 S098 S151 S167 S172 S180 S187 S215	User Experience Questionnaire (UEQ)	No	
S198	User-Reported Aesthetic Value	Yes	It is not a technique but a review of the literature on aesthetic value in interaction design.
S149	UTAUT	Yes	Final questionnaire not available.
S037	UX Curve	Yes	The main objective of the technique is to evaluate UX in the long term, which makes it difficult to apply it in the experiment.
S147	UX Evaluation Framework	Yes	Final questionnaire not available.
S208	UX Evaluation Questionnaire	Yes	The hedonic aspects evaluated by the technique are limited to the aesthetic aspects, which reduces its scope.
S143	UX Evaluation Questionnaire	Yes	The technique was developed to evaluate interactive galleries, having dimensions and attributes that do not have much relation with LMSs, such as the dimension "Immersion and presence", that presents attributes not very clear in their evaluation, such as time, consciousness and natural.
S069	UX Web Survey	Yes	Final questionnaire not available.
S124	VisAWI Questionnaire	Yes	Questionnaire related only to aspects of aesthetics, which limits its scope.

APPENDIX G – ANALYSIS OF THE TECHNIQUES SELECTED FOR THE FIRST EMPIRICAL STUDY

This appendix presents the result of the analysis of the identified and selected techniques from the systematic mapping carried out in Chapter 3 for the first empirical study. The table contains the original code of the publications where the technique was used, the name of the technique, if it was deleted, and the reason for the exclusion.

ID	Technique	Description	Reason for exclusion
S01	Questionnaire proposed by Koohang (2004)	It allows evaluating the usability perceived by users of the e-learning platform by menas of 19 characteristics, such as navigation, control, relevance of information, etc. It also makes it possible to verify the degree of importance perceived by users for each of these characteristics.	The technique does not present separation by factors, which makes it difficult to analyze further the attributes of usability and to compare the data with other techniques. The second part, where students describe the perceived importance of each factor can give the false impression that the system has no usability problems.
S16	Questionnaire proposed by Zaharias and Poylymenakou (2009)	The technique consists of a questionnaire to evaluate the usability of e-learning applications. It is possible to evaluate the platform considering 7 dimensions of usability: Content, Learning and Support, Visual Design, Navigation, Accessibility, Interactivity and Self-evaluation and Learnability. In addition, the technique also assesses the motivation to learn.	The questionnaires are complex, many related to pedagogical aspects of the course, which can be difficult for learners to evaluate.
S24	UseLearn technique proposed by Oztekin et al. (2010)	The technique is composed of 36 questions on a 5-point Likert scale, assessing: error prevention, visibility, flexibility, course management, 'interactivity, feedback and help', accessibility, 'consistency and functionality',	The technique requires learners to take lessons and carry out some tasks in the LMS prior to the evaluation.

ID	Technique	Description	Reason for exclusion
		evaluation strategy, memorization , completeness, aesthetics and reduced redundancy	
S39	Adapted TAM, proposed by Theng and Sin (2012)	The technique is composed by the following constructs and its variables: - ISO standards for usability (ISO 9241-11): Perceived Satisfaction (PS), Perceived Usefulness (PU) and Perceived Ease Of Use (PEOU); - Engagement in e-learning: Learning By Interaction (LBI) and Making Sense of Learning (MSL); - Self-efficacy: E-Learning Efficacy (ELE) and New Media Efficacy (NME); - E-learning Design and Support: Navigation Structure (NS), User Interface (US) and Personalization and Freedom of Control (PFC).	
S47	UX evaluation questionnaire proposed by Ssekakubo et al. (2014)	It consists of questionnaires used to evaluate the interface of a mobile application called mVULA. The technique verifies usability and utility through an online survey containing closed questions (5-point Likert scale) and open questions.	The scope of the questionnaire is limited to the evaluation of the functionalities provided by a mobile application designed to provide limited functionality options.
S48	Usability evaluation questionnaire proposed by Yusoff and Mat Zin (2011)	The questionnaire consists of 11 questions using a scale from 1 to 5 (1 - totally disagree, 2 - disagree, 3 - almost disagree, 4 - agree and 5 - strongly agree): ease of use, navigation, of the menus, visual, memorization, learning support, motivation, communication and collaborative learning.	Questionnaire without division by factors, which hinders a more in-depth analysis and comparison of data with other techniques. The questions used are also very general, which only indicate the perception of the user, for example: "The system is easy to use" and "the tasks are easy to execute".

ID	Technique	Description	Reason for exclusion
S51	Web-based Educational Environmental Attitude Scale applied by Torun and Tekedere (2015)	The technique evaluates participants' perceptions regarding the effectiveness of web instruction and its resistance to this type of application.	The technique does not allow identifying problems related to the usability of the platform, structure or its functionalities. In addition, the original language is in Turkish, which makes it difficult to carry out a more adequate translation.
S55	Questionnaire proposed by Alkhattabi (2015)	The technique uses checklists to evaluate the usability, practicality, and efficiency of the overall design and of the Tadarus Learning Management System. Each item in the checklist consists of a Likert scale that ranges from 1 (totally disagree) to 5 (totally agree). A comment session was also used for most items in order to obtain qualitative data about users' perceptions. The checklist consists of 3 sub-scales related to: general design, pedagogical design of materials and user satisfaction.	The questionnaire provided in the article presents the same questions for two different factors, which may indicate an error in the technique or in the publication, making its application unviable.
S58	Questionnaire proposed by Junus et al. (2015), adapted from Zaharias and Poylymenakou (2009)	It consists of questions that use a 5-point Likert scale, based on the eight categories proposed by Zaharias and Poylymenakou (2009): content, learning and support, visual design, navigation, accessibility, interactivity, self-assessment and learning ability, and motivation to learn.	Questionnaire adapted from the technique of Zaharias and Poylymenakou (2009) with a simplified language. The technique has, however, one less question than the original questionnaire, with no explicit reasons for its exclusion. The authors also did not present an analysis of factors and validity of the construct after the adaptations, in order to validate the technique.

APPENDIX H – TUXEL 1.0

Name:_____

Start time: _____:_____

Legend
1- Totally disagree / 2- Partially disagree / 3- Neither agree nor disagree / 4- Partially agree / 5- Totally agree

ITEM	NAVIGATION					
N1	The options in the LMS's navigation menus are visible and easy to remember.	1	2	3	4	5
N2	I can easily know where I am and where I can go in LMS.	1	2	3	4	5
N3	I can easily go anywhere on LMS.	1	2	3	4	5
N4	The information is organized hierarchically to facilitate the navigation.	1	2	3	4	5
N5	I can easily identify the links I've visited.	1	2	3	4	5
N6	The LMS provides a search engine for finding information.	1	2	3	4	5
ITEM	**LEARNABILITY**					
L1	The LMS is intuitive enough that I can carry out the activities without difficulties.	1	2	3	4	5
L2	The instructions provided by the LMS are clear and objective.	1	2	3	4	5
L3	I would be able to complete an activity/task even if no one was around to help me.	1	2	3	4	5
L4	I can easily know the functionality of each icon, button or link.	1	2	3	4	5
L5	It is clear what I should do in the LMS in case of doubt or difficulty.	1	2	3	4	5
L6	I would still know how to use the LMS even after a long time without using it.	1	2	3	4	5
L7	I can perform the activities in the LMS with the minimum of steps required.	1	2	3	4	5
L8	Overall, I find the LMS easy to learn to use.	1	2	3	4	5
ITEM	**CONSISTENCY**					
C1	The interface design (fonts, colors, themes, and buttons) is consistent throughout the LMS.	1	2	3	4	5
C2	The terminology (terms, words, and actions) is used consistently throughout the LMS.	1	2	3	4	5
ITEM	**VISUAL DESIGN**					
VD1	The interface of the LMS is aesthetically appealing.	1	2	3	4	5
VD2	The interface of the LMS does not display unnecessary or rarely used information.	1	2	3	4	5
VD3	The most important information is in visible places that catch my eye.	1	2	3	4	5
VD4	Related information is grouped in the LMS.	1	2	3	4	5
VD5	The colors used in the LMS allow a high contrast of the text, making reading easier.	1	2	3	4	5
VD6	The LMS reduces my cognitive load, that is, my mental effort, through the familiarization of items and sequences of actions.	1	2	3	4	5
VD7	The icons, symbols, and labels used in the LMS are intuitive.	1	2	3	4	5
ITEM	**PERSONALIZATION**					
P1	The LMS allows me to customize the interface by adding/removing elements or changing their position.	1	2	3	4	5
P2	The LMS allows me to use shortcut keys to improve my productivity.	1	2	3	4	5
P3	The LMS allows me to choose the language of my preference.	1	2	3	4	5

ITEM	LMS FEEDBACK					
LF1	The LMS always provides feedback when I perform an action.	1	2	3	4	5
LF2	The LMS informs me when mandatory fields are not filled and what are these fields.	1	2	3	4	5
LF3	The LMS informs me how to correctly fill in a particular field, e.g., stating that the date must be filled in DD/MM/YYYY format.	1	2	3	4	5
LF4	The LMS always asks me for a confirmation before performing any important actions, such as completing evaluative activities or deleting any files.	1	2	3	4	5
LF5	The LMS informs me about the progress of some action, e.g., through a bar indicating the percentage of sending a file.	1	2	3	4	5
ITEM	**HELP AND DOCUMENTATION**					
HD1	The LMS provides some kind of online or manual help.	1	2	3	4	5
HD2	It's easy to find what I need in the manual/online help.	1	2	3	4	5
HD3	The manual/online help provides simple and straightforward instructions to solve any problem or difficulty.	1	2	3	4	5
ITEM	**LEARNER'S CONTROL**					
LC1	I can undo/redo an action with ease.	1	2	3	4	5
LC2	The LMS allows me to cancel an action at any time.	1	2	3	4	5
LC3	The LMS allows me to leave at any time.	1	2	3	4	5
ITEM	**COLLABORATIVE LEARNING**					
CL1	The LMS allows me to carry out group activities (discussion or other collaborative activities).	1	2	3	4	5
CL2	The LMS allows me to communicate easily with other colleagues and/or teachers.	1	2	3	4	5
CL3	The LMS allows me to see what other colleagues have done in the system, such as the most readable content, most popular activities, etc.	1	2	3	4	5
CL4	The LMS allows me to share files, photos, videos and educational materials.	1	2	3	4	5
CL5	The LMS allows me to post questions and answer questions from other colleagues.	1	2	3	4	5
ITEM	**GOAL ORIENTATION**					
GO1	The LMS provides my score by performing the activities to track my performance.	1	2	3	4	5
GO2	The LMS allows me to track my progress in the discipline, such as showing the number of completed activities and topics, a comparative chart of performance against the class, etc.	1	2	3	4	5
ITEM	**INSTRUCTIONAL ASSESSMENT**					
IA1	The LMS makes it possible for me to perform evaluative activities to self-assess my learning.	1	2	3	4	5
ITEM	**ADDED VALUE**					
AV1	For me, it is more useful to learn through the LMS than in the traditional way in the classroom (e.g. through books or exercise books).	1	2	3	4	5
AV2	I prefer interacting with my colleagues through the LMS than face to face in the classroom.	1	2	3	4	5
AV3	I feel that I learn faster through the LMS than I usually do because it provides support when needed, e.g., through tips.	1	2	3	4	5
AV4	Through the LMS, I can learn a new content faster or easier, or recapitulate previous content.	1	2	3	4	5
AV5	The LMS allows me to carry out the activities more efficiently.	1	2	3	4	5
AV6	Overall, I find the LMS useful for learning.	1	2	3	4	5
ITEM	**MOTIVATION TO LEARN**					
ML1	I try to reach the highest possible score when I do my activities on the LMS.	1	2	3	4	5
ML2	The LMS is nice and interesting.	1	2	3	4	5

ML3	The LMS increases my motivation through the use of games or simulations	1	2	3	4	5
ML4	The LMS rewards me for completed tasks or achieved achievements, for example through symbolic medals, emblems or titles in the system.	1	2	3	4	5
ML5	The LMS meets my needs as a student.	1	2	3	4	5
ML6	Overall, I think the LMS motivates me to learn.	1	2	3	4	5
ITEM	**LMS FLEXIBILITY**					
LFL1	The LMS allows me to choose which parts of the course to access, the order and the pace of the study.	1	2	3	4	5
LFL2	I can consult the previous materials when there are questions.	1	2	3	4	5
ITEM	**INSTRUCTIONAL FEEDBACK**					
IF1	The LMS gives me, when possible, immediate feedback if the answer is right or wrong when responding to an activity.	1	2	3	4	5
IF2	When I fail the response in an activity, the LMS gives me specific feedback for this activity.	1	2	3	4	5
IF3	The LMS provides me with opportunities to seek additional feedback with my instructors/teachers when I have questions about any content, activity or issue.	1	2	3	4	5

End time: _____:_____

APPENDIX I – ADAPTED TAM (THENG AND SIN, 2012)

Name:_____

Start time: _____:_____

Legend
1- Totally disagree / 2- Partially disagree / 3- Neither agree nor disagree / 4- Partially agree / 5- Totally agree

ITEM	PERECIVED SATISFACTION					
PS1	I am willing to use the LMS to access course documents.	1	2	3	4	5
PS2	I am willing to use the LMS to communicate course topics with tutors and classmates for discussion of course topics.	1	2	3	4	5
PS3	I am willing to use the LMS to collaborate on group projects with classmates.	1	2	3	4	5
PS4	I am willing to use the LMS to submit course work and assignments.	1	2	3	4	5
PS5	I am willing to use the LMS to review and critique my friends' work.	1	2	3	4	5
PS6	I intend to use the LMS for self-directed learning purposes (i.e. not a course requirement).	1	2	3	4	5
ITEM	**PERECIVED UTILITY**					
PU1	Using the LMS would improve my academic performance.	1	2	3	4	5
PU2	Using the LMS would increase the productivity of my learning.	1	2	3	4	5
PU3	Using the LMS would enhance my effectiveness as a learner.	1	2	3	4	5
PU4	Using the LMS would be able to accomplish tasks more efficiently.	1	2	3	4	5
PU5	Overall, I find that the LMS is useful as a learning tool.	1	2	3	4	5
ITEM	**PERECIVED EASE OF USE**					
PEOU 1	I understand how to use the LMS.	1	2	3	4	5
PEOU 2	The tasks are performed with the minimum steps required.	1	2	3	4	5
PEOU 3	Using the LMS would not require a lot of my mental effort.	1	2	3	4	5
PEOU 4	Overall, I find that the LMS would be easy to use.	1	2	3	4	5
ITEM	**LEARNING BY INTERACTION**					
LBI1	I feel the LMS can promote inter-disciplinary learning.	1	2	3	4	5
LBI2	I feel the LMS can promote independent learning.	1	2	3	4	5
LBI3	I feel the LMS can promote knowledge sharing	1	2	3	4	5
LBI4	I am motivated to discuss course topics with classmates and tutors.	1	2	3	4	5
LBI5	I prefer online written interaction to face-to-face classroom interaction.	1	2	3	4	5
ITEM	**MAKING SENSE OF LEARNING**					
MSL1	I usually make good sense of classmates' online messages from discussion thread about course topics.	1	2	3	4	5
MSL2	I usually make good sense of such online messages from tutors.	1	2	3	4	5
ITEM	**E-LEARNING EFFICACY**					

EE1	I am fully aware of the functional capabilities of the LMS.	1	2	3	4	5
EE2	I know how to use all the functions in the LMS.	1	2	3	4	5
EE3	I could complete a task using the LMS even if no one was around to help me.	1	2	3	4	5
EE4	I could complete a task using the LMS even if I only have the instructional manual/online help facility for reference.	1	2	3	4	5
EE5	I could complete a task using the LMS if I could call someone for help.	1	2	3	4	5
EE6	I could complete a task using the LMS if someone helped me get started.	1	2	3	4	5
EE7	I would rate myself an expert working on personal computers.	1	2	3	4	5
EE8	I would rate myself an expert accessing information on the Internet / World Wide Web.	1	2	3	4	5
EE9	I would rate myself an expert working with learning portals.	1	2	3	4	5
EE10	I had used other E-learning systems previously so that I can perform the similar tasks easily.	1	2	3	4	5
EE11	I used the instruction manual or online help facility to familiarize with the system.	1	2	3	4	5
ITEM	**NEW MEDIA EFFICACY**					
NME1	I would rate myself an expert working with discussion threads.	1	2	3	4	5
NME2	I would rate myself an expert working with wikis.	1	2	3	4	5
NME3	I would rate myself an expert working with emails.	1	2	3	4	5
NME4	I would rate myself an expert working with blogs.	1	2	3	4	5
ITEM	**NAVIGATION STRUCTURE**					
NS1	The orientation for ‘Getting started with the LMS’ should be held for beginners.	1	2	3	4	5
NS2	Materials in the LMS are easily navigable and understandable.	1	2	3	4	5
NS3	You are always informed where you are in the LMS.	1	2	3	4	5
NS4	You can easily distinguish your recently visited links.	1	2	3	4	5
ITEM	**USER INTERFAEC**					
UI1	The UI color scheme provides ample contrast to aid in text readability.	1	2	3	4	5
UI2	The LMS is presented with distracting animated graphics or navigational elements.	1	2	3	4	5
UI3	All the menus appeared on the LMS are your most frequently used items.	1	2	3	4	5
UI4	Welcome message should be displayed with the student name on the main page.	1	2	3	4	5
UI5	The LMS prompts to response satisfactory feedback and confirmation after performing important tasks (e.g. submitting assignments, posting threads, etc.)	1	2	3	4	5
UI6	The content is presented clearly, in manageable “chunks” of information, for the learner.	1	2	3	4	5
ITEM	**PERSONALIZATION AND FREEDOM OF CONTROL**					
PFC1	Most frequently used contents can be modified to appear on my personal page in the LMS.	1	2	3	4	5
PFC2	I use “personal alerts” & “communication tools”.	1	2	3	4	5
PFC3	I customize the the LMS options, such as the "Student Panel".	1	2	3	4	5
PFC4	I use the "Recent Files" option when sending a file through the LMS.	1	2	3	4	5

End time: _____:_____

APPENDIX J – TUXEL 2.0

LOGIN	
ID	**Item**
LF2	The LMS does not inform the required fields to be filled, for example with a (*).
LF4	The LMS does not tell how to properly fill a given field. *(Make sure you are informed if the username/login field is an ID, e-mail, or username.)*
LF3	The LMS does not tell which field was not filled in. *(Try to log in without entering any data.)*
LF1	The LMS does not provide feedback when performing an action. *(Verify whether the LMS provides an error or success message when trying to login.)*
GENERAL INTERFACE	
ID	**Item**
VD5	Low color contrast that makes the text difficult to read.
N1	Navigation menu options are barely visible and/or difficult to remember.
C1	Interface design (fonts, colors, themes, and buttons) is not consistent. *(Verify if the interface design changes from one page to another.)*
C2	Terminology used (terms, words and actions) is not consistent. *(Check if different terms are used to refer to the same thing, for example: username -> login; send -> submit.)*
L4	Icon, button, label or link whose functionality or meaning is not so clear
VD2	Unnecessary content and/or information.
VD3	Important information in a place that is barely visible or does not attract attention.
N4	Content/information not organized/grouped logically. *(**Example:** Disciplines not arranged alphabetically or content not displayed in chronological order.)*
LC3	The page does not have an option to exit the system (log off) at any time.
N6	The page does not provide a search engine for finding information faster.
N5	The page does not differentiate the visited links from the unvisited ones. *(Access any link on the page and return to the previous page. Make sure that the link you visited appears, for example, in a different color from the other non-visited links.)*
N2	The page does not tell where in the LMS I am and where I can go. *(Check if the page shows the path to the current page. Example: Home → Courses → Computer Science)*
ASSIGNMENTS	
ID	**Item**
L2	Lack of instructions or the instructions are unclear and little objective. *(The LMS should tell, for example, how to carry out the activity [whether you have to type a text or attaching a file, for example].)*
LF4	It is not informed how to fill-in the field correctly. *(**Example:** If it is a file upload field, make sure the LMS shows which file types are allowed - PDF, DOC, etc.)*
LC2	I cannot cancel an action at any time. *(**Example:** If there is a field to load files, load it and see if it is possible to cancel sending the file in progress.)*
LF6	The LMS does not show me the progress of an action. *(Make sure the LMS shows, for example, a progress bar with the loading percentage when loading a file or playing a video.)*
LC1	I cannot undo/redo an action easily.

LF5	The LMS does not request a confirmation before taking any important action. *(The LMS should request a confirmation, for instance, when sending a task).*
LF1	The LMS does not provide feedback when performing an action. *(Verify whether the LMS provides an error or success message when sending a task.)*
ASSESSMENTS	
ID	**Item**
OO1	The LMS does not provide my score when performing the activities to track my performance. *(Make sure the LMS informs, after submitting the quiz, for instance, the score achieved.)*
IF1	The LMS does not provide immediate feedback, whenever it is possible, whether the answer is right or wrong when answering to an activity. *(Make sure the LMS shows you the questions you got it right or missed.)*
IF2	Feedback provided by the LMS during the activities is not relevant/appropriate to the problem or question being answered. *(When you missIf the LMS provides feedback, please check that it has additional information about the question, such as links to additional reading materials.)*

USABILITY INSPECTION – PROBLEM REPORTING TABLE

Please provide, in this table, details about the problem you identified by using the items of TUXEL 2.0. Whenever you identify a problem, please fill the table with these information:

- Item ID: refers to the ID of the TUXEL 2.0 item (e.g. LF2). You can assign more than one ID to the same problem, if you find it adequate;
- Location or Task: describe the location (e.g.: main page of the course) or the task (e.g. file upload) where the problem occurs;
- Problem description: provide details about the problem you identified (e.g. The red and blue colors used in menu impair text readability).
- Does the problem repeat? Inform whether the problem does not repeat, repeats system-wide or repeats in some tasks. If it repeats in only some tasks, specify in the field "In what tasks does it repeat?".

Item ID	Location or Task	Problem description	Does the problem repeat?	In what tasks does it repeat?
			( ) No. ( ) Yes, system-wide. ( ) Yes, in some tasks.	
			( ) No. ( ) Yes, system-wide. ( ) Yes, in some tasks.	
			( ) No. ( ) Yes, system-wide. ( ) Yes, in some tasks.	
			( ) No. ( ) Yes, system-wide. ( ) Yes, in some tasks.	
			( ) No. ( ) Yes, system-wide. ( ) Yes, in some tasks.	

GENERAL LMS EVALUATION CHECKLIST

This checklist evaluates general aspects of the LMS. Please read each item and put a ‘ ’ on the item(s) that relates to a problem you identified in the LMS. Feel free to make additional comments in the “Comments” field in order to detail the problem you faced/identified.

HELP AND DOCUMENTATION		
	ID	**ITEM**
	HD1	The LMS does not provide any kind of manual or online help.
If you **have not** marked the item above, evaluate the items below:		
	HD2	I cannot find what I need in the manual/online help with ease.
	HD3	Instructions provided in the manual/online help are not simple and straightforward to solve any problem or difficulty.
Comments:_____		
LMS LEARNABILITY		
	ID	**ITEM**
	L1	The LMS is not intuitive enough so I can do the activities without difficulties.
	L3	I could not complete an activity/task if no one was around to help me.
	L6	I would not know how to use the LMS after a long time without using it.
	L7	I cannot perform tasks on the LMS quickly, i.e., with the minimum of steps required.
	N3	I cannot easily go to anywhere on the LMS.
Comments:_____		
LEARNING THROUGH THE LMS		
	ID	**ITEM**
	CL1	The LMS does not allow me to carry out group activities (discussion forums, collaborative wikis, etc.).
	CL2	The LMS does not allow me to easily communicate with other colleagues and/or teachers.
	CL3	The LMS does not allow me to see what other colleagues have done in the system, such as the most read content, most popular tasks, etc.
	CL4	The LMS does not allow me to share files, photos, videos and educational materials.

	CL5	The LMS does not allow me to post doubts and answer questions from other colleagues.
	OO2	The LMS does not allow me to track my progress in the course. *(Check if there is a page showing, for example, your grades, the number of tasks and topics completed in the course, etc.)*

Comments:_____

LMS FLEXIBILITY		
	ID	**ITEM**
	FL1	The LMS does not allow me to choose which parts of the course to access, the order and pace of the study.
	FL2	I cannot consult previous materials when I have doubts.
	FI3	The LMS does not provide opportunities for me to seek additional feedback with my instructors/teachers when I have questions about any content, activity or issue.
	P1	The LMS does not allow me to customize the interface, for example, by adding/removing elements or changing their position.
	P2	The LMS does not allow me to use shortcut keys to improve my productivity.
	P3	The LMS does not allow me to choose its used language.
	AI1	The LMS does not allow me to carry out evaluative activities to self-assess my learning.

Comments:_____

USER EXPERIENCE EVALUATION

You will now evaluate the User eXperience (UX) conveyed by the LMS. Please mark, on each line, the circle that is closest to the adjective that better describes your UX, according to the example below:

Complicated	○	⊗	○	○	○	○	○	Easy

This response would mean that you rated the LMS as being more complicated than easy. Please, decide spontaneously. Don't think too much about your decision.

Remember: there is no right or wrong answers. It's your opinion that matters!

Unlikable	○	○	○	○	○	○	○	Pleasing
Uncomfortable	○	○	○	○	○	○	○	Comfortable
Does not meet expectations	○	○	○	○	○	○	○	Meets expectations
Obstructive	○	○	○	○	○	○	○	Supportive
Inneficient	○	○	○	○	○	○	○	Efficient
Impractical	○	○	○	○	○	○	○	Practical
Conventional	○	○	○	○	○	○	○	Innovative
Uncreative	○	○	○	○	○	○	○	Creative
Complicated	○	○	○	○	○	○	○	Easy
Confusing	○	○	○	○	○	○	○	Clear
Not Interesting	○	○	○	○	○	○	○	Interesting
Demotivating	○	○	○	○	○	○	○	Motivating

Please, leave your **criticisms** about the LMS you used, considering the ratings above:

Provide, below, **suggestions** to improve the LMS:

ANNEX A – UEQ QUESTIONNAIRE (LAUGWITZ ET AL., 2008)

Please make your evaluation now.

For the assessment of the product, please fill out the following questionnaire. The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression.

Example:

attractive	○	⊗	○	○	○	○	○	unattractive

This response would mean that you rate the application as more attractive than unattractive.

Please decide spontaneously. Don't think too long about your decision to make sure that you convey your original impression.

Sometimes you may not be completely sure about your agreement with a particular attribute or you may find that the attribute does not apply completely to the particular product. Nevertheless, please tick a circle in every line.

It is your personal opinion that counts. Please remember: there is no wrong or right answer!

Please assess the product now by ticking one circle per line.

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

ANNEX B – INTEGRATED EXPERIENEC MODEL QUESTIONNAIRE (VAN SCHAIK AND LING, 2011)

	Pragmatic Quality – I judge the web pages to be
PQ1	Complicated – simple
PQ2	Impractical – practical
PQ3	Confusing – clear
	Hedonic Quality-identification – I judge the web pages to be
HQI1	Amateurish – professional
HQI2	Gaudy – classy
HQI3	Unpresentable – presentable
	Hedonic Quality-stimulation – I judge the web pages to be
HQS1	Standard – creative
HQS2	Conservative – innovative
HQS3	Commonplace – new
	Beauty – I judge the web pages overall to be
Beauty1	Ugly – Beautiful
	Goodness – I judge the web pages overall to be
Goodness1	Bad – Good
	Perceived Enjoyment
PE1	I find using the Web site to be enjoyable
PE2	The actual process of using the Web site is pleasant
PE3	I have fun using the Web site
	Perceived Ease of Use
PEOU1	My interaction with the Web site is clear and understandable
PEOU2	It is easy for me to become skilful at using the Web site
PEOU3	I find the Web site easy to use
PEOU4	Learning to operate the Web site is easy for me
	Perceived Usefulness
PUSF1	I would find the Web site useful in finding information
PUSF2	Using the Web site would enable me to accomplish tasks more quickly when finding information
PUSF3	Using the Web site would enhance my effectiveness in finding information
PUSF4	Using the Web site would make it easier to find information