Applying Design Science Research to develop a Technique to Evaluate the Usability and User eXperience of Learning Management Systems

Walter T. Nakamura, Elaine H. T. de Oliveira, Tayana Conte

1Institute of Computing (IComp) – Federal University of Amazonas (UFAM) 69.080-900 – Manaus – AM – Brazil {walter,elaine,tayana}@icomp.ufam.edu.br

Abstract. As Internet-based technology becomes more accessible, learning institutions have been adopting platforms such as Learning Management Systems (LMSs) to support and also complement face to face learning. Their widespread adoption raises the importance to evaluate them regarding quality in use attributes such as usability and User eXperience (UX). This paper presents the application of Design Science Research (DSR) to develop a technique and a tool to evaluate the usability and UX of LMSs called TUXEL. By applying DSR, we ensure a high level of rigor in the development process, combining theory and practice. The empirical studies indicated that TUXEL is adequate, in addition to providing better results compared to other techniques.

1. Introduction

The widespread adoption of Internet-based technology led educational institutions to invest in new technologies to support teaching and learning process [Zaharias and Pappas 2016]. In this context, platforms such as Learning Management Systems (LMSs) have been adopted by educational institutions to provide teaching materials, online courses and also complement face to face learning. As the number of universities using LMSs grows, research on the quality in use of these platforms regarding quality attributes such as usability and UX has attracted considerable interest [Harrati et al. 2016].

Usability is related to the extent that a system, product or service can be used by users to accomplish their goals with effectiveness, efficiency and satisfaction in a specified context of use [ISO 9241-11 2017]. User eXperience (UX), on the other hand, is defined as person’s perceptions and responses, e.g., feelings and emotions, that result from the use or anticipated use of a product, system or service [ISO 9241-210 2010].

Usability and UX are quality in use attributes that have a great impact on teaching and learning process through LMSs. The former because the difficulty to use the platform may affect learners’ cognitive load. For instance, if an LMS does not provide a good usability, the learner may spend more time trying to use the platform rather than learning the educational content [Lanzilotti et al. 2006]. In addition to providing good usability, it is also important to consider how the learner feels when interacting with the LMS. More than doing tasks, it is important that the learner also feels motivated and engaged when learning through an LMS, which are, perhaps, some of the most important elements in every form of learning experience [Zaharias and Pappas 2016].

Differently from other contexts, an LMS should also meet the pedagogical objectives of the instructor, supporting people to learn in various contexts [Lanzilotti et
al. 2006]. Therefore, the evaluation of an LMS should integrate the assessment of the educational quality aspects of these platforms [Lanzilotti et al. 2006]. Due to these specificities, many authors consolidated interface usability with pedagogical aspects [Mtebe and Kissaka 2015]. However, there is still no widely accepted set of principles for the development of these techniques [Hovde 2015] and most of them are in early stages [Mtebe and Kissaka 2015], which reinforce the need for further studies. With regards to UX, despite its importance, few studies evaluated 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. According to Zaharias and Pappas (2016), UX evaluation has been neglected in e-learning field, which may explain the lack of specific techniques for evaluating the UX of LMSs.

Considering these gaps, we proposed a Technique for User eXperience Evaluation in e-Learning (TUXEL), in addition to a tool that supports the evaluation process. In order to develop TUXEL, we applied Design Science Research (DSR) methodology. By applying DSR, the researcher combines theory and practice while ensures a high level of rigor in the development of the artifact (something created for some practical purpose), developing it driven by theoretical assumptions and applying it to the problem context [Hevner and Chatterjee 2010]. Thus, the research becomes useful for both academic and practitioner audiences. Researchers can use our work as basis for developing new evaluation techniques, as well as use our artifact to evaluate LMSs. By using TUXEL, the researcher can identify usability, pedagogical usability and UX problems and then propose solutions to improve the quality of the LMS evaluated.

The remainder of this paper is as follows: Section 2 presents related work. Section 3 presents the development process of TUXEL, with all its phases. Finally, in Section 4, we conclude this paper by presenting our contributions and future work.

2. Development of Usability and UX evaluation techniques – Related Work

Usability and UX plays an important role in LMSs. In addition to providing good usability and a positive UX, an LMS should also support people to learn in various contexts, according to the pedagogical objectives of the instructor [Lanzilotti et al. 2006]. Given that general evaluation techniques cannot address these aspects, many authors developed techniques specific to evaluate LMSs. Mtebe and Kissaka (2015), for instance, expanded the set of interface heuristics by including educational criteria based on several works on the topic. Oztekin et al. (2010) developed a checklist to evaluate the usability of e-learning systems based on three other techniques. Zaharias and Pappas (2016) developed a survey instrument considering UX principles and psychological theories of intrinsic motivation to evaluate the UX of LMSs. Although these works presented the development of the technique, they did not evaluate it by obtaining feedback from the evaluators. Moreover, they did not clearly report the knowledge gained during its development, which would be a valuable contribution to the improvement or development of new techniques.

Gordillo et al. (2014) proposed a methodology by applying different usability/UX evaluation techniques according to the software development stage and gathered the perception of developers on the usefulness of each of these techniques. However, due to the application of different techniques, it is time consuming, making this approach costly. Additionally, some of them needs usability/UX experts to be applied. Finally, the authors employed general evaluation techniques, while not proposing improvements.
Although many research have been carried out to develop such techniques, there was no evidence of research performed by using a established methodology to develop them. It was also not common to develop techniques grounded on theoretical conjectures to support their decisions on the methodology used, e.g., to define its approach (inspection, test or inquiry). In this context, we propose the application of DSR to develop a technique to evaluate the usability/UX of LMSs called TUXEL. By employing DSR, we ensure a high level of rigor in the development process, providing empirical evidence grounded on theory. It also highlights the relevance of the problems addressed, making it clear the knowledge about how the proposed solutions may help [Storey et al. 2017].

3. Applying DSR to Develop a Technique to Evaluate LMSs

Design Science Research has been widely adopted in many research disciplines, such as information systems and management science, as a way to frame research by highlighting both the problem addressed and the intervention proposed, while supporting the researcher to convey how s/he build on and contribute to an existing knowledge base [Storey et al. 2017]. DSR consists in the design and investigation of artifacts that interact with a problem context and improve something in that context to solve a problem iteratively [Wieringa 2014]. Besides the experiences gained for performing the research, DSR may provide additions to knowledge, e.g., extensions to theories and methods made during the research [Hevner and Chatterjee 2010].

Figure 1 illustrates the development process of TUXEL. The rounded rectangles represent the steps carried out in DSR. The rectangles with arrows represent the input for the steps. The balloons refer to the reflections made during the artifact design or evaluation. Finally, the rectangles with a dashed line present the results from each step. DSR starts with the (A) problem and context definition and/or opportunities of improvement in a given context [Hevner and Chatterjee 2010]. In order to develop our proposal, we needed to gather information on the problem context. To do so, we carried out a systematic mapping (B) (Section 3.1) in order to obtain knowledge on the research topic, identify the problem and delimit the research scope. The results revealed research opportunities (C), which served as the basis for our proposal.

![Figure 1. DSR methodology cycle, based on the work of Hevner and Chatterjee (2010).](image-url)

After defining the problem and context, the research enters into the Design Cycle. In Design Cycle, researchers develop and specify one or more artifacts (proposed...
artifact) to interact with that context (D). Past knowledge and theoretical assumptions (E) may drive its development, denoting the rigor of DSR [Hevner and Chatterjee 2010]. Our scientific basis (F) was built by gathering the information about usability and UX evaluation techniques found in our systematic mapping and the results of a preliminary study. In the preliminary study (Section 3.2), our goal was to obtain empirical evidence on the existing UX evaluation techniques. The results of the systematic mapping and the preliminary study served as input for the development of TUXEL (G) (Section 3.3).

Finally, the artifact is evaluated by applying it to the problem and context (H), verifying whether it produces the desired effects and whether a new iteration over the Design Cycle is needed (I), while corroborating or questioning the validity of the theoretical assumptions (J) [Wieringa 2014]. In this step, we carried out two empirical studies. In the first study (Section 3.4), we verified the feasibility of TUXEL by applying it to evaluate an LMS while comparing it to another technique. The results indicated a need for improvements, which led us to perform a second iteration and refine it (Section 3.5). Finally, we carried out a second study to evaluate the refined technique (Section 3.6). Details of TUXEL and each empirical study are presented in Nakamura (2018).

3.1. Systematic Mapping Study

In order to define the problem and obtain knowledge on the research topic, we carried out a systematic mapping [Nakamura et al. 2017a]. This systematic mapping aimed to identify usability and UX evaluation techniques that have been applied to evaluate LMSs.

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 to be explored, for example: there was no evidence of a consolidated technique to evaluate both usability and UX of LMSs, and few studies evaluated the UX of these platforms. Regarding the latter, among these studies, all of them employed generic UX evaluation techniques. Due to the lack of studies and techniques designed to evaluate the UX of LMS, we needed more empirical evidence on whether generic UX evaluation techniques are enough for learners to convey their UX when using LMSs. Thus, we conducted a preliminary study.

3.2. Preliminary Study

We carried out a preliminary study [Nakamura et al. 2017b] with 34 students from the Federal University of Amazonas (UFAM), enrolled in Human-Computer Interaction class. We divided them into two groups, balanced according to the results of a characterization questionnaire. Most of the participants had low experience with UX evaluations and only 5 among the 34 participants had already used Edmodo. We evaluated and compared two UX evaluation techniques to evaluate an LMS called Edmodo: User Experience Questionnaire (UEQ) [Laugwitz et al. 2008] and Integrated Experience Acceptance Model (IEAM) [Van Schaik and Ling 2011]. We selected these techniques by applying exclusion criteria considering aspects such as evaluation period (before, during or after interaction), data source and availability.

The results revealed limitations of these techniques. Many students stated that they could not report the difficulties and justify their answers, as there is no field for comments. Additionally, they considered some adjectives inadequate to convey the UX of LMSs. However, both techniques were considered fast, easy to use and useful. When asked which technique they would prefer to use in the future, most students chose IEAM. They
stated that it is more focused, letting them know what aspect of the LMS is being evaluated. It may be due to its second part, which is composed by a set of questions about the use of the LMS. On the other hand, other participants considered UEQ easier and simpler, while providing a broader range of adjectives to express their UX.

Regarding Edmodo, the additional open question revealed that many students faced difficulties when performing the tasks on the platform. One participant, for instance, did not find the materials in Edmodo and searched for the content on the Internet. From a total of nine questions, he answered three of them incorrectly and one partially correctly. Although we cannot infer that he could have answered them correctly if he had found the materials, it would have been easier for him to answer the questions. Similarly, many students did not understand how to perform the matching quiz. This can directly impact the learning process, since it increases the cognitive load as students spend more time trying to use the LMS rather than learning the educational content [Lanzilotti et al. 2006].

The results of this study served as a starting point to develop TUXEL. The findings also provided empirical evidence on these techniques, while contributing to improve Edmodo and to develop/improve techniques to evaluate the UX of LMSs.

3.3. Development of the First Version of TUXEL

As stated before, the development of the artifact in DSR may be driven by theoretical conjectures, providing more rigor to the development process. Regarding UX evaluation techniques, due to the lack of studies in the context of LMSs, we carried out the preliminary study. On the other hand, studies on usability evaluation of LMSs have been carried out by many researchers [Nakamura et al. 2017a]. Most of the researchers of these studies employed questionnaire-based techniques due to their benefits of gathering the perceptions from participants’ point of view, while being fast, easy and low cost.

Given that usability/UX problems may interfere in the learning process, it is important that the platform be evaluated by those who will be using it to learn, i.e., the learner. Additionally, since learners may have never carried out a usability/UX evaluation before, the technique should be fast and easy to be applied and understood. Thus, we decided to develop TUXEL in form of a questionnaire.

The first step to develop our technique was to define its dimensions. To do so, we searched for the dimensions evaluated by the techniques returned in our systematic mapping. We identified two main dimensions: general usability and pedagogical usability. Additionally, we included the UX dimension due to its importance in the context of LMSs. Table 1 presents the dimensions and items evaluated by TUXEL.

The general usability criteria are composed by nine categories and their respective items, extracted from several works identified in our systematic mapping. One researcher categorized the items and two researchers, one of them usability expert and the other expert in computers and education, reviewed the categorization.

To better understand the concept of pedagogical usability, we first identified the papers that considered it in our systematic mapping. Then, we carried out a process known as backward snowballing, which consists in using a reference list to identify new papers to include. At the end of the process, we selected the work of Nokelainen (2006), one of the most cited works by the papers that considered pedagogical usability in our systematic mapping. Nokelainen (2006) defines that it relates to the assumption that the designers of
the learning platform were guided by either a conscious or subconscious idea of how its functions facilitate the learning of the material it is delivering. The author also provided a set of criteria to evaluate digital learning materials, divided in 10 dimensions, which we used as the basis for the pedagogical dimension of TUXEL. However, we did not consider the dimensions Learner activity, Applicability and Valuation of previous knowledge, given that they were related to the evaluation of the learning material, which is teacher dependent and not of the LMS itself. The selection and categorization process of the items for each category was the same of that performed for the general usability dimension.

![Table 1. Dimensions and items evaluated by TUXEL.](image)

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
<th>Categories – Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>It contains general usability criteria adapted to the context of LMSs.</td>
<td>Navigation (N), Learnability (L), Consistency (C), Visual Design (VD), Personalization (P), LMS Feedback (LF), Help and Documentation (HD), Learner Control (LC) and Error Prevention (EP).</td>
</tr>
<tr>
<td>Pedagogical Usability</td>
<td>It verifies the extent that the LMS was designed to facilitate and support the learning process.</td>
<td>Collaborative Learning (CL), Goal Orientation (GO), Instructional Assessment (IA), Added Value (AV), Motivation to Learn (ML), LMS Flexibility (LFL) and Instructional Feedback (IF).</td>
</tr>
<tr>
<td>User Experience (UX)</td>
<td>It evaluates learner’s perceptions regarding his/her experience when using the LMS.</td>
<td>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, motivates to learn/demotivates to learn.</td>
</tr>
</tbody>
</table>

We organized the general usability and pedagogical usability items in form of a questionnaire composed by a 5-point Likert scale, ranging from “strongly disagree” to “strongly agree”. Regarding UX dimension, we based it on UEQ, given that it was considered fast, easy and with many adjectives that cover a wide range of emotions. However, we selected only two adjectives from each of its dimensions in order to not cause cognitive overload. We also removed the adjectives that the participants from the preliminary study considered confusing, resulting in a total of 12 pairs of adjectives. Additionally, we added two optional comment fields in order to overcome the limitations of the original UEQ. In the first field, the learner can make criticisms considering his UX evaluation. In the second field, the learner can make suggestions for the LMS evaluated.

### 3.4. First Empirical Study

This study aimed to verify whether TUXEL identifies usability and pedagogical usability problems, in addition to being fast and easy to be applied by students. We carried out this study with 46 undergraduate students from UFAM, enrolled in Technical Texts in Computer Science course, most of them with low or no experience in usability evaluation. The participants evaluated an LMS called ColabWeb, which is based on version 3.1 of Moodle LMS. We performed a comparison between TUXEL and a technique developed by Theng and Sin (2012), chosen through a selection process that considered aspects such as scope (generic or specific for e-learning) and target audience (experts or novices). Details of this study and the selection process can be found in Nakamura (2018).

The results indicated that both techniques provide some information on usability and pedagogical usability of the LMS evaluated. However, some participants reported difficulty in remembering whether an aspect being evaluated exists in the platform or not. Moreover, the use of a scale made it difficult for us to interpret the results, although works...
such as by Oztekin et al. (2010) pointed out that their technique (also based on 5-point Likert scale) allowed them to determine usability problems and define relevant improvements. Some items had few differences between agreements, neutrals and disagreements, making it difficult to identify whether it was a problem or not. 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 to identify usability problems. Due to this, we decided to carry out a second iteration over the Design Cycle of DSR in order to refine the artifact considering the drawbacks and the suggestions provided by the participants.

3.5. Technique Improvement: Development of TUXEL 2.0

Our main concern was on the difficulty in identifying usability problems and the participants’ difficulty in remembering whether an evaluated item existed or not in the LMS. Given this, we decided to use an inspection-based approach. In inspection, the evaluation is usually performed by usability experts. However, since our goal is to make TUXEL applicable by students, it must provide guidance during the evaluation. Thus, we re-structured and changed it to a guided inspection approach in its second version.

In this version, we transformed all items in negative phrases to make it more intuitive for learners to find problems. We also included hints to items that evaluate specific aspects that novice inspectors may not know about. Additionally, we removed items that did not provide useful results in the first study, such as those from Added Value and Motivation category. We also grouped the items according to the task being inspected, in order to present only those that should be evaluated in each 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 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 aspects related to tasks, such as clarity of the instructions 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 task evaluated.

TUXEL 2.0 also has a checklist that evaluates general aspects of an LMS, divided in the following categories: Help and Documentation, LMS Learnability, Learning through the LMS and LMS Flexibility. Help and Documentation evaluates whether the help material is comprehensive and straightforward. LMS Learnability evaluates whether the LMS is easy to be learned. Learning through the LMS evaluates the extent that the LMS allows learners to learn through the platform. Finally, LMS Flexibility evaluates the extent that the LMS allows learners to have control over the platform and the learning process. The complete set of items can be found in Nakamura (2018).

To support the evaluation process and make it more interactive and intuitive to report problems, we also developed an evaluation tool, which consists in an extension for Google Chrome (see Figure 2). The tool provides instructions and the set of items from TUXEL 2.0 to evaluate the LMS according to the task being performed. It also provides a feature to take a screenshot and mark the area where the problem, in addition to allowing the inspector to detail it. At the end of the evaluation, it generates a report in a CSV format file with all the information of the inspection process, such as its duration, the ID of each violated item, the URL where the problem occurred and the commentaries.
3.6. Second Empirical Study

The second empirical evaluation aimed to evaluate TUXEL 2.0 in comparison to other techniques. Since we changed TUXEL to an approach based on guided inspection, we carried out a new technique selection process, selecting those specific to evaluate LMSs and that are based on inspection method. At the end of the process, we selected the Adapted Heuristic Evaluation (AHE) from Mtebe and Kissaka (2015).

We carried out a study with 22 students from two computer science courses at UFAM, divided into two groups, balanced according to their level of experience in usability/UX evaluations, each group using only one technique. Most of the students had medium or low experience in this type of evaluation. Only one participant (assigned to AHE group) had high experience. The object of this study was the ColabWeb LMS. Details of this study can be found in Nakamura (2018).

The results revealed a total of 103 unique problems, in which 58 were identified only by TUXEL, 29 only by AHE and 16 by both techniques. We analyzed the mean of the number of problems identified by each group according to the level of experience in usability evaluation. In all of them, the participants using TUXEL identified a higher number of problems. It may indicate that it provides a better guidance to inspectors in comparison to AHE. Additionally, our technique required, on average, less time to perform the inspection (49min. vs 54min.) and produced fewer false-positives (7.43% vs 16.59%). The statistical analysis on the efficiency and effectiveness with the Mann-Whitney non-parametric statistical test indicated that both efficiency and effectiveness of TUXEL were significantly higher than of AHE.

Regarding the tool, the participants perceived it, in general, as easy to use, flexible and useful, specially the screen marking feature. However, there are some improvement opportunities, e.g., in the marking feature itself, where some participants had difficulties.

4. Concluding Remarks

This paper presented the application of DSR to develop a technique and a tool to evaluate the usability and UX of LMSs called TUXEL. By employing DSR, we developed the technique grounded on theory and empirical evidence. In contrast to other works in the literature, we developed our technique by following a rigorous development process, aiming to solve the problem addressed while questioning or corroborating the theoretical conjectures. By doing so, we contributed by obtaining empirical evidence on different approaches used to evaluate the usability and UX of LMSs, making it possible to identify
how each of them works and to select those that are most adequate to evaluate these platforms. Additionally, DSR prompts researchers to clearly expose the knowledge about how each solution may help [Storey et al. 2017], benefiting practitioners who wishes to evaluate LMSs and improve their quality in use, or develop/improve other evaluation techniques from the insights obtained in this research. Our main contributions are:

**A secondary study on usability and UX evaluation of Learning Management Systems:** our systematic mapping provides an overview of the existing research on the topic and may serve as a starting point for new research in this area.

**A methodology for the development of artifacts based on Design Science Research:** the methodology followed to develop our technique may serve as the basis for the development of new artifacts, specially usability/UX evaluation techniques.

**A technique for evaluating usability and UX of LMS:** by using TUXEL, researchers may identify usability/UX issues, making it possible to develop solutions to improve the quality in use of LMSs and, consequently, the learning process through them.

**A tool to support the evaluation process:** the tool may be used to support the usability/UX evaluation, as well as a starting point for the development of new tools.

**Empirical evidence of some usability and UX evaluation techniques:** our empirical studies provide further information about the techniques compared. Their strengths and weaknesses may serve as basis for the development of new techniques.

As future works, we plan to evaluate TUXEL with different LMSs and student profiles, while improving it based on the results of the second empirical study.

**Acknowledgments**

We would like to thank the financial support granted by UFAM, FAPEAM through Edital 009/2017, CNPq through processes 423149/2016-4 and 311494/2017-0, and CAPES through process number 175956/2013.

**References**


