

Do Scale Type Techniques Identify Problems that Affect User eXperience?

User Experience Evaluation of a Mobile Application

Leonardo Marques, Walter Nakamura, Natasha Valentim, Luis Rivero and Tayana Conte

Institute of Computing
Federal University of Amazonas, UFAM
Manaus, Brazil

{lcm, walter, natashavalentim, luisrivero, tayana}@icomp.ufam.edu.br

Abstract—User experience (UX) is a quality aspect of an application that considers the emotions evoked by the system. There are several types of UX evaluation methods, such as scales, interviews, user monitoring, among others methods. However, there is still not enough information regarding if and in which contexts these methods are more suitable. This paper focuses on finding more information about the feasibility of applying scale-type methods for the UX evaluation in interactive systems. Thus, we carried out a UX evaluation on the Edmodo educational mobile application employing two scale-type methods: AttrakDiff and Hedonic Utility Scale. These methods were chosen after a selection process applying inclusion and exclusion criteria. The results indicate that it is possible to evaluate the quality of an application at a low cost. However, it is necessary to adapt these methods to provide a more complete report of the UX, allowing users to subjectively report their experiences and, consequently, identify the issues that affected the UX.

Keywords—User Experience; UX; AttrakDiff; Hedonic Utility Scale; Edmodo; comparative study;

I. INTRODUCTION

Technological evolution has enabled the emergence of new interaction paradigms, new technologies and new types of software [1]. This evolution has changed the way users interact and perceive interactive products. Hassenzahl and Tractinsky [2] have identified that pragmatic factors such as functionality and usability are not enough to satisfy the desire of customers seeking innovative products and products that meet their expectations. According to Bargas-Avila and Hornbæk [3], usability is too focused on the efficiency and the accomplishment of tasks, with a need for more comprehensive notions of quality. As a result, User eXperience has emerged as a concept of quality of use that, besides involving the aspects mentioned above, covers hedonic qualities or emotional aspects resulting from the use of an application [3].

Several UX evaluation methods have been proposed in the literature. Vermeeren et al. [4] carried out a survey on the UX evaluation methods used in the academy and in industry and identified 96 methods. Despite the large number of methods, there is a need to verify their applicability with regards to the resources, required skills to apply them, and the perceived ease of use of these methods [3].

This paper presents a comparison between two scale-type methods and the indications of their feasibility in the context of evaluating a mobile educational application called Edmodo. The UX evaluation of Edmodo was suggested by a University that was analyzing the possibility of using educational technologies in the classroom. Considering that UX is one of the aspects that can impact the acceptance of adopting teaching and learning technologies [5], we decided to carry out a UX evaluation to verify if Edmodo evoked a positive UX in its mobile version.

In order to select the UX evaluation methods, we carried out a selection process by means of inclusion and exclusion criteria, which resulted in two scale-type techniques: AttrakDiff [7][16] and the reduced form of the Hedonic Utility Scale (HED/UT) [8]. We compared these methods in terms of usefulness and ease of use, raising information to verify their feasibility for assessing mobile applications. The results allowed us to identify the positive and negative aspects of the methods employed, besides presenting opportunities for improvement in the application.

The remainder of this paper is organized as follows. Section II presents concepts related to UX. Section III describes the use of UX evaluation methods presenting the evaluated application, the selection process and the study carried out. Section IV presents the results obtained in this study and the discussions. Finally, Section V concludes the paper.

II. BACKGROUND

ISO 9241-210 [9] defines UX as "*person's perceptions and responses that result from the use and/or anticipated use of a product, system or service.*" Hassenzahl and Tractinsky [2] suggest that UX is a consequence of the user's internal state (predispositions, expectations, needs, etc.), the characteristics of the projected system (complexity, usability) and the context where the interaction occurs (organization, users, among others).

UX evaluation plays an important role in the development of interactive applications. Through the UX evaluation it is possible to identify how users apply, perceive and learn these applications, allowing the applications to evolve and adapt to user's expectations [10]. Thus, it is possible to identify potential problems in the use of applications and their causes, as well as to obtain suggestions for their improvement. There are several methods for evaluating UX, which can be categorized into three

types [11]: (i) written report, which consists of the evaluation of the experience through scales, forms and questionnaires; (ii) oral report, in which participants report their experiences through interviews or verbal methods; and (iii) observation/monitoring, in which participants are observed or use sensors to monitor their responses while perform activities related to the use of the system.

Written reporting methods, such as scale-type methods, have been widely used due to their low cost, ease of use, and the possibility of collecting data from both positive and negative experiences [12]. Examples of this type of method are questionnaires that use scales with semantic differentials, i.e., scales composed of pairs of words that are opposing adjectives, such as "simple/complicated" and "pleasant/unpleasant". However, there is little information regarding the positive and negative aspects of these methods. Thus, we selected methods within this category to: (a) analyze what type of outcomes are produced by these methods; (b) analyze if scales are sufficient to identify problems that have affected the UX; and (c) verify the feasibility of these methods to evaluate a mobile application.

III. APPLYING USER EXPERIENCE ASSESSMENT METHODS

A. Goals and Metrics

Table 1 presents the goal of the study according to the GQM (Goal Question Metric) paradigm, which allows to define and evaluate objectives in the stage of goal setting [13].

TABLE 1. GOAL OF THE STUDY ACCORDING TO THE GQM PARADIGM.

Analyze	The AttrakDiff and Hedonic Utility Scale UX evaluation methods
For the purpose of	characterizing.
With respect to	participants' perception in terms of Usefulness, Ease of Use and intention to use each method.
From the point of view	of users and researchers.
In the Context of	a UX evaluation of a real application in a Computer Science Introduction class.

The UX problems were verified through the number of tasks performed successfully in Edmodo as well as through difficulties experienced by the participants. The utility, ease of use and intention to use the methods were obtained through the TAM3 (Technology Acceptance Model) [14].

B. UX Method Selection Process

We applied the selection process on the UX evaluation methods list identified by Rivero and Conte [11], seeking to identify UX evaluation methods for mobile applications. This selection process consisted of two refinement steps.

The first refinement was based on the criteria described in Table 2. For each criterion, there is a description of what was considered for the exclusion. First, we considered only the methods available for consultation (EC2). After, we applied the other exclusion criteria, resulting in a set of 18 UX evaluation methods. The detailed specification of the criteria used for exclusion can be found in the technical report [6]

TABLE 2. CRITERIA FOR FILTERING THE UX EVALUATION METHODS.

Criteria	Description
EC1 (Type of method)	Methods characterized only as tools.
EC2 (Availability)	Methods not available for free or unavailable.
EC3 (Data source)	Methods whose data sources are not provided by users.
EC4 (Location)	Methods whose application is not possible in controlled environments.
EC5 (Type of Assessed Product)	Methods that cannot be applied to mobile app evaluation.
EC6 (Type of Assessed Artifact)	Methods whose evaluated artifact are not functional prototypes or final applications.
EC7 (Assessed Period of Experience)	Methods whose UX evaluation occurs before or during the use of the system.

Among the 18 selected methods, we queried each method in the AllAboutUX¹ Website in order to carry out the second refinement. This website presents, for each method, their characteristics and what is needed to use them. Some methods, such as the "Group-based expert walkthrough" [15] requires UX experts to be applied. Given that the UX of Edmodo would be assessed by users and not experts, we discarded methods that require experts to evaluate. We also discarded those developed for specific contexts, such as the "Attrak-Work Questionnaire" method [16], which was developed for the context of news and journalism. At the end of the second refinement, we selected three methods: AttrakDiff [7][16], Hedonic Utility Scale (HED/UT) [8] and Self-Assessment Manikin (SAM) [17].

We carried out a pilot study using the three selected methods to verify the data collected by each method and the outcomes that each one generated. The results indicated that SAM is a method that evaluates only the hedonic aspects of the experience, thus it would be unfair to compare it with AttrakDiff or HED/UT, which assess both pragmatic and hedonic aspects of the user experience. Thus, at the end of this process, we selected the AttrakDiff and HED/UT methods.

The AttrakDiff method consists of 28 pairs of words that evaluate pragmatic, hedonic and attractiveness aspects, whereas the HED/UT method, in its reduced form, consists of 12 pairs of words that evaluate the pragmatic and hedonic aspects. Regarding HED/UT, we applied its reduced version because the results obtained in [8] showed that its 12 pairs of words are sufficient to verify the quality, making it viable to measure the UX. The pairs of words of both methods are organized on a seven-point scale, in which the participant performs the UX evaluation by marking the closest point to the adjective that best characterizes his experience of use (see Figure 1).

C. Evaluated Mobile Application

We decided to evaluate the Edmodo application due to the university's suggestion to address its feasibility, while no studies related to the UX evaluation of this application were found. Edmodo is a Learning Management System (LMS) created in 2008 to manage learning activities. Its popularity among educational institutions has risen and it has more than 80 million users and more than 50 million downloads in its mobile version.

The evaluation of LMSs is a critical issue, as it can affect students' performance, making them spend more time trying to

¹ <http://www.allaboutux.org/all-methods>

understand how to use these environments than learning the educational content [18][19].

A	Human						X		Technical
	Isolating								Connective
	Pleasant								Unpleasant
B	Useful								Useless
	Impractical					X			Practical
	Necessary								Unnecessary

Figure 1. (A) AttrakDiff Questionnaire, (B) HED/UT Questionnaire.

D. Execution

We conducted the study with 38 volunteer students from the Federal University of Amazonas, who participated in an Introduction to Computing class. This class was partially online and used an LMS in the teaching/learning process.

For the execution of the study, we accommodated the participants in a laboratory and divided them into two groups, balanced according to their previous experience with the Edmodo app. One group used the AttrakDiff and the other group used HED/UT (details in the technical report [6]).

Initially, participants received a Consent Form, and then, they received: (a) a form for reporting their difficulties when using Edmodo, (b) the AttrakDiff and HED/UT, (c) the TAM3 questionnaire, and (d) a questionnaire with open questions related to the use of the UX evaluation methods.

Before performing the UX evaluation, we made a brief explanation about Edmodo and its functionalities. After, we asked the participants to download the application. All the participants used Android or iOS devices, in which a prior check was made to ensure that the execution flow of the activities in the application would not be changed.

After explaining the application, the participants received a schedule of activities. According to Nielsen [20], the basic rule for selecting a set of activities is that it must be chosen in such a way that they are as representative as possible. Thus, this script consisted of the following activities: (i) create a student account in Edmodo and enter the group through the access code, made available by the moderators during the study, (ii) update the profile photo, (iii) access the group library and download a file for reading, (iv) answer an activity containing two questions related to the text of the downloaded file and (v) attach a file and send it to the teacher.

At the end of the activities, the participants received a form about the difficulties encountered during the interaction with Edmodo in order to better understand the quantitative results of the UX evaluation. Then, each group received a UX evaluation method, in which each participant was instructed on how to perform the evaluation. The form about the difficulties and the number of tasks performed successfully in Edmodo was verified in order to identify the UX problems.

As the participants finished evaluating the application, they received the questionnaire based on the TAM3. The TAM3 is a model used to verify the acceptance of a technology that, among other dimensions, considers utility, ease of use and intention to

use. Participants were instructed to use this questionnaire to evaluate the UX evaluation method they used. We also attached to the TAM3-based questionnaire, a questionnaire containing five open questions related to their experiences with the UX evaluation method in order to better understand the aspects that made each method easy or difficult to use.

IV. RESULTS AND DISCUSSIONS

In order to make the comparison possible, we tried to equate both methods, since they evaluate the UX in different dimensions. AttrakDiff evaluates UX in four dimensions: Pragmatic Quality (PQ), Hedonic Quality/Stimulus (HQ/S), Hedonic Quality/Identity (HQ/I) and Attractiveness (ATT), while HED/UT evaluates two dimensions: Utility, which corresponds to the Pragmatic dimension, and Hedonic.

Given that the pragmatic dimensions of both methods are equivalent (evaluate the same experiences), we searched for definitions that qualified what was evaluated in the hedonic dimension of each method to verify if they were equivalent. According to Voss et al. [21], the hedonic dimension of HED/UT is the result of sensations derived from the experience of using products. According to Vääätäjä et al. [16], the HQ/S dimension of AttrakDiff is related to personal development, that is, curiosity, personal growth, skill development, and the proliferation of knowledge, i.e., feelings and sensations caused by the use of the application. Thus, the AttrakDiff HQ/S was considered equivalent to the HED/UT hedonic dimension. Thus, we considered only the Pragmatic and QH/E dimensions of AttrakDiff, and the Utility and Hedonic dimensions of HED/UT.

To compare the methods, we organized the data by factors. Factor 1 relates to the pragmatic dimension, while Factor 2 represents the hedonic dimension. The following subsections describe the results of the UX evaluation of Edmodo and the results regarding the methods used to evaluate the its UX.

A. Results and Analysis of the UX Evaluation on Edmodo

Table 3 presents the score for each Factor per method. According to Distefano et al. [22], when Factors are not defined by the same number of items, which is the case of the Factors of both methods, it is recommended to calculate the mean, making it possible to compare them with each other. The mean is also recommended by Sullivan and Artino Jr [23] when measuring less concrete concepts, such as satisfaction, where a single research item is not likely to capture the assessed concept completely. In order to obtain the Factor's scores, first we calculated the mean of each participant's scores per Factor. This mean was based on the score given by each participant in each item of the method's dimension. Then, we obtained the score of each Factor through the mean of the scores of each previously calculated participant.

Regarding Factor 1, the scores indicate that participants who evaluated UX using the HED/UT evaluated more positively the experience regarding the ease of use of Edmodo compared to those using AttrakDiff. Given that the scale ranges from 1 to 7, the UX would be positive if the scores were greater than or equal to 5. Thus, the results indicated that the participants considered that Edmodo provides a positive UX, since the lowest score was close to 5. Regarding Factor 2, the group that used AttrakDiff

felt it neutral, i.e., Edmodo was not considered bad, but it did not stimulate users so much. It indicates that Edmodo needs to implement improvements to stimulate and captivate users. The group that used HED/UT considered Edmodo's UX positive.

TABLE 3. FACTORS' SCORE ASSESSED BY THE ATTRAKDIFF AND HED/UT.

Factor 1 (Pragmatic)			Factor 2 (Hedonic)	
Method	AttrakDiff	HED/UT	AttrakDiff	HED/UT
Score	4,8	5,9	4,4	5,5

The responses provided by the participants in the form about the difficulties faced when using Edmodo reflected these scores. From a total of 38 participants, 18 reported having had difficulty finding the group's Library in Edmodo. This is a problem that affects the use of the application, being reflected in the UX evaluation, where most participants considered that Edmodo is very technical (AttrakDiff) and impractical (HED/UT). Other problems were also pointed out, such as the mix of words in Portuguese and English on the interface. These problems were only possible to be identified through the form on the difficulties faced in Edmodo, because the scales do not allow to identify the problems that affected the UX in this level of specification.

B. Results Regarding the UX Evaluation Methods

In order to verify the participants' perception regarding the usefulness, ease of use and intention to use, we applied the TAM3-based questionnaire. We used the median as a statistically significant measure for ordinal scales [24] with the same number of items. Table 4 shows the description of the items that compose each of the dimensions evaluated by TAM3.

TABLE 4. TAM3-BASED QUESTIONNAIRE ITEMS.

Description of the items on "Perceived Usefulness" (PU)	
PU1	Using the method improves my performance by reporting my experience with the application.
PU2	Using the method improves my productivity by reporting my experience with the application.
PU3	Using the method allows me to fully report the aspects of my experience.
PU4	I find the method useful for reporting my experience with the application.
Description of the items on "Perceived Ease of Use" (PEOU)	
PEOU1	The method was clear and easy to understand.
PEOU2	Using the method did not require much mental effort.
PEOU3	I think the method is easy to use.
PEOU4	I find it easy to report my experience with the application using the method.
Description of the items on "Intention to Use" (IU)	
IU1	Assuming I have access to the method, I plan to use it to evaluate my experience with an application.
IU2	Given that I have access to the method, I predict that I would use it to evaluate my experience with an application.
IU3	I plan to use the method to evaluate my experience with an app next month.

Table 5 shows the median values for each TAM3 item. Based on these data, we verified the items that had some variation, since these indicate which of the methods was better.

The items that had variations were PU2, PEOU1, PEOU4 and IU2. These items show that HED/UT had a better perception regarding AttrakDiff by the participants, indicating that short methods improve productivity (PU2), methods that use less formal terms are easier to use (PEOU1 and PEOU4), and these aspects influence intention to use (IU2).

TABLE 5. MEDIAN OF EACH ITEM PER METHOD.

	AttrakDiff	HED/UT
Perceived Usefulness (PU)		
PU1	5	5
PU2	5	6
PU3	5	5
PU4	5 and 6	6
Perceived Ease of Use (PEOU)		
PEOU1	5	6
PEOU2	6	6
PEOU3	6	6
PEOU4	5 and 6	7
Intention to Use (IU)		
IU1	5	5
IU2	5	6
IU3	4	4

The results of the questionnaire with open questions about the UX evaluation methods indicated some opportunities for improving them. In both methods, some participants reported that they were not able to express their experiences of use only through the scales. The methods do not allow them to write the problem that affected the UX, or in which part of the application they consider that there should be improvements, which indicates a limitation of the methods evaluated. One possibility of improvement would be, for example, the addition of a field so that the participant can report the difficulties that were not possible to be described only with the scales.

"I cannot describe the experience I had" – P08 (HED/UT).

"Not being able to express [the experience] in a more justified way" – P05 (AttrakDiff).

Regarding AttrakDiff, the participants reported the difficulty in understanding some terms, considering them formal (see quote from P12). This may impact the UX report, since the participant can point out in any way when evaluating. The results of Table 5 showed that HED/UT had a better perception of users than AttrakDiff, because it used less formal terms. This could be an indication that terms that are used daily by users must be used instead.

"[There were] some formal words that I did not know what they meant" – P12

Regarding HED/UT, some participants indicated that the available options were insufficient to evaluate the experience satisfactorily (see quote from P02). Others reported that just having to point out an "X" makes the method simple and easy (see quote from P03).

"Only more options to point out" – P02

"It's simple and easy" – P03

Thus, there are some gaps with regards to the methods used in the study. Based on the reports, scale methods should be complemented with questionnaires or a comment field, allowing the evaluator to describe the difficulties faced and what aspects were enjoyable when using the application. The lack of these fields makes it difficult to implement improvements in the evaluated application, since it is possible only to know that the application needs improvements, but not which problems users have faced. In addition, it is recommended to use less formal

terms, in order to make the method more comprehensible, as shown in Table 5.

V. CONCLUSIONS

Performing a UX assessment is important for gaining end-user insight about a particular application. In this paper, we showed that the use of scale-type methods allows to perform the UX evaluation quickly, not making the evaluation process tiring for the user. In addition, few resources are required to evaluate the quality of an application, reducing the cost of evaluation, which makes the use of these methods attractive.

However, this type of method has the limitation of not collecting qualitative data of the evaluation, i.e. the subjective information that describes the difficulties encountered by the users and that could point out the problems of the evaluated application. This may indicate that only using scale-type methods may not produce such detailed results, making it difficult to precisely identify which aspects affected the UX during the use of the application. For a holistic assessment, complementation of the scales with the open questions was positive, making it possible to obtain the positive aspects and the aspects that need to be improved in the application.

The results of the UX evaluation showed that Edmodo has a positive UX and that it can be used by teachers as a tool to support the teaching/learning process. However, some improvements are needed, such as facilitating the access to the library and fixing the mix of languages on the interface. By doing this, the application can have a greater acceptance and become easier and more enjoyable to use, important aspects to have a competitive advantage over other applications.

We hope that the results from this study contribute to the development of UX evaluation techniques that make use of the positive aspects found in the scale-type methods and that provide the negative aspects, such as the lack of a field where the participants can detail their experiences, to obtain a more complete and detailed UX report. In addition, we hope that suggestions for improvements can contribute to the improvement of the Edmodo application.

ACKNOWLEDGMENT

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

REFERENCES

- [1] C. Rusu, V. Rusu, S. Roncagliolo and C. González, "Usability and User Experience: What should we care about?", *International Journal of Information Technologies and Systems Approach*, v.8, n.2, p.1-12, 2015.
- [2] M. Hassenzahl and N. Tractinsky, "User experience-a research agenda", *Behaviour & information technology*, v.25(2), p.91-97, 2006.
- [3] J. A. Bargas-avila and K. Hornbæk, "Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience", In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 2689-2698, 2011.
- [4] A. P. Vermeeren, E. L. C. Law, V. Roto, M. Obrist, J. Hoonhout and K. Väänänen-Vainio-Mattila, "User experience evaluation methods: current

- state and development needs", In *Proc. of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, p. 521-530, 2010.
- [5] N. Harrati, I. Bouchrika, A. Tari and A. Ladjailla, "Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis", *Computers in Human Behavior*, 61, 463-471, 2016.
- [6] L. Marques, W. T. Nakamura, L. Rivero, N. Valentim and T. Conte, "TR-USES-2017-0012. Supporting Material for Evaluating User Experience in a Mobile Education Application". *Technical Report of Usability and Software Engineering Group (USES)*, 2017 Available in <http://uses.icomp.ufam.edu.br/relatorios-tecnicos>.
- [7] M. Hassenzahl, M. Burmester and F. Koller, "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität", In *Mensch & Computer*, p. 187-196, 2003.
- [8] H. Van Der Heijden and L. S. Sørensen, "Measuring attitudes towards mobile information services: An empirical validation of the HED/UT scale", In *Proceedings of the European Conference on Information Systems (ECIS)*, p. 765-777, 2003.
- [9] ISO DIS 9241-210:2010. "Ergonomics of human system interaction -Part 210: Human-centred design for interactive systems". *International Standardization Organization (ISO)*. Switzerland.
- [10] A. Moreno, A. Seffah, R. Capilla and M. Sanchez-Segura, "HCI practices for building usable software", *IEEE Computer*, (4), p. 100-102, 2013.
- [11] L. Rivero and T. Conte, "A systematic mapping study on research contributions on UX evaluation technologies", In *Proc. of the 16th Brazilian Symposium on Human Factors in Computing Systems (IHC)*, Joinville, Brazil, 2017.
- [12] P. Desmet, "Measuring emotion: Development and application of an instrument to measure emotional responses to products", In *Funology*, Springer Netherlands, p. 111-123, 2005.
- [13] V. Basili, G. Caldiera and H. Rombach, "Goal question metric paradigma", *Encyclopedia of Software Engineering (1)*, John Wiley & Sons, New York, p. 528-532, 1994.
- [14] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions". *Decision sciences*, 39(2), 273-315, 2008.
- [15] A. Foelstad, "Group-based Expert Walkthrough", In *R³UEMs: Review, Report and Refine Usability Evaluation Methods*, edited by Scapin, D. and Law, E., p. 58-60, 2007.
- [16] H. Väättäjä, T. Koponen and V. Roto, "Developing practical tools for user experience evaluation: a case from mobile news journalism", In *European Conference on Cognitive Ergonomics: Designing beyond the Product-Understanding Activity and User Experience in Ubiquitous Environments*, p. 23, 2009.
- [17] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential". *Journal of behavior therapy and experimental psychiatry*, 25(1), p. 49-59, 1994.
- [18] R. Lanzilotti, C. Ardito and M. F. Costabile, "eLSE methodology: A systematic approach to the eLearning systems evaluation", In *Educational Technology & Society*, v. 9, n. 24, p. 42-53, 2006.
- [19] H. B. Santoso, M. Schrepp, R. Isal, A. Y. Utomo and B. Priyogi, "Measuring User Experience of the Student-Centered e-Learning Environment", In *Journal of Educators Online*, v. 13, n. 1, p. 58-79, 2016.
- [20] J. Nielsen, "Usability Engineering", In *Morgan Kaufmann Publishers Inc*. San Francisco, CA, USA, 1993.
- [21] K. E. Voss, E. R. Spangenberg and B. Grohmann, "Measuring the hedonic and utilitarian dimensions of consumer attitude", In *Journal of marketing research*, v. 40 (3), p. 310-320, 2003.
- [22] C. Distefano, M. Zhu and D. Mindrila, "Understanding and using factor scores: Considerations for the applied researcher", In *Practical Assessment, Research & Evaluation*, v. 14 (20), p. 1-11, 2009.
- [23] G. M. Sullivan and A. R. Artino Jr, "Analyzing and interpreting data from Likert-type scales", In *Journal of Graduate Medical Education*, v. 5, n. 4, p. 541-542, 2013.
- [24] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell and A. Wesslén, "Experimentation in software engineering", *Springer Science & Business Media*, 1st edition, 2012.