Journal of the Brazilian Computer Society, 2025, 31:1, doi: 10.5753/jbcs.2025.4306

Learner Experience Evaluation: a Feasibility Study and a Benchmark

Gabriela Corbari dos Santos [**Federal University of Paraná** | ***gabrielacorbari@ufpr.br***]
Deivid Eive dos S. Silva [**Federal University of Paraná** | ***dessilva@inf.ufpr.br***]
Natasha M. C. Valentim [**Federal University of Paraná** | ***natasha@inf.ufpr.br***]

Department of Informatics, Federal University of Paraná, R. Evaristo F. Ferreira da Costa, 383-391 - Jardim das Américas, Curitiba, PR, 81530-090, Brazil.

Received: 23 March 2024 • **Accepted:** 24 June 2024 • **Published:** 20 February 2025

Abstract Learner eXperience (LX) is a concept derived from User eXperience (UX) and it can be defined as the perceptions, answers, and performances of learners interacting with Digital Communication and Information Technologies (DICTs). Evaluating the LX to obtain experiences that support and facilitate learning and knowledge mastery is important. Thus, we developed the LEEM to assess and improve the learner's experience using DICTs during learning. The LEEM is a generic evaluation model and can be used for any level of education; it can be used independently of the discipline and used with any educational technology. Therefore, this article presents a feasibility study to evaluate the LEEM steps and sentences from the perspective of potential users. Nineteen teachers from different levels of education participated in this study. The study results were analyzed and generated in a new version of LEEM. The results showed positive points of LEEM, such as a practical, objective, easy-to-use, and useful model. In addition, opportunities for improving some items and sentences of LEEM were obtained. The teachers also suggested adding a description at the ends of the scales to facilitate the response to the items. This study contributes to creating a body of knowledge about LEEM and analyzing its use, feasibility, and evolution. Moreover, realizing the lack of content and synthesized characteristics about the technology that evaluates the LX, we carried out a benchmark on the LX evaluation technologies identified from a Systematic Mapping Study (SMS) to compare them with LEEM, in addition to presenting important characteristics to be analyzed in these types of technology, such as the elements and types of LX assessment and whether there is tool support for this assessment.

Keywords: Learner eXperience, Evaluation, Model, LX, Qualitative analysis, Quantitative analysis, Student eXperience

1 Introduction

Society is full of Digital Information and Communication Technologies (DICTs), which are gradually being inserted into people's lives to facilitate and assist in their daily activities. When applied to the school environment, DICTs can provide interactive and collaborative experiences and influence learners' self-esteem [Imamura and Baranauskas, 2019; Muriana and Baranauskas, 2021].

Soloway *et al*. [1994] state that there is a need to support learners in learning experiences in activities that are out of their reach, seeking protagonism in the learning process. In this sense, Soloway *et al*. [1994] introduces the term Learner, which is used to refer to students who are constantly learning or to professionals who put themselves in a position to learn, subjecting themselves to all the challenges faced when using DICTs.

Based on this, the term Learner eXperience (LX) emerged, which is a concept derived from User eXperience (UX) [Kawano *et al*., 2019]. The definition of UX according to ISO 9241-210 is user perceptions and responses resulting from using or anticipating a system, product, or service. According to Schmidt and Huang [2022], LX focuses on a specific class of users (the learner) engaged in a particular task (related to learning) while using a distinct type of technology (a technology tool designed for learning). In addition, Shi [2014] conceptualizes LX as a learner's perceptions and responses, resulting in learner behavior, attitudes, beliefs, sensations, and emotional responses obtained, among others. Therefore, there is a need to create experiences that facilitate learning and mastering knowledge [Correa *et al*., 2021].

Huang *et al*. [2019] believe that LX with DICTs can be evaluated and improved through different elements, such as Value, Usability, Adaptability, Confortability, and Desirability. The elements of LX refer to the components that guide the process of evaluating LX, making it possible to verify various characteristics of the experience, including feelings and emotions in learning. In this way, the elements of LX can vary depending on the objective of the evaluation, the type of artifact chosen, and the learning theory adopted.

From a Systematic Mapping Study (SMS) [dos Santos *et al*., 2022], different forms of evaluation (focus group, word pairs, records/notes of feelings, and others) and different elements of LX were identified. Based on the results of this SMS, no approach was identified that considers different forms of assessment and allows for a pluralistic and more meaningful LX for learners considering different elements. However, Huang *et al*. [2019] recommend that LX be assessed comprehensively, considering the diversity of learners and their learning preferences, so that all aspects of the experiences are evaluated. It is believed that it will be possible to contemplate and evaluate various aspects of the expe-

rience using DICTs. LEEM (Learner Experience Evaluation Model) was developed based on this need.

LEEM aims to evaluate and improve LX in the use of DICTs. The LEEM has three stages of evaluation: a pre-evaluation, a during-evaluation, and a post-evaluation of the LX, due to the possibility of continuously monitoring and recording the progress of the LX. At each stage, it is possible to use different ways of evaluating the LX, thus generating the possibility of evaluating more elements of the LX. The LEEM consists of three checklists and a focus group. It is hoped that the LEEM will enable the learner to play a leading role through self-assessment, autonomy, and independence in the learning process.

Therefore, this article presents a feasibility study with the LEEM to analyze its structure and content, including stages, checklists, and items. Nineteen teachers, including elementary, high school, higher education, and postgraduate teachers, participated in this study. The participants' responses were analyzed quantitatively and qualitatively, looking at ease and difficulty of use, perceived usefulness, intention of future use, and suggestions for improvement. The study's main aim is not to find a definitive answer but to create a body of knowledge about the proposed technology [Shull *et al*., 2004].

In addition, this article presents a benchmark for analyzing and characterizing the evaluation technologies identified in the SMS [dos Santos *et al*., 2022]. We noticed great difficulty in finding synthesized information on the content and characteristics of the LX evaluation technologies. This difficulty may exist for other researchers and interested parties. We, therefore, benchmarked evaluation technologies such as questionnaires and models that assess LX. We analyzed the characteristics of these evaluation technologies, such as the elements of LX, the types of LX assessment, the stages at which LX was assessed, the variations in sentences/questions, the number of sentences/questions, and the scales used. We also analyzed whether there is support to assess LX and, if so, what support is. In addition, we analyzed the public who evaluated a technology that can be used and whether there are automated versions of it. We hope that this synthesis and analysis can facilitate the process of choosing the most appropriate LX evaluation technology for the context of educators.

The study presented in this article contributes to the field of Human-Computer Interaction (HCI) as it presents an LX evaluation model that seeks to assess and improve the use and interaction of learners with DICTs. Understanding LX during the learning process is fundamental, as it provides insights for educators into the importance of reviewing and, if necessary, adapting the DICTs used according to the individually needs of the learners [Martinelli and Zaina, 2021]. This is also important to avoid situations where learners feel uncomfortable with the DICTs, promoting more effective interaction and a more positive learning experience [da Silva and Ziviani, 2018]. In addition, the proposed model can help educators rethink their teaching and learning strategies when they notice that learners have reported difficulties with the resources adopted. Educators can also observe whether learners remain motivated in the educational activity and what could be improved. Finally, educators can also benefit from the Benchmark results by identifying relevant characteristics of LX assessment technologies and which of these technologies best suit their context.

This article is an extended and revised version of the paper published at IHC 2023 [dos Santos *et al*., 2024]. It is organized into sections: Section 2 presents related work. Section 3 introduces LEEM. Section 4 presents the planning and execution of the feasibility study. Section 5 presents the quantitative analysis. Section 6 presents the qualitative analysis and improvements made to the LEEM. Section 7 presents the Benchmark. Section 8 presents the discussions. Section 9 presents the limitations. Section 10 presents the final considerations and future work.

2 Related Works

Models are valuable tools for understanding, analyzing, and creating complex systems [Ruiz and Snoeck, 2018]. The following are two papers that present a model for evaluating LX. Both works were identified in the SMS [dos Santos *et al*., 2022] and evaluated LX in different scenarios. For example, for the work by Ruiz and Snoeck [2018], the model is aimed at evaluating adapted Kirkpatrick training.

The work by Tabares *et al*. [2021] uses the model for a set of feedback rules. Subsequently, the work of Martinelli and Zaina [2021] will be presented, which evaluates the learning experience of HCI learners using the Flipped Classroom methodology.

In Ruiz and Snoeck [2018]'s work, the adapted Kirkpatrick model is determined by the learning environments where didactic tools support teaching. Adaptation consists of using concrete metrics and instruments for each level of Kirkpatrick's model. The metrics are measured using the elements of LX: Reaction, Learning, Behaviour, and Results. Through the results of the LX elements of the different levels of the model, it is possible to properly assess the complexity of training programs and their effectiveness. The levels are considered the way to build and evaluate the evidence. The evidence makes it possible to assess the measure of training that contributed to the results and whether the results correspond to expectations. The model has a guide to evaluation questions and the appropriate criteria for evaluating the learning process.

In turn, the work of Tabares *et al*. [2021] presents the Feedback Rules Model. For example, a rule in this model comprises four elements: Scenario, Event, Condition, and Action. The Scenario defines the behavior of the context in a learning environment, a predictive model provides the Event, the Condition evaluates the events, and the Action provides the learner feedback. Tabares *et al*. [2021] carried out a controlled experiment with higher education learners to capture the data and define the feedback rules. Thus, it was possible to identify the contribution of this approach to diagnosing weaknesses in the content and identifying the context that can influence the assessment results [Tabares *et al*., 2021].

Martinelli and Zaina [2021], in their work, investigated learning experiences in the context of the Flipped Classroom in online teaching. The authors proposed evaluating LX throughout the learning process and at the end of the edu-

cational activities. In this way, Martinelli and Zaina [2021] used a questionnaire and an interview, respectively. In this study, LX was assessed based on the three elements of Satisfaction, Motivation, and Feeling about learning. This work made it possible to understand the human factors that influence the learning experience, especially in remote classes [Martinelli and Zaina, 2021].

The studies cited have limitations. For example, in the work by Ruiz and Snoeck [2018], the evaluation results are limited to data collected only during training. The ideal scenario would be for learners to continue using the tool for a longer period after training and to provide feedback. In Tabares *et al*. [2021]'s work, data and the learner's context are collected and analyzed during the learning process. However, for feedback to be available before or during the learning process, it will be necessary to have prior knowledge of the context and scenarios that lead to learner success or failure in assessment processes. And Martinelli and Zaina [2021]'s work took place with one class of learners. However, the authors believe that the ideal scenario would be for two or more classes of learners to take part. In the assessment of LX, the ideal is for the assessment to take place longitudinally, i.e., before, during, and after an educational activity, as this allows for continuous monitoring and recording of the progress of the learner's experience. In addition, it is important to evaluate LX in different ways and consider different elements, as this can generate more effective learner experiences, which according to Huang *et al*. [2019], result in engaging and memorable educational experiences. However, no model in the literature met all these requirements. To address this limitation, LEEM was proposed, which will be discussed below.

3 The LEEM

The LEEM aims to evaluate and improve LX in the use of DICTs. The LEEM has three evaluation stages: pre-evaluation, during-evaluation, and post-evaluation (Figure 1 - in yellow) to enable continuous monitoring and recording of LX's progress during an educational activity [Lykke *et al*., 2015]. Each stage evaluates and checks as many elements of LX as possible. The different types of evaluating LX are presented in Figure 1 in orange. And the elements of LX are presented in Figure 1 in purple. The LEEM considers the maximum of possible elements and types of LX validation to allow the educator to be adaptable at the moment of the assessment, which enables customization to meet different objectives and needs.

3.1 LEEM construction process

LEEM was created through six steps, as seen in Figure 2. In step I, it was decided that an evaluation model would be created, as this type of technology offers an approach that allows subjective elements to be captured, making it an advantageous option [Ruiz and Snoeck, 2018] [Lima *et al*., 2021].

In step II, it was decided that the LEEM model would be organized into three assessment stages: (1) pre-assessment, (2) assessment during, and (3) post-assessment, as illustrated in Figure 1, to offer continuous monitoring and recording of LX progress during an activity [Nygren *et al*., 2019].

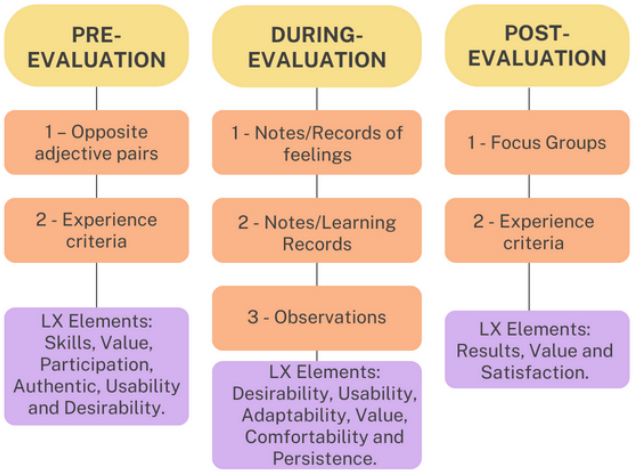

Figure 1. Learner Experience Evaluation Model.

In step III, different types of assessment were selected to be integrated into LEEM to capture various aspects of LX throughout an educational activity using a computer system. The types of assessment considered were: Opposite Adjective Pairs and Experience Criteria for the pre-assessment, Feeling and Learning Notes/Records and Observation for the during-assessment, and Focus Groups and Experience Criteria for the post-assessment. These assessment technologies were identified in the SMS and chosen based on the objectives and nature of these stages of the LEEM model. The motivations behind choosing these technologies are presented in the pre-, during-, and post-assessment steps.

In step IV, it was decided that different LX elements would be considered at each stage of the LEEM, to guide the types of LX assessment to be considered in the LEEM. Initially, the elements of the basic LX reference, which is Huang *et al*. [2019], were included. Subsequently, all the LX elements identified in the SMS were analyzed. It was then checked which of these elements was most suitable for the different ways of assessing LX in LEEM, taking into account the goals of each stage of the model.

In step V, artifacts were defined in the form of checklists and a set of questions to help collect the feedbacks in the LX assessment. Finally, in step VI, the LEEM model was created to support the evaluation of LX during the use of DICTs. One of the main advantages of the LEEM is its adaptability, which allows it to be customized to suit different goals. In this way, teachers can use the steps and elements they consider most relevant according to the educational goals of the subject. The LEEM model, including its checklists and assessment types, will be presented in greater detail in the following Subsection. It is important to emphasize that all the decisions made in constructing the LEEM were reviewed with the advisor and co-advisor. These researchers have experience in the area of HCI and Informatics in education. More details of the LEEM construction process are available in dos Santos *et al*. [2023].

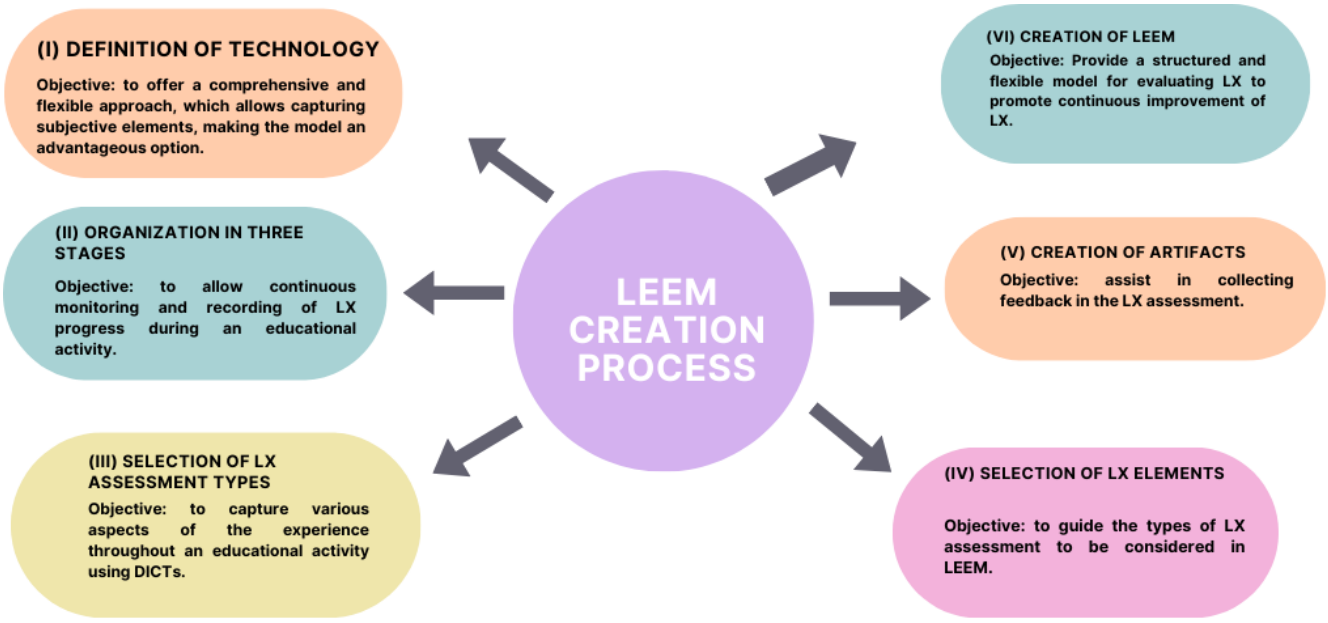

Figure 2. LEEM construction process.

3.2 Stages of the LEEM

The first stage, called the pre-evaluation, aims to identify the profile of the learners, considering their difficulties and needs regarding their learning experience [dos Santos *et al.*, 2023]. At this stage, it is also possible to analyze the results to help form groups for collaborative activities at school or university. It is suggested that the teacher apply group formation techniques according to their familiarity. Twelve pairs of words were drawn up for the LEEM model's pre-evaluation checklist [dos Santos *et al.*, 2023], as partially shown in Figure 3. For example, in the item "I prefer to work like this", the learner can select the option closest to the word they want, "practice" or "theoretical".

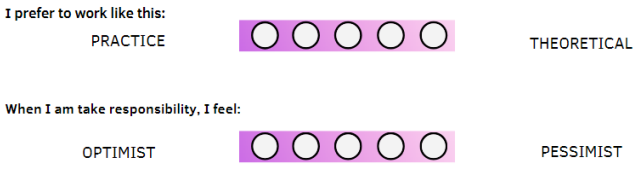

Figure 3. Part of the LEEM pre-evaluation checklist.

In the second stage, called evaluation-during, LX's evaluation occurs during the educational activity. This evaluation seeks to record the feelings and information about learning from both the learner's and the teacher's point of view [dos Santos *et al.*, 2023]. For this reason, the checklists that the learner (Figure 4) and the teacher (Figure 5) will answer have similar sentences to assess their perspective on the learning process. In Figure 4, the learner will report their feelings using the emoticons of the Self-Assessment Manikin (SAM) scale [Lang, 1980] about finding support material. On the other hand, the teacher will rate on a five-point Likert scale whether they agree that the learner enjoyed accessing the support material (Figure 5). Applying these checklists after 50% of the educational activity has been completed is recommended because both the learner and the teacher will already have had some experience to give feedback [dos Santos *et al.*, 2023]. For example, if the educational activity lasts two weeks, these checklists should be applied at the end of the first week. It should be noted that the teacher can answer a checklist for each learner, or if they are working in a group, they can answer a checklist for each group, or they can answer a checklist for the whole class, depending on the educational objective.

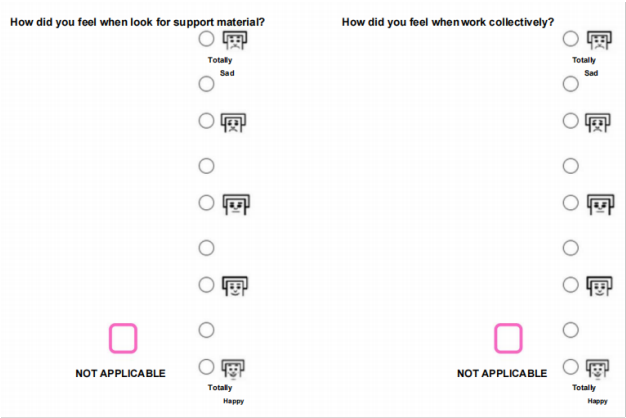

Figure 4. Part of the LEEM during-evaluation (learner) checklist.

The third and final stage, called post-evaluation, aims to evaluate the learner's experience of completing an educational activity using DICT. This stage allows learners to reflect on and self-assess their learning and have the freedom to be active and critical subjects in their learning [dos Santos *et al.*, 2023]. For the third stage, the teacher and the learner must complete the evaluation checklists during the activity. It is also necessary for all learners to have completed the activity proposed by the teacher. Figure 6 shows part of the open questions that can be used to direct the focus group. To carry out this stage, to the teacher is suggested to act as a

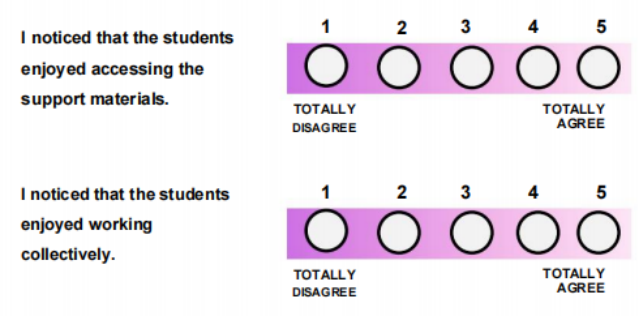

Figure 5. Part of the LEEM during-evaluation (teacher) checklist.

moderator during the discussions with the learners. Learners can express themselves randomly, and everyone doesn't need to answer all the questions. The current full version of LEEM is available on Figshare[1].

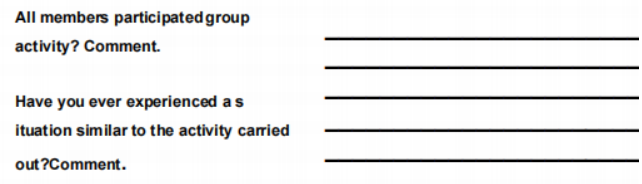

Figure 6. Part of the LEEM post-evaluation.

3.3 Step by step of the application of LEEM

A step-by-step guide has been drawn up to help teachers apply the LEEM, as shown in Figures 7 and 8. The step-by-step is organized into Mandatory Steps (MS) and Optional Steps (OS). MS consists of applying the LEEM checklists; collecting and analyzing the data obtained by checklists; the steps linked to the educational activity with DICTs; and also the step of the checklist that the teacher must answer. OS is composed of the steps of working in pairs/trios or larger groups, as well as working individually.

Step-01 (MS) is the application of the pre-assessment checklist, before explaining the educational activity with DICTs to the learners (Figure 7); then Step-02 (MS) is carried out, analyzing the data obtained from the pre-assessment checklist. If necessary, it is recommended to make adaptations to the educational activity with DICTs; if working individually, Step-03 (OS) is carried out, the explanation of the educational activity with DICTs can be carried out after analyzing the data from the pre-assessment checklist; and if working in pairs/trios or larger groups, Step-03 (OS) is carried out, grouping the learners using the data obtained from the pre-assessment checklist and group formation techniques according to the teacher's familiarity; after this, the explanation and start of the educational activity with DICTs take place, which is Step-04 (MS).

The definition of the half of the educational activity for the application of the assessment-during checklist must be carried out by the teacher, characterizing Step-05 (MS); for Step-06 (MS) the teacher is responsible for carrying out continuous monitoring of the educational activity with DICTs; and also, the teacher must answer its during-assessment checklist, being Step-07 (MS); and apply the during-assessment checklist with learner, being Step-08 (MS); after applying the during-assessment, Step-09 (MS), the data from the during-assessment checklists are analyzed to adjust the needs/difficulties of the learners for the other 50% of the educational activity.

Step-10 (MS) represents the continuous monitoring of the educational activity with DICTs to complete the other 50% (Figure 8); if working individually, Step-11 (OS), suggests finalizing the educational activity and applying the post-assessment checklist individually; if working in pairs/trios or larger groups, Step-11 (OS), suggests finalizing the educational activity by applying the post-assessment checklist with the focus group technique so that the learners can respond as a group; Step-12 (MS), the teacher may use the data for this activity or future activities. Thus, it is suggested to analyze the data obtained in the post-assessment and compare it with the data obtained in the pre-assessment, as well as check how the learners performed in this educational activity; Finally, the LEEM application is concluded.

4 Study planning and execution

The feasibility study was carried out as suggested by Shull *et al*. [2004] to check whether the objectives of the proposed technology can be met before it is applied in a real context. Therefore, in this article, we first sought to create a body of knowledge of the LEEM model from the teachers' perspective. This feasibility study can be considered operational [Rosa *et al*., 2020], as it sought to identify teachers' perceptions of the items and sets of questions on the LEEM checklists. The Ethics Committee of the Federal University of Paraná approved the study with the CAAE: 64733822.0.0000.0102.

4.1 Instruments

For this LEEM feasibility study, teachers were first asked to sign an Informed Consent Form (ICF) if they agreed to participate. Subsequently, the researcher briefly presented the LEEM along with its purpose and functionalities in a meeting via MEET[2]. After this, the teachers were instructed to complete the participant characterization forms and LEEM evaluation questionnaires.

For the characterization questionnaire, ten questions were asked. The data obtained from these questions is described in the following subsection. The LEEM evaluation questionnaire was divided into six parts corresponding to: (1) Pre-evaluation; (2) Learner during-evaluation; (3) Teacher during-evaluation; (4) Post-evaluation; (5) The elements of LX and (6) General evaluation of LEEM. The teachers' participation consisted of evaluating the LEEM in its entirety, regardless of who would answer (teacher or learner), and also without applying it in a real context.

It is worth noting that the LEEM evaluation questionnaire contained the three indicators of the Technology Acceptance Model (TAM) [Venkatesh and Bala, 2008], namely (a) Ease

[1] `https://figshare.com/s/188c4b881ed1e4072244`

[2] `https://meet.google.com`

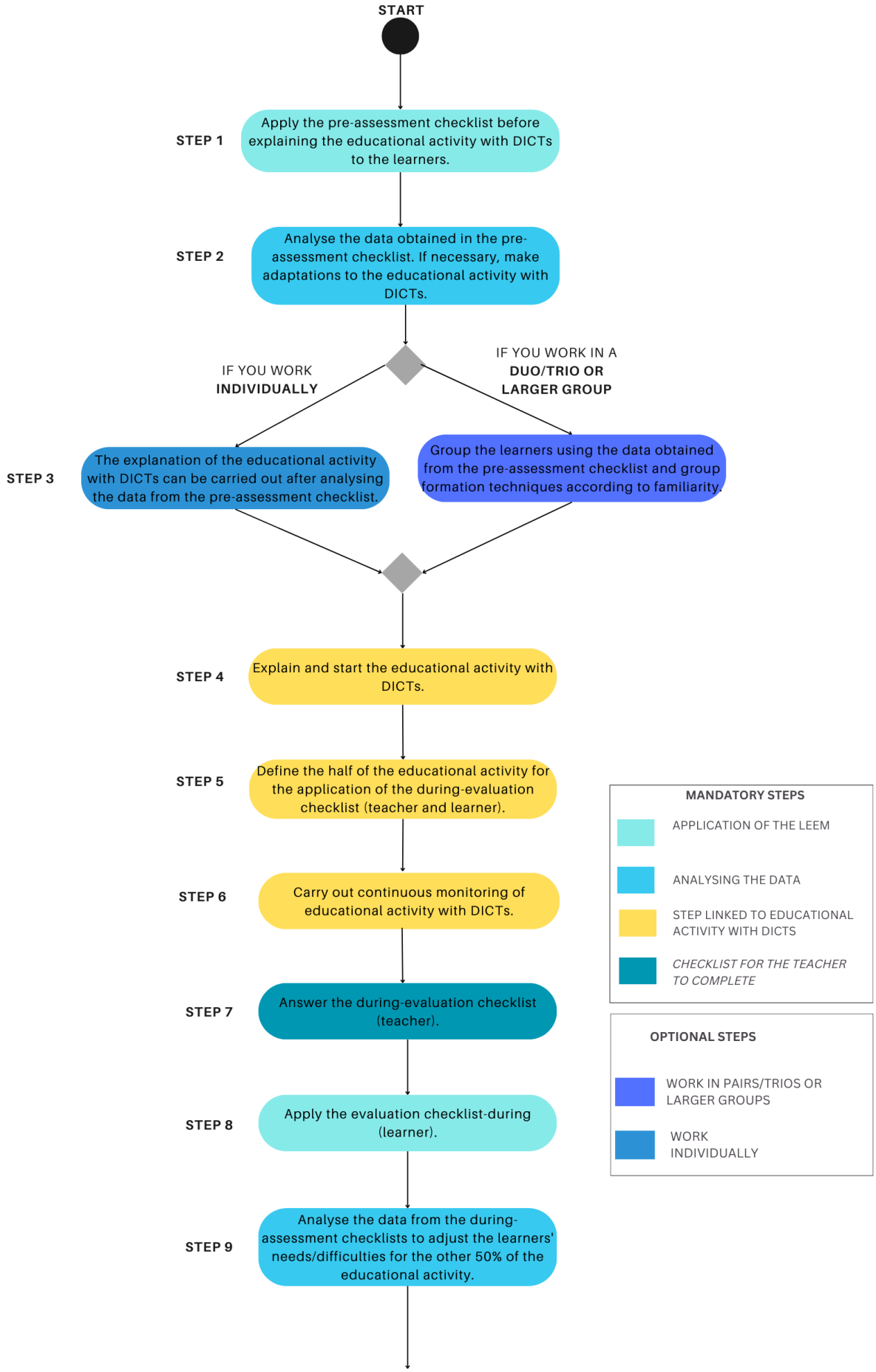

Figure 7. Part of step-by-step of the application of LEEM for the teacher.

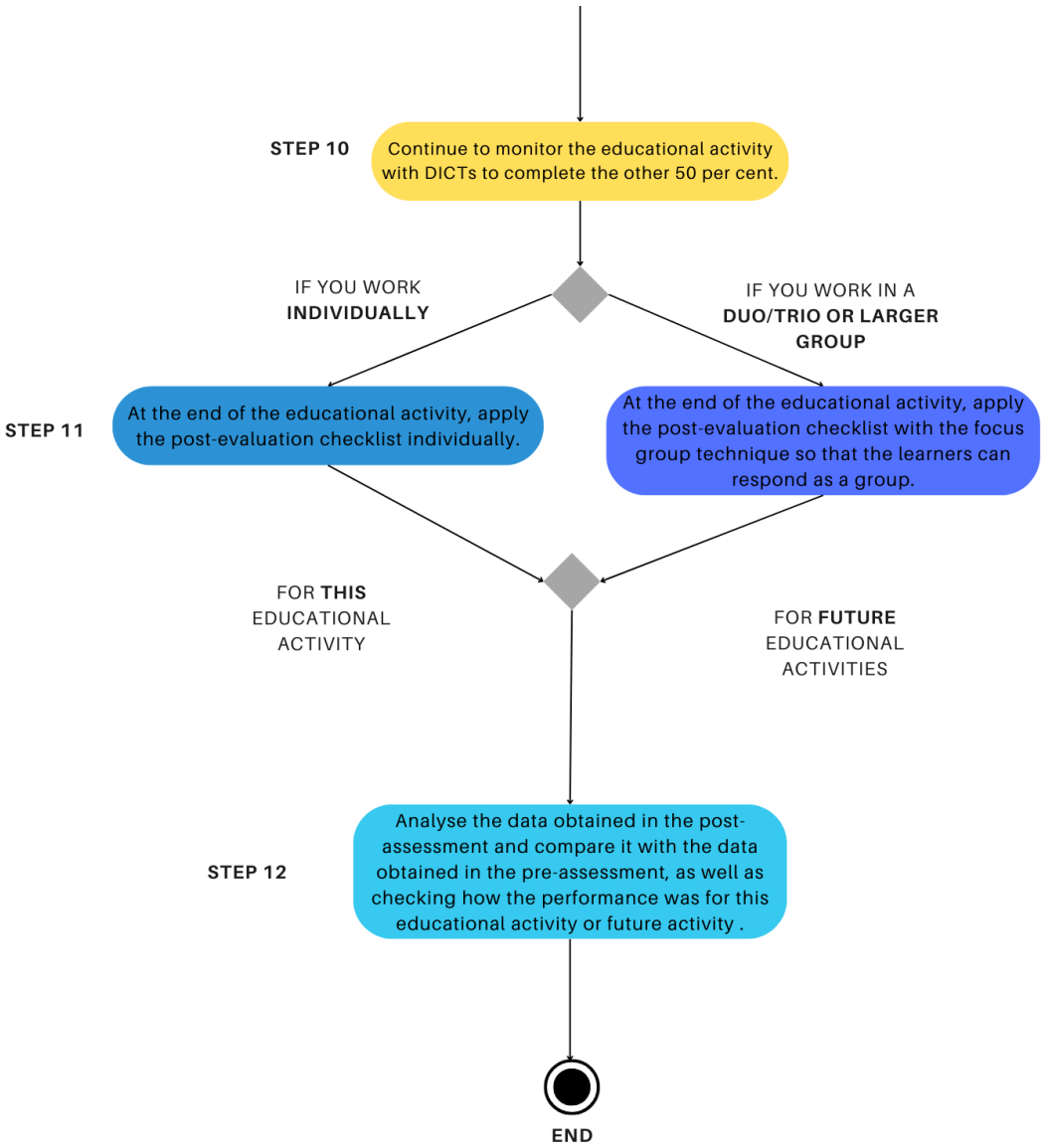

Figure 8. Continued step-by-step of the application of LEEM for the teacher.

of Use, (b) Perceived Usefulness, and (c) Intention of Future Use. In addition, this questionnaire had open questions so that participants could detail their assessment and suggest improvements for the LEEM. All the instruments used in this feasibility study can be found at the following link[3].

4.2 Population and Sample

It was decided that the target audience for this study would be teachers from different levels of education, including primary and secondary, undergraduate and postgraduate. This was due to the knowledge and experience of these participants with their teaching public. In addition, it was decided to select teachers from different levels, as LEEM is generic and needs to be simple to cater to the diversity of learners.

Teachers were invited to take part by contact via social media (Facebook and LinkedIn) and email institutional. During the first contact, the research project was presented, and the teachers who agreed to participate voluntarily were given the guidelines for participating in the study via an online meeting. This meeting took place via the Meet platform. At this meeting, the study's data collection instruments were presented, and they will be presented in Subsection 4.1.

The characterization data of the 19 teachers participating in this feasibility study is presented below. Regarding the gender of the teachers (Figure 9), 63.16% (N = 12) are female and 36.84% (N = 7) are male.

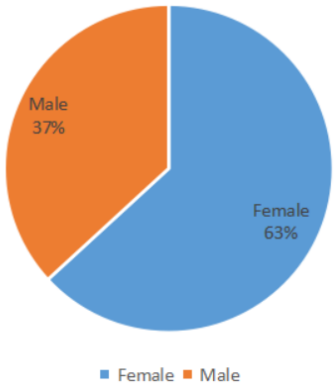

Figure 9. Gender of the teachers.

Regarding the age of the teachers (Figure 10), 42.11% (N = 8) are between 51 and 60 years old, 26.32% (N = 5) are between 31 and 40 years old, 21.05% (N = 4) are between 41 and 50 years old, 5.26% (N = 1) are between 20 and 30 years old and 5.26% (N = 1) are over 61 years old.

Regarding the region where they work as teachers (Figure 11), 63.16% (N = 12) work in the South, 26.32% (N = 5) in the Southeast, 5.26% (N = 1) in the North, and 5.26% (N = 1) in the Northeast of Brazil. The fact that the South region stands out is due to the predominance of contacts the researchers have with teachers in this region.

About the level(s) of education at which these teachers work (Figure 12), 47.37% (N = 9) work at the undergraduate level, 10.53% (N = 2) work at Graduate School (specialization degree), 36.84% (N = 7) work at Elementary Education, 31.58% (N = 6) work at High School, 15.79% (N = 3) work at Middle School, 15.79% (N = 3) work at Graduate School (master's degree) and 10.53% (N = 2) work at Graduate School (PhD degree).

[3] https://figshare.com/s/188c4b881ed1e4072244

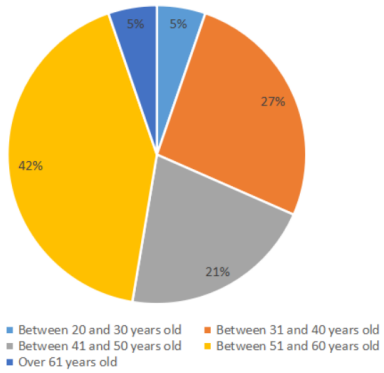

Figure 10. Age of the teachers.

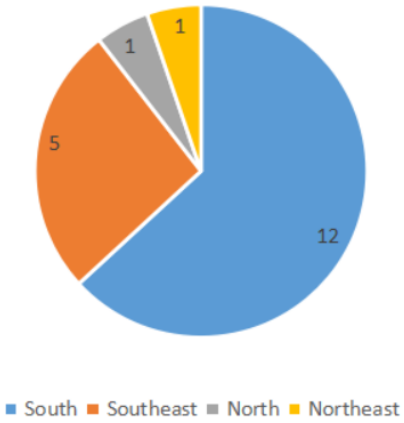

Figure 11. Region where they work as teachers.

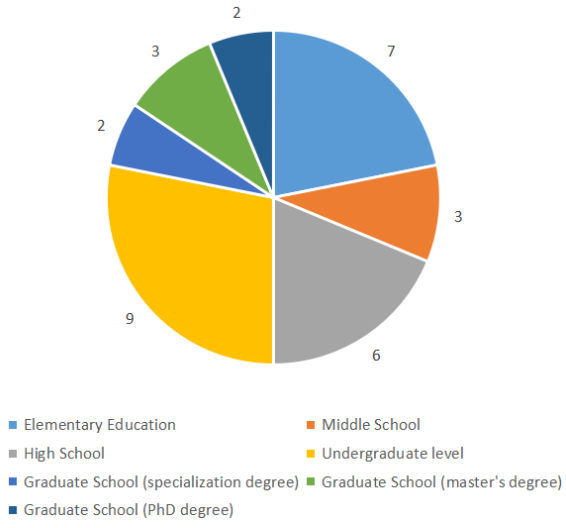

Figure 12. Levels of education at which these teachers work.

About the type of institution where the teachers work (Figure 13), 36.84% (N = 7) are from the municipal system, 36.84% (N = 7) are from the federal system, 21.05% (N = 4) are from the state system and 15.79% (N = 3) are from the private sector, showing that the majority of respondents

work in public education: municipal and federal (N = 14).

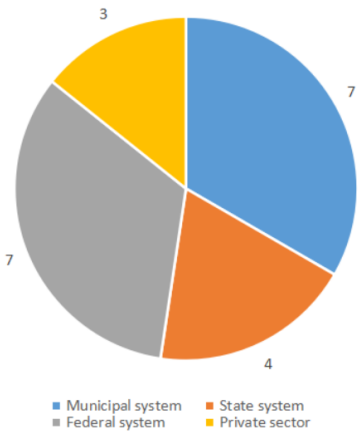

Figure 13. Type of institution where the teachers work.

Regarding the class, grade, or period that the teachers work in, 26% (N = 5) work in the 1st year of Elementary School, 26% (N = 5) work in undergraduate courses between the 1st and 5th periods, 21% (N = 4) work in undergraduate courses between the 6th and 10th periods, 16% (N = 3) work in the 5th year of Elementary School, 16% (N = 3) work in the 2nd and 3rd year of High School, 11% (N = 2) work in the 2nd, 3rd and 4th year of Elementary School, 11% (N = 2) work in the 6th year of Middle School, 11% (N = 2) work in the 1st year of High School and the remaining 5% (N = 1) each work with special education and work in the 7th, 8th and 9th year of Middle School. Regarding the subject(s) the teachers teach, 37% (N = 7) teach Maths, 26% (N = 5) teach Portuguese, 21% (N = 4) teach Science, 16% (N = 3) teach Computer Science, 16% (N = 3) teach Agronomy, 11% (N = 2) teach Arts and the rest correspond to 5% (N = 1) each: Physics, Physical Education, Philosophy, Geography, Sociology, Basic Sanitation, Technical Drawing and Materials Selection.

Regarding the use of DICTs to support teaching and learning (Figure 14), 84% (N = 16) use DICTs as support, 11% (N = 2) do not use DICTs as support, and 5% (N = 1) sometimes use DICTs as support. As for how long the teacher has been using the DICTs for teaching/learning, 47% (N = 9) have been using them for more than two years, 16% (N = 3) have been using them for more than ten years, 16% (N = 3) have been using them for more than a year, 11% (N = 2) did not say how long they had been using them, and the rest correspond to 5% (N = 1) each, namely: using them for less than two months and using them for between six months and a year.

About the DICTs that teachers usually use (Figure 15), 84% (N = 16) use Google Drive, 21% (N = 4) use Quizziz, 16% (N = 3) use Kahoot! and the others correspond to 5% (N = 1) each: Google Class, Construct, Geogebra, Moodle, Youtube, MindMeister, Canva, Padlet, Google Earth, Winplot, OneNote (Microsoft), Jamboard, Openboard, Xournal++, Scratch, Matlab and did not inform.

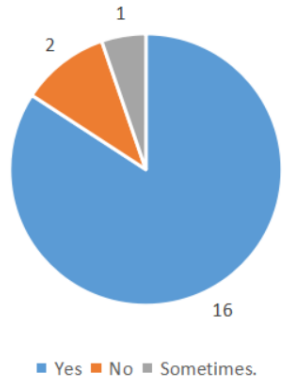

Figure 14. Use of DICTs to support teaching and learning.

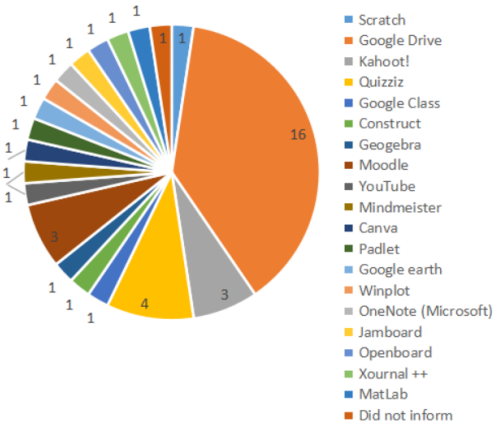

Figure 15. The DICTs that teachers usually use.

4.3 Preparation and execution

Before the study, an individual meeting with each participant was held, in which the ICF, the characterization form, the LEEM checklists, and the LEEM evaluation questionnaire were presented. This presentation lasted around 30 minutes. Participants were given one week to complete the LEEM evaluation questionnaire.

The study was conducted online because we were looking for participants from different levels of education and regions. We held a meeting with each teacher via Meet. After these meetings with the participants, instructions were sent by e-mail containing the link to the data collection instruments. The participants carried out the activities, and their answers were stored on the Survey Monkey[4] platform. Participants undertook the following stages in the study: (1) participants who agreed to take part in the study signed the ICF; (2) they completed the characterization questionnaire; (3) they read and evaluated the LEEM; and (4) they answered the evaluation questionnaire. The average time taken to answer the LEEM evaluation questionnaire was 43 minutes.

4.4 Data Analysis

The data obtained from the LEEM evaluation questionnaire was analyzed quantitatively and qualitatively. Before analysis, the data was cleaned, coded, and then organized. The quantitative data was analyzed using descriptive statistics,

[4] `https://www.surveymonkey.com`

looking at the participants' responses about the TAM indicators. The qualitative data was analyzed using the Grounded Theory (GT) method, defined by Corbin *et al*. [1990]. GT has three stages in the coding process: open coding (1), axial coding (2), and selective coding (3). The first stage is open coding, in which the data is coded according to the answer given by each participant to each question. Subsequently, in axial coding, the codes were grouped according to their properties and related to each other, thus forming categories that represent their characteristics. Selective coding was not carried out as the intention was not yet to create a theory. This is because the open and axial coding stages were sufficient to understand the feasibility of using the LEEM and make the necessary adjustments to its content based on the teachers' perceptions.

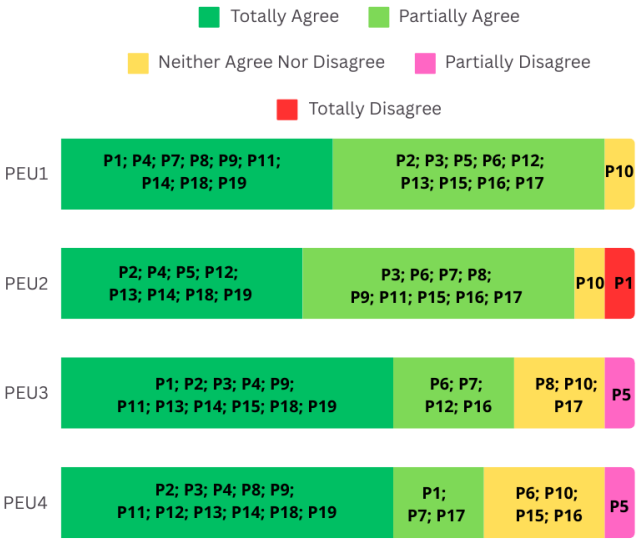

Figure 16. Ease of Use of LEEM.

5 Quantitative Analysis

In the LEEM evaluation questionnaire, the answers were obtained for the three indicators of the TAM3 model, which assessed acceptance of the LEEM. The answers were obtained using a five-point Likert scale (totally agree, partially agree, neither agree nor disagree, partially disagree, and totally disagree). The results of the TAM3 are described below.

The Ease of Use indicator defines the degree to which a person believes that using a specific technology would be easy using the following questions: (PEU1) My interaction with LEEM was clear and understandable, (PEU2) Interacting with LEEM does not require too much of my mental effort, (PEU3) I find LEEM easy to use and (PEU4) I find it easy to use LEEM to assess the learner's experience of using DICTs. Figure 16 shows the participants' perception of the Ease of Use indicator. The graph's vertical axis represents the affirmative of the indicator, and the horizontal axis refers to the degree of acceptance of the participants. Codes have been added to the bars to represent the participants (P1 to P19) and their respective evaluations. In Figure 16, it can be seen that for P1, interacting with LEEM required mental effort (PEU2). Because P1 works with elementary school learners, the LEEM is believed to need further adjustments to make it more suitable for this target audience. It was also identified that P5 partially disagreed with sentences PEU3 and PEU4, demonstrating difficulty using and evaluating LX using DICTs. It is believed that P5 experienced difficulties due to the length of the LEEM, which may have caused some discomfort when evaluating it. Overall, even in the face of some difficulties, most of the participants were optimistic about the acceptance of LEEM, demonstrating that it was easy to use.

The Perceived Usefulness indicator defines the degree to which a person believes that technology can improve their performance through the following questions: (PU1) Using LEEM can improve the performance of the assessment of the learner experience in the use of DICT, (PU2) Using LEEM can allow me to increase productivity in the evaluation of learner in the use of DICT, (PU3) Using LEEM has increased my effectiveness in the evaluation of the learner experience in the use of DICT and (PU4) I find LEEM useful in supporting the assessment of the learner experience in the use of DICT. Figure 17 shows the participants' perception of the Perceived Usefulness indicator. In general, this indicator shows a higher level of agreement in statements PU2 and PU4, which indicates that LEEM can be more productive in evaluating and useful in assessing LX in the use of DICT. In statement PU1, participants P5 and P17 partially disagreed, indicating that using LEEM may not improve the performance of the evaluation of LX in the use of DICTs very much, demonstrating a need to improve LEEM in this regard. It is believed that P5 and P17 felt some difficulty when evaluating LEEM, mainly because, in this first study, they had not had the opportunity to apply it. For statement PU2, participant P5 partially disagreed, indicating that using LEEM did not increase productivity in assessing learners' use of DICTs. It is believed that P5 made this assessment because he had not applied LEEM in a real scenario. For statement PU3, participant P6 partially disagreed about LEEM allowing for increased effectiveness in assessing LX using DICTs. It is believed that because P6 reports that he has a busy day-to-day life, he suggests using a tool that would need to be accessible on smartphones, for example, to increase effectiveness in assessing LX in the use of DICTs.

Finally, the Intention of Future Use indicator defines the degree to which a person believes they would use the technology in the future through the following questions: (IU1) Assuming I have access to the LEEM, I intend to use it, and (IU2) Considering that I have access to the LEEM, I anticipate that I will use it at other times. Figure 18 shows the teachers' feedback on the Intention to Use in the Future. Figure 18 shows that most participants are interested in using LEEM. In this indicator, only P6 and P17 partially disagreed with these sentences, showing doubts about the future use of LEEM. Thus, it is believed that if adjustments are made to improve the LEEM, these teachers will be able to use it.

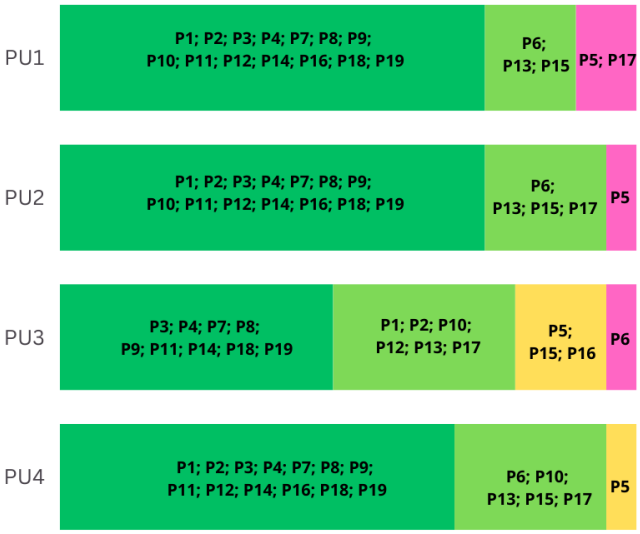

Figure 17. Perceived Usefulness of LEEM.

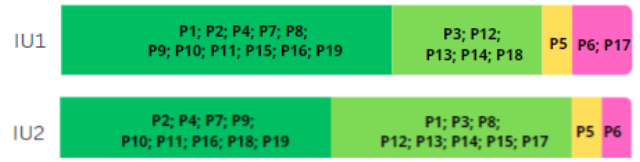

Figure 18. Intention of Future Use of LEEM.

6 Qualitative Analysis and Improvement

The qualitative data obtained from the open questions in the LEEM evaluation questionnaire was analyzed using the first two stages of the GT [Corbin *et al*., 1990] method. In the first stage of the GT (open coding), codes were created according to participant feedback. Next, the codes were grouped according to their properties, creating concepts that represent categories. Finally, these codes were related (axial coding - second stage). The qualitative analysis in this study aims to evaluate and improve the LEEM items and sentences through the participants' suggestions. Atlas.ti version 7[5] was used for this analysis. The categories and their respective codes identified in this analysis are presented in the following subsections.

6.1 Difficulties in understanding

This category represents the difficulties in understanding the LEEM items. The suggestions were to improve the visualization and understanding of the items, such as the word pairs used (see the 1st quote from P5), clarifying whether the advantage is for the teacher or the learner (see the quote from P17), the quality of the SAM scale images (see the 2nd quote from P5), clarifying whether the DICTs are effective in the teaching and learning processes or in carrying out the proposed activity (see the quote from P2), changing terms for better understanding (see the 3nd and 4rd quotes from P5). Other suggestions for improvement include clarifying the preference for using DICTs for educational activity (see the 6th quotes from P5), and clarifying the preference for working collectively and/or individually (see the 7th quote from P5).

"Challenging is always motivating?" (P5)

"Advantage about what? For the teacher or the learner?" (P17)

"The drawing is small, making it difficult to identify the expression" (P5)

"Where are they effective, in teaching/learning or the activity" (P2)

"Do you remember or did you contribute?" (P5)

'What does it mean to be authentic?" (P5)

"Did you prefer it over not using it?" (P5)

"It's difficult to determine whether it was exclusively due to one factor, such as the use of DICTs unless the teaching was only carried out in this way" (P5)

"Another point: what if they didn't like working together? Will the teacher switch to individual activities or give them the option?" (P5)

For P5's 1st quote, the word pair in item 7 of the pre-evaluation checklist was changed from "challenging (versus) demotivating" to "motivating (versus) demotivating". For the quote from P17, item 1 of the during-evaluation checklist (teacher) was changed from "It was advantageous to carry out the activity with the DICTs" to "I realized that it was advantageous for the learners to carry out the activity with the DICTs". Based on this quote, we added the phrase "I realized that for the learners (...)" at the beginning of each item in this checklist. For the 2nd quote from P5, the image quality of the SAM scale used in the during-evaluation (learner) checklist was improved. For the quote from P2, the instructions for the during-evaluation checklists (learners and teacher) were supplemented by adding the following phrase "The checklist should be answered halfway through the educational activity. For example, this checklist should be completed at the end of the first week of a two-week activity. The sentences on this checklist should be answered considering the activity in question".

For P5's 3nd quote, item 3 of the focus group for post-evaluation was changed from "Do you remember what your contribution to the activity was?" to "What was your contribution to the activity? Please comment". For P5's 4rd quote, item 4 of the focus group was changed from "Do you consider yourself authentic in carrying out the project activities?" to "Did you contribute your ideas (were you authentic) in carrying out the project activities? Please comment".

No changes were made to P5's 5th quote, as the suggestion is directed at item 8 of the assessment-during (learner) checklist, which is "Did you prefer to use DICTs for the activity?", so it is not clear what the suggestion to change is. The 6th quote from P5 was not made, because to use LEEM it is compulsory to use DICTs to carry out the educational activity and also, the aim is to evaluate the learner's experience during an educational activity only in this way.

No changes were made to P5's 7th quote because item 1 assesses the learner's preference for working individually or

[5] https://atlasti.com/

collaboratively for the LEEM pre-assessment stage. This allows the teacher to assess the learner's preference and even change the planned activity if they feel the need. It is up to the teacher to define how the educational activities in their class will be configured, whether individually or group. LEEM only provides a way for the teacher to know the learner's preference. Therefore, we decided to keep this item as it was.

6.2 Teachers' perception of possible doubts that learners might have

This category represents the inquiries about LEEM items that teachers had to improve learner understanding. The participants suggested bringing up the meaning of some terms (see the 1st quote from P17), improving the writing of some sentences to make them simpler and more understandable (see the quote from P5 and the 2nd from P17), checking for biased questions (see the 3rd quote from P17), analyze the focus (see the 2nd quote from P5), and analyze the term used (see quote from P15).

"If you understand what is effective, OK. If not, it needs to be defined" (P17)

"How will the learner know this?" (P5)

"What needs?" (P17)

"Will anyone answer no?" (P17)

"Is it only with the DICTs or do you have the teacher/monitor/other material?" (P5)

"Did you like it or did you find it easy? " (P15)

Through P17's 1st quote, a change was made to the during-evaluation checklists (learners and teacher). In item 12 of the during-evaluation checklist (learner), the item "I found the DICTs effective" was changed to "I found the DICTs sufficient (effective) to complete the activity". In the during-evaluation checklist (from the teacher), item 7 was changed from "The learner(s) found the DICTs effective" to "I realized that the learner(s) found the DICTs sufficient (effective) to complete the activity". In the same way, the teacher believes that the word "efficient" may be difficult for learners at the elementary school level to understand. Therefore, item 13 of the during-evaluation checklist (learner) was changed from "I found the DICTs efficient" to "I found the DICTs adequate (efficient) to carry out the activity within the allotted time" and item 6 of the during-evaluation checklist (teacher) was changed from "I found the DICTs efficient" to "I found the DICTs adequate (efficient) to carry out the activity within the allotted time".

For the quote from P5, item 15 of the during-evaluation checklist (learner) was changed from "Did you feel that the DICTs met your needs?" to "I felt that the DICTs met my needs to carry out the activity". For the 2nd quote from P17, a change was made to item 9 of the during-evaluation checklist (from the teacher), which was changed from "Did the DICTs meet the needs of the learner(s)" to "I felt that the DICTs met the needs of the learner(s) when carrying out the activity". For P17's 3rd quote, item 6 of the focus group for post-evaluation was changed from "Did you feel responsible while carrying out the activity?" to "Did you feel responsible/involved while carrying out the activity? Please comment".

For P5's 2nd quote, in which the participant questions the focus of items 17 and 18 of the during-assessment checklists (learner), it is justified that the focus is always on the DICTs and that this is clear in the instructions of the LEEM checklists. For P15's quote, no changes were made to item 2 of the during-assessment checklist (teacher). This is because this item analyses whether the learners liked looking for extra material and/or even additional information. The aim is to identify how the learners reacted when they did this. Therefore, the term used is "liked it". Also, as the LX element assessed in this item is Desirability, we have kept the verb liked, as it makes more sense in this context.

6.3 Negative points

This category represents the negative points of the LEEM identified by the teachers. In this sense, one of the participants believes that the LEEM may not obtain the full veracity of the learner's answers (see the quote from P13), and another participant was concerned about the size of the LEEM (see the 1st quote from P6). In addition, one participant found the LEEM labor-intensive to apply (see the 2nd quote from P6). Another participant reported that the post-assessment checklist lacked an answer option (see quote from P5).

"This type of model doesn't give us the full picture of the learner's real situation about the use of DICTs in the teaching and learning process because we don't have control over the veracity of the learner's answers." (P13)

"Long questionnaires can discourage teachers from answering other questionnaires." (P6)

"I would hardly use a labor-intensive tool to assess my learners' learning performance." (P6)

"An answer option was missing from all the questions in this section" (P5)

About Q13, as with any type of evaluation carried out with learners and teachers, the reliability and veracity of the information obtained depends directly on who is answering. To reduce the chance of the participant not answering exactly about their experience, we tried to provide a formative evaluation in LEEM. Formative assessment is an ongoing method that helps teachers monitor learners' progress and identify any challenges they face as they learn [Villas Boas, 2006].

About P6's 1st and 2nd quotes, the LEEM has three stages, three checklists, and a set of sentences. The only checklist the teacher is responsible for answering is the assessment-during checklist (by the teacher). For the other checklists, the teacher will guide and lead the learners in completing them. LEEM was set up in checklist format precisely to make it easier to fill in the checklists more quickly and to serve as a learner learning diary. In addition, the idea was to give this responsibility to the learners so that the teacher would have more time to support them and monitor their progress. Even so, if the teacher doesn't want to carry out the during-assessment stage (the teacher's), it is recommended that the LEEM be partially applied, i.e., only the checklists for the learners. It's also worth noting that the LEEM checklists are not completed all at once but at different times to minimize fatigue.

Regarding P5's quote, the post-evaluation checklist brings questions that can be used in a focus group. It is therefore

hoped that the focus group discussion will be held to obtain the information and describe it in the items on the post-evaluation checklist.

6.4 Redundancy of items

This category presents redundancies identified in LEEM items, such as redundancy on the comfortability assessed in the during-evaluation checklist (learner) (see the 1st quote P15), redundancy when comparing the during-evaluation checklist (learner) with the during-evaluation checklist (teacher) (see the 2nd quote P15), and redundancy on the comfortability assessed in the during-evaluation checklist (teacher) (see the quote P6).

"Redundant to the above question" (P15)

"Isn't it redundant with question 2?" (P15)

"Some questions seem redundant and of the same scope, comfortable and at ease, for example" (P6)

About the first quote from P15, the reason is that the three items correspond to the Comfortability element, which is being analyzed in different spheres. For example, item 20 analyses whether the learner felt comfortable accessing the DICTs, item 21 whether the learner felt comfortable accessing the support materials, and item 22 whether the learner felt at ease in the learning environment when using the DICTs. Therefore, different aspects are being analyzed in each item of this element. It was, therefore, decided not to make any changes to these items.

For the 2nd quote from P15, the idea is to compare the learners' checklists with the teacher's, thus evaluating the experience from both points of view. It should be emphasized that, in a real context, these checklists will be answered simultaneously by different people. For P6's quote, it is justified that the two items cited are present in the Comfortability element, and analyze different spheres of this element. Therefore, no changes were made either.

6.5 Positive points

This category presents the LEEM's positive points, such as the fact that it is good (see a quote from P1), that it supports the assessment process (see a quote from P2), that it is practical and objective (see a quote from P12), that it is an interesting model (see a quote from P14) and that it is complete and suitable for its purpose (see a quote from P16).

"I think it's very good" (P1)

"I believe that LEEM will help a lot in the assessment processes" (P2)

"It's practical and objective" (P12)

"It's an interesting model" (P14)

"It seemed to me to be a very complete and appropriate tool for the type of analysis that is proposed" (P16)

It is worth emphasizing that LEEM will allow LX to be assessed, taking into account the diversity and learning preferences of the learners, as well as allowing this to be recorded at different times during the experience of educational activities. Thus, it is believed that it will be possible to contemplate and evaluate various aspects of LX.

6.6 Indicators for Future Use

This category refers to suggestions for the future use of LEEM, such as use in a real context (see a quote from P2), dependence on changes and improvements to using it in the future (see a quote from P5), and use in an academic research context (see a quote from P6).

"I would like to see it working, in practice" (P2)

"It depends on the changes, the extension, and the inclusion of other aspects of teaching in the LEEM" (P5)

"If I'm doing an academic research activity, then I would use it" (P6)

Regarding the quote from P2, it is reported that the study of LEEM in a real context is already planned and will be carried out soon. About the quote from P5, it should be noted that the LEEM has changed in line with the teachers' requests in this study, generating a new version. Furthermore, further changes may be made once the study has been carried out in a real context. About the quote from P6, it is worth emphasizing that the LEEM is a generic model, and it is up to the teacher to decide which audience to apply it to.

6.7 Improvement suggestions

6.7.1 General improvement suggestions for LEEM

The suggestions for improvement were to change the order of the word pairs to standardize them from positive to negative (see a quote from P13), add a description of the extremes on the SAM scale and Likert scale (see the 1st and 2nd quotes from P2), adjust the number of questions for the same item (see the 1st and 2nd quotes from P17) and change the term used (see quote from P5).

"I think in ascending order from left to right" (P13)

"The learner will have doubts when answering the question" (P2)

"The same goes for the other questions with a Likert scale" (P2)

"Question p 3 answers" (P17)

"Which of the two?" (P17)

"Once again, very general" (P5)

By quoting P13, the order of all the pairs of words was changed to maintain a pattern. For example, item 2 of the pre-assessment checklist was changed from "reactive (versus) proactive" to "proactive (versus) reactive", leaving the word pairs ordered from positive to negative. For the 1st quote from P2, the description of the extremes on the SAM scale was added, being Totally Sad and Totally Happy in the evaluation-during checklist (learner). For the 2nd quote, the description on the Likert scale was added, with Strongly Disagree and Strongly Agree on the evaluation-during checklists (learner and teacher).

For the 1st quote from P17, a change was made to item 9 of the post-evaluation for the focus group from "Did the activity broaden your horizons? Did it create new challenges? Or did you remain in your comfort zone?" to "Did the activity generate new challenges or remain in your comfort zone? Please comment". For P17's 2nd quote, item 2 of the post-evaluation for the focus group was changed from "Did you find the activity interesting? Surprising?' to 'Did you find

the activity interesting, stimulating, or thought-provoking? Please comment".

For the quote from P5, item 14 of the during-evaluation checklist (learner) was changed from "Did you find the DICTs adaptable to your context?" to "Did you find that the DICTs applied (adapted) well to your context?". For this same quote, item 8 of the during-evaluation checklist (teacher) was changed from "Did the learner(s) find the DICTs adaptable to their context?" to "I realized that the learner(s) found that the DICTs applied (adapted) well to their context".

6.7.2 Suggested improvements to the pre-evaluation items

The suggestions were to change the word pairs in item 1 to team/group (see quote P5), item 2 to efficiency and response (see 1st quote from P1), item 4 to persuaded and motivated or pessimistic and optimistic (see quote P14), item 5 to favor learning or have no impact on learning (see quote P15), item 8 for discouraging (see the 1st quote from P16), item 9 for adapting the language to the age group (see the quote from P7), item 10 for shy or introverted (see the 2nd quote from P16) and item 11 for true and uncertain (see the 2nd quote from P1).

"From collaborative to team/group" (P5)

"Efficiency and responsiveness" (P1)

"Persuaded - Motivated or Pessimistic - Optimistic" (P14)

"Favour learning or "have no impact" (P15)

"Swap "monotonous" for "discouraging"" (P16)

"Appropriate language for the age group" (P7)

"Replace "shy" with "shy" or "introverted"" (P16)

"True and uncertain" (P1)

For P5's quote, the word pair in item 1 was changed from "individually (versus) collaboratively" to "individually (versus) group". For P1's 1st quote, item 2 was changed from "proactive (versus) reactive" to "acting (versus) reacting". For the same quote, the question was changed from "I consider myself:" to "I prefer:". For P14's quote, item 4 was changed from "persuaded (versus) optimistic" to "pessimistic (versus) optimistic".

For P15's quote, item 5 was changed from "contribute (versus) harm" to "contribute (versus) not impact". For the 1st quote from P16, item 8 was changed from "encouraging (versus) monotonous" to "encouraging (versus) discouraging". For P7's quote, item 9 was changed from "contextualize (versus) distract" to "concentrate (versus) distract".

For P16's 2nd quote, item 10 was changed from "participative (versus) sheepish" to "participative (versus) shy". For P1's 2nd quote, item 11 was changed from "authentic (versus) inauthentic" to "true (versus) uncertain". For the same quote, the question was changed from "I consider myself" to "I consider my ideas".

6.7.3 Suggestion for improvements to the during-evaluation items (learner)

The suggestions were to change item 1 to how the learner feels emotionally (see the quote from P10), item 7 to how the learner feels when making improvements to the activity (see the 1st quote from P5), item 16 to extract the feeling referring to the use of DICTs (see the 1st quote from P15), item 17 to change to the term "by means" (see the 2nd quote from P5) and item 19 to improve understanding of the content when interacting with group mates (see the 2nd quote from P15).

"How do you feel emotionally today?" (P10)

"How did you feel when you had to make the improvements requested by the teacher?" (P5)

"Learner can express how they feel about learning, but in the question, they won't be able to extract the feeling when using DICT" (P15)

"Replace through with "by means"" (P5)

"Were you able to improve your understanding of the content by interacting with your group mates?" (P15)

For the quote from P10, item 1 was changed from "How do you feel today?" to "How do you feel emotionally today?". For the 1st quote from P5, item 7 has been changed from "How did you feel about making improvements to the activity requested by the teacher?" to "How did you feel about having to make the improvements requested by the teacher?". For P15's 1st quote, item 16 was changed from "Were you able to learn the content in its entirety?" to "Did you feel that you learned all the content of the activity?".

For the 2nd quote from P5, item 17 was changed from "Were you able to understand the content through the suggestions for improvements in the activity that the teacher suggested?" to "Were you able to understand the content by means the suggestions for improvements given by the teacher in the activity?". For P15's 2nd quote, item 19 was changed from "Were you able to understand the content when you talked to your group mates about it?" to "Were you able to improve your understanding of the content when you interacted with your group mates?".

6.7.4 Suggestion for improvements to the during-evaluation items (teacher)

Suggestions for improvement were to change item 11 to assess whether learners were able to learn the content satisfactorily (see quote from P7) and item 13 to talk (see quote from P4).

"The learners managed to learn the content satisfactorily" (P7)

"To talk" (P4)

For the quote from P7, item 11 was changed from "The learner(s) understood the content through the suggestions for improvement requested by you" to "The learner(s) understood the content satisfactorily through the suggestions for improvement requested by you". For P4's quote, item 13 was changed from "The learner(s) understood the content when they talks to colleagues in the group about the subject" to "The learner(s) understood the content when they talked to colleagues in the group about the subject".

6.7.5 Suggested improvements to the post-evaluation items

Suggestions for improvement were for item 1 to make it clear that all members of the group participated actively (see a

quote from P16), item 2 to make the intention clearer (see a quote from P13), and item 10 whether they learned from the activity carried out (see quote from P10).

"Did everyone in the group actively participate? " (P16)

"If the intention is to find out what situation was experienced, I suggest rephrasing the question along the lines of question 5" (P13)

"You learnt from the activity carried out" (P10)

For the quote from P16, item 1 was changed from "Did your group participate actively?" to "Did everyone in the group participate actively?". For the quote from P13, item 2 was changed from "Does the activity developed relate to a situation you have experienced?" to "Have you ever experienced a situation similar to the activity developed?". For the quote from P10, item 10 was changed from "Do you think that the activity made it easier for you to understand potential situations or intentions in your life?" to "Do you think that the learning achieved in the activity carried out will help you in situations in your life?".

6.7.6 Suggested improvements to LX elements

Suggestions for improvement were to replace the Authentic element with Desirability for the pre-evaluation checklist (see the 1st quote from P15) and add the elements Persistence, Interest, Effort, and Engagement for the during-evaluation checklists (learner and teacher) (see the 2nd quote from P15).

"I have doubts about the Authentic Element, as I don't have an exact understanding of its meaning. I would substitute and add: Desirability" (P15)

"I would add: Persistence, Interest, Effort, Engagement (the latter is fundamental in gamification, for example)" (P15)

P15's 1st quote replaced the Preference element with the Desirability element, as the Desirability element also assesses Preference. The Authentic element refers to learners' authenticity when doing activities with DICTs. This element has been retained, as it is hoped to obtain information on how the learner self-assesses their activities.

For the 2nd quote from P15, the Persistence element was added, as suggested. To this end, the item "Persisted (did not give up) in the face of the obstacles that arose during the activity" was added to the during-evaluation checklist (learner), and the item "I realized that the learner(s) persisted (did not give up) in the face of the obstacles that arose during the activity" was added to the during-evaluation checklist (teacher). The other suggested elements, Interest, Effort, and Engagement, were not added because they are associated with other elements such as Usability and Desirability.

7 Benchmark

Due to the various technologies to evaluating LX identified in SMS [dos Santos *et al*., 2022], a benchmark also was carried out to characterize and analyze these LX evaluation technologies. A benchmark is a method used to measure, compare, define best practices, implement, and improve a product or software [Anand and Kodali, 2008]. We considered evaluation technologies such as questionnaires and models. These technologies have evaluated LX in different contexts, so nine evaluation technologies were considered for this benchmark, as shown in Table 1, codified as II to X. These evaluation technologies were selected to make comparisons with the LEEM model (codified as I), which was presented in Section 3.

Ten types of evaluation technologies were cataloged, and some criteria were defined to analyze the LX evaluation technologies. The criteria were named Contribution, Elements of LX, Types of LX evaluation, and Evaluation steps, present in Table 2. Looking at the ten evaluation technologies, we noticed some variations in the sentences, the number of sentences, and the scales used. Therefore, we felt the need to catalog these results. We also looked at whether there is any support for using the LX evaluation technology and, if so, what that support is. The public that evaluated each technological technology was cataloged. It was also checked whether the evaluation technologies had automated versions that could facilitate collecting and analyzing responses. The criteria mentioned are shown in Tables 3 and 4.

Regarding the types of technologies (a) present in the evaluation technologies, there were two model type and the eight questionnaire type (Table 2). It is believed that the questionnaire was the most used due to its practicality of application, ease of collecting data, and the possibility of maintaining the anonymity of the participants. However, if a more complementary LX assessment is desired, the model is the type of technology recommended, because it provides questions that guide the LX assessment, including the criteria that should be assessed [Ruiz and Snoeck, 2018].

The number of LX elements identified (b) in the technologies varies between one (III, VI, VII, and VIII) and 11 (I) LX elements. These elements directly imply the evaluation of the components that guide the LX evaluation process [Huang *et al*., 2019]. Thus, the greater the number of LX elements, the more aspects will be assessed, resulting in a more comprehensive and robust LX assessment technology. For example, the LEEM model (I) uses eleven different LX elements. However, when the objective is a faster and less in-depth LX evaluation, it is believed that fewer LX elements may be useful. The other evaluation technologies present only one element of LX, being Learning, Motivation, Affective learning, and Attitudes to learning.

About the types of LX evaluation found in the technologies (c), it was noted that there is a variation between one (III and V) and six (I) types of evaluating LX. These types of evaluation are directly related to the LX elements, as they guide the LX evaluation process. It is, therefore, believed that the greater the number of LX assessment types, the proportionally greater the number of LX elements. This allows for a more holistic LX assessment, as it considers as many aspects as possible that are present in the scenario [Huang *et al*., 2019]. However, if the objective is a quicker assessment, without considering most of the aspects, fewer assessment technologies may be employed.

The LX evaluation occurred at different times during an educational activity (d). In three technologies identified, LX was only assessed after an educational activity had been completed (II, III, and VIII). In four technologies, the evaluation of LX occurs before and after an educational activity ends

Table 1. Authors who evaluate LX based on SMS.

Number	Evaluation technologies
I.	LEEM model [dos Santos *et al*., 2023]
II.	Questionnaire regarding experiences with e-texts [Chapman *et al*., 2016]
III.	Post-intervention questionnaire about learners' perceptions of the Communication Idol contest [Dune *et al*., 2016]
IV.	Learner satisfaction questionnaire and post-test [El Mawas *et al*., 2020]
V.	Questionnaire to measure learner experiences [Kawano *et al*., 2019]
VI.	Semantic differential questionnaire to picture of learners' experiences [Lykke *et al*., 2015]
VII.	Evaluation model of the INBECOM project [Nygren *et al*., 2019]
VIII.	Factor analysis questionnaire [Reyna and Meier, 2018]
IX.	Pre and post-test questionnaire of Stanley and Zhang [2018]
X.	Questionnaire to assess attitudes and negotiations in learning [Yeh and Chen, 2019]

Table 2. Result of the benchmark.

Number	a. Contribution	b. Quantities of LX elements	c. Quantities of evaluation types	d. Evaluation steps
I.	Model	11	6	Before, during and after
II.	Questionnaire	4	3	After
III.	Questionnaire	1	1	After
IV.	Questionnaire	3	2	Before and after
V.	Questionnaire	3	1	Before and after
VI.	Questionnaire	1	5	Before, during and after
VII.	Model	1	4	Before, during and after
VIII.	Questionnaire	1	2	After
IX.	Questionnaire	3	2	Before and after
X.	Questionnaire	4	5	Before and after

(IV, V, IX, and X). In the other technologies, the LX assessment occurs before, during, and after an educational activity (I, VI, and VII). Evaluating LX at different times during an educational activity is important to allow for continuous monitoring. It also makes it possible to record LX throughout the educational activity [Lykke *et al*., 2015]. Evaluating LX at just one moment in the educational activity can directly affect the results obtained because it is not possible to make the necessary improvements when encountering a problem halfway through the activity that could impact the end of the LX.

About the public that used LX assessment technology (e), 70% (N = 7) were higher education learners: I, II, III, VI, VIII, IX, and X (Table 3). There are LX technologies that secondary school learners, such as IV and VII technologies, assessed. There are other roles, such as elementary, high,h and higher school learners (I) and software engineers (V).

It was also investigated whether the assessment technology offers support (f and g). For example, Deodato [2015] argues that a guide or step-by-step approach is important for a more comprehensive evaluation and to enable better practices. Thus, of the ten technologies that evaluate LX, only three (I, VII, and X) have this type of support, allowing for a more rigorous and complete evaluation. Thus, model LEEM (I) in turn offers step-by-step instructions on how to apply it and guidelines with suggestions on how to answer it.

About the scales used in the LX evaluation technologies (h), it was noted that more than one type of scale was used. The majority of technologies, 70% (N = 7), use the 5-point Likert scale (I, II, III, IV, V, IX, and X). The technology uses a 4-point Likert scale (VIII). The LEEM model (I) also uses the 9-point SAM scale. In addition, there are word pair scales with varying points, for example, the LEEM (I) model uses a 5-point scale, and the Semantic differential questionnaire (VI) to picture learners' experiences uses a 9-point scale.

The number of sentences (i) found varies between five (V) and 63 (I) (Table 4). The number of sentences implies how much information can be provided by the learner. Thus, technologies wishing to extract more information should consider that more questions may be favorable. On the other hand, LX technologies that extract less information should consider that they will have a quicker and less in-depth evaluation. It is believed that the number of sentences directly implies how many elements and types of LX can be evaluated. We investigated how sentences (j) are in the technologies that evaluate the LX. Thus, it was found that the majority use affirmative-type sentences (90% | N = 9) and question-type sentences (90% | N = 9). For example, for the LEEM (I) model, 75 of the 63 sentences are affirmatives, and 11 are questions. We also checked whether the LX evaluation technologies had automated versions that could facilitate collecting and analyzing responses (k). However, we identified that no technology that evaluates LX offers an automated version.

Of the 10 LX evaluation technologies, 70% (N = 7) assessed LX at more than one time during the educational activity: LEEM Model (I), Learner satisfaction questionnaire and post-test (IV), Questionnaire to measure learner experiences (V), Semantic differential questionnaire to picture of learners' experiences (VI), the evaluation model of the INBECOM project (VII), Pre- and post-test questionnaire of Stanley and Zhang [2018] (IX), and Questionnaire to assess attitudes and negotiations in learning X). These seven LX evaluation technologies are presented below.

Table 3. Result of the benchmark.

Number	e. Public who evaluated	f. Is there support?	g. What is the support?	h. What is the scale?
I.	Teachers of elementary, high school and higher education	Yes	Instructions and step by step	Word pairs of 5 point, scale SAM of 9 points, and scale Likert of 5 points
II.	Learners of higher education	No	-	Scale Likert of 5 points
III.	Learners of figher education	No	-	Scale Likert of 5 points
IV.	Learners of secondary education	No	-	Scale Likert of 5 points
V.	Learners of software engineers	No	-	Scale Likert of 5 points
VI.	Learners of higher education	No	-	Word pairs of 9 points
VII.	Learners of secondary education	Yes	Step by step	-
VIII.	Learners of higher education	No	-	Scale Likert of 4 points
IX.	Learners of higher education	No	-	Scale Likert of 5 points
X.	Learners of higher education	Yes	Step by step	Scale Likert of 5 points

Table 4. Result of the benchmark.

Number	i. Number of sentences	j. What are the sentences like?	k. Automated version?
I.	63	Open question and affirmative sentences	No
II.	19	Open question and affirmative sentences	No
III.	18	Open question and affirmative sentences	No
IV.	18	Yes/no question, open questions and affirmative sentences	No
V.	5	Affirmative sentences	No
VI.	12	Opens question	No
VII.	22	Opens question and affirmative sentences	No
VIII.	33	Opens question and affirmative sentences	No
IX.	48	Opens question and affirmative sentences	No
X.	23	Opens question and affirmative sentences	No

The LEEM model (I) that assesses LX before, during, and after the educational activity was presented in Section 3. The Learner satisfaction questionnaire and post-test from El Mawas *et al*. [2020] (IV) evaluates LX before and after the educational activity. Thus, two questionnaires, a pre-test and a post-test, were developed and applied. Each questionnaire consists of nine questions, four of which are affirmative and five of which are open-ended. The questionnaires are designed to assess the learners' knowledge, which is why Geography teachers drew them up. The questionnaires were drawn up and presented to the learners in printed form and then digitized, but we couldn't find them in the publication to replicate it.

A questionnaire to measure learner experiences of Kawano *et al*. [2019] (V) was developed to assess LX at the end of each learning unit. Therefore, the data from the experiments were considered LX values. The values are used to tabulate the experiences through indices using a five-point Likert scale. For the evaluation questionnaire applied before the educational activity, the learners' expectations were considered, that is, what they expected of the educational activity. For the evaluation questionnaire after the educational activity, the LX achieved after the educational activity was considered. An example taken from the post-assessment questionnaire is how well learners understood the content of the learning unit.

Lykke *et al*. [2015] developed a semantic differential questionnaire to portray learners' experiences (VI). The questionnaire is characterized by its semantic differential. The aim is to capture learners' connotative perceptions and attitudes toward learning design at different moments of the educational activity. For the pre-assessment, the semantic differential questionnaire was used to analyze the learners' perceptions before the educational activity. The during evaluation, it was observed whether the learners worked actively in the educational activity. A questionnaire was used to guide the execution of the focus group for the post-evaluation. The questionnaires were developed based on the 10 criteria of "positive experiences" [Jantzen *et al*., 2011]. Therefore, the pre-assessment questionnaire has 12 pairs of opposing adjectives for learners to consider. The adjectives were related to negative or positive emotional states that a learner may have experienced during the educational activity [Lykke *et al*., 2015]. A nine-point scale was organized for the questionnaire, which could be graded from -4 to 0 to +4. It was stipulated that the pairs would have a positive and negative value, such as "Sad X Happy".

The evaluation model of the INtegrating BEhaviorism and COnstructivism in Mathematics (INBECOM) project by Nygren *et al*. [2019] (VII) aims to explore learners' affective learning experiences during educational activities. The model has five stages, being (1) the construction of an educational model; (2) the development, experimentation, and evaluation of the game; (3) content and translations, experimentation, and evaluation of the tutor system; (4) presenting a course; and (5) generating a theory. For the third phase of the INBECOM model, a pre-questionnaire, an intervention (during-questionnaire) and a post-questionnaire were carried out. The intervention (questionnaire-during) includes interviews, diaries/learning records, observations, questionnaires and diaries/records of feelings. For interviews, it is suggested to carry them out with a group of learners, that is, two

to three per interview, and there is an example of a question used: "How did the sticks help you solve problems?". For the observations, the learning diaries/records, and the feelings diaries/records, it is suggested that the teacher record the teaching and the learners' experiences in parallel with the educational activity. There are no examples of the questions used for the pre-test and post-test questionnaires.

The evaluation of LX carried out by Stanley and Zhang [2018] took place through two questionnaires (IX), a pre-test and a post-test. The pre-test questionnaire assesses factors relevant to how class activity affects each learner's results during an activity. The post-test questionnaire focused on the learners' behavior and perceptions after the educational activity. The questionnaires had 10 questions and 38 sentences, where for the pre-test, they were question-type, and for the post-test, they were affirmative-type sentences. An example of a question for the pre-test is 'How many hours do you work per week on average? For the post-test, an example sentence is "The class was an enriching experience". Other examples are present in the Table 5. A five-point Likert scale was used for the answers to the questionnaires, ranging from totally disagree to totally agree.

The Questionnaire to assess attitudes and negotiations in learning of Yeh and Chen [2019] (X) was designed to investigate learners' attitudes to the experiences of the educational activity carried out. A pre-questionnaire was carried out to obtain information from participants on whether they owned smartphones and their preferences. The post-evaluation questionnaire was developed based on the Davis [1989] constructs: Perceived ease of use, Perceived usefulness, Attitude towards use and Intention to use. The questionnaire contains 15 questions, which learners can answer on a five-point Likert scale. Additionally, there are 8 open-ended question phrases.

8 Discussions

8.1 Feasibility study discussions

Through this feasibility study made with LEEM, it can be affirmed that the LX assessment is essential for verifying learning experiences with DICTs. In this sense, Tabares *et al*. [2021] emphasized that the LX assessment contributed to diagnosing weaknesses in learning. This LX assessment makes it possible to identify learner behavior and content that still needs to be better understood, using elements of LX that help to identify aspects that can influence the results of this assessment. In the work of Ruiz and Snoeck [2018], it was observed that learners learn better when a model is used and it is appropriate to their needs and interests. It is, therefore, believed that the proposed LEEM model will provide satisfactory results and a meaningful learning experience mediated by DICTs.

Unlike the work of Ruiz and Snoeck [2018], Tabares *et al*. [2021] and Martinelli and Zaina [2021], which evaluated LX during and/or at the end of the activities, LEEM follows the recommendation of Huang *et al*. [2019] to be as comprehensive as possible to collect the learning experience at various times (before, during and after learning). This type of evaluation promotes continuous monitoring and recording of LX's progress throughout the learning process. In this way, this feasibility study sought to improve the LEEM based on teachers' feedback. With this, it is hoped to contribute to the HCI community using a robust LX assessment model, organized into three stages and providing different types of assessment to promote better interaction with DICTs in learning.

In short, the main improvements to the LEEM, through this study, are in improving the LEEM items, as detailed in Section 6. There were generic terms, some items with more than one question, redundant items, and the need to add a new LX element. Therefore, it can be said that the objective of the feasibility study was achieved, as it created a body of knowledge about the LEEM and, consequently, allowed it to improve and evolve.

8.2 Benchmark discussion

The benchmark presented in this article aims to characterize the LX evaluation technologies identified in the SMS [dos Santos *et al*., 2022]. This characterization occurred through a need to find synthesized information about the LX evaluation technologies.

The results showed that 40% (N = 4) of technologies use only one element to perform the LX assessment. On the other hand, the LEEM model includes 11 distinct LX elements. It is believed that the more LX elements are used in the assessment, the more aspects will be perceived related to the learners' difficulties and needs that can impact their learning process.

Another outcome identified is whether there is support and what that support is. Thus, only 30% (N = 3) of technologies presents a step-by-step guide for the use and/or application of the LX evaluation. In addition, one technology offers instructions for applying and filling out the model (I). Therefore, this orientation is considered a differentiator, as it allows for a more comprehensive assessment and the application of best practices.

9 Limitations

Some limitations can affect the validity of the study. Therefore, this study categorized the limitations according to Wholin *et al*. [2012] and classified as internal, external, and conclusion.

For the internal limitations, we considered two limitations that could lead to an inadequate interpretation of the study results, namely (1) training effects and (2) time. To avoid training effects (1), all participants received the same type of training and followed the same scope and instructions. In addition, about evaluation time (2), each participant was given seven days to evaluate the LEEM and answer the evaluation questionnaire.

The external limitation was that all the participants were teachers. In other words, it was not carried out with learners. At this initial stage, the study focused on validating the LEEM to improve it before it was applied in a classroom, and the results are considered to be indicative rather than conclusive.

Table 5. Part of questions and sentences used in the study of Stanley and Zhang [2018].

Test	Number	Sentences
Pre.	01.	With which gender do you identify?
Pre.	02.	How old are you?
Pre.	03.	Do you identify as Hispanic?
Pre.	04.	What was the highest school completed by your mother or parent 1?
Pre.	05.	What was the highest school completed by your father or parent 2?
Pre.	06.	Are you eligible, or have you received a Pell Grant?
Pre.	07.	How many hours do you work per week on average?
Pre.	08.	How many credits did you take last semester?
Pre.	09.	Is English your first language?
Pre.	10.	How many previous online courses have you taken?
Post	01.	I am able to easily access the Internet as needed for my studies.
Post	02.	I am comfortable communicating electronically.
Post	03.	I am willing to actively communicate with my classmates and instructors electronically.
Post	04.	I feel that my background and experience will be beneficial to my studies.
Post	05.	I am comfortable with written communication.
Post	06.	I possess sufficient computer keyboarding skills for doing online work.
Post	07.	I feel comfortable composing text on a computer in an online learning environment.
Post	08.	I am motivated by the material in online activities.
Post	09.	Learning is the same in class and at home online.
Post	10.	I feel that I can improve my listening skills the same working online as in an-person class

For the limitation of conclusion, it was considered that the study was carried out remotely. Therefore, it was impossible to control for bias caused by external factors. Some external factors of the environment may have compromised the results, such as noise in the classroom while using the LEEM or interruptions during the study. However, based on the results, it was considered that the participants fulfilled all the tasks requested in the study and contributed to improving the LEEM.

Threats that could not be controlled about the benchmark may have influenced our research results. For example, (i) insufficient information in article to be extracted; (ii) non-standard terms of LX aspects; (iii) lack of detailing of LX evaluation technologies. However, we believe that the peer-review process reduced the impact of these threats.

10 Conclusions and Future Work

This article aimed to present LEEM, a model that allows LX to be assessed when interacting with DICTs, encompassing different assessment forms and as many elements as possible. Through the results obtained with LEEM, teachers will be able to better prepare their next lessons with feedback from learners at different times in their learning.

This article also presented the feasibility study to improve and evolve LEEM. The quantitative results of the TAM indicators showed that most participants found LEEM easy to use and useful and that they intend to use it in the future. It was also noted that teachers considered that the LEEM may require mental effort to answer. To improve this aspect, changes were made to the terms in the LEEM items.

The qualitative analysis revealed the difficulties identified by the teachers in the LEEM and the possibilities for improvement, such as making the checklists' items clearer and more cohesive according to the objective of each element of the LX assessed. The participants also reported the size of the LEEM as a negative point. On the other hand, it is believed that applying it according to the recommendations (before, during, and at the end of the educational activity) won't cause so much fatigue when filling it out. However, it will only be possible to assess fatigue when LEEM is applied in a real context. Overall, based on the qualitative analysis, it was also realized that the LEEM is a practical, interesting, objective, and appropriate assessment model for the type of analysis it is intended to carry out.

The study's main technologies reported in this article point to the need to improve the items, scales, and LX elements present in the LEEM. The items on the pre-assessment checklist were organized gradually, from the most positive to the most negative. On the other hand, the items on the during-assessment checklist (learner) had to be adjusted according to the feeling scale (SAM scale) or the learning scale (Likert scale). In addition, the Persistence element had to be added to the during-evaluation checklists (learner and teacher), as it allows us to assess how the learner reacted to obstacles during the educational activity [Fotaris *et al*., 2016]. Overall, through this feasibility study, it has been possible to create a body of knowledge through the evolution of LEEM and also to determine its viability.

The current version of the LEEM contains six different ways of assessing LX (Opposite Adjective Pairs, Experience Criteria, Feelings Notes/Records, Learning Notes/Records, Observations, and Focus Groups) and 12 elements of LX (Skill, Value, Participation, Authentic, Usability, Preference, Desirability, Adaptability, Comfortability, Persistence, Results and Satisfaction). This version can be found at link[6]. The benchmark carried out showed that there are no automated LX evaluation technologies of those identified in the SMS that are of the questionnaire and model type. The bench-

[6] `https://figshare.com/s/188c4b881ed1e4072244`

mark allowed us to compare LX evaluation technologies, identifying that most of them evaluate LX in more than one moment of educational activity. In addition to allowing you to compare the number of elements and ways to evaluate the LX. In future work, we intend to conduct a case study with LEEM in two different classes with learners and their respective teachers on the Computer Science course at the Federal University of Paraná. This study will evaluate LEEM in a real context and identify points for improvement.

Declaration

Acknowledgements

We thank the teachers who participated in the study. We are grateful for the financial support from the Coordination for the Improvement of Higher Education Personnel (CAPES) - Program of Academic Excellence (PROEX).

Funding

This study was financed in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) - Program of Academic Excellence (PROEX).

Authors' Contributions

Gabriela Corbari dos Santos, is the main contributor of this article. Deivid Eive dos S. Silva: Supervision, Writing-Review Editing. Natasha M. C. Valentim: Supervision, Writing-Review Editing. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The authors confirm that the data supporting the findings of this study are available within the article.

References

Anand, G. and Kodali, R. (2008). Benchmarking the benchmarking models. *Benchmarking: An international journal*, 15(3):257–291. DOI: 10.1108/14635770810876593.

Chapman, J. R., Seeley, E. L., Wright, N. S., Glenn, L. M., and Adams, L. L. (2016). An empirical evaluation of a broad ranging e-text adoption with recommendations for improving deployment success for students. Available at:`https://eric.ed.gov/?id=EJ1167339`.

Corbin, J. *et al*. (1990). Basics of qualitative research grounded theory procedures and techniques. Available at:`https://scholar.google.com/scholar_lookup?hl=en&publication_year=1990&author=J+Corbin&author=A+Strauss&title=Basics+of+Qualitative+Research%3A+Grounded+Theory+Procedures+and+Techniques`.

Correa, C. M., de Freitas, G. V. M., dos Santos Eberhardt, A. L., and Silveira, M. S. (2021). From now on: Experiences from user-based research in remote settings. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*, IHC '21, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3472301.3484334.

da Silva, E. J. and Ziviani, H. E. (2018). Desenho e música no ensino de ihc: relato de experiência de uma aula sobre conceitos básicos da engenharia semiótica. In *Anais Estendidos do XVII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais*, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/ihc.2018.4210.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology.

Deodato, J. (2015). Evaluating web-scale discovery services: a step-by-step guide. *Information Technology & Libraries*, 34(2):19–75. DOI: 10.7282/T3K64M3R.

dos Santos, G. C., dos S. Silva, D. E., and C. Valentim, N. M. (2024). Feasibility study of a model that evaluates the learner experience: A quantitative and qualitative analysis. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3638067.3638119.

dos Santos, G. C., Silva, D. E., and Valentim, N. M. (2022). Um mapeamento sistemático da literatura sobre iniciativas que avaliam a experiência do aprendiz. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 621–633, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbie.2022.224673.

dos Santos, G. C., Silva, D. E., and Valentim, N. M. (2023). Proposal and preliminary evaluation of a learner experience evaluation model in information systems. In *Proceedings of the XIX Brazilian Symposium on Information Systems*, SBSI '23, page 308–316, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3592813.3592919.

Dune, T., Bidewell, J., Firdaus, R., and Kirwan, M. (2016). Communication idol: Using popular culture to catalyse active learning by engaging students in the development of entertaining teaching and learning resources. *Journal of University Teaching & Learning Practice*, 13(5):15. DOI: 10.53761/1.13.5.3.

El Mawas, N., Tal, I., Moldovan, A.-N., Bogusevschi, D., Andrews, J., Muntean, G.-M., and Muntean, C. H. (2020). Investigating the impact of an adventure-based 3d solar system game on primary school learning process. Available at: `https://eric.ed.gov/?id=EJ1263189`.

Fotaris, P., Mastoras, T., Leinfellner, R., and Rosunally, Y. (2016). Climbing up the leaderboard: An empirical study of applying gamification techniques to a computer programming class. Available at: `https://eric.ed.gov/?id=EJ1101229`.

Huang, R., Spector, J. M., and Yang, J. (2019). *Educational Technology a Primer for the 21st Century*. Springer. DOI: 10.1007/978-981-13-6643-7.

Imamura, R. E. M. and Baranauskas, M. C. C. (2019). A framework for socio-enactive educational systems: Link-

ing learning, design, and technology. In *IHC '19: XVIII Brazilian Symposium on Human Factors in Computing Systems*, IHC '19, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3357155.3358443.

ISO 9241-210 (2019). Ergonomics of human-system interaction — part 210: Human-centred design for interactive systems. Available at:`https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en`. Acessed in 23/02/2022.

Jantzen, C., Vetner, M., and Bouchet, J. (2011). Oplevelsesdesign. Available at:`https://scholar.google.com/scholar_lookup?hl=en&publication_year=2011&author=C.+Jantzen&author=M.+Vetner&author=J.+Bouchet&title=Oplevelsesdesign`.

Kawano, A., Motoyama, Y., and Aoyama, M. (2019). A lx (learner experience)-based evaluation method of the education and training programs for professional software engineers. In *Proceedings of the 2019 7th International Conference on Information and Education Technology*, ICIET 2019, page 151–159, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3323771.3323789.

Lang, P. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. Available at:`https://cir.nii.ac.jp/crid/1574231875645651456`.

Lima, D. T., Zacarias, R. O., de Souza, K. E. S., dos Santos, R. P., and da Rocha Seruffo, M. C. (2021). Analytical model for classifying areas of interest in interactive systems. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*, IHC '21, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3472301.3484357.

Lykke, M., Coto, M., Jantzen, C., Mora, S., and Vandel, N. (2015). Motivating students through positive learning experiences: A comparison of three learning designs for computer programming courses. *Journal of Problem Based Learning in Higher Education*, 3(2):80–108. DOI: 10.5278/ojs.jpblhe.v0i0.1130.

Martinelli, S. R. and Zaina, L. A. M. (2021). Learning hci from a virtual flipped classroom: Improving the students' experience in times of covid-19. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*, IHC '21, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3472301.3484326.

Muriana, L. a. M. and Baranauskas, M. C. C. (2021). Affecting user's self-esteem: Analysis under the self-determination theory perspective and design recommendations. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*, IHC '21, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3472301.3484331.

Nygren, E., Blignaut, A. S., Leendertz, V., and Sutinen, E. (2019). Quantitizing affective data as project evaluation on the use of a mathematics mobile game and intelligent tutoring system. Available at:`https://www.ceeol.com/search/article-detail?id=804185`.

Reyna, J. and Meier, P. (2018). Using the learner-generated digital media (lgdm) framework in tertiary science education: a pilot study. *Education Sciences*, 8(3):106. DOI: 10.3390/educsci8030106.

Rosa, J. C. S., Rêgo, B. B. d., Garrido, F. A., Valente, P. D., Nunes, N. J., and Matos, E. (2020). Interaction design and requirements elicitation integrated through spide: A feasibility study. DOI: 10.1145/3424953.3426498.

Ruiz, J. and Snoeck, M. (2018). Adapting kirkpatrick's evaluation model to technology enhanced learning. In *MODELS '18: ACM/IEEE 21th International Conference on Model Driven Engineering Languages and Systems*, MODELS '18, page 135–142, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3270112.3270114.

Schmidt, M. and Huang, R. (2022). Defining learning experience design: Voices from the field of learning design & technology. *TechTrends*, 66(2):141–158. DOI: 10.1007/s11528-021-00656-y.

Shi, L. (2014). Defining and Evaluating Learner Experience for Social Adaptive E-Learning. In *2014 Imperial College Computing Student Workshop*, volume 43 of *OpenAccess Series in Informatics (OASIcs)*, pages 74–82, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. DOI: 10.4230/OASIcs.ICCSW.2014.74.

Shull, F., Mendoncça, M. G., Basili, V., Carver, J., Maldonado, J. C., Fabbri, S., Travassos, G. H., and Ferreira, M. C. (2004). Knowledge-sharing issues in experimental software engineering. *Empirical Software Engineering*, 9(1):111–137. DOI: 10.1023/B:EMSE.0000013516.80487.33.

Soloway, E., Guzdial, M., and Hay, K. E. (1994). Learner-centered design: The challenge for hci in the 21st century. *Interactions*, 1(2):36–48. DOI: 10.1145/174809.174813.

Stanley, D. and Zhang, J. (2018). Do student-produced produced videos videos enhance enhance engagement engagement and learning learning in the online online environmentenvironment. *Online Learning*, 22(2). DOI: 10.24059/olj.v22i2.1367.

Tabares, M. S., Vallejo, P., Montoya, A., Sanchez, J., and Correa, D. (2021). Seca: A feedback rules model in a ubiquitous microlearning context. In *DATA'21: International Conference on Data Science, E-learning and Information Systems 2021*, DATA'21, page 136–142, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3460620.3460745.

Venkatesh, V. and Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39(2):273–315. DOI: 10.1111/j.1540-5915.2008.00192.x.

Villas Boas, B. M. d. F. (2006). Avaliação formativa e formaçãoo de professores: ainda um desafio. Available at:`http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S1981-04312006000100006&nrm=iso`.

Wholin, C., Runeson, P., Host, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2012). Experimentation in software engineering: an introduction. *Springer Berlin, Heidelberg*, page 236. DOI: 10.1007/978-3-642-29044-2.

Yeh, S.-W. and Chen, C.-T. (2019). Efl learners' peer negotiations and attitudes in mobile-assisted collaborative writing. *Language Education & Assessment*, 2(1):41–56. DOI:

https://dx.doi.org/10.29140/lea.v2n1.100.