Applying Pedagogical Usability for Designing a Mobile Learning Application that Support Reading Comprehension [†]

Gustavo Zurita [1],*, Nelson Baloian [2], Sergio Peñafiel [2] and Oscar Jerez [3]

[1] Management Control and Information Systems Department, Faculty of Economics and Business, Universidad de Chile, Diagonal Paraguay 257, Santiago 833-0015, Chile

[2] Department of Computer Science, Universidad de Chile, Santiago 837-0456, Chile; nbaloian@dcc.uchile.cl (N.B.); spenafie@dcc.uchile.cl (S.P.)

[3] Teaching and Learning Centre. Economics and Business Faculty, Universidad de Chile, Diagonal Paraguay 257, Santiago 833-0015, Chile; ojerez@fen.uchile.cl

* Correspondence: gzurita@fen.uchile.cl.

† Presented at the 13th International Conference on Ubiquitous Computing and Ambient Intelligence UCAmI 2019, Toledo, Spain, 2–5 December 2019.

Abstract: The pedagogical usability is an important characteristic of applications that support learning as it relates to the added value students perceive while using it for learning. A good pedagogical usability means that an application has more chances to be accepted and used by the students thus raising the possibilities that students actually will learn with it. However important, this concept tends to be neglected by many authors. In this work we show how this concept can be applied to evaluate an application by presenting a real example of an application that has been re-designed to improve its usability, thus showing how pedagogical usability can be operationalized to be applied in general. The application shown in this work is called RedCoMulApp (Reading Collaborative Multiple-option Application) and its goal is to raise the reading comprehension ability of 12th grade high school students. In order to capture the pedagogical usability, we used 12 metrics to design a questionnaire with 26 questions to be answered with a 5-point Likert-scale, plus two open-ended questions to capture aspects that were positively valued, and those that should be improved. The results for the pedagogical usability allow us to validate that the design of the application was perceived by the students as pedagogically useful to learning about reading comprehension.

Keywords: smartphones in the classroom; collaborative user Interfaces for education; pedagogical usablity

1. Introduction

Usability is without a doubt a critical aspect in any software application. Experts recognized this long ago [1,2]. In [2] authors state that "the evaluation of the use of educational software must take into account usability as well as learning and, crucially, the integration of issues concerned with usability and learning". In [3,4], authors state that it is important to perform usability testing to guide the development and design process of any computer-based learning application. According to [3,5,6], there are two relevant related concepts for measuring the appropriateness of a computer-based learning application: the first one is the degree to which learners can easily and efficiently use a learning application to satisfy their goals and requirements, named as technical usability. The second one is the learner perception by which they select, organize, and recognize their sensory impressions in order to interpret and understand their learning environment, named as pedagogical usability.

While technical usability is concerned with usability of the application, i.e., the user interface of the software application, pedagogical usability is concerned with whether the application, its content and tasks support learners in learning in a certain context according to select pedagogical objectives. A well-designed computer-based learning application with acceptable levels of technical or pedagogical usability, from different perspectives, increases the likelihood of learners having a successful learning experience.

One of the recurrent learning topics when developing computer supported collaborative learning tools has been supporting activities for improving reading skills of children from elementary school levels [7]. However, although statistics show illiteracy rates around the world sharply declining over recent years [8], it seems that reading comprehension skills lag behind acceptable levels. For the above reason, in [9] the authors developed a pedagogical activity supported by a mobile learning application called RedCoApp (Reading Comprehension Collaborative learning application) in order to train the reading comprehension of high school students. During this activity, students had to read a text and identify the key words for understanding its message by marking them. Nevertheless, although this is a very natural gesture when reading a paper text and highlighting words with a marker, using the most popular software, it turned out to be a difficult one when using tablets [9]. Furthermore, this experience has also shown that the approach of using constructed development responses has its disadvantages, especially when applied to mass courses. Some of them are the following:

- It is difficult for students to develop a constructed response using mobile learning applications, mainly because HCI problems related to the screen's size and interaction mode.
- It is difficult for the teacher to evaluate all the answers developed by students.
- It is difficult to monitor the degree of progress of the students' work, how many have responded, and if they responded well.
- It is difficult to manage the interactions inside the classroom in order to achieve learning.

Due to these difficulties, we came to the conclusion that a learning activity based on questions that are answered through the selection of multiple alternatives would make the whole process much easier, i.e., with a better technical usability; since the evaluation of the correctness of the answers are easy to validate. In fact, the advantages of multiple-choice tests are that are easy to apply, reliable, standardized and objective.

Although some authors [10,11] criticized these kinds of evaluation tests, mainly because a constructed response is supposed to require higher skills from the student, therefore, allowing a student to perform a more elaborate learning activity. However, in [12] the authors show that there is an equivalence between constructed responses to those of alternatives. The justification for the above is that students must first build an answer and then verify/check against possible alternative responses. In [11] authors argue that students should not only mark an alternative but also justify it.

In Chile, there is a single standardized admission test for almost all universities called PSU, that is psychometrically validated and provided by DEMRE (Department of Evaluation, Measurement and Educational Registration) and independent unit of the University of Chile [13]. An important part of this test consists of reading comprehension, measured by multiple-choice questions. This is a summative evaluation. Based on [14], we can state that it is possible to convert an activity for a summative evaluation into a formative one, if the evaluation is used as feedback for the student to reflect on and reformulate their original answers. Therefore, we develop a new version of a pedagogical activity supported by a mobile learning application named RedCoMulApp (Reading Comprehension Collaborative learning application based on Multiple-choice questions), in order to train the reading comprehension of middle school students, that was already reported in [15].

Given the technical usability problems that we had in the first version of the application RedCoApp [9], in this work we report a finished study about the pedagogical usability based on students perception found in a new redesigned version of RedCoMulApp that use multiple-choice questionnaires, instead of marking key words. Although multiple-choice questionnaires are associated to summative evaluations, the implemented learning activity uses them within a collaborative learning activity in which students have to justify, first individually then collaboratively,

their choice with a short text; thus, converting it in a formative learning activity. The developed system was used and evaluated in a real learning situation. Furthermore, an important aspect is that we use the multiple-choice questions that normally are provided each year by DEMRE, since the PSU is designed for having all problems being pedagogically validated. Therefore, the main goal of this work is to report on a study that measure the students perceptions of the pedagogical usability of the collaborative mobile learning application RedCoMulApp described in [15].

2. Technical Usability Issues in RedCoApp

RedCoApp was developed in 2017 [9]. Figure 1 shows a screenshot of the application in which the student had to identify 1 to 3 key words, which they think represent the main ideas exposed in the piece of text and write a short comment that justifies their selection.
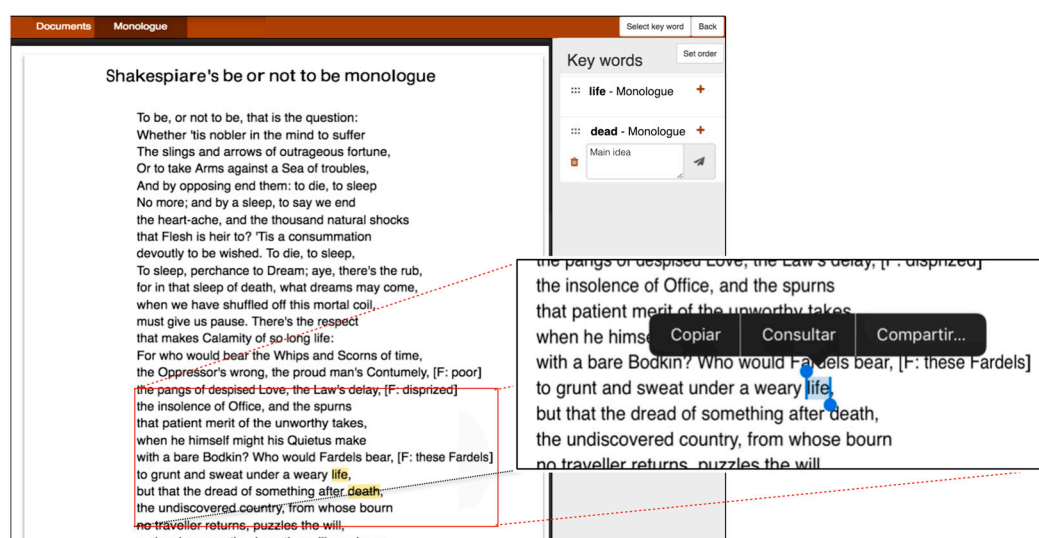

Figure 1. Shows a screenshot of the main interface of RedCoApp when seen in horizontal position on an iPad. It shows the text a student has to read, over which 2 words have been marked as keywords. These are "life" and "dead", which are highlighted in yellow. The enlarged portion of the screenshot corresponds to a close-up showing the state of the iPad's interface after selecting the Word "life" in order to mark it as keyword.

During the experimental evaluation of RedCoApp, we found the following technical usability problems. To mark a keyword using RedCoApp running on an iPad (provided by us for all students), a text selection action must be performed. According to how this process works in the IOS operating system for mobile devices, a user should touch a word on the screen, then adjust two pointers that mark the beginning and end. The selected word(s) appears in a different color that highlights it (see Figure 1 bottom part). Then, some of the next technical usability problems can occur: Although the process of selecting a word is not difficult, the first technical usability problem happens, if by mistake, the student does not mark the word he or she wanted. In this case, to correct the error, it is necessary to start the process again by touching the correct word on the screen, after 2 s. It is here where the first usability problem occurred: students tried to mark another word almost instantaneously, so the incorrect word marked previously, remains still selected, while the student persistently presses another word without managing to correct the error.

The second technical usability problem comes from managing the start and end of the selection, which involves dragging the lines that initially marked the selected word with the balls up and down, so that other words can be added to the left (start of the selection) or right (end of the selection). This process generated the following usability problem: if the student does do not have enough motor skills to take the line with the ball up or down to increase the selection left or right, the selection of the initial word is modified, and the selection modification process leaves marked words (or parts of words) that were not meant to be selected.

The third technical usability problem arose when students wanted to increase the selection to the left (beginning of the selection): if in addition to moving the line to the left, the line is moved slightly downwards, and the last character of the word initially selected remains marked.

The previous mentioned technical usability problems were frequent in users without experience using IOS. Almost all the 45 students who used RedCoApp had Android cell phones, and only 25% had an iPhone (similar to an iPad) with IOS. It should be considered that the students had to repeat the process of marking three key words for each text in the three phases of the learning activity, i.e., "individual", "anonymous" and "teamwork" phases. This issue introduced a significant delay in the activity, making it practically impossible to finish the learning activity within the term of the available class period.

3. Brief Description of the New Version of RedCoMulApp

We introduced a modified version of RedCoApp, named RedCoMulApp [15], based on multiple-choice questions in order to solve the technical usability problems previously described. RedCoMulApp can be used under two roles: teacher and student.

The teacher's role allows for creating the learning activity, and a real-time monitoring of the task development. This monitoring allows the teacher to review the progress of the learning activity that students are performing during "individual", "anonymous" and "teamwork" phases. To create the learning activity a teacher performs actions, for instance: (a) uploads a text used as a context for the multiple-choice questions; (b) introduces the multiple-choice questions with their corresponding right answers; or (c) assigns the task to work teams: each one composed of two or three students. For the real-time monitoring of the task development, the teacher has access to relevant visual information during the execution of the learning activity in order to review the progress of the students in each phase.

The student's role, let to develop the "individual", "anonymous" and "teamwork" phases which students should go through with RedCoMulApp in order to accomplish the reading comprehension learning activity in that order. In the "individual" phase, each student individually reads the text provided by the teacher. As each student answers the questions, the teacher can see the answers. In the "anonymous" phase, students do the same as in the previous phase (reading the text and answering the multiple-choice questions) with the difference that each student can see the answers of the other two classmates of their groups without knowing who they are. Students have then to confirm their previous answer or change them based on what their classmates have answered. Finally, in the "teamwork" phase, students see the names of their classmates and they meet face-to-face to choose a single option together. They can talk to each other, exchange opinions and discuss their disagreements. All students of the group have to select the same option as answer; otherwise, they receive a message from the system to do so.

4. Usability for Mobile Applications in Learning Environments

According to [16,17], mobile technology is recognized to be one of the most significant directions of current knowledge-based society concepts, with research outcomes in mobile learning studies being significantly positive [17]. The growth in the use of mobile devices such as, smartphones and tablets in everyday life demonstrate an increasing demand to access educational information, instruction and software applications using mobile technology [16,17]. Many studies have given encouraging results for using mobile technologies to support students in the teaching and learning process [17]. Along with this high penetration of mobile learning applications, there are usability problems of mobile devices and learning, which should be considered when developing and testing different types of interfaces and mobile applications [16], and not only consider experimental methods focused on learning outcomes and experimental evaluations [17].

Usability is defined as the ability of a product to be used by specific users to achieve specific goals with effectiveness, efficiency, and satisfaction, in a specific context of use [18]. The concept (generic) deals with the facility (or quality) of use, and should not be neglected. It prevents the risk of frustrating the users expectations, which is essential in the educational environment [19]. The

definition is primarily concerned with a software product; however, it can be applied to mobile learning software taking into consideration features specific to mobile devices and digital supported learning aspects. Usability has a 'technical' or 'design' component that addresses the software application context and permeates the planning and implementation proper to systems and virtual environments. The essence of literature in design usability area, reinforces the implication that, if a virtual interface is properly designed and managed, with a high usability level, its users will easily learn to use it, and be pleased to have their needs supplied. On the other hand, although the implementation of educational applications supported by mobile devices have appropriate design usability levels, it does not mean that their functionalities support pedagogical processes that guarantee student learning. It is also necessary to ensure compliance with the educational objective with efficiency, effectiveness and satisfaction. In this way, it is necessary to incorporate a concept in education, considered to be emerging and innovative for some authors, whose bases are replicated or adjusted to different areas of knowledge, including human-computer interaction and interaction design: pedagogical usability [5,16,19,20].

A comparison of usability evaluation methods for computer learning applications that were analyzed in [6], identifies two main types: technical usability evaluation, and other specific usability evaluation methods, like pedagogical usability. The first type is categorized by analytical and empirical methods, being oriented to identify usability problems and improve the usability of an interface design, i.e., they have a more technical focus. The second is distinguished according their three usability constructs: the context of use, the user and her/his goals; mainly based in exploring users' perception, satisfaction and motivation to learn.

While the technical usability is concerned with the interaction of the user with a computational application interface, the so-called pedagogical usability is concerned with "whether the tools, content, interface and the tasks of the computational application, support learners in learning in a particular context and according to selected pedagogical objectives" [6]. According to [16], in order to develop or improve a mobile learning application it is not sufficient to evaluate whether one can use it or not (technical usability); they must also want to use it [20]. The technological usability is one of the user interfaces, centered on mobile device factors; whereas the pedagogical usability it the one which includes criteria related with learning factors supported by the mobile device.

4.1. Usability in the Context of Teaching and Learning

To capture pedagogical issues that are particularly fundamental to collaborative learning, the usability concept needs to be extended beyond technical usability [5]. That means, usability has to be considered in a different way when it is being evaluated in the context of teaching and learning, and the concept of pedagogical usability, already presented by first time by [1], can be helpful when considering the close relationship between usability and pedagogical design [5,16].

Usability is seen as a determinant factor of success of an educational platform and it should be given special attention when talking about accessing it on mobile devices that were not designed for educational purposes. Based on the fact that the device's and UI usability issues influence in different ways the user experience [21], authors in [16] consider that the usability of a mobile learning platform has to be regarded from four different perspectives: device usability, user interface usability, educational content usability, and pedagogical usability.

4.2. Pedagogical Usability

Pedagogical usability is the analysis of the way an educational application (including tools, content, tasks and interface) supports students in their learning process within various learning contexts according to given learning objectives [5,16]. It should especially consider educational aspects, such as the learning process, learning purposes, user's needs, the learning experience, learning content and learning outcomes [21]. Improving the technical usability of a mobile learning application is not enough for its users to use it, they must also want to use it [16], thus pedagogical usability is the most appropriate method to identify how much users want to use the learning application, [5,20].

Pedagogical usability is defined as a sub-concept of utility, as technical usability is a component of usability; thus, in addition to the dialogue between a user and a computer based application captured by the device usability, and user interface usability perspectives, [5]. According to [16,22], pedagogical usability is the analysis of the way a mobile learning application (tools, content, tasks and interface) supports students in their learning process within various learning contexts according to learning objectives. It should be especially concerned with educational aspects such as the learning process, purposes of learning, user's needs, and the learning experience. Pedagogical usability can be evaluated in the context of educational environment, occasion in which the content is consumed by students, who wish to achieve specific educational purposes, effectively and efficiently, and getting satisfaction in their learning.

The study presented in [5], proposed ten pedagogical usability metrics, associated with questions proposed for them, in order to capture the subjective end-student perception and utility satisfaction with the pedagogical aspects:

- **Learner Control**. Minimizing working memory load, as users have limited memory capacity, usually 7 +/− 2 units. Control of the technology should be taken away from the teachers and instructional designers and given to the learners. Information must be presented in meaningful, interconnected entities, and not in separate pieces that are hard to understand.
- **Learner Activity**. A teacher's role depends on underlying pedagogic assumptions. Learning material should gain learners attention. Learners should feel that they own the goals of action and thus the results.
- **Cooperative/Collaborative Learning**. The Constructivist view is based on social learning and knowledge sharing via collaborative tasks. Learners are able to discuss and negotiate on various approaches to the learning task. Tools might support asynchronous or synchronous social navigation.
- **Goal Orientation**. Instructivists emphasize few clearly defined goals, constructivist goals should also be clear, but set by the learners themselves. Task decomposition can be inductive or deductive.
- **Applicability**. Authentic activities and contexts: examples should be taken from authentic situations. Transferring learned knowledge or skills are useful to other contexts. Human development should be considered in a way that the material is relevant for target population's developmental stage.
- **Added Value**. When mobile devices and digital learning material are used in a learning situation, it is expected that this is done to introduce identifiable added value to the learning in comparison to, for example, printed material, and material produced by the teacher or the students themselves.
- **Motivation**. Motivation affects the whole learning process. Intrinsic (need for deep understanding) and extrinsic (need for high grades) motivation should be considered.
- **Valuation of previous knowledge**. Prerequisites—needed to accomplish learning tasks. Meaningful encoding (elaboration), learner is encouraged to make use of his/her previous knowledge.
- **Flexibility**. Pretesting and diagnostics help adapt learning material for different learners. Task decomposition, small and flexible learning units is required.
- **Feedback**. The computer-based application or learning material used should provide the student with encouraging and immediate feedback. Encouraging feedback increases learning motivation; immediate feedback helps the student understand the problematic parts in his/her learning.

In order to evaluate the pedagogical usability of a mobile interface, [16] proposed nine pedagogical usability metrics:

- **Instruction**. Whether the application's instruction is clear or if it needs a lecturer's intervention.
- **Learning content relevance**. The extent to which the content of the application supports students in their learning.

- **Learning content structure**. The degree to which the content of the application is organized in a clear, consistent and coherent way that supports learning.
- **Tasks**. The extent to which the tasks performed on the application help students achieve their learning goals.
- **Learner variables**. The degree to which learner variables are considered in the application.
- **Collaborative learning**. The extent to which the application allows students to study in groups.
- **Ease of use**. The provision by the application of clear directions and descriptions of what students should do at every stage when questions to the mobile class activities are answered.
- **Learner control**. The characteristics possessed by the application that allows students to make instructional choices. The extent to which the learning material is broken down into meaningful units. The extent to which the learning material in the application is so interesting to students that it compels them to participate.
- **Motivation**. The degree to which the application motivates students.

4.3. Metrics for Pedagogical Usability Applied to RedCoMulApp

To measure the pedagogical usability of the mobile learning application RedCoMulApp described in [15], see Section 3 for a brief description; we combine the metrics presented in [5] and [16]. For each metric, a set of 5-point Likert-scale questions (from "strongly disagree" to "strongly agree") were presented to students after performing the learning activity. These questions took in consideration the criteria that had been operationalized in [5]. Therefore, considering the characteristics of the instructional design of the learning activity supported by RedCoMulApp, 12 metrics were applied. The first 8 (learner control, learner activity, collaborative learning, goal orientation, applicability, add value, motivation, feedback) correspond to the ten metrics described in [5], and the next 4 (learning content relevance, learning content structure, ease to use, instruction) to the nine metrics proposed in [16], being of these 9 metrics, 4 in common (learner control, motivation, learner activity with tasks, and learner activity with learner variables) with the proposals in [5]. In total, 26 questions for the measurement of the 12 metrics were designed (see Table 1.) Question Q3. was added because it is related to one of the most relevant characteristics of the application evaluated. For metric all questions Q4. to Q6. were designed according to the activities students perform using the application described in [15]. Question Q8. was added because it is related to a fundamental aspect of the application supporting collaboration. Questions Q1., Q2., Q7., and Q9. to Q17. were slightly modified from [5], replacing the word "program", or "learning material" by "application". Questions Q18. to Q26. were elaborated based on the definition of the metrics described in [16].

In addition, to avoid responses with the same tendency, the 26 questions were alternated with 13 positive questions and 13 negative questions.

5. Applying Pedagogical Usability to RedCoMulApp

5.1. Subjects and Settings

The evaluation took place at the Faculty of Economics and Business of the University of Chile, with 12th grade students from nine Santiago mid-income high schools, from July to October, 2017, consisting of six 90-min sessions in total. There were 78 students (62% female) in total, with ages between 15 and 16 years old.

5.2. Procedure

The students worked during a regular 80-min class in each of the six sessions of the language and communications subject. During the first session, the teacher gave 10-min basic instructions about the collaborative activity. Students performed a first test activity for 15 min in order to learn how to use the application. This activity consisted of reading a simple short text and answering three questions. After this preparation, students performed the activity.

Table 1. Questionnaire to capture students' perceptions of pedagogical usability of RedCoMulApp**.**

Pedagogical Usability	
Metrics	**Questions**
M1. **Learner Control**	**Q1**. When I worked on this assignment, I felt that I, not the application, held the responsibility for my own learning. **Q2**. The application does not present information in a format that makes it easy to learn. **Q3**. Seeing anonymously the answers and their justifications of my classmates, makes it easier for me to be sure of my answer, and/or correct it.
M2. **Learner Activity**	**Q4**. Reviewing my answers at the end of the activity does not allow me to be aware of my learning progress. **Q5**. Being able to answer once again an erroneous answer given, supports my learning. **Q6**. Elaborating justifications to my answers, do not allow me to be sure of my selection of the answers that I am giving.
M3. **Collaborative** **learning**	**Q7**. When I am using the application, it lets me know that other users are doing. **Q8**. Exchanging points of view, opinions and comments face-to-face with classmates, does not support my learning. **Q9**. This application lets me talk with my classmates. **Q10**. This application does not support teamwork. **Q11**. It is pleasant to use the application with another student, using my own device.
M4. **Goal Orientation**	**Q12**. The application does not evaluate my learning achievements. **Q13**. This application tells me how much progress I have made in my learning.
M5. **Applicability**	**Q14**. I feel that in the future, I could not use what I learned with the support of this application.
M6. Add value	**Q15**. I learn more with this application, than with other conventional methods.
M7. Motivation	**Q16**. I do not try to achieve a score as high as I can in this application
M8. Feedback	**Q17**. This application provides me with immediate feedback of my activities.
M9. Learning **content relevance**	**Q18**. The text that should be read in the application does not facilitate my task of correctly answering multiple-choice questions. **Q19**. The multiple-choice questions are organized in a clear and coherent way, and support my learning.
M10. **Learning content** **structure**	**Q20**. Following the sequence of three phases of the application does not allow to answer multiple-choice questions effectively. **Q21**. Being able to revise my answers and having the possibility of changing them together with justification, facilitates the achievement of my learning.
M11. **Ease of use**	**Q22**. The application clearly does not support me in what I have to do at each phase of the learning activity. **Q23**. In the application it is easy to follow the sequence of the "individual phase", followed by "anonymous" and "teamwork" collaboration re elaboration phases. **Q24**. The multiple-choice questions to e answered and the justifications to be made have an appropriate format.
M12. **Instruction**	**Q25**. The indications to use the application are clear, and therefore, the intervention of the teacher is not necessary to be able to carry out the learning activity. **Q26**. In general, the task performed with the application was hard to complete.

5.3. Activity Description

The activity was designed according to the goals in reading comprehension for students of 12th grade defined by the language area of the Chilean Ministry of Education. Before starting the activity, the teacher explained the methods and techniques of text reading comprehension. Later, during the experimental sessions, students received an 800 to 900-word text that they had to read and then answer a multiple-choice questionnaire in order to evaluate their level of comprehension of the text they read. In each of the next five sessions, they received a new text to read. The six chosen texts were extracted from the curricular content of the Chilean Ministry of Education along with the corresponding multiple-choice questions.

In each session, the activity started by providing the students with iPads and network access to the RedCoMulApp application on which they had to log in with their personal account and password. The teacher then started the activity on an iPad, activating the "individual" phase for approximately 20 min, where students individually accessed and read the text. After that, they answered 12–14 associated multiple-choice questions. Once all the students finished this phase, they continued to the "anonymous" phase for nearly 30 min, where again each student accessed individually the text to answer the same multiple-choice questions as before, but having access to the answers of their two classmates. At this phase, students could confirm their previously selected options or change their answers. Finally, in the "teamwork" phase, lasting approximately 30 min, the students received from RedCoMulApp the names of their colleagues, who in the previous phase saw

their answers, and with whom they now meet face to face to answer the same multiple-choice questions as a group. In this phase, students are expected to talk to each other, exchange opinions, discuss and sort out their disagreements in order to reach a group answer.

5.4. Instruments

After two days of the sixth session, the students were required to answer the questionnaire from Table 1. The time given to answer this survey was 24 h, and all the students answered it. Similar to other technical usability questionnaires, to help answer the questionnaire, a 5-point adjective-anchored Likert-scale, from "strongly disagree" to "strongly agree", was added to each question.

Additionally, two open-ended questions were added to the questionnaire, in order to capture students' comments regarding aspects of the tool that were positively valued: "What is the best thing you would highlight in this application?" and those that should be improved: "Do you think something could improve in this application?"

5.5. Results

The results analyzed in this Section come from 78 student responses from the questionnaire to capture the perception of the pedagogical usability of RedCoMulApp, together with the analysis of the categories of the 2 additional open-ended questions.

Table 2 shows the frequency distribution of the students answers to the questionnaire from Table 1, along with the mean for each of the 26 questions (Q1. to Q26.), and the mean for each of the 12 metrics (M1. to M12.). Since each question is associated with a 5-point Likert-scale from "strongly disagree" to "strongly agree", the positive responses were scored 1, 2, 3, 4, 5; and the negative questions were scored as 5, 4, 3, 2, 1. Therefore, the Likert-scale applied to the questions of pedagogical usability, quantifies the opinions and attitudes of the students regarding the achievement of the tool in supporting of the educational objective.

The five metrics to measure the pedagogical usability of RedCoMulApp presented in Table 1, which obtained the highest values according to Table 2 are the following:

(a) M3. Collaborative Learning (4.71). Teamwork was perceived as the most outstanding aspect. In particular, questions Q8., Q9. and Q11. are the ones with higher averages, which denote students strongly agreeing with exchanging points of view and opinions through face-to-face conversations.

(b) M6. Add value (4.65). To answer multiple-choice questions similar to those they will have to answer in the PSU test was perceived as a relevant benefit. They also appreciated that in the following two phases, they could learn more, by reworking their answers based on the comments of the answers of others ("anonymous" phase), and by reworking their answers collaboratively with their peers ("teamwork" phase). Students usually prepare for the PSU, by conducting tests consisting of answering only multiple-choice questions. That is, RedCoMulApp turns a summative activity into a formative one.

(c) M1. Learner Control (4.63). This means that according to students, the application enables them to manage the information required to facilitate their learning process.

(d) M11. Ease to use (4.61). According to students' opinion it is easy to perform the tasks of answering to the multiple selection questions and justifying their answers during the three phases of the activity.

(e) M7. Motivation (4.56), and M4. Goal Orientation (4.56). This evaluation shows a high motivation of the students to perform the activity and obtain the highest possible results. Students perceive that RedCoMulApp supports them in their effort of obtaining correct results, informing them about the achieved progress level at the end of the learning activity.

Additionally, the analysis of the answers to the open question "What is the best thing you would highlight in this application?" (see Figure 2), identifies 3 categories with the highest frequencies: "to review justifications of the answers of their classmates", "to interact with classmates" and "to share answers with classmates" (19, 17 and 12 times respectively). This corroborates the results seen for M3., Collaborative Learning. In a similar way, the categories "re-elaboration of answers in phases", "to elaborate a justification for the answer" and "activity divided in phases" with frequencies 13, 10 and 8 respectively corroborate the results for metric M6. Add value.

Table 2. Results of the pedagogical usability. Frequencies of responses to the 5-point Likert-scale. Mt. = Metrics, Q. = Questions. 1 = "strongly disagree". 5 = "strongly agree".

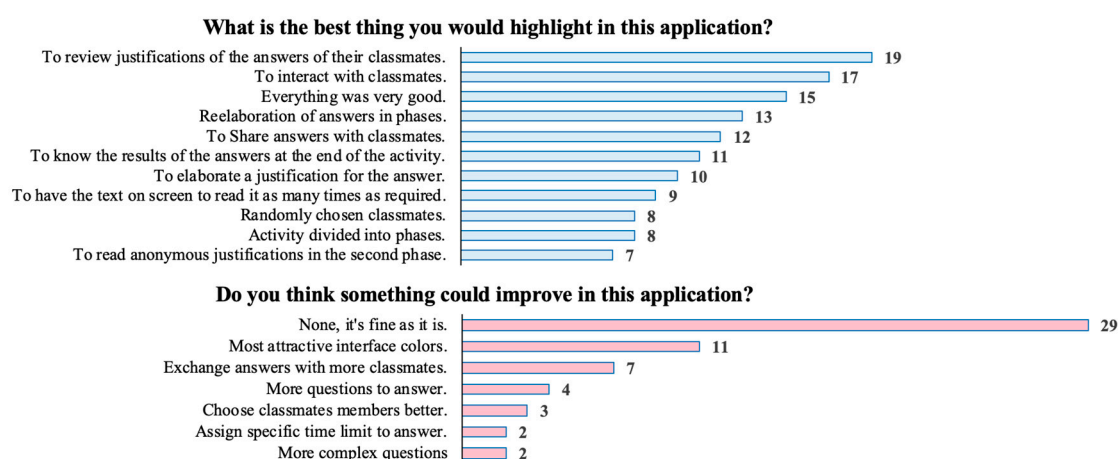| Metrics	Q.	Likert Scale (Frequency)					Mean Q.	Mean Mt.
		1	2	3	4	5		
M1.	**Q1.**	0	0	14	17	47	4.42	4.63
Learner Control	**Q2.**	52	21	5	0	0	4.60	
	Q3.	0	0	2	7	69	4.86	
M2.	**Q4.**	52	21	5	0	0	4.60	4.55
Learner Activity	**Q5.**	0	2	7	22	47	4.46	
	Q6.	49	21	3	1	0	4.59	
M3.	**Q7.**	0	1	4	17	55	4.64	4.71
Collaborative Learning	**Q8.**	65	11	2	0	0	4.81	
	Q9.	0	0	1	16	61	4.77	
	Q10.	53	22	3	0	0	4.64	
	Q11.	0	0	3	16	59	4.72	
M4.	**Q12.**	46	27	4	1	0	4.51	4.56
Goal Orientation	**Q13.**	0	0	2	27	49	4.60	
M5. Applicability	**Q14.**	51	19	5	3	0	4.51	4.51
M6. Add Value	**Q15.**	0	1	6	12	59	4.65	4.65
M7. Motivation	**Q16.**	47	28	3	0	0	4.56	4.56
M8. Feedback	**Q17.**	0	0	7	26	45	4.49	4.49
M9.	**Q18.**	40	21	12	3	2	4.21	4.21
Learning Content Relevance	**Q19.**	0	2	18	19	39	4.22	
M10.	**Q20.**	42	28	8	0	0	4.44	4.52
Learning Content Structure	**Q21.**	0	0	6	19	53	4.60	
M11.	**Q22.**	61	16	1	0	0	4.77	4.61
Ease to use	**Q23.**	0	1	6	21	50	4.54	
	Q24.	48	22	8	0	0	4.51	
M12.	**Q25.**	0	0	9	16	53	4.56	4.54
Instruction	**Q26.**	49	21	8	0	0	4.53	
							Total	**4.55**

Figure 2. Responses to open-ended questions. We counted a total of 129 aspects to highlight because of the frequency students mentioned them, which were grouped in 11 categories and 58 aspects to improve, which whose frequencies were categorized in 7 groups.

From the rest six metrics, four of them (M2. Learner Activity, M5. Applicability, M10. Learning Content Structure, and M12. Instruction) have a metric over 4.5 which is a high (considering that the highest is 5.0). M8. Feedback, has a value of 4.49, which is close to the other ones.

The lowest value corresponds to the M9. metric, Learning Content Relevance. We attribute this result to circumstantial facts, since in a certain way this value depends on the level of complexity that

students perceive through the texts they have to read, and the multiple-choice questions chosen by the teacher based on the contents proposed by the Ministry of Education. Figure 2 shows the frequencies of the category's analysis of student responses to the open-ended of the questionnaire.

We identified 129 aspects classified in 11 categories for the answers given to the question "What is the best thing you would highlight in this application?" The most frequent ones were the review of the justification of the answers given by the classmates (19 times) and the possibility of interacting face-to-face with them during the "teamwork" phase (17 times). Fifteen students mentioned that in general, the whole application was well designed without mentioning a particular aspect. The re-elaboration of the answers in three phases obtained the 4th highest value with 13 mentions. Other frequently mentioned aspects were the possibility of sharing answers with their peers in both "anonymous" and "teamwork" phases (12 times), knowing at the end of the activity the correct and incorrect answers given in each phase (11 times), elaborating a justification for each response given in each phase (10 times), and the possibility of reading the base text anytime on which the questions were answered (9 times). Less frequent mentioned aspects, each with a frequency close to 10% of the total sample: the random choice of co-workers (8 times), the activity divided into three phases "individual", "anonymous" and "teamwork", and the possibility of reading the justifications given by the teammates anonymously (7 times).

We identified 58 aspects for the answers given to the question "Do you think something could improve in this application?" Half of the answers (29) fell under the category of "No, it is fine as is". From the remaining 29 aspects, we identified six categories. The most mentioned was using more attractive colors for the interface. The next one was the suggestion of incorporating and option for exchanging answers with more classmates, which can be easily implemented since the teacher configures the number of participants for each group at the beginning of the activity, which can range between one and the whole class. For the future, we will experimentally evaluate implications regarding the number of members that a work team can introduce in the learning activity.

Finally, less mentioned categories were: (a) add more questions; (b) choose other teammates; (c) assign a specific time span for answering the questions, and (d) introduce questions that are more complex. All these aspects can be implemented with the current version of RedCoMulApp since:

(a) In the creation of the educational activity, the teacher can add as many questions as wanted.
(b) In addition to the random criterion to form the members of the team works, other criteria can be used: based on the responses collected in the "individual" phase, either between members with the same answers, or with different answers; or based on the academic performance of the students.
(c) The teacher can verbally inform the students of the time they have in each phase. Once this time has elapsed, the teacher activates the next phase. The application does not specifically implement a functionality that establishes a time limit to answer the questions in each of the phases.
(d) Teacher chooses the complexity of the activity at its creation time.

Given the above, the analysis of categories of the open questions allows us to conclude that RedCoMulApp has a highly positive perception on the part of the students. The improvements they suggested, except for the redesign of the colors of the interface, are aspects that can be handled by the current version of RedCoMulApp and have to do more with the contents and configurations of the activity than with changes in the application.

Therefore, according to the results of Table 2 analyzed above and the categories analysis of the two open-ended questions, we can state that the pedagogical usability perception of the students is positive for the 12 metrics described in Table 1. This means that RedCoMulApp, according to the students' perception, is a tool that has a satisfactory pedagogical utility for which it was designed.

6. Conclusions and Future Work

This study presents an approach for evaluating pedagogical usability aspects of a mobile learning application RedCoMulApp, designed for support reading comprehension of 12th grade high school students, see Section 3 for a brief description of the application, and [15] for more details. In

order to capture the pedagogical usability, we used metrics that were combined from [5,16], with tailored questions for each metric that consider the criteria that had been operationalized in [5].

Therefore, we applied the reading comprehension learning activity supported by RedCoMulApp with real users, and report results of this study measuring the perception of pedagogical usability that 78 students have about RedCoMulApp. Considering the design characteristics of this application, a questionnaire was designed with 26 questions to be answered with a 5-point Likert-scale (from "strongly disagree" to "strongly agree"), based on 12 metrics, plus two open-ended questions to capture aspects that were positively valued, and those that should be improved. The results obtained for the pedagogical usability validates that students find the design of the application pedagogically useful for learning about reading comprehension. This validation is important, since the method based on multiple-choice with 3 phases of re-elaboration ("individual", "anonymous" and "teamwork") is usually a methodology to easily develop formative and not only summative evaluations. That is, we can show that we manage to convert a summative evaluation into a formative one that from the point of view of the students, has an added value.

As future work, we envisage using RedCoMulApp to perform experimental tests with content other than reading comprehension, such as mathematics, and history, geography and social sciences, and with different levels of difficulty in the questions. Likewise, we will carry out tests with different sizes of members of each work team, to identify the implications that this can introduce in learning, and pedagogical usability. For example, we expect to perform tests with work teams with a small, large, or with specific sizes of team members: 3, 4, 5, 10, etc. Additionally, we would like to carry out tests with different team works grouping criteria; for example, based on the correct results given to multiple-choice questions in the "individual" phase: those who answered the same correct answers, with those who answered incorrectly, or well between those who responded correctly, or among those who responded incorrectly.

References

1. Nielsen, J. Evaluating hypertext usability. In *Designing Hypermedia for Learning*; Springer: Berlin/Heidelberg, Germany, 1990; pp. 147–168.
2. Squires, D.; Preece, J. Usability and learning: Evaluating the potential of educational software. *Comput. Educ.* **1996**, *27*, 15–22.
3. Gould, D.J.; Terrell, M.A.; Fleming, J. A usability study of users' perceptions toward a multimedia computer-assisted learning tool for neuroanatomy. *Anat. Sci. Educ.* **2008**, *1*, 175–183.
4. Koohang, A.; du Plessis, J. Architecting usability properties in the e-learning instructional design process. *Int. J. E Learn.* **2004**, *3*, 38–44.
5. Nokelainen, P. An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *J. Educ. Technol. Soc.* **2006**, *9*, 178–197.
6. Vukovac, D.P.; Kirinic, V.; Klicek, B. A comparison of usability evaluation methods for e-learning systems. *DAAAM Int. Sci. Book* **2010**, 271–288. Available online: https://www.daaam.info/Downloads/Pdfs/science_books_pdfs/2010/Sc_Book_2010-027.pdf (accessed on 18 November 2019).
7. Cheung, A.C.; Slavin, R.E. Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Read. Res. Q.* **2013**, *48*, 277–299.
8. Van Staden, S.; Zimmerman, L. Evidence from the progress in international reading literacy study (PIRLS) and how teachers and their practice can benefit. In *Monitoring the Quality of Education in Schools*; Brill Sense: Leiden, The Netherlands, 2017; pp. 123–138.

9. Zurita, G.; Baloian, N.; Jerez, O.; Peñafiel, S. Practice of skills for reading comprehension in large classrooms by using a mobile collaborative support and microblogging. In Proceedings of the CYTED-RITOS International Workshop on Groupware, Saskatoon, SK, Canada, 9–11 August 2017; pp. 81–94.

10. DiBattista, D.; Sinnige-Egger, J.-A.; Fortuna, G. The "none of the above" option in multiple-choice testing: An experimental study. *J. Exp. Educ.* **2014**, *82*, 168–183.

11. Park, J. Constructive multiple-choice testing system. *Br. J. Educ. Technol.* **2010**, *41*, 1054–1064.

12. Bennett, R.E.; Rock, D.A.; Wang, M. Equivalence of free-response and multiple-choice items. *J. Educ. Meas.* **1991**, *28*, 77–92.

13. Santelices, M.V.; Catalán, X.; Horn, C. University admission criteria in Chile. In *The Quest for Equity in Chile's Higher Education: Decades of Continued Efforts*; Rowman & Littlefield: Lanham, MD, USA, 2018; p. 81.

14. Dixson, D.; Worrell, F.C. Formative and summative assessment in the classroom. *Theory Pract.* **2016**, *55*, 153–159.

15. Zurita, G.; Baloian, N.; Jerez, O.; Peñafiel, S.; Pino, J.A. Findings when converting a summative evaluation instrument to a formative one through collaborative learning activities. In Proceedings of the International Conference on Collaboration and Technology, Costa de Caparica, Portugal, 5–7 September 2018; pp. 1–16.

16. Ivanc, D.; Vasiu, R.; Onita, M. Usability evaluation of a LMS mobile web interface. In Proceedings of the International Conference on Information and Software Technologies, Kaunas, Lithuania, 13–14 September 2012; pp. 348–361.

17. Sung, Y.-T.; Chang, K.-E.; Liu, T.-C. The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Comput. Educ.* **2016**, *94*, 252–275.

18. ISO/TC 159/SC 4 Ergonomic of human-system interaction (Subcommittee). *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs).: Guidance on Usability International Organization for Standarization.* 1998; p. 45. Available online: https://www.sis.se/api/document/preview/611299/ (accessed on 18 November 2019).

19. Junior, F.S.; Ramos, M.; Pinho, A.; Rosa, J.S. Pedagogical Usability: A theoritical essay for e-learning. *Holos* **2016**, *32*, 3–15.

20. Khomokhoana, P.J. Using Mobile Learning Applications to Encourage Active Classroom Participation: Technical and Pedagogical Considerations. Ph.D. Thesis, University of the Free State, Bloemfontein, South Africa, 2011.

21. Hussain, A.; Kutar, M. Apps vs devices: Can the usability of mobile apps be decoupled from the device? *IJCSI Int. J. Comput. Sci.* **2012**, *9*, 3.

22. Kukulska-Hulme, A. Mobile usability in educational contexts: What have we learnt? *Int. Rev. Res. Open Distrib. Learn.* **2007**, *8*, doi:10.19173/irrodl.v8i2.356.