

Quality Evaluation of Mobile Learning Applications

Gustavo Willians Soad
University of São Paulo (USP)
Institute of Mathematics and
Computer Science - ICMC
São Carlos-SP, Brazil
Email: gustavo.soad@usp.br

Nemésio F. Duarte Filho
Federal Institute of São Paulo (IFSP)
campus São João da Boa Vista
São João da Boa Vista-SP, Brazil
Email: nemesiofreitas@gmail.com

Ellen Francine Barbosa
University of São Paulo (USP)
Institute of Mathematics and
Computer Science - ICMC
São Carlos-SP, Brazil
Email: francine@icmc.usp.br

Abstract—Mobile learning has emerged as a new and promising learning modality, providing more interactivity and flexibility to learners, tutors and teachers in carrying out educational activities and practices. However, its adoption, still poses problems and challenges, such as the quality evaluation of the resulting learning applications. In this paper, we discuss the establishment of a method for evaluating the quality of mobile learning applications. The development of the method was based on a systematic review of quality characteristics of mobile learning applications, which resulted in a quality characteristics catalog for this type of application. The proposed quality method comprises: (i) a quality model, which establishes a well-defined set of criteria and requirements; (ii) metrics, i.e., the quality requirements are mapped into attributes that can be measured. To make the measurements was developed a checklist, which comprises 80 questions, classified into eight different groups responsible for the establishment of quality criteria; and (iii) judgment criteria, i.e., for each defined checklist answer a punctuation is assigned, allowing the question to receive a value characterizing its level of quality. The method was applied in the evaluation of three mobile applications for English teaching: Duolingo, Winguia, and Rosetta Stone. In order to evaluate these three applications, they have been tested by a specialist in the software quality area and for each of the applications, the proposed checklist was answered. From the results obtained, we can conclude that the method proposed is adequate for evaluating the quality of mobile learning applications.

I. INTRODUCTION

The advent of mobile technology has contributed to the appearance of a new type of learning modality, called mobile learning (m-learning) [1], [2], [3]. This kind of learning is characterized by providing a strong interaction among learners, teachers and tutors. Allowing them to contribute, participate and access the learning content using mobile devices (mobile phones, tablets, laptops, smartphones, among others), at any time and anywhere.

Despite the benefits provided, mobile learning also presents weaknesses and challenges to be considered. Concern the weaknesses, we can cite techniques that are related to the hardware, for instance: (i) low storage capacity; (ii) battery life; (iii) keyboard size; and (iv) screen size [4], [5], [6]. In addition, there are non-technical weaknesses, for instance: (i) difficulty in handling mobile technology; (ii) motivation to study outside of class time; and (iii) avoid distracting points because mobile devices offer many features for its users [4], [7], [8].

The quality evaluation of m-learning applications is one of the issues to be addressed, especially due to lack of methods to behold specific quality aspects of a m-learning application. Although we can find different quality evaluation methods for software products [9], [10], [11], they are still very generic, making it difficult to evaluate the m-learning applications in deep.

Motivated by this scenario, in this work, we aim at establishing a quality evaluation method for mobile learning applications. The method has been developed based on a set of ISO/IEC standards [12], [13], [9] and on the Duarte Filho and Barbosa work [14], which established a specific process for evaluating the quality of m-learning environments. It is important to notice that our domain of research is broader, since m-learning environments are a particular type of m-learning applications. We have applied the proposed method into the quality evaluation of three mobile applications for the teaching of the English language. The preliminary results achieved suggest the practical feasibility of adopting the method proposed in the quality evaluation of mobile learning applications.

This paper is organized as follows. In Section II the background for our work is summarized. In Section III we describe the main aspects of our quality evaluation method. In Section IV we discuss the use of the method for evaluating the quality of mobile applications. Finally, in Section V, we present our conclusions and perspectives for future work.

II. BACKGROUND

In this section, we briefly present the concepts related to m-learning as well as the existing quality evaluation methods for software products in general.

A. Mobile Learning

Mobile learning can be defined as any type of education or learning that occurs when the learner is not in somewhere predetermined or fixed place, or when the individual takes advantage of learning opportunities offered by mobile technologies, thus combining the concepts of technology and mobility [2].

The use of m-learning applications allows individuals be more productive when they consume, create or interact with information through mobile and portable devices, because

these devices accompanying the individual on a regular way from the start to finish of their daily tasks [3]. In another related perspective, m-learning can be defined as the ability of using handheld devices to access learning resources [15].

With the evolution of technology and digital inclusion, it is clear the importance of m-learning for current society. The benefits offered by m-learning are beyond the accessibility, convenience and communication. Through mobile devices, learners can use specific learning environments, collaboration tools, social networks, e-books, didactic content, among others.

There are several researches investigating the benefits that m-learning can bring as well as the limitations found in their use. Some studies suggest that m-learning causes the interest of the learners, encourages social interaction and stimulates discussion about the course materials [4], [8], [6].

On the other hand, m-learning still has several limitations to be addressed. For instance, technical limitations may be related to hardware such as: low storage capacity and battery autonomy, which limits the use of the mobile device for a long period of time, or even the size of the screens and keyboards, which creates limitations on the display and data entry in the application [5], [4], [6].

In addition to the technical limitations mentioned, there are non-technical limitations as well. For instance, the need for motivating learners to study outside the educational setting so that the lack of interest can be overcome, does not prevent the completion of the learning activities proposed [4].

In the scope of our work, the concern for quality becomes an important issue to be addressed. Quality, in this setting, refers not only to the technological view, but also to pedagogical, socio-cultural and socio-economic aspects [16]. Despite its relevance, few studies about quality guidelines for m-learning applications can be found in the literature. Thus, it is difficult to identify well-defined and widely used supporting mechanisms for the quality evaluation of mobile learning applications.

B. Quality Evaluation Methods

A quality model consists of a set of characteristics that relate to each other and which provides the basis for specifying quality requirements and quality evaluation [12].

In short, the quality of a software product can be evaluated from the perspective of the development process, which relates to activities performed during the stages of the life cycle, and from the perspective of the product, which refers to the product characteristics after its development.

To establish ways to ensure the quality of a software product, different models and standards have been established. Among the existing quality models, two of them are considered seminal in context of quality models: the McCall model [10] and the Boehm model [11]. These models define several characteristics that are still presented in the current models, such as reliability and portability.

Additionally, there are the models proposed by ISO (International Organization for Standardization) and IEC (International

Electrotechnical Commission), which also aim at establishing ways to ensure the quality of a software product. The series of standards ISO/IEC 9126 [12] defines a quality model in which product quality characteristics are presented. Moreover, the series of standards ISO/IEC 14598 [13] presents a framework for definition of an evaluation process for software products.

More recently, ISO/IEC has launched the ISO/IEC 25000 [9], also called SQuaRE (Software Product Quality Requirements and Evaluation) model. ISO/IEC 25000 unifies the ISO/IEC 9126 and ISO/IEC 14598 standards, creating a logically organized set of standards, which involves two main processes: (i) the requirements specification; and (ii) the software quality evaluation. ISO/IEC standards have been considered in the establishment of our quality evaluation method for m-learning applications, described next.

III. A METHOD FOR QUALITY EVALUATION OF M-LEARNING APPLICATIONS

The evaluation of mobile learning applications was made by adapting the quality evaluation process for m-learning environments, proposed in Duarte Filho and Barbosa [14] for the m-learning applications context. The adaptation was necessary, because there are differences between the evaluation of m-learning environment and a m-learning application, such as example: (i) issues related to recovery environment in case of failures; (ii) characteristics related to tax and legal audits performed in the environment; and (iii) generally it consists of more than one system and has a staff to keep all the infrastructure operating.

To assist in updating and adapting the evaluation proposal for the mobile applications context, a quality characteristics catalog for m-learning applications defined in Soad et al. [17] was used. In this catalog we propose a set of characteristics divided into technical (functionality, performance, usability, security, portability), educational (pedagogical), sociocultural (communication) and socioeconomic (support). This catalog was developed, based on a systematic review that aimed to find work related to the models and quality characteristics for m-learning applications. Figure 1 represents the catalog divided between the characteristics that were used for the quality evaluation.

To assist in understanding the model defined, some characteristics are briefly described next:

- **Pedagogical:** the pedagogical characteristic of a mobile application are intends to address the issues related to teaching and learning.
 - Knowledge at the right time: the content should be offered at the right time, avoiding offer content to teach something that depends on a content that the learner have not learned yet;
 - Reuse of learning content: application's ability to provide the reuse of content;
 - Cognitive effort of activities: it refers to the cognitive effort spent by the user in using the application and the acquisition of knowledge.;

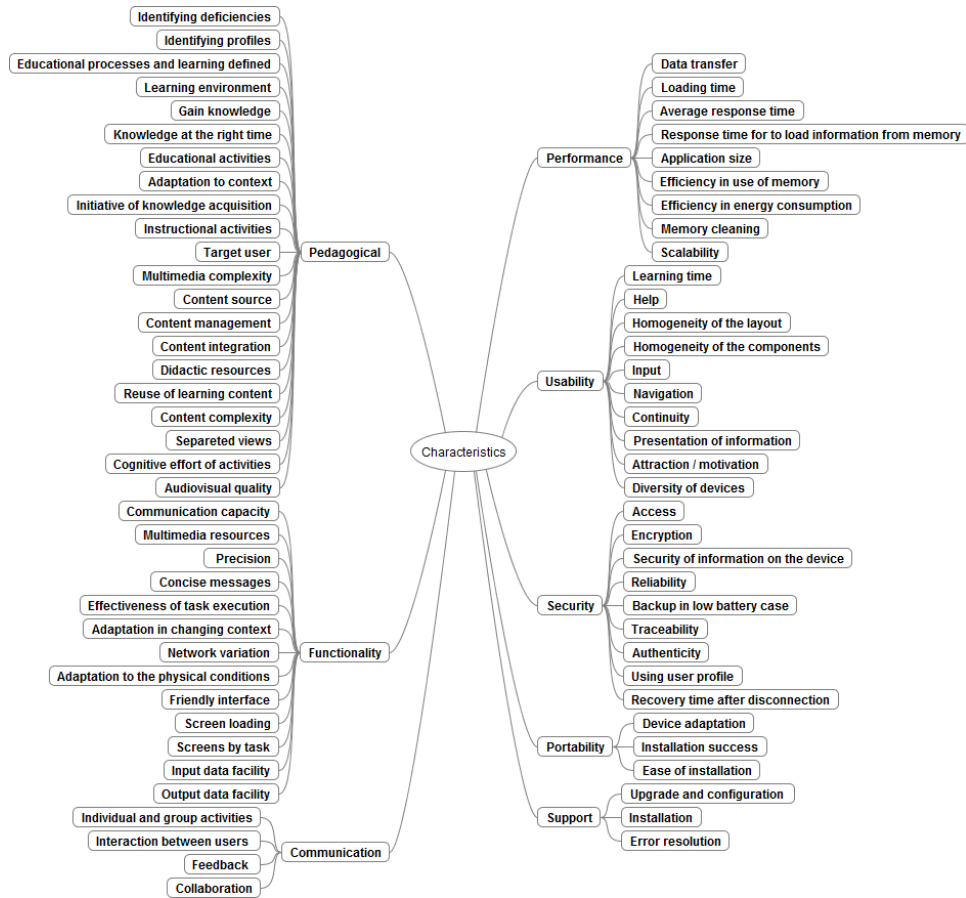


Fig. 1. Quality characteristics catalog for m-learning applications

- **Multimedia complexity:** the application should not have many multimedia elements, as these can cause distraction points for the user;
- **Separated views:** refers to the separation of the views of the student and tutor, or according to each user profile the application should provide different views;
- **Content complexity:** contents should not have too much information on lessons and should be offered gradually, thus preventing a complex content can affect the user's understanding;
- **Content integration:** it refers to the integration of content, providing a defined flow between them;
- **Content management:** it refers to the control, management and monitoring of the content offered;
- **Functionality:** it refers to the availability of functions that meet the implicit and explicit requirements for the application.
 - **Effectiveness of task execution:** ability to facilitate the execution of tasks, minimizing the time, cost of connection and the amount of data transmitted;
 - **Communication capacity:** data transmission capabilities between devices;
 - **Screens by task:** number of screens used during the execution of a task;
 - **Precision:** ability to provide accurate results;
- **Multimedia resources:** ability to support different types of media such as video, text, audio, among others.
- **Communication:** it refers to the ability to provide communication and interaction among users.
 - **Individual and group activities:** the activities should provide the possibility of learning individually or in group;
 - **Feedback:** application's ability to allow reciprocal feedback between learners and tutors;
 - **Interaction between users:** the application must allow interaction among its users, providing this feature through blogs, forums, groups, among others.
- **Performance:** it refers to performance relative to the amount of resources used under stated conditions, i.e., it refers to optimize the use of available resources.
 - **Loading time:** it refers to the waiting time until the requested information be fully loaded;
 - **Efficiency in energy consumption:** Due to the limitations of the batteries of the devices, their use efficiently is essential thus avoiding energy waste;
 - **Memory cleaning:** it refers to the unnecessary data storage on the device, such as files that are no longer used by the application;
- **Usability:** refers to the application's ability to be used by specific users in a specific context, offering ways to enable the

application to be understood, learned and used.

- **Learning time:** time spent by the user to learn how to use the application;
 - **Homogeneity of the layout:** related to the standardization of components, maintaining their characteristics for all application screens;
 - **Homogeneity of the components:** related to various types of components that provide the same functionality for this reason it is suggested the minimum variation among the components;
 - **Input:** it should not be required the entry of a lot of data, since there are limitations on the screen and keyboard;
 - **Presentation of information:** capacity limitation of the information according to the screen display area;
 - **Diversity devices:** it refers to the variety of existing devices, allowing exclusive usability improvements according to the device used.
- **Security:** it refers to the application's ability to protect information and data, protecting them from unauthorized access and guaranteed access according to the different authorization levels;
- **Reliability:** ability to provide access only to authorized users and applications;
 - **Traceability:** refers to storage of records of events occurring in the application, and can thus track the user-made paths;
 - **Recovery time after disconnection:** application's ability to re-establish the connection after a disconnection and shall maintain its operation due to these unexpected interruptions;
 - **Authenticity:** it refers to the guarantee of authenticity of the actions of user;
- **Portability:** refers to the application of capacity of being transferred from one hardware or software to another.
- **Device adaptation:** application's ability to adapt to different types of mobile devices;
 - **Installation success:** capacity of the application to be installed and uninstalled on different mobile devices.
- **Support:** refers to the support services offered by the application, aiming to offer an efficient support to users.
- **Upgrade and configuration:** the application must be updated and configured quickly and automatically.;
 - **Error resolution:** the application should provide help (by the application itself or by contacting the application's support) for error correction.

Figure 2 provides an overall view of the quality evaluation method for mobile learning applications we have worked on.

In the establishment of our method, we have considered the quality evaluation process for m-learning environments, proposed in Duarte Filho and Barbosa [14]. In short, we have followed this process in the context of m-learning applications (which is broader than m-learning environments). The process is composed of six steps, as following:

- **Define objectives and contextualization of the evaluation:** this step aims at establishing the goals of the quality evalua-

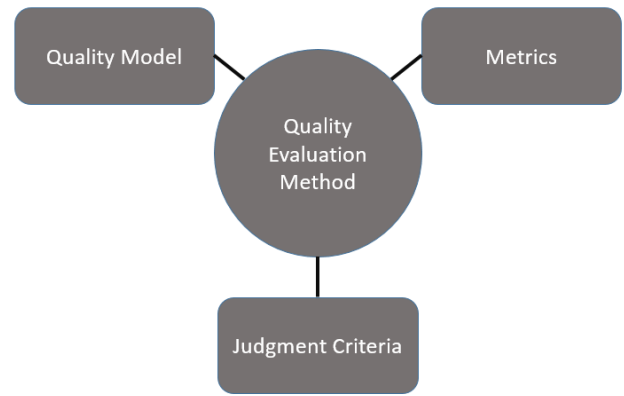


Fig. 2. Quality evaluation method for m-learning applications

tion, as well as the domain of the mobile learning applications considered for evaluation.

- **Define and configure the quality model:** the quality model is responsible for establishing a well-defined set of criteria and requirements for m-learning applications. The model used herein is based on the best practices and knowledge established by ISO/IEC 14598 [13] and ISO/IEC 25000 [9], and on the Duarte Filho and Barbosa work [18], updated and adapted according to a systematic review performed in the context of quality characteristics for mobile learning applications which was presented earlier. More detailed information on the systematic review conducted can be found in Soad et al. [17].

- **Identify the evaluation metrics:** in this step, the quality requirements are mapped to attributes that can be measured. In short, measurable attributes are set based on the knowledge of experts, through creating simple and direct questions, that are evaluated and scored by a checklist. For our evaluation proposal, we used the checklist defined in Duarte Filho and Barbosa [14], adapted to the context of m-learning applications. We highlight that the changes made are also in accordance with the results achieved in the systematic review performed by Soad et al. [17].

For example, the checklist for environments has the following question about the quality characteristic: "Does the environment offers to the system administrator, functionalities about the statistics of product usage, such as user management, created items, notifications, among others?". This question is not well suitable for m-learning applications, because the application is for personal use and the system administrator role overtake the boundary of applications. For this reason this question is used in the environments context and not in application context. Therefore, the question was replaced by: "Does the application provides data about its use? (e.g. performance in class, hours of use, scores, etc)". The above question can be evaluated in the application context, thereby changing the checklist to adapt to the proposed scenario.

The resulting checklist comprises a total of 80 questions, classified into eight different groups, which establish the quality criteria for mobile learning applications. Table I summarizes the quality criteria defined as well as their main

objectives.

TABLE I
QUALITY CRITERIA FOR M-LEARNING APPLICATIONS

Quality Criteria	Main Objectives
Functionality	Questions of this group evaluate functionality aspects, such as the accuracy and ease of performing tasks.
Security	Questions of this group evaluate security aspects, such as confidentiality, authenticity, traceability and integrity.
Performance	Questions of this group evaluate performance aspects, such as the behavior with respect to time, capacity and resource utilization.
Pedagogical	Questions of this group evaluate pedagogical aspects, such as didactic content, interactivity and learning.
Usability	Questions of this group evaluate aspects of usability, ensuring that the m-learning application can be attractive to the learner.
Support	Questions of this group evaluate aspects related to support, such as incident handling and assistance in resolving errors.
Communication	Questions of this group evaluate the application communication aspects, such as notifications and collaboration tools.
Portability	Questions of this group evaluate aspects of portability, such as ease of adaptation and installation to different systems.

Table II illustrates some of checklist questions related to the quality criteria proposed.

- Define score levels and judgment criteria: to define the score levels for the checklist questions, firstly we have to specify the type of nominal answers that will be applied to the questions. For each defined answer, a score is assigned that allowed a question receive the minimum value 0 and maximum 10. Table III shows the possible answers to the questions and their respective scores.

TABLE III
DEFINED SCORES FOR THE POSSIBLE ANSWERS

Options	Score
Yes	10
No	0
Efficient	10
Acceptable	7.5
Regular	5
Deficient	2.5
Inefficient	0

The judgment criteria were defined based on the Martinez' work [19], in which three levels of judgment were established: Superior, Medium and Lower. To achieve the Superior level, the application must be in conformance with at least 80% of the specified quality requirements, i.e., the final result can vary between 80 and 100, on a scale from 0 to 100. Similarly, for the Medium level the application must be in conformance with at least 50% of the quality requirements, and for the Low level the application has a conformance less than 50%.

- Design the evaluation: the evaluation should be designed according to the specified guidance of system domain. In our case, the evaluation was designed for the mobile learning applications context.

- Execute the evaluation: in the last stage, the evaluator should perform three main tasks: (i) collect quality measures; (ii) compare the measures with regard to predefined quality

criteria; and (iii) evaluate the data obtained during the evaluation.

IV. APPLYING THE QUALITY EVALUATION METHOD

In order to conduct a preliminary validation of the proposed quality evaluation method, we have selected three mobile applications, all of them used in the teaching of English. The applications were chosen through searches in *Google Play*¹, at the educational category. The selection criteria considered were the learning domain (teaching of English) and the number of application installations by users. At the end of the selection process, the following applications were selected (Figure 3): (i) Duolingo² (released version April 14, 2015); (ii) Wlingua³ (released version April 28, 2015); e (iii) Rosetta Stone⁴ (released version March 4, 2015).

A. Duolingo Evaluation

In order to evaluate Duolingo, we firstly applied the checklist (available in <https://goo.gl/c3hYqy>), which was answered (80 questions) by a specialist. The specialist responsible for evaluating the applications has a deep knowledge about software quality and is involved in research in this area. Table IV present a set of questions related to Usability and Table V shows the scores obtained for each question related to the Usability characteristic.

TABLE IV
USABILITY CHECKLIST QUESTIONS

Number	Questions
1	How the standardization of the application can be evaluated? (photos, letters, colors, menus, etc.)
2	Are the names of titles, menus and labels in accordance with their representation in the application context?
3	Does the application have some demonstration of its functions?
4	Does learners with different types of experiences or ages have equal opportunity to succeed with the established functions?
5	What is the degree of learning that users have with respect to the application?
6	Does the application provide help through tutorials and manuals or is the functions intuitive allowing the discard of this kind of help?
7	Does the application present an effective separation of its action components (eg. buttons, checkbox, menu, etc.) to avoid that they can be triggered accidentally?
8	Can the application be used and controlled easily and efficiently?
9	What is your evaluation in relation the "friendly" interface of the application?
10	How are evaluated the error messages generated by the application?

From the total of 100 points related to the usability characteristic, Duolingo achieved 90 points. The mobile application did not achieve the full score for usability because of the lack of demonstrations of its functionalities. Even though, Duolingo was classified at the Superior level with respect to Usability.

¹Google Play: <http://play.google.com>

²Duolingo: <https://goo.gl/Ho3CGW>

³Wlingua: <https://goo.gl/fcLG4C>

⁴Rosetta Stone: <https://goo.gl/BRNAaA>

TABLE II
CHECKLIST QUESTIONS (FULL CHECKLIST IS AVAILABLE IN [HTTPS://GOO.GL/C3HYQY](https://goo.gl/C3HYQY))

Quality Criteria	Question
Functionality	Do the present functions in the application meet the specific needs of tutors and learners?
	Does the application have a search engine, facilitating the search for specific functions?
	...
Security	Does the application make a count and lock on the amount of times the user tried to login and failed?
	In case of failures over a transaction (eg. disconnection), is the information entered into the application kept?
	...
Performance	Is the application loaded in a satisfactory time?
	How does the application behave with respect to performance of response time? (How long does it take to process the execution of a feature that runs on the local environment?)
	...
Pedagogical	Can the activities and educational materials used in the application be developed through text, images, graphics, multimedia, games, sound, among others?
	Does the m-learning application enable greater interactivity using mobile devices as interactive agents (eg., using photos, sensitivity, location, graphics processing, sound, etc.)?
	...
Usability	How the standardization of the application can be evaluated? (photos, letters, colors, menus, etc.)
	Are the names of titles, menus and labels in accordance with their representation in the application context?
	...
Support	Does the application provide help in solving problems, such as a list of common problems and their solutions, thus allowing the user find the solution of the problem?
	Does the application present some mechanism against voluntary / involuntary user errors? For example, enter invalid data in certain fields.
	...
Communication	Does the application provide performance feedback for learners instantly (through score tests, comparative responses, among others)?
	Does the application have specific tools for collaboration tasks (through wikis, games, microblogs and forums)?
	...
Portability	Can the m-learning application be used in an efficient and correct way in different mobile devices?
	How does the application behave running and accessed from different mobile devices? (eg. tablets, smartphones, cell phones, etc.)



Fig. 3. Selected mobile learning applications for English teaching

TABLE V
SCORES OBTAINED IN USABILITY EVALUATION OF DUOLINGO

Question	Score	Question	Score
1	10	6	10
2	0	7	10
3	10	8	10
4	10	9	10
5	10	10	10
Total	90		

Figure 4 summarizes the results obtained for the complete evaluation of Duolingo, considering the total set of quality characteristics.

Based on the overall results, Duolingo was classified at the Medium level of quality. The highest scores were obtained for Usability, Performance and Portability. The lowest score refers to the Support characteristic. Table VI summarizes the strengths and weaknesses of each quality criterion evaluated for Duolingo.

B. Overall Results

Our quality evaluation method was applied to the other selected m-learning applications, thus evaluating their quality in accordance with the characteristics previously defined. To compare the answers about usability, Table VII and Table

TABLE VI
DUOLINGO QUALITY EVALUATION: STRENGTHS AND WEAKNESSES

Criteria	Weaknesses	Strengths
Functionality	Lack of a search system to find specific features in the application; lack of a mechanism for adapting content according to available resources.	Use of different media features (e.g. video, audio, images, etc) and data presentation about the application usage (e.g. performance in class).
Security	The application does not provide any kind of lock against indiscriminate login attempts without the correct password by the user. Also, the user is not informed about how safe the application is.	The application was available throughout the evaluation process; even when simulating a failed connection, the information entered in the application was kept.
Performance	The application does not have any connection measurement mechanism to warn the user if the connection does not support the application.	The application presented a satisfactory response time throughout the evaluation process, even during the communication with the server.
Pedagogical	The application can not automatically detected contextual information linked to the user's context, thus providing a generic learning material.	The structure of the learning context enables more interactivity using the mobile device as an interactive agent and provides different types of learning activities.
Usability	the application does not present demonstrations about its features.	It has a friendly interface, with standardized and intuitive features.
Support	When simulating a serious error, the application does not behave satisfactorily (the error is not reported, the current job is not recorded and the current version of the application is not displayed).	The application can be modified according to user needs.
Communication	The application does not provide tools for collaborative tasks. Also, it does not have mechanisms for complaint messages with offensive content.	The application provides feedback on the learners' performance, displays event notifications to the user, and enables and disables notifications by the application itself.
Portability	The application does not allow off-line browsing, i.e., it becomes unusable without an Internet connection.	The application is constantly updated and the new version is configured automatically. Also, it is compatible with different operating systems.

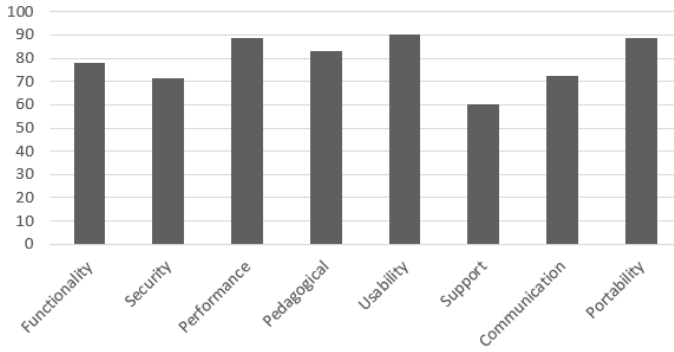


Fig. 4. Duolingo: scores achieved for each quality criterion

VIII shows the answers of each question. Through the data in the tables, its possible identify the questions about quality that influence to the Usability characteristic not reached the maximum rating of the evaluation.

TABLE VII
SCORES OBTAINED IN USABILITY EVALUATION OF WLINGUA

Question	Score	Question	Score
1	10	6	0
2	0	7	10
3	10	8	10
4	10	9	7.5
5	10	10	7.5
Total		75	

TABLE VIII
SCORES OBTAINED IN USABILITY EVALUATION OF ROSETTA STONE

Question	Score	Question	Score
1	0	6	10
2	0	7	10
3	0	8	0
4	10	9	10
5	7.5	10	10
Total		57.5	

Additionally to the of the characteristics presented, the Figure 5 summarizes the results achieved for each quality characteristic, for each m-learning application.

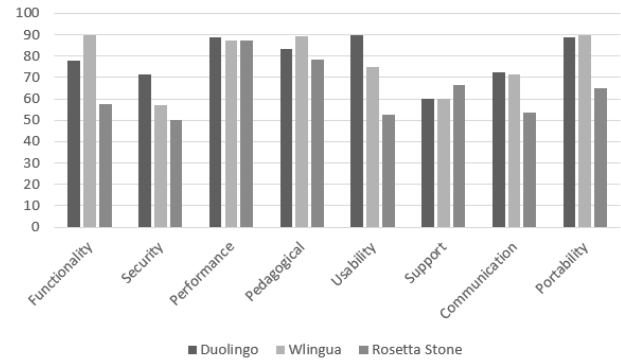


Fig. 5. Overall scores achieved for each quality criterion

From the results obtained, we observed that Duolingo and Wlingua have similarities in the scores for each characteristic, while for the Rosetta Stone application we noticed a greater difference among the characteristics evaluated. Figure 6 presents the overall results of the quality evaluation for such applications.

Duolingo, Wlingua and Rosetta Stone were all classified in the range of Medium quality mobile learning applications. However, Rosetta Stone presented a Lower level with respect to the others. Next we summarize the main advantages and limitations of the applications evaluated.

- **Duolingo:** Advantages: (i) free use; (ii) its interface is pleasant, intuitive and with standard components; and (iii) it allows a greater interactivity using various types of activities. Disadvantages: (i) the security level is not informed to register the password; (ii) there is no mechanism to report offensive messages; and (iii) some features do not have demonstrations of use.

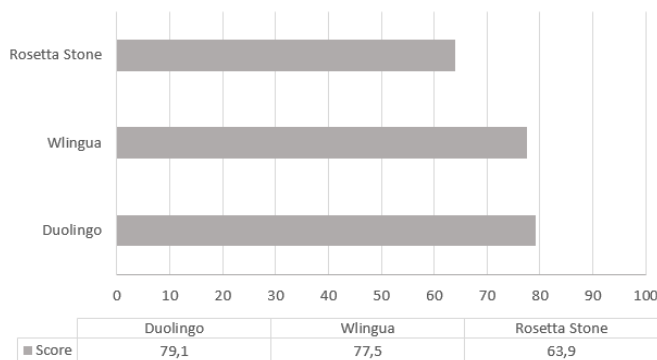


Fig. 6. Final scores achieved for each m-learning application

- **Wlingua:** Advantages: (i) free version (limited for one month); (ii) in case of instability in the connection, the application is able to adapt itself in order to keep its use; and (iii) information storage avoids data loss. Disadvantages: (i) the security level is not informed to register the password; and (iii) the application can not be modified according to the user needs.

- **Rosetta Stone:** Advantages: (i) technical support; (ii) many resources that enable a greater interactivity; and (iii) good performance in communicating with external resources. Disadvantages: (i) buttons with confused features; (ii) it does not support using the application in the portrait mode; and (iii) keyboard is not adjusted in accordance with the content context (e.g., when you type an e-mail, the keyboard does not display the symbol “@”).

Besides the specific disadvantages pointed out, there are some common problems among the applications as well: (i) they do not allow the use of the application in off-line mode; (ii) the collaboration is not encouraged; also, there are no available resources for this; (iii) there is no blockage when trying to repeatedly authenticate with incorrect password; and (iv) assistance and manuals mechanisms are not provided.

V. CONCLUSIONS AND FUTURE WORK

In this paper we presented a method for the quality evaluation of m-learning applications, considering technical, pedagogical, economic and socio-cultural aspects. In order to validate our proposal, the method was applied in the quality evaluation of three m-learning applications for English teaching. The preliminary results obtained suggest that the method is adequate for evaluating the quality of mobile learning applications.

Despite the promising results observed, we highlight the need of a more formal and complete validation. This validation has been planned and will require efforts to evaluate a considerable set of m-learning applications, for different knowledge domains and considering distinct types of users.

We also intend to evolve the method, prioritizing the quality characteristics according to the type of users, such as applications developers, teachers, tutors and learners, for instance.

ACKNOWLEDGMENT

The authors would like to thank the Brazilian funding agencies - FAPESP (Process 2014/03389-9), CAPES (Procad 071/2013) and CNPq.

REFERENCES

- [1] K. C. Nah, P. White, and R. Sussex, “The potential of using a mobile phone to access the internet for learning eff listening skills within a korean context,” *ReCALL*, vol. 20, no. 03, pp. 331–347, 2008.
- [2] C. O’Malley, G. Vavoula, J. Glew, J. Taylor, M. Sharples, and P. Lefrere, “Guidelines for learning/teaching/tutoring in a mobile environment. mobilearn project deliverable,” *Retrieved March*, vol. 24, p. 2013, 2003.
- [3] S. Wexler, J. Brown, D. Metcalf, D. Rogers, and E. Wagner, “Mobile learning: What it is, why it matters, and how to incorporate it into your learning strategy,” *Guild Research*, 2008.
- [4] R. Picek and M. Grcic, “Evaluation of the potential use of m-learning in higher education,” in *Information Technology Interfaces (ITI), Proceedings of the ITI 2013 35th International Conference on*, June 2013, pp. 63–68.
- [5] Y.-M. Huang, Y.-T. Lin, and S.-C. Cheng, “Effectiveness of a mobile plant learning system in a science curriculum in taiwanese elementary education,” *Computers & Education*, vol. 54, no. 1, pp. 47–58, 2010.
- [6] M. Al-Ani, S. Hameed, and L. Faisal, “Students’ perspectives in adopting mobile learning at university of bahrain,” in *e-Learning “Best Practices in Management, Design and Development of e-Courses: Standards of Excellence and Creativity”*, 2013 Fourth International Conference on, May 2013, pp. 86–89.
- [7] C.-H. Su and C.-H. Cheng, “A mobile game-based insect learning system for improving the learning achievements,” *Procedia - Social and Behavioral Sciences*, vol. 103, no. 0, pp. 42 – 50, 2013, 13th International Educational Technology Conference.
- [8] S. Fazlina, A. Manap, and R. Rias, “Mobile learning awareness among students at higher learning institutes: A case study,” in *Informatics and Creative Multimedia (ICICM), 2013 International Conference on*, Sept 2013, pp. 226–229.
- [9] . ISO/IEC 25000, *International Organization for Standardization/ International Electrotechnical Comission ISO/IEC 25000: Systems and software Quality Requirements and Evaluation (SQuARE)*. ISO/IEC, 2005.
- [10] J. P. Cavano and J. A. McCall, “A framework for the measurement of software quality,” *ACM SIGSOFT Software Engineering Notes*, vol. 3, no. 5, pp. 133–139, 1978.
- [11] B. W. Boehm, J. R. Brown, and M. Lipow, “Quantitative evaluation of software quality,” in *Proceedings of the 2nd international conference on Software engineering*. IEEE Computer Society Press, 1976, pp. 592–605.
- [12] . ISO/IEC 9126, *ISO/IEC 9126. Software engineering – Product quality*. ISO/IEC, 2001.
- [13] . ISO/IEC 14598, *ISO/IEC 14598: Information Technology - Evaluation of Software Products*. ISO/IEC, 1998.
- [14] N. F. Duarte Filho and E. F. Barbosa, “A contribution to the quality evaluation of mobile learning environments,” in *Frontiers in Education Conference, 2013 IEEE*, Oct 2013, pp. 379–382.
- [15] S. J. Kinshuk, E. Sutinen, and T. Goh, “Mobile technologies in support of distance learning,” *Asian Journal of Distance Education*, vol. 1, no. 1, pp. 60–68, 2003.
- [16] A. A. Economides, “Requirements of Mobile Learning Applications,” *International Journal of Innovation and Learning*, vol. 5, 2008.
- [17] G. W. Soad, N. F. Duarte Filho, and E. F. Barbosa, “Uma contribuição ao estabelecimento de características de qualidade para aplicações educacionais móveis,” 2015.
- [18] N. F. Duarte Filho and E. F. Barbosa, “A requirements catalog for mobile learning environments,” in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, ser. SAC ’13. New York, NY, USA: ACM, 2013, pp. 1266–1271.
- [19] M. Martinez, G. Azevedo, S. Lopes, P. Pagliuso, R. Colombo, M. Rodrigues, and M. Jino, “The software product evaluation database-supporting mede-pros,” in *Software Engineering Standards, 1999. Proceedings. Fourth IEEE International Symposium and Forum on*, 1999, pp. 182–191.