An Empirical Investigation on a Pedagogical Pattern Language for Mobile Learning Applications

Maria Lydia Fioravanti*, Camila Dias de Oliveira*, Lilian Passos Scatalon*, Ellen Francine Barbosa*
*Institute of Mathematics and Computer Science, University of São Paulo (ICMC-USP), São Carlos, SP – Brazil
{mlfioravanti, camila_oliveira, lilian.scatalon}@usp.br, francine@icmc.usp.br

***Abstract*—This Research to Practice Full Paper presents an empirical investigation of a pedagogical pattern language for mobile learning apps. Mobile learning (m-learning) applications can provide several benefits to learners. However, there are still several problems to be further investigated, such as the challenge to keep the learner motivated while using the application during educational activities. The application design can help to solve this kind of problems. There is a need to understand such problems in the pedagogical level and then properly eliciting requirements that would address those problems. Pedagogical patterns try to capture expert knowledge regarding the practice of teaching and learning in a way that it is possible for others to reuse this experience. Therefore, pedagogical patterns can be a tool to assist in the design of new m-learning applications as well as to the improvement of the existing ones. Aiming to bridge this gap, a pedagogical pattern language entitled MLearning-PL has been proposed to guide the requirements elicitation of m-learning applications projects. We conducted an empirical investigation in order to evaluate the proposed pattern language by means of a experimental study, whose main purpose was to evaluate the solutions provided using MLearning-PL in comparison to an ad hoc approach. A feedback questionnaire was also applied to understand the participants' perceptions about MLearning-PL on its clearness and completeness. The obtained results provided preliminary evidences of the applicability, effectiveness and efficiency of MLearning-PL.**

I. Introduction

Computational learning applications play a key role in educational activities, both in academia and in industry. In this scenario, mobile learning has emerged as a new and promising learning modality [1], [2], providing more interactivity and flexibility to learners, tutors and teachers in carrying out educational activities and practices. However, despite providing several benefits and facilities, mobile learning applications also present problems and challenges that need to be better investigated. These problems and challenges are not only limited to technical aspects, but also include pedagogical issues, such as keeping the learner motivated to avoid dropouts, dealing with different learning styles (e.g., sensory, intuitive, visual, auditory and so forth) [3], guiding the learner in self-learning, among others.

In another research line, the term "pattern" has the meaning initially given by Christopher Alexander for architectural patterns [4], [5]: "each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way you can use this solution a million times over, without ever doing it the same way twice". In other words, a pattern is an abstract solution to a recurring problem in a given context. In this sense, a pattern language is a structured collection of patterns that rely on each other to transform requirements and constraints into an architecture [6].

Patterns and pattern languages are mechanisms to describe best practices, good designs, and capture experience in a way that it is possible for others to reuse this experience. In a related perspective, pedagogical patterns try to capture expert knowledge regarding the practice of teaching and learning [7].

Aiming to address the problems and challenges presented in mobile learning applications, a pedagogical pattern language entitled MLearning-PL [8] has been proposed to guide the requirements elicitation phase of mobile learning applications projects.

In order to validate the proposed pattern language, we have conducted an experimental study, according to Wohlin et al. [9] guidelines. The main idea was to evaluate the effectiveness, efficiency and applicability of MLearning-PL, in the context of mobile learning applications requirements elicitation, in comparison to an ad hoc approach. The subjects of this experimental study were people involved in educational activities, i.e., teachers and tutors. They had to solve an activity containing different situations involving pedagogical problems (some of them using MLearning-PL as a support, and some of them using an ad hoc approach). In general, the results suggested that MLearning-PL provides better results than using an ad hoc approach.

The remainder of this paper is organized as follows. In Section II related work is presented. An overview of MLearning-PL is presented in Section III. In Section IV, we discuss the planning of the experimental study and the pilot study conducted. The conduction of the experimental study is briefly discussed in Section V. The results are presented in Section VI. Then, Section VII analyzes the threats to validity. Finally, our conclusions and perspectives for future work are drawn in Section VIII.

II. Related Work

Besides the technical aspects, it is important to consider the pedagogical issues related to mobile learning when defining the requirements of such applications. According to Bergin et al. [7], most educators and trainers are not taught how to teach, since the undergraduate courses do not address important pedagogical aspects to form a teacher. Rather, they often find themselves teaching by accident [7]; in other words, typically

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

a person with a skill that is in demand will be asked to teach it. However, knowing the subject matter is very different from knowing how to teach it.

In this context, pedagogical patterns, also known as learning patterns, try to capture expert knowledge regarding the practice of teaching and learning [7]. The goal is to capture the essence of the practice in a compact form that can be easily communicated to those who need such knowledge. Presenting this information in a coherent and accessible form can mean the difference between the need of every new instructor to relearn what is already well known by senior faculty. In this sense, pedagogical patterns can easier knowledge transfer of teaching within the community. Therefore, pedagogical patterns can be important tools to assist in the design of new teaching and learning applications as well as to the improvement of the existing ones.

Joseph Bergin [10] presented 14 pedagogical patterns for the development of courses in Computer Science. The patterns have varying degrees of application, from the overall organization of a course to classroom practices. The patterns are presented in order of scale, starting with those involving the planning of semester courses to daily activities in classrooms. Afterwards, *The Pedagogical Patterns Project*[1] put together several educators interested in contributing to the formalization of those pedagogical patterns, in addition to others that have been incorporated into a pattern language.

In a related perspective, researchers at Keio University [11] conducted a project entitled *Learning Patterns Project*, in which they developed learning patterns to support the learning of university students. Other initiatives in the research of pedagogical patterns, especially related to teaching with Learning Management Systems [12], [13], can also be found in the literature.

III. MLearning-PL: an Overview

Despite the initiatives previously mentioned, to the best of our knowledge, there was no pedagogical pattern language focusing on the mobile learning context. Aiming to bridge this gap, we proposed MLearning-PL[2][8], a pedagogical pattern language whose main idea is to provide support on pedagogical issues to help analysts in avoiding or diminishing already known pedagogical problems. In a real scenario, suppose a geography teacher who wants to create a mobile application to teach geography to her/his students. Since he/she takes on the role of requirements analyst, MLearning-PL is helpful during this process to support the teacher in eliciting the pedagogical requirements of the mobile application. He/she can consult MLearning-PL to mitigate the risk of not considering some important pedagogical aspects.

In general terms, MLearning-PL is a pedagogical pattern language for mobile learning applications, to assist in the definition of mobile applications in order to keep learners motivated and committed to using such applications, considering their different learning styles and an effective knowledge acquisition. MLearning-PL comprises 14 patterns, related as shown in Figure 1.

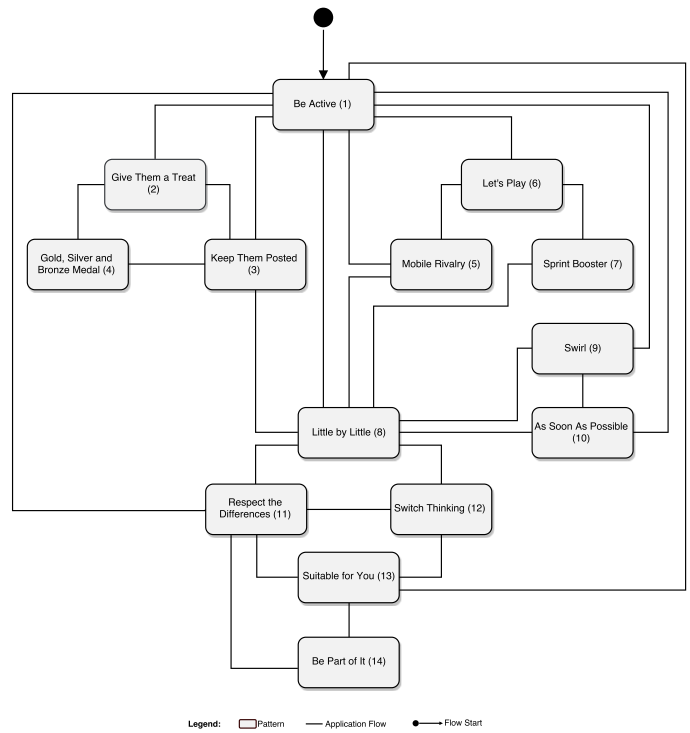


Fig. 1: MLearning-PL (Extracted from Fioravanti and Barbosa (2017)[8])

All patterns address some strongly interrelated aspects: *Engagement*, *Motivation*, *Learning style* and *Knowledge effectiveness*. *Keep Them Posted*, for instance, gives a solution to keep the learner motivated to learn while using the app. *Little by Little* helps in the context of avoiding the situation of students becoming bored and disinterested.

IV. Planning of the Experimental Study

In this section, we report the planning of the experimental study regarding goals, hypotheses, variables and analysis procedure.

Lab packages[3] containing more information and details on the experimental study were created and are available. A lab package describes an experiment providing materials for its replication, highlights opportunities for variation, and builds a context for combining results of different types of experimental treatments [14].

A. Goals

The main goal of the experimental study was to evaluate MLearning-PL in relation to an ad hoc approach, i.e., no specific approach is used, but previous background. Thus, the object of study is MLearning-PL, our pedagogical pattern language.

[1]http://www.pedagogicalpatterns.org/
[2]https://sites.google.com/usp.br/mlearning-pl/
[3]https://sites.google.com/usp.br/mlearning-pl/experimental-studies

The Goal/Question/Metric (GQM) method was proposed to support the definition of quantifiable goals and interpretation of collected measurement data [15]. It is a goal-oriented approach to derive metrics from measurement goals to ensure the collected data are usable and serve a purpose. The scope of the experimental study can be summarized by the GQM template, as follows:

Analyze MLearning-PL
For the purpose of Evaluation
With respect to its Effectiveness and Efficiency
From the point of view of the Requirements Analysts
In the context of Teachers and/or Tutors solving pedagogical problems

Effectiveness is defined as the degree to which something is successful in producing a desired result. In our experimental study, it is related to whether MLearning-PL can successfully help the identification of solutions to pedagogical problems.

Efficiency concerns the quality or property of being efficient, i. e., producing at minimum waste and expense, or unnecessary effort. In our experimental study, it is related to use of only the time necessary for solving the problems.

B. Hypotheses Formulation

Once the aim was the comparison of two different approaches for pedagogical problems solving, the research questions were formalized into the following hypotheses, so that statistical tests could be conducted:

RQ1. Does the use of MLearning-PL help analysts to provide better pedagogical solutions to m-learning problems?

- Null hypothesis: The effectiveness of MLearning-PL is lower than or equal to that of an ad hoc approach. $H1_0$: Effectiveness(MLearning-PL) $\leq$ Effectiveness (ad hoc)
- Alternative hypothesis: The effectiveness of MLearning-PL is higher than that of an ad hoc approach. $H1_a$: Effectiveness (MLearning-PL) $>$ Effectiveness (ad hoc)

RQ2. Does the use of MLearning-PL lead analysts to solve pedagogical problems faster?

- Null hypothesis: The efficiency of MLearning-PL is lower than or equal to that of an ad hoc approach. $H2_0$: Efficiency (MLearning-PL) $\leq$ Efficiency (ad hoc)
- Alternative hypothesis: The efficiency of MLearning-PL is higher than that of an ad hoc approach. $H2_a$: Efficiency (MLearning-PL) $>$ Efficiency (ad hoc)

C. Variables

Some aspects of the experimental study were kept constant for both approaches used; the approach itself was the aspect that varied. Considering the constant aspects, i.e., the parameters of the experiment [16], both groups had a minimum background in educational activities, they received a training about mobile learning and also solved the same activities.

An experimental study must be carefully designed for the drawing of meaningful conclusions. In our case, the design that best suited the experimental study was *one factor with two treatments*. The independent variable was the *problem resolution approach* and the two treatments applied were:

- *MLearning-PL*, when the subject had the pattern language to support the problem solving process; and
- *ad hoc*, when the subject had no further artifact to support the process, but only previous background.

The subjects' performance was evaluated according to two different aspects, which were the dependent variables, namely *effectiveness* and *efficiency*. Effectiveness was measured with three different metrics:

(i) *correctness* – average percentage of problems solved correctly, calculated as

$$\frac{\sum_{i=1}^{n} \frac{s_i}{m} 100}{n}$$

where
s_i is the number of problems solved by subject i;
m is the total number of problems to be solved; and
n is the number of subjects in the group.

(ii) *completeness* – average score of solutions' completeness, calculated as

$$\frac{\sum_{i=1}^{n} \frac{\sum_{p=1}^{m} x_i}{m}}{n}$$

where
x_i is the score of completeness;
p is the number of the problem;
m is the total number of problems to be solved; and
n is the number of subjects in the group.

The score of completeness for each solution is measured according to an adaptation of Degree of Closeness (DoC) Evaluation Criteria [17] shown in Table I:

TABLE I: Degree of Completeness (Adapted from McCracken et al. [17])

Score	Interpretation
5	Touchdown. The solution is complete.
4	Close but something missing. The solution is satisfactory, but some details are missing.
3	Close but far away. The solution can be applied, but it is far from being complete.
2	Close but even farther away. The solution is not complete enough.
1	Not even close. The solution shows that the subject had no idea about how to solve the problem.

(iii) *complexity* – average score of solutions' complexity, calculated as

$$\frac{\sum_{i=1}^{n} \frac{\sum_{p=1}^{m} x_i}{m}}{n}$$

where

x_i is the score of complexity;

p is the number of the problem;

m is the total number of problems to be solved; and

n is the number of subjects in the group.

The score of complexity for each solution also follows an adaptation of Degree of Closeness (DoC) Evaluation Criteria [17] shown in Table II:

TABLE II: Degree of Complexity (Adapted from McCracken et al. [17])

Score	Interpretation
5	Touchdown. The solution has a high degree of complexity.
4	Close but something missing. The solution has a medium degree of complexity.
3	Close but far away. The solution has a low degree of complexity.
2	Close but even farther away. The solution is simple.
1	Not even close. The solution shows that the subject had no idea about how to solve the problem.

Metric *efficiency* was then defined to be measured as the average time spent on the solving of all problems, calculated as

$$\frac{\sum_{i=1}^{n} x_i}{n}$$

where

x_i is the time spent in solving the problems by subject i and

n is the number of subjects in the group.

D. Analysis Procedure

The analysis procedure included both quantitative and qualitative components. Quantitative data were the first source for the testing of the hypotheses, whereas the qualitative data were analyzed in order to complement the quantitative analysis, since a small sample of subjects was available.

Considering the quantitative analysis, univariate analyses of the dependent variables *Effectiveness* and *Efficiency* were performed to test the hypotheses. Mann-Whitney tests were performed for all dependent variables, with a level of significance set to $\alpha = 0.05$. In short, the Mann-Whitney test is a non-parametric test that compares two independent samples of the same size or unequal, whose scores are measured by, at least, an ordinal level [18].

A qualitative analysis was also conducted, once it could potentially offer additional complementary evidences. To perform this analysis, a *feedback form* was applied and was further analyzed for a better understanding of the subjects' experience.

E. Pilot Study

According to Mendonça et al. [19], observational studies or pilot studies are important to reduce the risks associated with defining a new controlled experiment, since they aim at improving the experimental material. Observational studies are run to sanity test the experimental design, test and improve the experiment material, and to check the timing of the experimental activities [19].

Therefore, prior to performing our experimental study, two pilot studies were conducted using the first version of MLearning-PL[4].

The first pilot study aimed at validating the experimental instruments (forms, activities, artifacts) and also to train the instructors. It was conducted with four subjects, providing input for the improvement of the experimental instruments.

After, a second pilot study was conducted. It involved the participation of 8 subjects and aimed at improving the experimental design and also the experimental object. More details on the protocol and the results obtained can be found at the lab package[5]. According to the results of the hypothesis test, MLearning-PL provided better results than the ad hoc approach, maybe due the systematization introduced in the process by means of using patterns, since systematic approaches are usually better than ad hoc approaches and, particularly, patterns can be used to solve common problems. However, such results were preliminary and only could be concluded with an experimental study involving a larger number of participants and an improved version of the pattern language, refined using the results of this pilot.

According to the obtained results, subjects were enthusiastic and positive about MLearning-PL and its importance, which indicates positive evidence on its use over an ad hoc approach for supporting the pedagogical problem-solving process. Such results, associated with patterns experts' opinion, helped to improve and refine the pattern language that was used to conduct the experimental study, presented next.

V. Experimental Study

As discussed in Section IV-E, pilot studies provided inputs for the improvement of MLearning-PL. The pattern language was substantially changed and the version used in the experimental study[6] comprised 14 patterns, i. e., additional seven new patterns. The experimental study was conducted in this scenario. In this section, we report on the details of its conduction.

A. Experimental Subjects

As the experimental study aims to analyze the results from the point of view of requirements analysts for mobile learning applications, individuals previously or currently involved in educational activities were recruited, no matter if such activities involved face-to-face, blended or distance learning. 15 Computer Science students from the Institute of Mathematics and Computer Science at University of São Paulo (ICMC-USP), who fullfilled these requirements, performed the tasks in an afternoon. From the 15 participants, eight carried out

[4]https://goo.gl/i1N3L5
[5]https://sites.google.com/usp.br/mlearning-pl/experimental-studies
[6]https://goo.gl/Bprqyk

the activities using MLearning-PL pattern language, while the other participants did not use any further artifact.

A *subject characterization form*[7] was applied prior to the experimental study; Table III shows the information gathered from the participants. The sampling of the population was a *stratified random sampling* and the subjects' experience was considered for their division into two groups.

The presence of subjects with no experience in both groups is important, once the results show whether experience is a determining factor in the use of MLearning-PL.

B. Experimental Objects

The subjects performed six activities in a fixed order, which were our experimental objects. The activities consisted of situations involving pedagogical problems to be solved.

Firstly, a general context was given for a better understanding of a real situation. The initial text was: *"Suppose you are a specialist consultant of your research topic and have been hired to assist in the implementation of a mobile educational application to teach concepts related to your area. Developer analysts have already listed the technical requirements of the application, however they rely on your advice and pedagogical experience to clearly specify what pedagogical requirements the mobile application must meet. It is known that motivation and engagement in activities is a major issue for learners when using distance education tools, thus you should assist them in their doubts using your experience in tutoring and/or teaching in your area of expertise. In this scenario, how would you help them solve the following problems?"*.

After that, the participant was asked to solve each one of the following activities.

- **Activity 1:** The initial motivation on using the application is temporary and learners may feel bored, resulting on dropouts. How can you mitigate this problem in the mobile app?
- **Activity 2:** Keeping learners engaged in the learning process when there is no on-site periodic engagement can be difficult. Therefore, it is necessary that they continue to exercise the knowledge acquired to achieve the established goal, besides knowing what stage they are at before reaching that goal. How can this be handled in the mobile app?
- **Activity 3:** When a topic takes longer to be presented than the learner's concentration time, this may bore him/her and lead to difficulties in the learner's understanding, especially when the learner has almost achieved his/her goal. How can this problem be avoided in the mobile app?
- **Activity 4:** When learning a new subject, the learner must exercise the new knowledge and skills acquired. However, if the learner considers the proposed activity too difficult, he/she may feel frustrated and discouraged. On the other hand, if the activity is considered too simple by some learners, they may have the feeling that they are not learning any new content. How to deal with this situation so that all learners carry out the activities and remain engaged?
- **Activity 5:** You want to ensure the effectiveness of the knowledge acquired by the learners, but the theory may not be as simple as it seems due to the difficulties that arise during this process. Even though you want to enable students to solve meaningful problems as early in the course as possible, you cannot overload them because they must remember the most important ideas and concepts. How to deal with this situation without overloading learners?
- **Activity 6:** You are teaching learners with different backgrounds and characteristics, which means they have different learning styles. How to deal with this situation in the mobile app to provide content that is suitable for several kinds of learners?

[7]https://goo.gl/rbyJTz

VI. Results

A. Quantitative Analysis

The two hypotheses were tested regarding the effectiveness and efficiency of MLearning-PL in comparison to the ad hoc approach. The calculations performed provided the following measures for each established metric, shown in Table IV.

The Mann-Whitney test was applied and the U-values and p-values obtained are shown in Table V.

Since U is not smaller than 1 in any of the cases, the null hypotheses cannot be rejected. The U-test is not statistically significant in this case, because the p-value is higher than α for the two metrics, therefore, neither of the null hypotheses could be rejected.

According to Table IV, considering correctness, the median of MLearning-PL (100%) is higher than that of the ad hoc approach (83.33%) (Figure 2), which indicates subjects that used MLearning-PL could solve more problems correctly than those who used the ad hoc approach.

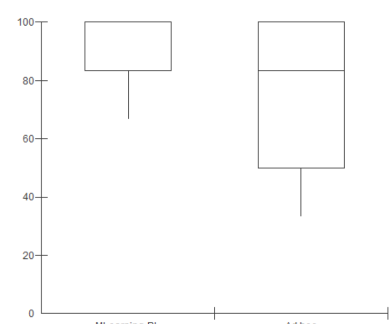


Fig. 2: Percentage of problems solved correctly

The results are similar for completeness and complexity. When MLearning-PL was used, the medians were 3.67 and 3.29 (Figures 3 and 4) for completeness and complexity of the solutions, respectively, whereas the use of the ad hoc approach obtained 2.5. The first experimental study showed differences of 0.35 and 0.3, respectively, for the completeness

TABLE III: Experimental Study: Subjects' Background

Subject	What is your experience as a teacher?	What is your experience as a tutor or teaching assistant?	What is your experience with mobile learning applications?	Approach used
1	None	Between 1 and 6 months	None	Ad hoc
2	More than 5 years	More than 3 years	Up to 1 year	Ad hoc
3	Up to 1 year	Between 1 and 6 months	Up to 1 year	Ad hoc
4	Between 3 and 5 years	Between 1 and 6 months	None	Ad hoc
5	None	Between 6 months and 1 year	None	Ad hoc
6	Up to 1 year	Between 1 and 3 years	None	Ad hoc
7	None	Between 1 and 3 years	None	Ad hoc
8	None	Between 1 and 6 months	Up to 1 year	MLearning-PL
9	None	None	None	MLearning-PL
10	None	Between 1 and 3 years	None	MLearning-PL
11	Up to 1 year	Between 1 and 3 years	None	MLearning-PL
12	Up to 1 year	Between 1 and 6 months	None	MLearning-PL
13	None	Between 1 and 6 months	Up to 1 year	MLearning-PL
14	None	Between 1 and 6 months	None	MLearning-PL
15	More than 5 years	Between 1 and 6 months	None	MLearning-PL

TABLE IV: Experimental Study: Measures

	Subject		Effectiveness		Efficiency
		correctness	*completeness*	*complexity*	*efficiency*
Ad hoc	1	100	3.33	2.83	47.43
	2	33.33	1.75	1.67	20.35
	3	50	2.08	2.08	11.61
	4	83.33	2.50	2.50	39.37
	5	100	3.00	2.75	38.81
	6	100	2.92	2.83	34.36
	7	50	2.33	2.08	26.38
	Median	83.33%	2.50	2.50	34.36
MLearning-PL	8	100	4.17	3.75	35.03
	9	83.33	3.58	3.25	44.53
	10	100	3.92	3.42	32.77
	11	100	3.33	2.83	34.04
	12	83.33	3.75	3.33	23.73
	13	100	3.58	3.25	42.59
	14	100	4.08	3.75	28.50
	15	66.67	2.58	2.83	23.57
	Median	100%	3.67	3.29	33.41

TABLE V: Experimental Study: Values for Mann-Whitney test

Metric	U-value	p-value
correctness	18.5	0.1358
completeness	3.5	0.0023
complexity	2	0.0013
efficiency	27	0.4539

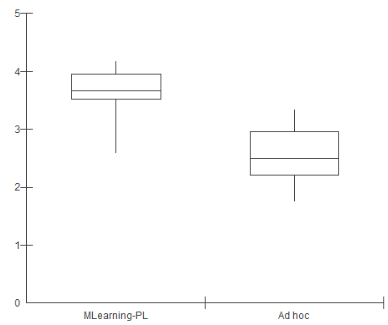


Fig. 3: Solutions' completeness

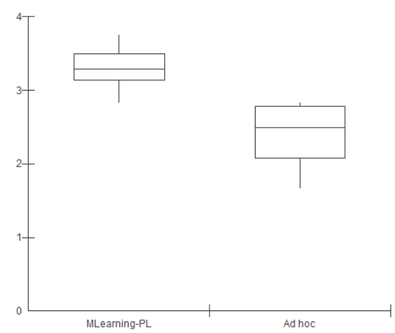


Fig. 4: Solutions' complexity

and complexity of the solutions obtained with the use of MLearning-PL in comparison to the ad hoc approach. The differences are, respectively, 1.17 and 0.79, which are higher than those previously obtained, and reinforce MLearning-PL assists in the obtaining of more complete and complex solutions.

Finally, concerning the time spent on the tasks, Figure 5 shows the median for MLearning-PL is 33.41 minutes against 34.36 for the ad hoc approach. The difference is much smaller than that previously obtained, probably due to the larger number of patterns available, since the pattern language was substantially changed.

According to the results of the hypothesis test, MLearning-PL provided better results than the ad hoc approach, maybe due the systematization introduced in the process by means of using patterns, since systematic approaches are usually better than ad hoc approaches and, particularly, patterns can be used to solve common problems.

B. Qualitative Analysis

The answers provided by the participants and their perceptions regarding the proposed activities were considered for the qualitative analysis of the results.

The subjects that used MLearning-PL were more inclined to reach a solution closer to that expected from the *Activities*

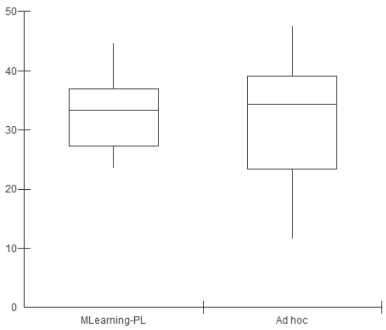


Fig. 5: Time to solve all the problems (in minutes)

proof template[8]. The time spent on the solving of the problems was similar in both groups, probably because: (i) the subjects that used the ad hoc approach did not know how to answer the questions in detail and did not take longer time detailing the answer; (ii) the subjects that used MLearning-PL were more careful and analytical to answer the questions; or (iii) the use or not of an extra artifact did not influence the execution time of the task.

Regarding the *feedback questionnaires*, the subjects' opinion about the difficulty in performing each activity is summarized in Figure 6. In general, we could not infer the use or not of a given approach facilitated the resolution of each problem. The opinions are balanced (with slight differences) between both treatments for most activities. In the first activity (Figure 6a), for instance, two subjects considered it hard, even with the use of MLearning-PL as a supporting artifact. Similarly, in the second and third activities (Figure 6b and Figure 6c), a subject that used MLearning-PL considered the activity very hard. Figure 6d shows balanced opinions between both treatments concerning the fourth activity. On the other hand, the scenario is different for the fifth and sixth activities, shown in Figure 6e and Figure 6f, respectively. The large majority of subjects that used MLearning-PL as a supporting mechanism considered the activities easy, whereas those who used an ad hoc approach were divided.

It is worth mentioning that previously discussed difficulty levels of the activities are only subjects' perceptions on it, which may vary due several reasons that are not related to the experimental study. In general, more experienced subjects were prone to consider the activities easier than those with less or no experience, regardless of the approach used, which suggests the difficulty is more related to each subject's background.

The participants who performed the activities with the ad hoc approach were asked if any further artifacts would be useful in the problem-solving process. Figure 7a shows that five subjects agreed and two neither agreed, nor disagreed on the usefulness of an artifact, which indicates the problem-solving activity is not trivial and may benefit from the use of additional artifacts.

[8]https://goo.gl/dBzLt3

The participants who used the MLearning-PL approach were asked about the pattern language used. Firstly, we asked how helpful it was to support the performed activities and most of them agreed or strongly agreed that MLearning-PL helped in the the problem-solving process (Figure 7b). We also asked their opinions on the completeness and clearness of the pattern language. Regarding completeness, the results were not unanimous (Figure 7c), i. e., 50% (four subjects) agreed it was complete, 25% neither agreed, nor disagreed and 25% disagreed. Although the subjects considered MLearning-PL not complete enough, they did suggest improvement points to the pattern language, such as new patterns or even modifications of the existing ones. Finally, regarding clearness (Figure 7d), MLearning-PL was pointed out as clear and easy to understand, since 87,5% of the subjects answered "agree".

The subjects were enthusiastic and positive about MLearning-PL and its importance, which indicates positive evidences on its use to support the pedagogical problem-solving process. The experimental results were used as input to improve and refine the pattern language.

VII. Threats to Validity

We discuss the threats to validity of the experimental study based on the guidelines proposed by Wohlin et al. [9].

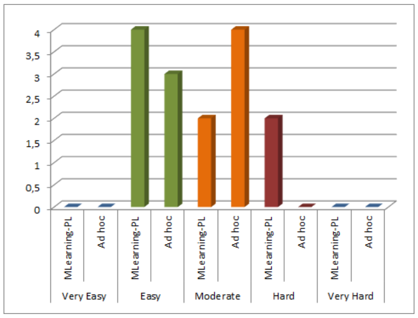
Conclusion validity is concerned with the relationship between the treatment and the outcome. A potential threat is the size of the available sample may not be sufficient to provide significant results; therefore, we included descriptive statistics and qualitative analysis towards mitigating this threat.

The subject's involvement in educational activities (as a teacher or tutor) might be another threat; therefore, we used the characterization form to gather information on the subjects' background and offered a tutorial on mobile learning.

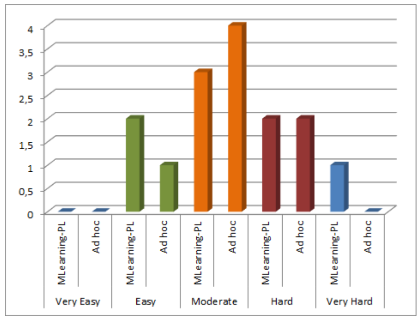
Internal validity is concerned to whether the treatment causes the outcome. The following potential threats were identified:

(i) Each participant's time of experience in education may be a systematic variation in the experimental study. To reduce it, we adopted stratified random sampling to balance the groups.
(ii) The selection of problems to be solved by the subjects may not represent the problems faced in educators' daily basis. To mitigate it, the problems were validated by a specialist.
(iii) The supporting tool used during the experimental study to answer the questionnaires and activities could represent a threat since not all the participants were familiar with it. To reduce it, we explained all the necessary steps to the subjects.
(iv) Dropout subjects is a potential threat that was mitigated asking the recruited participants to find a day and time suitable for everyone to perform the experimental study.

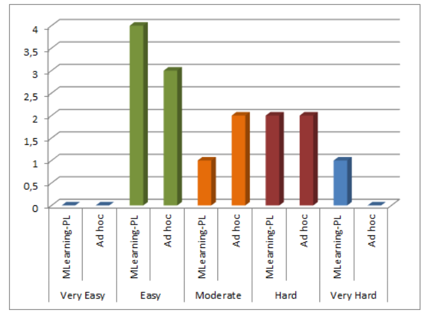
Construct validity is concerned with the relation between theory and observation. A potential threat is that the activities of both groups should have the same level of difficulty, which



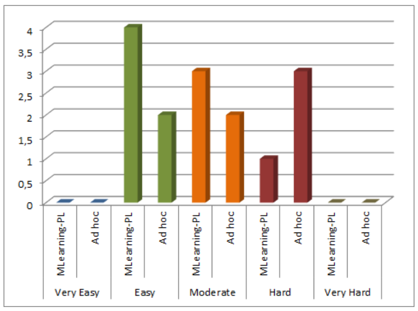
(a) Activity 1



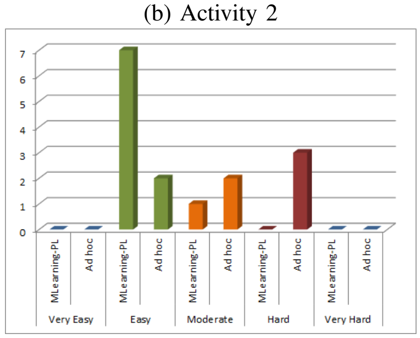
(b) Activity 2



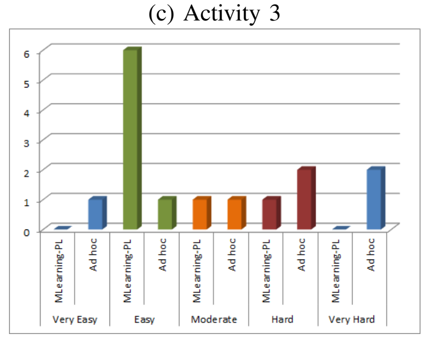
(c) Activity 3



(d) Activity 4

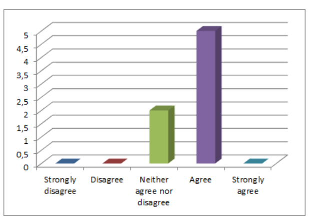


(e) Activity 5

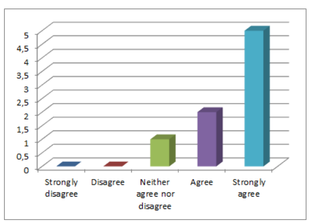


(f) Activity 6

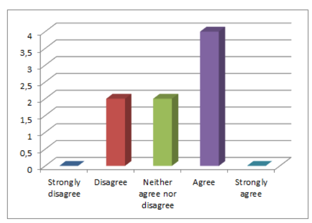
Fig. 6: How difficult was it for you to carry out the activities?



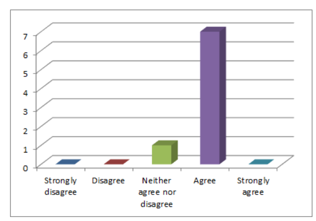
(a) Do you believe that the use of further artifacts would help in the problem-solving process?



(b) Do you believe that MLearning-PL helped in the process of solving pedagogical problems?



(c) Do you believe that MLearning-PL is complete?



(d) Do you believe that MLearning-PL is clear and easy to understand?

Fig. 7: Answers to the feedback questionnaire

was mitigated by the running of a pilot for refining the experimental study.

External validity is concerned with generalization. The sample may not be representative of the population, which was mitigated by the application of the *characterization form*.

VIII. Conclusions and Future Work

This paper presented an experimental study conducted for the evaluation of the pedagogical pattern language for mobile learning applications, entitled MLearning-PL. We analyzed the percentage of problems correctly solved by the subjects, the completeness and the complexity of each solution, and the time to solve all the problems. We compared the results of two approaches: MLearning-PL and ad hoc.

Although the results were not statistically significant, they were complemented by a qualitative analysis, which revealed MLearning-PL provides better results than the ad hoc approach. In general, users were enthusiastic and positive about MLearning-PL and its importance, which evidences its use over an ad hoc approach to support the pedagogical problem-solving process.

As future work, we point out the need of conducting further experimental studies involving more subjects and also considering a greater diversity of experimental subjects in relation to the area of expertise and experience in teaching. Furthermore, case studies or surveys might be conducted for a more detailed qualitative analysis.

Acknowledgements

The authors would like to thank the Brazilian funding agencies – FAPESP, CAPES (Procad 071/2013) and CNPq. The authors are also grateful to professors Paulo Cesar Masiero and Rosana Teresinha Vaccare Braga for their valuable contributions to this work and also to all those who helped by participating in the studies.

REFERENCES

[1] 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.

[2] 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.

[3] R. M. Felder, L. K. Silverman *et al.*, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988.

[4] C. Alexander, *A Pattern Language: Towns, Buildings, Construction*. New York: Oxford University Press, 1977.

[5] ——, *The Timeless Way of Building*. New York: Oxford University Press, 1979.

[6] J. O. Coplien, "The patterns handbooks," L. Rising, Ed. New York, NY, USA: Cambridge University Press, 1998, ch. Software Design Patterns: Common Questions and Answers, pp. 311–319.

[7] J. Bergin, J. Eckstein, M. Volter, M. Sipos, E. Wallingford, K. Marquardt, J. Chandler, H. Sharp, and M. L. Manns, *Pedagogical Patterns: Advice For Educators*. Joseph Bergin Software Tools, 2012.

[8] M. L. Fioravanti and E. F. Barbosa, "A Pedagogical Pattern Language for Mobile Learning Applications," in *Proceedings of 24th Conference on Pattern Languages of Programs (PLoP 2017)*, Vancouver, BC, Canada, oct 2017.

[9] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in software engineering*. Springer Science & Business Media, 2012.

[10] J. Bergin. (2007, Mar.) Fourteen Pedagogical Patterns. [Online]. Available: http://csis.pace.edu/~bergin/PedPat1.3.html

[11] T. Iba, T. Miyake, K. Shimonishi, T. Kato, Y. Kobayashi, N. Yotsumoto, M. Hanabusa, M. Iida, and M. Sakamoto, *Learning Patterns: A Pattern Language for Creative Learning*, 1st ed., ser. 1. CreativeShift Lab, may 2014.

[12] P. Avgeriou, A. Papasalouros, S. Retalis, and M. Skordalakis, "Towards a Pattern Language for Learning Management Systems," *Educational Technology & Society*, vol. 6, no. 2, pp. 11–24, Apr. 2003.

[13] J. Coutinho Anacleto, A. Talarico Neto, and V. P. d. A. Neris, "Cog-learn: An e-learning pattern language for web-based learning design," *eLearn*, vol. 2009, no. 8, Aug. 2009. [Online]. Available: http://doi.acm.org/10.1145/1595390.1595437

[14] R. E. Garcia, E. N. Höhn, E. F. Barbosa, and J. C. Maldonado, "An Ontology for Controlled Experiments on Software Engineering," in *Proceedings of the Twentieth International Conference on Software Engineering & Knowledge Engineering (SEKE'2008)*, San Francisco, CA, USA, Jul. 2008, pp. 685–690.

[15] V. R. Basili, "Applying the goal/question/metric paradigm in the experience factory," *Software Quality Assurance and Measurement: A Worldwide Perspective*, pp. 21–44, 1993.

[16] N. Juristo and A. M. Moreno, *Basics of software engineering experimentation*. Springer Science & Business Media, 2013.

[17] M. McCracken, V. Almstrum, D. Diaz, M. Guzdial, D. Hagan, Y. B.-D. Kolikant, C. Laxer, L. Thomas, I. Utting, and T. Wilusz, "A multi-national, multi-institutional study of assessment of programming skills of first-year cs students," *ACM SIGCSE Bulletin*, vol. 33, no. 4, pp. 125–180, 2001.

[18] S. Siegel and N. J. Castellan, *Nonparametric statistics for the behavioral sciences*, 2nd ed. New York : McGraw-Hill, 1988.

[19] M. Mendonça, D. Cruzes, J. Dias, and M. C. F. de Oliveira, "Using observational pilot studies to test and improve lab packages," in *Proceedings of the 2006 ACM/IEEE International Symposium on Empirical Software Engineering*, ser. ISESE '06. New York, NY, USA: ACM, 2006, pp. 48–57. [Online]. Available: http://doi.acm.org/10.1145/1159733.1159743