

Evaluating Usability of M-Learning Application in the Context of Higher Education Institute

Aijaz Ahmed Arain¹, Zahid Hussain¹(✉), Wajid H. Rizvi²,
and Muhammad Saleem Vighio¹

¹ Science and Technology, Quaid-E-Awam University of Engineering,
Science and Technology, Nawabshah, Pakistan

{aijaz, zhussain, saleem.vighio}@quest.edu.pk

² Institute of Business Administration, Karachi, Pakistan
wrizvi@iba.edu.pk

Abstract. In a system development process, usability evaluation is a crucial part of the process. The purpose of this paper is to test the usability of new developed application called “DARSGAH” for getting optimum outcomes from m-learning (Mobile Learning) technology. This application provides learning opportunities to University students. Initially, pilot test was conducted with 10 participants. After slight word refinement and minor changes in the application full scale usability test was conducted with other 100 participants who were selected randomly from the University by generating random numbers (student IDs) using SPSS. The data was collected using System Usability Scale (SUS) questionnaire immediately after the test, to obtain students’ overall perception about the application. The findings revealed that the application is user friendly, effective, and efficient and the users were satisfied as the overall average score of SUS is 84. The total task completion rate was 100 % which means that the application was very effective.

Keywords: M-learning · Usability testing · System Usability Scale (SUS) · Usefulness · DARSGAH · Mobile application

1 Introduction

1.1 M-learning

M-learning technology provides learning opportunities offered by mobile technologies when user is not at predetermined location. By ubiquitous technology user can take advantages to learn from anywhere at any time with the use of mobile devices [1].

The schools, colleges, universities and other educational institutions have been motivated to switch towards m-learning due to omnipresent access through wireless technologies to enhance learning and teaching methods [2].

GSM/2.5G technology had a limitation as lower speed caused slower adoption of the m-learning technology. But recently 3G & 4G technology has been introduced in

Pakistan so now there is a strong need to utilize the higher speed for adopting the m-learning technology.¹

The rest of the paper is divided into following sections: Subsects. 1.2 and 1.3 describe usability and SUS, respectively. Section 2 is about related work while Sect. 3 provides details about the application. Methodology has been described in Sect. 4 while results and statistical analysis are given in Sect. 5 and finally in Sect. 6 conclusion is given.

1.2 Usability

Usability is a degree of ease to use software technology and having user friendly interface (UI). Common qualities regarding usability that define the degree of ease are labeled as ease of use, efficiency, effectiveness, learnability, and user satisfaction.

For developers, the purpose of usability testing is to point out the interface characteristics of the product, minimize the cost of development and support, and raise its market attractiveness [3].

There is empirical evidence to suggest that usability testing is very compulsory to eliminate usability problems [4–6].

1.3 System Usability Scale

The measurement of the usability is very complex; however, a simple, quick and dirty usability scale has been developed by Brooke [7]. There is a suggestion that the SUS survey questionnaire should be conducted immediately after the usability test. The SUS scale consists of 10 usability items; these items are used to assess the ease of use, efficiency, learnability, and satisfaction of an existing system. Each item of SUS scale consists of 5 point Likert Scale. All odd items are phrased positively and to obtain the result 1 has to be subtracted from participant's response and all even items are negatively phrased questions where to obtain the result participant's response has to be subtracted from 5, which scales each item from 0 to 4. The total score is multiplied by 2.5 to provide a score out of 100, which is interpreted as a percentile ranking and not as a percentage.

Tullis and Stetson state that 68 SUS average score is threshold. If the score is above 68, it will be considered more than average and if the score is below 68, then score is considered less than average [8].

2 Related Work

Fetaji et al. have introduced “Mobile Learning Usability Attribute Testing” (MLUAT) which is the usability evaluation method and they performed the comparison of MLUAT with two other usability evaluation methods. The goal was to test e-learning

¹ www.pta.gov.pk/annual-reports/ptaannrep2013-14.pdf (Last accessed on: 20-11-2015).

and m-learning. The results of comparison showed that MLUAT is more efficient usability evaluation method [9].

Alelaiwi and Hossain have conducted a practical usability evaluation study on specific e-learning tools. They formed two groups of real users, one group has HCI knowledge and other without HCI knowledge. Usability evaluation questionnaires were distributed to both groups. They concluded that the group without HCI knowledge was more satisfied than the group with HCI knowledge because HCI knowledge group has more expectations with e-learning tool [10].

Two alternative approaches were adopted by Summers and Watt are: 1. Mobile applications are created and tested by students; and 2. Existing products are modified by students using instructions/documentation rather than creating a new one from the scratch. The Results were successfully achieved through both approaches and it was recommended that both approaches could be easily adopted and adapted [11].

Lee and Salman have used Mobile Collaborative Learning (MCL) which is a mobile learning application and it has more importance in educational environment. They have presented both theoretical and technical fundamentals for designing and developing an effective Mobile Collaborative Learning environment. They have also introduced a new approach for building a mobile learning application. Finally, using Android operating system they have designed and constructed a prototype with a suggestive infrastructure for MCL [12].

The impact of a screen size of mobile phones on users' efficiency, effectiveness and ease of use was measured by Raptis et al., using System Usability Scale. They conducted an experiment using same brand (Samsung) with three different screen sizes. Sixty participants were involved in the experiment [13].

Baillie and Morton have tested two mobile applications; both applications are designed on same concept, one application is designed with HCI other is designed without HCI principles. Both applications were tested with users in a field to see which was the simplest and most intuitive to use. Its usability was measured by using System Usability Scale [14].

3 Introduction to DARSGAH Application

The name of the m-learning application developed is "DARSGAH" which means "Place of Learning" in English. DARSGAH is a web based application and it is accessible with native Android based mobile app. The core purpose for developing the DARSGAH application is to test if the learning outcomes of (Higher Education Institute) students can be enhanced. The application has been developed at Quaid-e-Awam University, Nawabshah, Pakistan. The application has five features: Lecture wise videos, Lecture wise notes, Lecture wise MCQ (Multiple Choice Questions) tests, Chat room for group chatting and Forum for discussion about videos, Notes and MCQ tests.

Figure 1 shows three screens, first screen (a) is a login screen for authentication, second screen (b) is a video lecture screen with Like, Rating and Total viewed options. Users can also make notes about video and save permanently in their account for future reference. Users can write comments about current video for other users. Third screen (c) shows complete lecture wise notes.

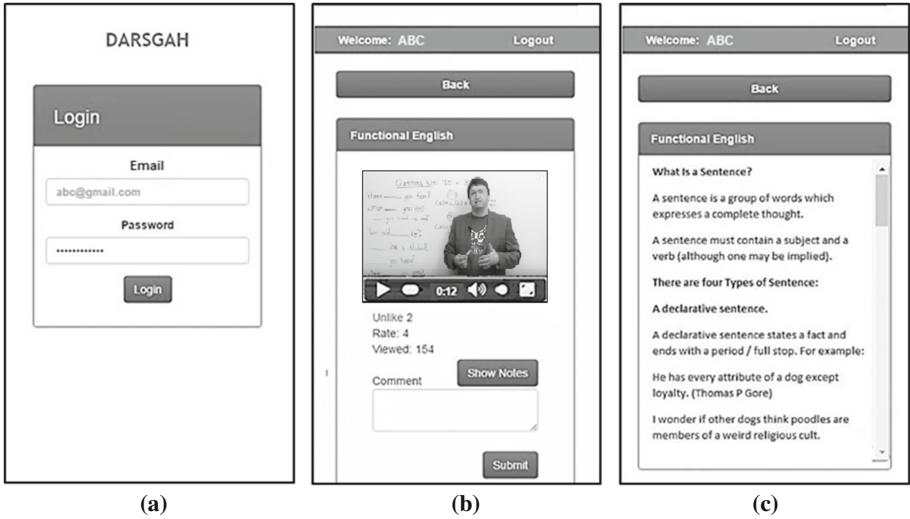


Fig. 1. Screenshots of (a) Login screen (b) Video lecture (c) Lecture wise notes

The application also has lecture wise MCQ tests for practice and students can rate each question and they can write comments as well. The application also has chat room for group chatting of online users. All chatting messages will be stored in database and can be viewed by any student for a reference. Another feature of the application is a forum, where Students can put questions and can also give answers for any question.

4 Methodology

Before conducting usability test of DARS GAH the m-learning application, pre and post-test questionnaires were developed. After the pilot study by ten representative users, minor changes were incorporated in the application and minor word refinement was also carried out in the questionnaires. In addition to that System Usability Scale questionnaire was used to get user perception on the m-learning application. In the full scale test, hundred (n = 100) participants from the university students were selected randomly having smartphones where 70 % are male and 30 % are female participants who took part in the usability test. All the participants were university students having smartphone usage experience and most of them had experience of internet use on smartphone.

Among several usability evaluation methods, formal experiment was used due to its wide spread use. Before performing usability tasks, all participants filled pre-test questionnaire for collecting demographic information. During usability test all the participants performed 5 given usability tasks one by one in the office of the university lab on same smartphone and on same internet bandwidth. The model of smartphone was Samsung Galaxy Grand Prime. The time duration and the tasks completion rate were recorded for each participant. After performing the usability tasks, the test participants

filled the post-test questionnaire to collect their opinion about the DARSGAH m-learning application and System Usability Scale questionnaire has also been filled by the same participants for general perception about the DARSGAH application.

5 Results

The results are presented below in the subsections.

5.1 Demographic Data

Figure 2(a) shows that all the participants are graduate studying in different departments and there is no participant studying in post-graduate studies. Figure 2(b) shows three age groups from range 16–20 years having 65 % participants, from range 21–25 years having 35 % participants and from range 26 years and above having no participant at all. Figure 2(c) shows male and female ratio of the participants.

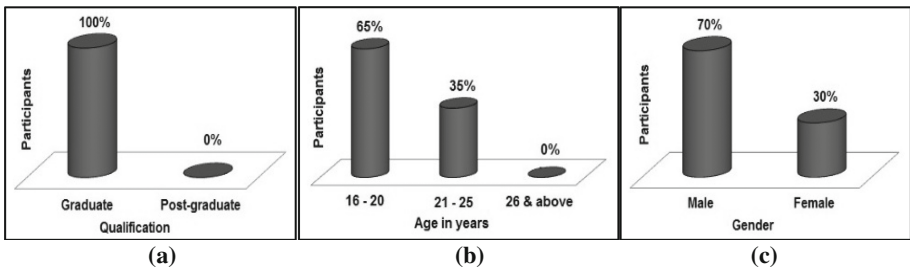


Fig. 2. (a) Qualification (b) Age in years (c) Gender of participants

Figure 3(a) shows 100 % participants using smartphone where 73 % participants using smartphone from 1 to 5 years, and 27 % from 6 to 10 years. Similarly Fig. 3(b) shows 51 % test participants using internet on daily bases, 29 % weekly, 4 % monthly, 14 % occasionally and only 2 % participants never used the internet on their smartphones.

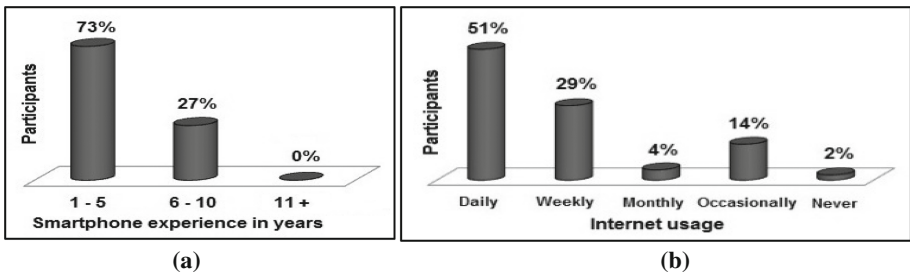


Fig. 3. (a) Smartphone experience in years and (b) Internet usage frequency

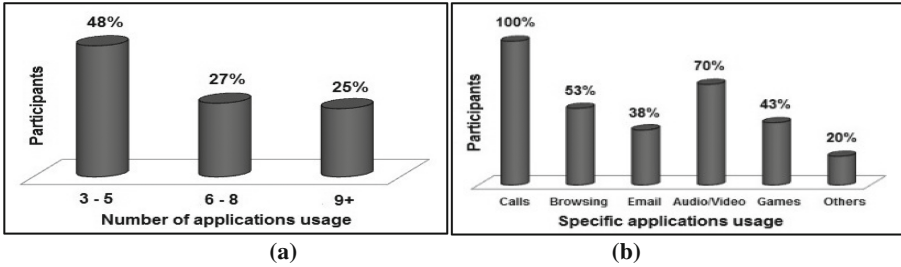


Fig. 4. (a) Number of apps usage frequency and (b) Specific apps usage frequency

Figure 4(a) shows three groups of users (1) 48 % users are using 3 to 5 applications on daily basis, (2) 27 % users are using 6 to 8 applications and (3) 25 % users are using 9 or more than 9 applications daily. Figure 4(b) shows 100 % participants use their smartphones for calls, 53 % for browsing, 38 % for emails, 70 % for Audio/Video (Entertainment), 43 % for games and 20 % participants used other applications on their smartphones.

5.2 Usability Test

In the usability test there were total five tasks to be performed by the participants. Effectiveness was measured by the task completion rate. The total task completion rate was 100 % which means that the application was so easy and effective that all the participants completed all the five given tasks.

Efficiency was measured by task completion time, Fig. 5(a) shows male and female participants’ performance time in seconds and Fig. 5(b) shows the overall result of the tasks completion time with 5 given usability tasks. The performance ratio of male and female participants is approximately equal. Task No. 2 has depth level 1 that’s why the participants took less time to complete it while Task No. 1, 3 & 5 have same depth (depth level 2) that’s why the participants took almost the same time for completion. Task No. 4 has depth level 3 that’s why the participants took more time than other tasks which shows that overall the application is efficient.

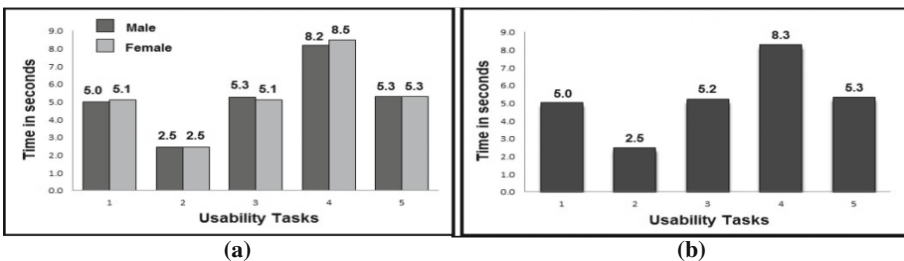


Fig. 5. (a) Average time of tasks completion by male & female (b) Total average time of tasks completion

5.3 System Usability Scale

System Usability Scale (SUS) is most widely useable scale in HCI world.

Figure 6(a) shows average score of male which is 85, while for female is 83 and mean is 84 which is more than average SUS score (68). Figure 6(b) shows 100 participants' score graph. It means m-learning application is very easy to use for new users. This figure also shows participants may use this m-learning application efficiently and almost all participants are satisfied.

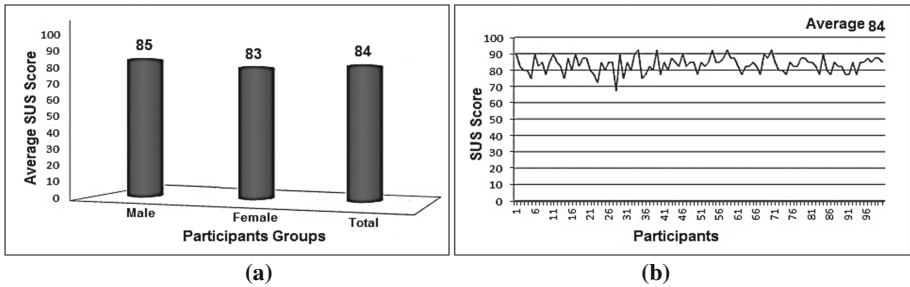


Fig. 6. (a) Average SUS score of male, female & total (b) SUS score of 100 participants

Table 1 shows that the 91.7 % participants' given response is 3 or 4 score (4 is the best) that means the application is user friendly and efficient. Only 3 % participants' response is towards 0 or 1, and 5.3 % participants' response are neutral.

Table 1. Ten SUS items' response from score 0 to 4

SUS score from 0-4	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Participants' response in %
0	1	0	0	1	0	1	0	0	0	0	0.3 %
1	0	4	0	10	0	9	0	1	2	1	2.7 %
2	0	12	0	6	7	14	2	3	1	8	5.3 %
3	51	50	34	35	69	52	32	32	53	53	46.1 %
4	48	34	66	48	24	24	66	64	44	38	45.6 %

Results of Q1 & Q9 of SUS questionnaire show that almost all the participants are satisfied. The result of Q5 shows that the application's design is good and the application is well integrated. Q6 shows that the application is consistent; Q4 & Q10 show application's learnability which is good and Q2, 3 & 8 show that the application is very user friendly.

5.4 Statistical Analysis

SUS Score. The results were extracted using SPSS v.20. Table 2 shows group differences on SUS scores.

Table 2. Group differences

Group	Category	Mean and SD
Smartphone use	1 to 5 years	(M = 83.34, SD = 4.24)
	6 to 10 years	(M = 83.42, SD = 6.01)
Number of apps use	3 to 5 Apps	(M = 83.96, SD = 6.96)
	6 to 8 Apps	(M = 82.67, SD = 5.39)
	9 + Apps	(M = 83.96, SD = 6.96)

Table 3 shows effects of smartphone use and number of Apps use on SUS scores.

Table 3. Independent T-test and one-way ANOVA

Independent T-test	Smartphone use years	t(90) = -0.07	p = 0.944
One-way ANOVA	Number of Apps use	F (2,97) = 0.610	p = 0.545

The result shows that there was no effect of smartphone use in years and number of apps used on a smartphone on perceived usability. There was no statistical significant difference across the groups on SUS scores. Since SUS mean score across the groups was around 84 suggesting that the perceived usability is high across the group.

Task Completion Time. Table 4 shows mean and standard deviation of completion time of each task.

Table 4. Descriptive statistics

Task	Mean	Std. deviation
1. Play video of lecture no. 2	5.0300	1.00960
2. Enter into the chat Room	2.4800	0.62732
3. Open Notes of lecture no. 3	5.2100	1.02784
4. Add a New question into the forum	8.2900	1.42343
5. Take test of lecture no. 4	5.3300	1.03529

Table 5 shows effect of smartphone use in years and number of Apps used on total task completion time.

Table 5. Independent T-test and one-way ANOVA

Independent T-test	Smartphone use years	t(98) = -0.1.8	p = 0.06
One-way ANOVA	Number of apps use	F (2,97) = 1.009	p = 0.368

The results are consistent to the SUS scores; there was no statistically significant difference across the groups on total task completion time as well.

6 Conclusions

This study investigated the DARSGAH application for testing the efficiency, effectiveness and user satisfaction with two usability evaluation methods (a) Formal experiment (b) System Usability Scale questionnaire. Both the methods show that the application is very user friendly, efficient and effective and can be utilized for the m-learning purpose. The total task completion rate for all the participants was 100 % which means that the application was very effective. The application was equally user friendly and efficient for the across the groups (smartphone use in years/number of Apps use) on both SUS score and total task completion time.

References

1. Ko, E.H., Chiu, D.K., Lo, P., Ho, K.K.: Comparative study on m-learning usage among LIS students from Hong Kong, Japan and Taiwan. *J. Acad. Librariansh.* **41**(5), 567–577 (2015)
2. Althunibat, A.: Determining the factors influencing students' intention to use m-learning in Jordan higher education. *Comput. Hum. Behav.* **52**, 65–71 (2015)
3. Gong Chao.: (2009) Human-computer interaction: the usability test methods and design principles in the human-computer interface design. In: 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009, pp. 283–285, 8–11 August 2009
4. Bandi, A., Heeler, P.: Usability testing: a software engineering perspective. In: 2013 International Conference on Human Computer Interactions (ICHCI), pp. 1–8. IEEE (2013)
5. Cheng, J., Zhu, Y., Zhang, T., Zhu, C., Zhou, W.: Mobile compatibility testing using multi-objective genetic algorithm. In: 2015 IEEE Symposium on Service-Oriented System Engineering (SOSE), pp. 302–307. IEEE (2015)
6. Zhang, T., Gao, J., Cheng, J., Uehara, T.: Compatibility testing service for mobile applications. In: IEEE Symposium on Service-Oriented System Engineering, pp. 179–186 (2015)
7. Brooke, J.: SUS-A quick and dirty usability scale. *Usability Eval. Ind.* **189**(194), 4–7 (1996)
8. Tullis, T., Stetson, J.: A comparison of questionnaires for assessing website usability. In: Proceedings of the Usability Professionals Association (UPA) 2004 Conference, pp. 7–11 (2004)
9. Fetaji, B., Ebibi, M., Fetaji, M.: Assessing effectiveness in mobile learning by devising MLUAT (Mobile Learning Usability Attribute Testing) methodology. *Int. J. Comput. Commun.* **5**(3), 178–187 (2011)
10. Alelaiwi, A., Hossain, M.S.: Evaluating and testing user interfaces for e-learning system: blackboard usability testing. *J. Inf. Eng. Appl.* **5**(1), 23–30 (2015)
11. Summers, S., Watt, A.: Quick and dirty usability testing in the technical communication classroom. In: Professional Communication Conference (IPCC), 2015 IEEE International, pp. 1–4. IEEE (2015)
12. Lee, K.B., Salman, R.: The design and development of mobile collaborative learning application using android. *J. Inf. Technol. Appl. Educ.* **1**(1), 1–8 (2012)

13. Raptis, D., Tselios, N., Kjeldskov, J., Skov, M.B.: Does size matter?: investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. In: Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 127–136. ACM (2013)
14. Baillie, L., Morton, L.: Designing quick & dirty applications for mobiles: making the case for the utility of HCI principles. In: Proceedings of the ITI 2009 31st International Conference on Information Technology Interfaces, ITI2009, pp. 293–298. IEEE (2009)